Genome-Wide Studies on the Molecular Functions of Pax7 in Adult Muscle Satellite Cells

Vincent Punch

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

Abstract

Pax3 and Pax7 belong to a family of conserved transcription factors that play important and diverse roles in development. In the embryo, they carry out similar roles in neural and somite development, but Pax7 fails to compensate for critical functions of Pax3 in the development of limb musculature. Conversely, in the adult, Pax7 is necessary for the maintenance and survival of muscle satellite cells, whereas Pax3 cannot effectively fulfill these roles in the absence of Pax7.

To identify the unique roles of Pax7 in adult muscle cells, we have analyzed global binding of Pax3 and Pax7 by ChIP-Seq. Here, we show that despite highly homologous DNA-binding domains, the majority of binding sites are uniquely recognized by Pax7 and are enriched for homeobox motifs. Genes proximal to conserved, unique Pax7 binding sites cluster into specific functional groups which may reflect the unique biological roles of Pax7. Combining Pax7 binding sites with gene expression data, we describe the regulatory networks directed by Pax7 and show that Pax7 binding is associated with positive gene regulation. Moreover, we show *Myf5* is a direct target of Pax7 and identify a novel binding site in the satellite cell control region upstream of *Myf5*.

Acknowledgements

I would like to thank my fellow lab mates for providing an enjoyable atmosphere and constructive and creative environment. You have each contributed to this project in your own way, by stimulating thought provoking discussions, providing technical assistance, or merely social support. I especially thank Dr. Michael Rudnicki for allowing me the opportunity to undertake this research project, for providing scientific guidance and a superb environment in which to carry out this work. I would like to thank Theodore Perkins, Gareth Palidwor, Chris Porter, and Dr. Hang Yin for useful discussions and brainstorming sessions on how to best analyze our vast amount of data. I thank my advisory committee, Dr. Bernard Jasmin, Dr. David Picketts, and Dr. Luc Sabourin, for their guidance along the way. Finally, I thank my parents for their love and support over the years. You encouraged me to keep going when I wanted to give up and helped me every step of the way.

Contributions of Collaborators and Co-Authors

All technical work and preparations of the thesis and manuscripts were carried out by the thesis author except as otherwise specified. Solexa sequencing of ChIP products was carried out in full by the technical staff of Dr. Frank Grosveld at Erasmus University and Medical Center in Rotterdam, Netherlands. Gareth Palidwor and Chris Porter carried out bioinformatics analyses on ChIP-Seq data and provided the corresponding figures used throughout Chapter 2. Dr. Yoichi Kawabe provided RNA samples for the microarray analysis and all baculovirus generated protein used in these studies, in addition to contributing western blot data for Chapter 3. Dr. Vahab Soleimani provided technical assistance with sample preparation for ChIP-Seq experiments. *In vivo* analyses of the B195APZ BACs were carried out by Dr. Jaime Carvajal of Dr. Peter Rigby's group at the Division of Gene Function and Regulation at the Institute of Cancer Research in London, UK. The author carried out experiments and provided data for Figures 4b, 4f, 4g, and 4h in Appendix A.

Contents

List of Tables

List of Figures

List of Abbreviations

BAC	bacterial artificial chromosome
bFGF	basic fibroblast growth factor
bHLH	basic helix-loop-helix
BMP	bone morphogenic protein
BSA	bovine serum albumin
ChIP	chromatin immunoprecipitation
ChIP-Seq	high throughput ChIP sequencing
DMEM	Dulbecco's modified Eagle's medium
DML	dorsomedial lip
DNA	deoxyribonucleic acid
dpc	days post coitum
ECR	evolutionary conserved region
EDL	extensor digitum longus
eGFP	enhanced green-fluorescent protein
EMSA	electrophoreitic mobility shift assay
EMT	epithelial-mesenchymal transition
ER	estrogen receptor
FBS	fetal bovine serum
FDR	false discovery rate
FKHR	FoxO1
GAPDH	glyceraldyde-3-phosphate dehydrogenase

GO	gene ontology
H3K4	histone 3 lysine 4
hbox	homeobox
HGF	hepatocyte growth factor
HMT	histone methyltransferase
HPLC	high performance liquid chromatography
MACS	model analysis of ChIP-Seq
MAPK	mitogen-activated protein kinase
mdx	Duchenne/Becker muscular dystrophy (*dystrophin* mutant)
MEF	myocyte enhancer factor
MRF	myogenic regulatory factor
NO/NOS	nitric oxide/nitric oxide synthetase
PBS	phosphate-buffered saline
PCP	planar cell polarity
PCR	polymerase chain reaction (quantitive, qPCR)
PFA	paraformaldehyde
prd/pbox	paired box
puro	puromycin
RDA	representational difference analysis
RIPA	radioimmunoprecipitation assay
RMS	rhabdomyosarcoma (alveolar, ARMS; embryonal, ERMS)
RNA	ribonucleic acid
RT	reverse transcription
SAM	significance analysis of microarray

SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
Shh	sonic hedgehog
SP	side population
Sp	*Splotch* (*Pax3* mutant)
TA	tibialis anterior
TAP	tandem affinity purification
TCID	transcipt cluster identifier
TEV	tobacco etch virus
TF	transcription factor
TSS	transcription start site
UTR	untranslated region
VLL	ventrolateral lip
WS	Waardenburg syndrome
YFP	yellow fluorescent protein

Chapter 1 – General Introduction

Introduction

1.1. An Overview of Skeletal Muscle

Skeletal muscle, as its name implies, is a striated muscle that, attached to bones, is responsible for skeletal movements. Skeletal muscle makes up over half of the body's mass and is controlled by the somatic nervous system. Attachments to bones are mediated by collagen fibers called tendons. Individual muscles are comprised of bundles of muscle fibers, long cylindrical cells generated from the fusion of many individual muscle progenitor cells, and as such, contain multiple nuclei. Fibers are structured repeats of basic functional units called sarcomeres, composed of actin and myosin fibrils necessary for muscle contraction. Contractions are thought to operate according to the sliding filament model, whereby "thick" filaments of myosin slide over "thin" filaments of actin in a ratchet-like motion to alter the length of sarcomeres. This process is stimulated by the arrival of neural impulses at neuromuscular junctions, which triggers a release of Ca^{2+} from the sarcoplasmic reticulum. Ca^{2+} induces an allosteric change in troponin proteins, which unblock myosin binding sites on actin filaments.

Skeletal muscle is remarkable in its ability to regenerate, even following extensive damage. Complex cascades of cellular responses can repair or generate new muscle tissue that is properly vascularized, innervated, and most importantly, contractile. Skeletal muscle makes up nearly half of the body's mass, therefore understanding the molecular and biological mechanisms responsible for its growth and repair, especially in the context of injury or disease, is of supreme importance to biologists. Furthermore, skeletal muscle is an ideal model for the study of many cellular and molecular events. Muscle progenitor cells are well-studied in their molecular programs regulating proliferation, commitment, and

differentiation, both in the context of embryonic development and adult regeneration. The ease by which these programs can be extrinsically controlled makes them ideal for the study of cellular differentiation processes, as well as transcriptional regulatory networks from a systems biology perspective.

1.2. Stem Cells Regulate the Growth and Regeneration of Muscle

Our understanding of skeletal myogenesis is built on a thorough understanding of the mechanisms regulating the proliferation, commitment, and differentiation of muscle stem cells. During development, skeletal muscle progenitors are derived from the somites, transient structures of mesodermal cells that give rise to muscle, dermis, and vertebrae. Precursor cells, upon activating the myogenic program, give rise to proliferating myoblasts that express muscle-specific genes. These cells ultimately withdraw from the cell cycle and fuse with each other to form multinucleated myotubes that mature to become muscle fibers.

Limb muscles are formed following successive waves of migration of multipotent stem cells from the somite into the developing limb buds. The development of trunk and limb musculature is well understood in comparison to the development of craniofacial muscles. Somitomeres, developed from cephalic paraxial mesoderm, in combination with occipital somites, and prechordal mesoderm, collectively form the voluntary craniofacial muscles, and contribute to the formation of neck and jaw muscles (Asakura and Rudnicki, 2002; Noden, 1991; Trainor et al., 1994).

In the adult, the nature of stem cells that participate in skeletal muscle regeneration is somewhat less clear. Satellite cells, a heterogeneous population of stem cells and committed progenitors, reside underneath the basal lamina of mature muscle fibers and are responsible for postnatal muscle growth and regeneration. In their resting state, they are mitotically

quiescent and comprise only a fraction of all myonuclei. However, they re-enter the cell cycle in response to muscle injury and proliferate as myoblasts that fuse to repair damaged tissue.

Both muscle and non-muscle resident cell populations have also emerged as effectors of muscle regeneration, including endothelial cells associated with embryonic limb muscles (Le Grand et al., 2004), mesangioblasts (De Angelis et al., 1999; Esner et al., 2006; Minasi et al., 2002), stem cells derived from bone marrow (Bittner et al., 1999; Ferrari et al., 1998), and pluripotent cells found within muscle-derived side population (SP) cells (Asakura et al., 2002; Gussoni et al., 1999).

An underlying theme common among myogenic stem cells is the activation of a "master switch" that induces commitment and terminal differentiation along the myogenic lineage. The diversity among muscle stem cell populations demonstrates that the regulation of myogenesis is under the precise control of a number of signaling and transcription networks that translate a variety of extrinsic signals from different environments into myogenic activation.

1.3. Molecular Mechanisms of Myogenic Determination

Early studies in the field of skeletal myogenesis (Devlin and Emerson, 1978; Konieczny and Emerson, 1984; Konigsberg, 1963; Yaffe, 1968) led to the discovery of the myogenic regulatory factors (MRFs) – MyoD, Myf5, myogenin, and MRF4 (Myf6) (Braun et al., 1990a; Braun et al., 1989; Davis et al., 1987; Edmondson and Olson, 1989; Rhodes and Konieczny, 1989; Wright et al., 1989). The MRFs form a family of transcription factors that function as nodal points for the flow of myogenic information (Weintraub et al., 1991). Activation of the MRF pathway drives commitment and initiates differentiation along the

myogenic lineage, as is evident in its remarkable ability to convert a variety of nonmuscle cell types, such as 10T½ fibroblasts, into muscle (Choi et al., 1990; Davis et al., 1987). Each member of the MRF family is capable of activating the other members and maintaining its own expression through robust auto- and cross-regulatory networks (Braun et al., 1989; Edmondson and Olson, 1989; Thayer et al., 1989). Collectively, these pathways drive high levels of MRF expression upon activation, in a fail-safe manner that locks the cell into myogenic commitment. The MRFs share a homologous basic helix-loop-helix (bHLH) domain responsible for DNA binding and dimerization with members of the E-protein family, including E12, E47, and HEB (Hu et al., 1992; Murre et al., 1989; Parker et al., 2006). MRF-E-protein dimers bind core CANNTG ("E-box") elements found in the promoters of key muscle-specific genes (Rudnicki and Jaenisch, 1995). Despite their similarities, there is limited functional redundancy owing in part due to unique sequences in the N- and C-termini (Braun et al., 1990b; Weintraub et al., 1991).

Myogenesis is further regulated by the myocyte enhancer factor 2 (MEF2) family of MADS-box transcription factors, which act as MRF cofactors to synergistically activate muscle-specific genes (Molkentin and Olson, 1996). MEFs are expressed in many different tissues, but only activate gene transcription in muscle cell types (Naya et al., 1999). MEF2 proteins stimulate gene expression by binding AT-rich domains within the promoters of muscle specific genes. *Myogenin* and *MRF4* both require MEF binding in addition to MyoD or Myf5 for transcriptional activation (Black et al., 1995; Buchberger et al., 1994).

1.4. Mechanisms of Embryonic Myogenesis

During development, somites form from segmented paraxial mesoderm in an anterior-posterior fashion. Their differentiation, occuring along the dorsal-ventral axis, is

stimulated by signals emanating from the neural tube and notochord, giving rise to a dorsally-located epithelial layer called the dermomyotome around 8.0 days post coitum (dpc) and a ventrally-located mesenchymal layer called the sclerotome (reviewed in (Parker et al., 2003a)). Scleretomal cells migrate around the neural tube and extend laterally to form the bone and cartilage that comprise the vertebrae and ribs. The dermomyotome, however, maintains a proliferating epithelial structure that develops to form the dermis and skeletal muscle of the limbs and trunk (Kaufman, 1992).

The dermomyotome is further subdivided into the dorsomedial epaxial and ventrolateral hypaxial components. By 8.5 dpc, cells from the dorsomedial lip (DML) cease proliferation and migrate to the underside of the dermomyotome forming the first differentiated myocytes in the epaxial myotome (Williams and Ordahl, 1997, Christ and Ordahl, 1995). This process is mirrored at the ventrolateral lip (VLL) at 9.5 dpc, as cells migrate and differentiate to form the hypaxial myotome (Kaufman, 1992, Asakura, 2002). The epaxial myotome eventually gives rise to deep back muscles, while the hypaxial myotome forms the trunk muscles. At the hypaxial lip of the dermomyotome, some cells undergo an epithelial-mesenchymal transition, delaminate, and migrate to sites of muscle development (Asakura and Rudnicki, 2002; Buckingham et al., 2003). Successive waves of migration and differentiation comprise the basis for the formation of limb muscles, diaphragm, ventral trunk muscles, and tongue (Figure 1).

1.5. Molecular Networks Regulating Embryonic Myogenesis

Delamination and migration of skeletal progenitors from the hypaxial dermomyotome is dependent on the expression of c-met, a receptor tyrosine kinase and its ligand, HGF/scatter factor (Figure 1, panel 1). HGF signaling induces epithelial-to-mesenchymal

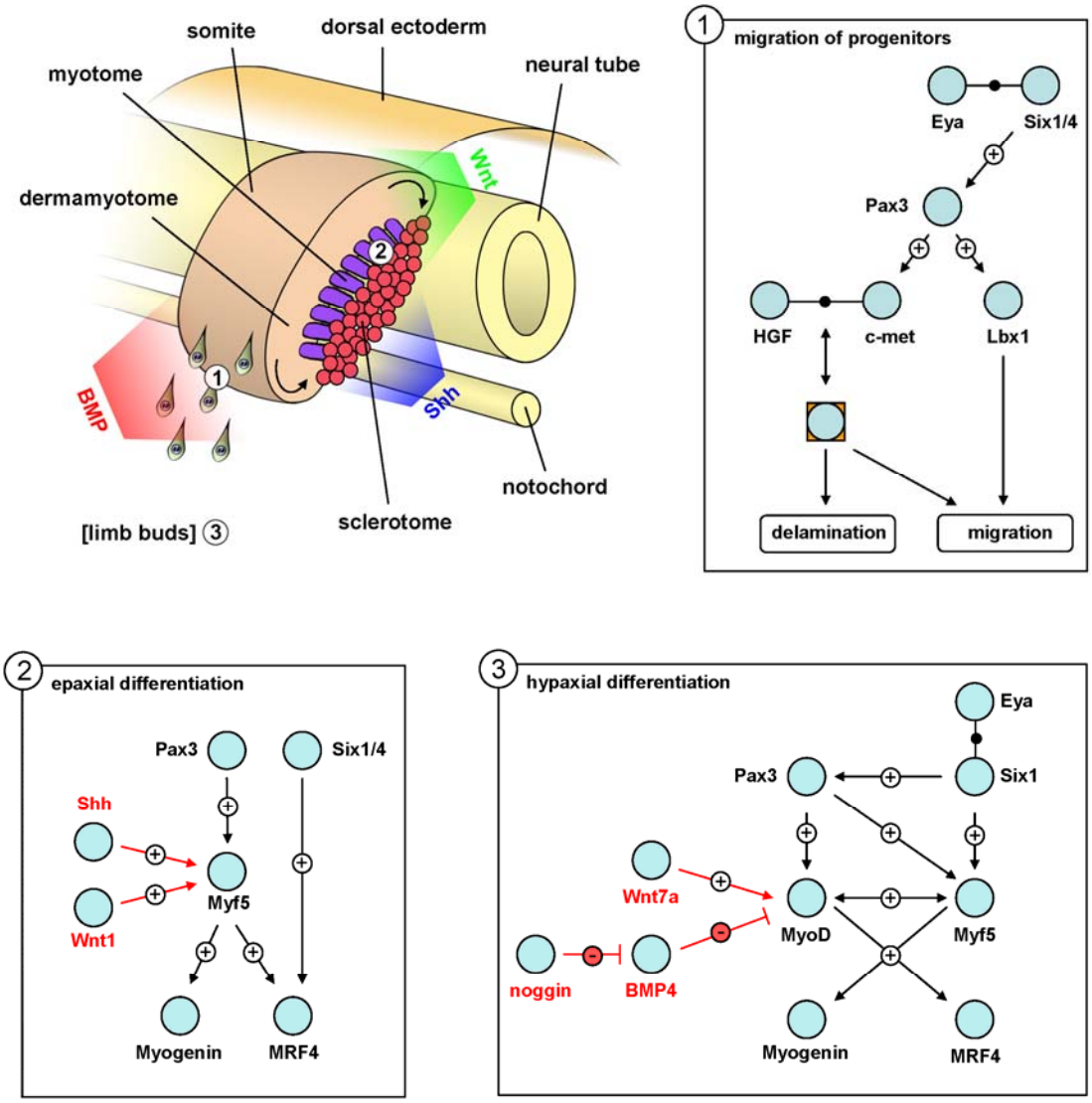

somite
dorsal ectoderm
myotome
neural tube
dermamyotome
Wnt
2
1
BMP
Shh
notochord
sclerotome
[limb buds] ③
1
migration of progenitors
Eya
Six1/4
Pax3
HGF
c-met
Lbx1
delamination
migration
2
epaxial differentiation
Pax3
Six1/4
Shh
Myf5
Wnt1
Myogenin
MRF4
3
hypaxial differentiation
Eya
Pax3
Six1
Wnt7a
MyoD
Myf5
noggin
BMP4
Myogenin
MRF4

Figure 1. Transcriptional Networks of Embryonic Myogenesis

Somites on either side of the neural tube give rise to muscle progenitor cells. Cells from the dorsomedial and ventrolateral regions of the dermomyotome migrate underneath the dermomyotome to form the epaxial and hypaxial components of the myotome, respectively. Hypaxial cells from the ventrolateral lip migrate to sites of future muscle development, forming the trunk and limb muscles, while epaxial cells ultimately form deep back muscles.
(1) Survival of myogenic cells within the hypaxial dermomyotome requires Pax3, whose expression is dependent on Six1 or Six4. Pax3, in turn, regulates the expression of c-met and Lbx1. Delamination and migration of myogenic cells from the hypaxial dermomyotome to limb buds is dependent on the expression of c-met, a tyrosine kinase receptor, and its ligand, HGF. Lbx1 is required for proper movement along migratory routes.
(2) Within the myotome, Pax3 can activate the expression of Myf5. Differentiation proceeds in the absence of MyoD expression, however, MRF4 expression is dependent on Six1 or Six4. Myf5 expression is stimulated by Sonic hedgehog (Shh) signals produced by the notochord and floor plate and Wnt1 signals produced by the neural tube.
(3) In limb buds, Six1 can activate Pax3 and Myf5, and Pax3 can drive the expression of MyoD and Myf5. The expression of either Myf5 or MyoD is sufficient to initiate the myogenesis. MyoD expression is tightly regulated by Wnt and BMP pathways. Wnt7a, produced by the dorsal ectoderm, activates MyoD in hypaxial cells, while BMP4, produced by the lateral plate mesoderm, inhibits the expression of MyoD. Noggin expression at the dorsal-medial lip of the hypaxial dermamyotome blocks BMP signaling facilitating the fine control of MyoD expression in specific domains.

conversion of cells in the dermomyotome by inducing Snail family transcription factors and PI3K/Akt signaling (Chang et al., 2011; Leroy and Mostov, 2007). HGF is also produced by mesodermal cells along migratory routes and at target sites for myogenic progenitors (Dietrich et al., 1999). The transcription of *c-met* is directly regulated by Pax3 (Epstein et al., 1996). Consequently, Pax3 mutant (*Splotch*) embryos are deficient for c-met expression and form no limb musculature due to myogenic progenitors failing to populate developing limb buds (Bober et al., 1994; Daston et al., 1996; Goulding et al., 1994). Additionally, Pax3 is required for the expression of Lbx1, a homeobox-containing gene expressed in migrating progenitors. Progenitor cells in Lbx1 mutant mice delaminate appropriately, but display aberrant migration patterns and fail to reach the limb buds, suggesting Lbx1 functions in interpreting signals along migratory routes (Brohmann et al., 2000; Gross et al., 2000; Schafer and Braun, 1999) and induction of EMT (Yu et al., 2009).

In the somite, Myf5 expression precedes myogenic differentiation in the dermomyotome. Epaxial cells located within the myotome activate myogenin and MRF4 following Myf5 expression, and differentiation occurs in the absence of MyoD (Bober et al., 1991; Ott et al., 1991; Sassoon et al., 1989) (Figure 1, panel 2). Migrating cells, on the other hand, do not express any myogenic factors until they have reached the limb (Tajbakhsh and Buckingham, 1994). In progenitor cells, Myf5 and MyoD are the first expressed MRFs and are responsible for establishing commitment to the myogenic lineage. As they share partial functional redundancy, either Myf5 or MyoD is required for the normal development of skeletal muscle. The loss of both genes results in a total ablation of skeletal muscle progenitors and total loss of skeletal muscle. In this scenario, somitic cells normally destined to become myogenic adopt other cell fates (Kablar et al., 1999; Rudnicki et al., 1993; Tajbakhsh et al., 1997). In *MyoD*$^{-/-}$ mice, myogenic cells compensate by upregulating Myf5

(Rudnicki et al., 1992), however differentiation is somewhat delayed, especially in branchial arches and limb muscles, suggesting that Myf5 is initially insufficient for the progression of myogenesis (Kablar et al., 1997). Conversely, *Myf5*$^{-/-}$ mice are delayed in myotome formation until the activation of MyoD, but are otherwise normal (Braun et al., 1994; Braun et al., 1992; Tajbakhsh et al., 1997).

Myogenin and MRF4 are required for myoblast fusion and terminal differentiation events (Rudnicki and Jaenisch, 1995). *Myogenin*$^{-/-}$ mice contain normal numbers of myoblasts and express MyoD, however are arrested in their myogenic progression prior to fusion of mononuclear cells and therefore show a dramatic reduction in skeletal muscle mass (Hasty et al., 1993; Megeney and Rudnicki, 1995). *MRF4*$^{-/-}$ mice develop normal skeletal muscle, but express higher levels of myogenin, which suggests myogenin may compensate for the loss of MRF4 (Zhang et al., 1995). Conversely, in *MyoD*$^{-/-}$*MRF4*$^{-/-}$ mice, *myogenin* levels are normal, but phenotypes are identical to *myogenin*-null mice (Rawls et al., 1998). In this scenario, presumably the levels of myogenin are insufficient to compensate for the loss of MRF4. The similarity between *myogenin*$^{-/-}$ and *MyoD*$^{-/-}$*MRF4*$^{-/-}$ phenotypes reinforces the idea that myogenin and MRF4 have overlapping functions.

Pax3 is further implicated in regulation of myogenic cell fate, through regulation of MRF activity. Pax3 is critical for the survival of the hypaxial dermomyotome, and plays a critical role in the activation of MyoD during embryonic development (Tajbakhsh and Buckingham, 2000; Tajbakhsh et al., 1997) (Figure 1, panel 3). Pax3 acts in a synergistic regulatory network which includes Six, Dach, and Eya family members. While Six1 is necessary for MyoD and myogenin expression in limb buds, studies with *Myf5-LacZ* mice have demonstrated that Pax3 and Six1 can also direct *Myf5* expression through adjacent Pax3 and MEF3 binding sites located within an embryonic enhancer located 57.5kb upstream of

Myf5 (Bajard et al., 2006; Giordani et al., 2007; Laclef et al., 2003). Eya proteins act as cofactors of Six proteins and function to activate Six target genes, including *Pax3*. Loss of Eya proteins results in an absence of myogenic progenitors in developing limbs as a consequence of Pax3 downregulation (Grifone et al., 2007). Dach2 and Pax3 function together in a positive feedback loop to regulate each other's expression. Moreover, Dach2 can synergize with Eya2 to drive expression of Pax3 and induce MRF gene expression (Heanue et al., 1999; Ridgeway and Skerjanc, 2001).

Numerous *in vivo* studies have investigated the signaling networks regulating myogenesis and have identified that the neural tube, dorsal ectoderm, notochord, and mesoderm produce signals which regulate myogenic activity, implicating Sonic hedgehog (Shh), Wnt, and BMP as major signaling pathways involved in this process (Figure 1). Shh and Wnt signals are documented for their combinatorial effect on activating Pax3/7 and consequently Myf5 and MyoD (Maroto et al., 1997; Munsterberg et al., 1995). In particular, Myf5 is regulated by Shh through a Gli1 binding site in the *Myf5* epaxial enhancer (Borycki et al., 1999; Gustafsson et al., 2002; McDermott et al., 2005). Wnt1 is expressed by the neural tube and preferentially activates Myf5 in the epaxial myotome. Wnt7a, expressed by dorsal ectoderm, preferentially activates MyoD in hypaxial lineages. Other Wnts, including Wnt4, Wnt5a, and Wnt6, Wnt6 can mimic signals from the neural tube and dorsal ectoderm, activating both MyoD and Myf5 in explants of paraxial mesoderm (Tajbakhsh et al., 1998). BMP signaling inhibits expression of MyoD, blocking myogenesis downstream of Pax3 (Amthor et al., 1998; Reshef et al., 1998). The expression of the BMP antagonist Noggin in the dorsal-medial lip of the dermomyotome facilitates MyoD expression in Pax3-positive cells (Hirsinger et al., 1997). The spatial control of BMP signaling suggests BMPs function to prevent premature differentiation of myogenic cells in the dermomyotome.

1.6. Satellite Cells and Adult Muscle Regeneration

Adult regenerative myogenesis is largely governed by the satellite cell compartment. Satellite cells were first described by Mauro (Mauro, 1961) and reside on the surface of muscle fibers, between the basal lamina and sarcolemma (Bischoff, 1994). In their resting state, they are mitotically quiescent and are characterized by low cytoplasmic volume and high degree of heterochromatin (Schultz, 1976). In response to environmental stresses, such as exercise or injury, quiescent satellite cells are activated, whereby they re-enter the cell cycle and proliferate extensively forming myoblasts that differentiate and fuse to repair muscle tissue.

Until recently, the origins of satellite cells were somewhat unclear. The discovery of Pax7 as a specific marker for satellite cells facilitated new research in lineage tracing of satellite cell progenitors (Seale et al., 2000). In the somite, myogenic progenitors express Pax3 and Pax7, transcription factors belonging to the paired-homeodomain (Pax) gene family (Horst et al., 2006) and are absent in their expression of MRFs. Satellite cells first appear on muscle fibers around 16.5 dpc in the mouse at their sublaminal positions and similarly, Pax7 is expressed by sublaminal cells after 16.5dpc (Relaix et al., 2005; Seale et al., 2000). Pax3, by contrast, is expressed sporadically in satellite cells of some muscle groups (Kassar-Duchossoy et al., 2005; Relaix et al., 2004). At 10.5 dpc, Pax3+/Pax7-, Pax3-/Pax7+, and Pax3+/Pax7+ can all be found within the dermomyotome, however $Pax3^{+}/Pax7^{+}$ cells form a myogenic domain believed to be a major source for myogenic expansion (Kassar-Duchossoy et al., 2005; Relaix et al., 2005).

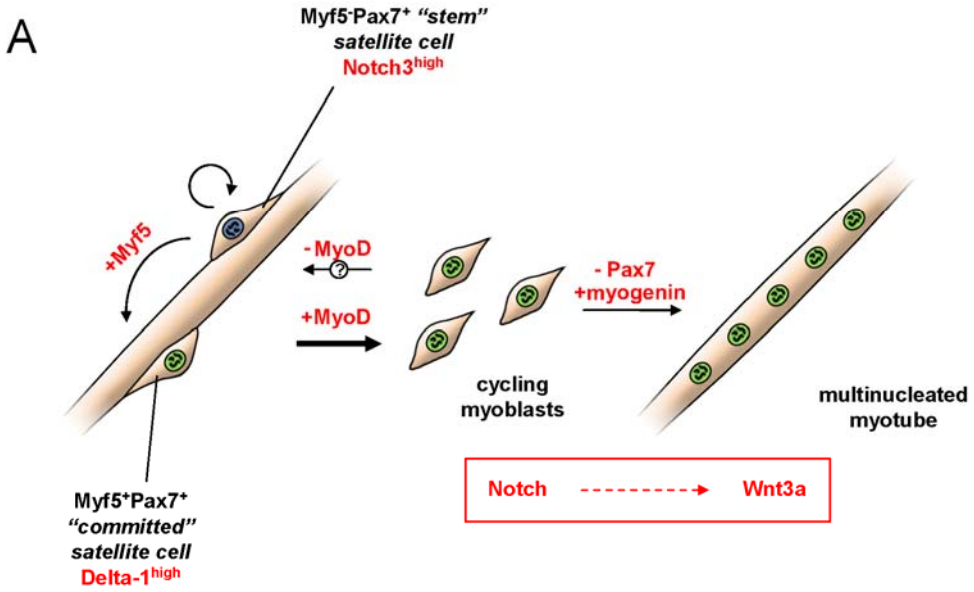
Lineage tracing experiments with Cre knock-in alleles have demonstrated that descendants of Pax7- and Pax3-expressing cells can contribute to satellite cell populations

and muscles (Hutcheson et al., 2009; Lepper and Fan, 2010; Schienda et al., 2006). Recent work by Lepper et al expanded upon this notion by demonstrating with a Pax7Cre-ER inducible allele that Pax7-expressing cells at 9.5 dpc give rise to dorsal trunk muscles, dorsal dermis, brown adipose tissue, and diaphragm muscles, but not ventral body wall or limb muscles. By contrast, cells labeled after 12.5 dpc are muscle-restricted, take up sublaminal satellite cell positioning by 16.5 dpc, and contribute to satellite cell pools, and are competent to regenerate muscle (Lepper and Fan, 2010). These works clearly define an origin for satellite cells from within the developing somite.

Satellite cells ultimately account for 2-7% of sublaminar muscle nuclei in adult mice. The number and frequency of satellite cells associated with fibers varies by age and across different muscle groups. Oxidative, slow-twich (Type I) fibers (i.e. Soleus) generally have higher numbers of satellite cells associated with them than glyolytic, fast-twich (Type II) fibers (i.e. EDL) (Charge and Rudnicki, 2004). While satellite cell numbers tend to decrease with age (Gibson and Schultz, 1983; Snow, 1977), the number of satellite cells tends to remain constant over multiple cycles of degeneration and regeneration, demonstrating an implicit capacity for self-renewal (Collins and Partridge, 2005; Kuang et al., 2007; Seale and Rudnicki, 2000).

1.7. Self-Renewal and Commitment of Satellite Cells

For muscle to maintain its regenerative potential there must be a continual replenishing of satellite cells. Following activation, cell fate decisions specify satellite cell self-renewal or differentiation. Most proliferating myoblasts will downregulate Pax7 and induce expression of myogenin (Figure 2A). A small subset of cells, however, maintain Pax7 expression and are thought to represent self-renewing cells that will return to


A
Myf5-Pax7+ "stem" satellite cell
Notch3high
+Myf5
- MyoD
+MyoD
cycling myoblasts
- Pax7
+myogenin
multinucleated myotube
Myf5+Pax7+ "committed" satellite cell
Delta-1high
Notch
Wnt3a

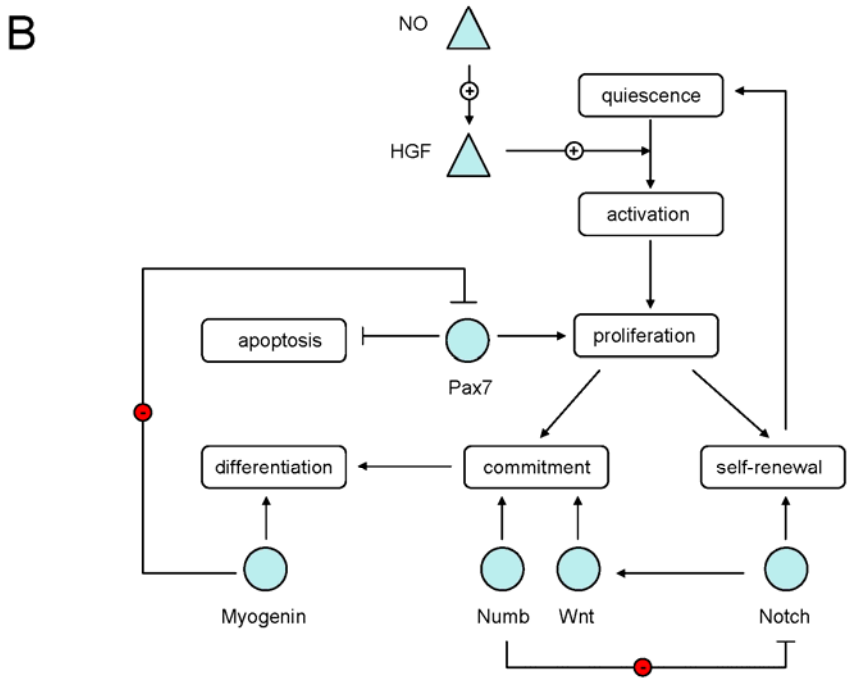

B
NO
HGF
quiescence
activation
proliferation
apoptosis
Pax7
differentiation
commitment
self-renewal
Myogenin
Numb
Wnt
Notch

Figure 2. Transcriptional Networks Regulating Satellite Cell Behavior

(A) Satellite cells are heterogeneous for Myf5 expression. Pax7+/Myf5- cells undergo symmetric cell divisions to replenish the satellite cell pool. They also give rise to Pax7+/Myf5+ cells through asymmetric divisions which preferentially undergo myogenic commitment. Most committed cells down regulate Pax7 and express myogenin, however some may instead lose MyoD expression and return to quiescence. Notch3 is asymmetrically expressed by Myf5- cells, while the Notch-activating ligand Delta-1 is highly expressed by Myf5+ cells. Notch signaling prevents early activation of MyoD and positively regulates proliferative expansion of myoblast progeny. A transition from Notch to canonical Wnt signaling acts as a switch to halt proliferation and initiate terminal differentiation in myoblasts, which fuse to repair or form new myotubes at sites of injury.
(B) Damage to muscle fibers stimulates the production of nitric oxide which, in turn, triggers a release of HGF from the extracellular matrix. These signals activate nearby satellite cells, which re-enter the cell cycle and undergo many rounds of proliferation producing myoblasts. Myoblasts are faced with a decision between self-renewal or myogenic commitment which may be guided, in part, by signals from the niche.

quiescence and repopulate the satellite cell compartment (Olguin and Olwin, 2004; Zammit et al., 2004b; Zammit et al., 2006). It has been suggested this fate choice is actively decided by asymmetric mitotic events, as evidenced by the fact that activated satellite cells are able to contribute to regenerating fibers as well as the satellite cell compartment (Collins and Partridge, 2005; Montarras et al., 2005).

Satellite cells are heterogeneous for Myf5 expression. Work by Kuang et al. demonstrated by use of a Myf5-Cre-induced reporter, that roughly 10% of sublaminar Pax7-expressing cells have never expressed Myf5 (Kuang et al., 2007). Furthermore, it was shown that, through asymmetric cell divisions occurring in the apical-basal plane, that $Pax7^{+}Myf5^{-}$ cells give rise to both $Pax7^{+}Myf5^{-}$ and $Pax7^{+}Myf5^{+}$ daughter cells. Conversely, planar divisions give rise to identical daughter cells, underscoring a role for the niche in satellite cell identity. Importantly, this heterogeneity reflects inherent differences in satellite cell behavior; Pax7+/Myf5- cells extensively contribute to replenishment of the satellite cell compartment, as evidenced through transplantation studies into Pax7 null mice, while Pax7+/Myf5+ cells precociously differentiate (Figure 2A). These observations define a hierarchical composition in which $Pax7^{+}Myf5^{-}$ cells represent self-renewing stem cells, while the $Pax7^{+}Myf5^{+}$ cells represent committed progenitors (Kuang et al., 2007).

1.8. Extrinsic Signals Regulating Satellite Cell Activation

Comparative microarray studies by Fukada et al. describe a host of genes upregulated in quiescent satellite cells, including many negative regulators of cell cycle progression and myogenic inhibitors (Fukada et al., 2007). Presumably, many of these are potentially targeted for downregulation through the process of activation. Conversely, activated satellite

cells upregulate specific gene sets, including those involved in DNA and RNA synthesis and cell cycle progression. But what controls the switch from quiescence to activation?

A number of cell-surface markers are thought to be involved in activation, including c-met, FGFR1, FGFR4, syndecan-3, and syndecan-4 (Cornelison et al., 2001; Cornelison et al., 2004; Flanagan-Steet et al., 2000; Tatsumi et al., 1998). Mechanical stress on muscle fibers triggers satellite cell activation through the release of nitric oxide (NO) and HGF (Anderson and Pilipowicz, 2002) (Figure 2B). Fiber damage stimulates NO production by Nitric oxide synthase (NOS), triggering a release of HGF from the extracellular matrix, which then rapidly associates with its receptor (Tatsumi et al., 2002; Tatsumi et al., 2006). NOS activity is also required for the maintenance of quiescence in satellite cells, and therefore NO may play a dual role regulating satellite cell behavior (Wozniak and Anderson, 2007). Although the role of c-met activation in satellite cells is unknown, HGF and FGF molecules can stimulate MAPK signaling pathways known to regulate proliferation and regulation in many cell types. It has recently been observed that signaling by p38 α/β MAPK is required for satellite cell proliferation (Jones et al., 2005), despite previously described roles for p38α/β in promoting differentiation of myogenic cells (Puri et al., 2000; Wu et al., 2000; Zetser et al., 1999).

The cell-surface receptor Notch mediates communication between Notch-expressing cells and cells expressing Notch-activating ligands such as Delta-1. Notch expression and signaling is increased following satellite cell activation, whereas myogenic differentiation occurs following inhibition of Notch signaling (Conboy and Rando, 2002). While all satellite cells express Notch1, Notch3 is asymmetrically distributed in quiescent satellite cells, with high expression found within $Myf5^-$ cells (Figure 2A). Conversely, Delta-1 is highly expressed in $Myf5^+$ cells (Fukada et al., 2007; Kuang et al., 2007) and interferes with

the expression of MyoD and may serve to prevent precocious myogenic differentiation. Delta-1 expression has no deleterious effect on Pax3 or Myf5 expression (Delfini et al., 2000). Moreover, it has been suggested that following asymmetric division, expression of Delta-1 by the committed daughter cell serves to activate Notch signaling in its sister self-renewing cell (Kuang et al., 2007). Inactivation of Notch signaling by deletion of RBP-J, a primary nuclear mediator and repressor of transcription, causes a reduction of embryonic myogenic progenitors and a complete ablation of adult satellite cells, while deletion of Delta-1 reduces the population of embryonic progenitors (Schuster-Gossler et al., 2007; Vasyutina et al., 2007). Furthermore, it has been demonstrated that the age-related reduction in regenerative potential is due to insufficient Notch signaling and reduced Delta expression in old muscles (Conboy et al., 2003; Conboy and Rando, 2005). Numb, an inhibitor of Notch signaling influences cell fate choices by promoting differentiation (Shinin et al., 2006). Numb is asymmetrically segregated into satellite cell derived myoblasts and has been correlated to myogenic differentiation in culture (Conboy and Rando, 2002). Asymmetric Numb expression is also evident in mitoses occurring within the epaxial dermomyotome (Venters and Ordahl, 2005) and treating somites with MyoD or signals that induce epaxial myogenesis increases Numb protein levels (Holowacz et al., 2006). Furthermore, the deletion of Notch inhibitor Stra13 results in increased Notch signaling and has been demonstrated to compromise satellite cell differentiation (Sun et al., 2007a). Similarly, treatment with DAPT, an inhibitor of γ-secretase activity required for the cleavage of Notch's intracellular domain upon activation, inhibits satellite cell proliferation and self-renewal (Kuang et al., 2007). Taken together, these observations suggest that Notch signaling is a key mediator of self-renewal, by regulating the inhibition of myogenic differentiation and promoting the stem-cell state within satellite cells.

In satellite cells, Wnt signaling appears to be involved in fate regulation. Conflicting observations implicate Wnt signaling in both specification (Polesskaya et al., 2003; Seale et al., 2003) and repression (Brack et al., 2007) of the myogenic lineage, while some studies refute the involvement of Wnt signaling entirely during regeneration due to expression of antagonistic sFRP proteins which silences the pathway (Zhao and Hoffman, 2004). These differences may be attributed to the diversity of Wnt proteins and their antagonists, and alternate signal transduction pathways (canonical vs. non-canonical). In the canonical Wnt signaling pathway, Wnt ligands bind to cell surface frizzled receptors, activating a signal cascade that represses the degradation of intracellular β-catenin, which converts TCF from a transcription repressor to activator. By contrast, non-canonical pathways operate independently of β-catenin and transduce Wnt signals through various coreceptors including those that activate planar cell polarity (PCP) and calcium pathways (Semenov et al., 2007).

In vitro studies indicate that Wnt signaling plays a number of roles in myogenesis. Recent work by Le Grand et al. demonstrated that Wnt7a is expressed during regeneration and acts through a non-canonical pathway that includes its receptor Fzd7 and Vangl2, a component of the PCP pathway, to induce the expansion of satellite stem cells and significantly augment muscle regeneration (Le Grand et al., 2009). Conversely, other studies implicate Wnt signaling in pro-differentiation roles during myogenesis (Petropoulos and Skerjanc, 2002; Ridgeway et al., 2000), consistent with embryonic models describing Wnt-induced activation of MRFs. However, whether these systems more appropriately model embryonic myogenesis or regeneration is debated (Zhao and Hoffman, 2004). Wnt signaling is sufficient to induce the myogenic commitment of adult stem cells following acute injury (Polesskaya et al., 2003; Torrente et al., 2004). Recent work has illuminated the interplay of Notch and Wnt signaling during regeneration. Brack and colleagues describe the

involvement of Notch in regulating proliferative expansion of myoblasts in adult myogenesis (Conboy and Rando, 2002), but implicate a transition from Notch to Wnt signaling as a mediator for the switch between myoblast proliferation and induction of differentiation in these progenitors (Figure 2). Moreover, they propose a model in which Notch and Wnt signaling converge on determining the state of activation of GSK3β, which they suggest is a pivotal determinant of muscle stem cell fate (Brack et al., 2008).

1.9. Roles of Pax3 and Pax7

Although clear parallels exist between the mechanisms governing development and regeneration, it is becoming increasingly clear that regeneration does not recapitulate development. Whereas Pax3 is critical for survival of muscle stem cells in the embryo, Pax7 is essential for satellite cell maintenance. *Pax7*$^{-/-}$ mice have a severe deficit in muscle regenerative capacity and drastically reduced muscle mass compared to wild type mice, the result of an ablation of satellite cell pools in early postnatal life (Mansouri et al., 1996; Oustanina et al., 2004; Seale et al., 2000). As such, Pax7 has been implicated in lineage survival though the promotion of satellite-cell propagation and inhibition of apoptosis (Relaix et al., 2006; Zammit et al., 2004b). In contrast, Pax3 does not appear to play a critical role in regeneration, as its expression is absent from most satellite cells (Kassar-Duchossoy et al., 2005; Montarras et al., 2005). However, in the embryo, Pax3 and Pax7 have overlapping expression domains and share some functional redundancy; Pax7 can substitute for Pax3 in somite development and activation of myogenesis, but cannot compensate for Pax3 in regulating migration events or efficient activation of *c-met* (Relaix et al., 2004).

Pax7 dependency in satellite cells has been suggested to be limited to a critical juvenile period when satellite cells are transitioning to a quiescent state (Lepper et al., 2009). By using tamoxifen to induce the Cre-mediated recombinations of Pax3 and Pax7, Lepper and colleagues showed that the conditional elimination of Pax7 does not impact the function or survival of satellite cells beyond P21. Moreover, this is not attributable to compensation by Pax3, suggesting that alternative mechanisms could arise to compensate for the loss of Pax7 postnatally. Some critics have argued, however, that these observations may also arise as the result of incomplete Pax7 recombination. Collectively, these data together argue that Pax3 mediates the migratory phase of the lineage whereas Pax7 is required to achieve their myogenic potential.

We have recently undertaken a systems approach to investigate mechanisms of Pax7 function, combining mass spectrometry, chromatin immunoprecipitation, and comparative genomics. We identified several novel and strongly regulated direct target genes of Pax7, including *Myf*5. Furthermore, we observed that Pax7 associates with a histone methyltransferase (HMT) complex which includes Ash2L, Wdr5, and MLL2. Importantly, we show that Pax7 recruits this complex to Pax7 target genes, and functions to establish transcriptionally active domains within chromatin through epigenetic modifications (McKinnell et al., 2008).

Interestingly, it has also been suggested that Pax7 may antagonize myogenic progression, as its overexpression has been observed to downregulate MyoD and impair myogenic progression (Olguin and Olwin, 2004; Zammit et al., 2006). One explanation is that satellite cells may be highly sensitive to Pax7 levels, and Pax7 may both activate and inhibit myogenic progression, by regulating additional target genes at higher thresholds of

expression. Unsurprisingly, a comprehensive understanding of Pax7 target genes and functional groups is a major goal of ongoing research efforts.

1.10. Alternative Splicing of Pax3 and Pax7

Pax3 and *Pax7* undergo alternative splicing events that occur within the highly conserved paired box DNA binding domain. At the exon 2/3 boundary, both proteins can undergo alternative splicing resulting in the addition of a glutamine (Q+) residue (Vogan et al., 1996). *Pax7* can also be alternatively spliced at the exon 3/4 boundary resulting in the inclusion of glycine and leucine (GL+) residues, yielding four possible isoforms (Pax7a, Q-/GL+; Pax7b, Q+/GL+; Pax7c, Q-/GL-; Pax7d, Q+/GL-) (Ziman and Kay, 1998). In the case of *Pax3*, the exclusion of glutamine (Q-) results in an increased binding affinity to paired domain consensus motifs (Vogan et al., 1996), while in the case of *Pax7*, the GL modification can alter the expression of target genes (Ishibashi, 2006). Alternative splicing may play a role in the selection of target sites.

1.11. Human Phenotypes of Pax3/7 Mutations

Mutations in Pax3 are associated with the development of Waardenburg syndrome, characterized by varying degrees of hearing loss and alterations in skin and hair pigmentation, but no muscle phenotype is evident. Pax3 is involved in developmental pathways that give rise to neurons and melanocytes. Accordingly, mutations in Pax3 cause the auditory-pigmentory symptoms in WS-I patients and may play a role in the pathogenesis of melanoma or neuroblastoma (Wang et al., 2008).

Chromosomal translocations causing gene fusions between FKHR and Pax3 or Pax7 are characteristic of alveolar rhabdomyosarcoma (ARMS), a pediatric soft tissue cancer

derived from the muscle lineage. The translocation events fuse the transactivation domain of FHKR to the DNA binding domain of Pax3 or Pax7, leading to increased transcription from Pax3/7 binding sites (Wexler and Helman, 1994). These chimeric proteins are expressed at high levels in ARMS tumors. Histologically, the tumors contain collections of poorly differentiated tissue, and weak evidence of muscle differentiation as marked by scant MyoD and desmin staining. Studies on the transcriptional behaviour of Pax3-FKHR and Pax7-FKHR suggest that the chromosomal translocations exaggerate the normal function of Pax3 and Pax7 in myogenic progenitor cells, leading to dysregulation of growth, apoptosis, differentiation and motility (Barr, 2001; Bennicelli et al., 1996; Keller et al., 2004). More recently, it was shown that growth factor receptors and homeobox transcription factors are among gene groups induced by Pax3-FKHR (Cao et al., 2010).

1.12. Systems Biology Approaches to Understanding Myogenic Regulatory Networks

Extensive studies have been carried out to describe the regulatory networks controlled by these gene families, many of which have effectively combined gene expression profiling, chromatin pulldowns, and computational prediction to elucidate binding sites of individual MRF and MEF2 proteins (Blais et al., 2005; Huang et al., 2004; Liu et al., 2004; Sandmann et al., 2006). Studies such as those carried out by Blais et al. demonstrate the diversity and complexity of gene networks controlled by these key regulators. Their approach combining using ChIP-chip analysis and expression profiling described targeted processes of MRF and MEF2 proteins beyond the expected systems of muscle development and differentiation, signal transduction, and chromatin remodeling. This work, in particular, highlighted their roles in synapse formation and transmission at neuromuscular junctions and stress response pathways. Other studies have applied similar methods to describe general

functional behaviors of these proteins. Work by Cao et al. illustrates that MyoD and myogenin bind a shared set of promoters, however carry distinct roles; MyoD drives histone acetylation, whereas myogenin synergizes with MyoD to drive expression of a set of genes expressed late in myogenesis (Cao et al., 2006).

Upstream of MRF activity, complex mechanisms govern the functions of muscle stem and progenitor cells, many of which converge on MRF regulation. Although numerous pathways and processes impacting MRF activity have been described, we know little about the networks that control muscle stem cell behavior and fate decision. In part, this may be due to difficulties in isolating large numbers of muscle stem cells for analysis or maintaining their identity and functional capacities in culture. For example, quiescent satellite cells represent a small percentage of myonuclei and become activated when removed from their niche, a major obstacle when studying mechanisms of satellite cell activation. Furthermore, the diversity among muscle stem cells is reflected in an alternative organization of many network components, making it difficult and confusing to generalize pathways and processes across different muscle stem and progenitor cell types.

1.13. Peripheral Networks

A host of other networks are also implicated in regulation of satellite cell function (Kuang et al., 2008). Our lab recently reported a role for the transmembrane protein Megf10 in satellite cell activation. Megf10 is expressed by all satellite cells and its overexpression drives increased cell proliferation. *Megf10* silencing results in a depletion of satellite cells on fibers and reduced cell number in culture, concomitant with an inhibition of Notch signaling and precocious differentiation (Holterman et al., 2007). Other studies have identified *FoxO* genes as regulators of satellite cell proliferation/self-renewal. Mice lacking *Foxk1*, an

antagonist of *FoxO* genes, show impaired muscle regeneration and decreased numbers of satellite cells. *Foxk1*$^{-/-}$ satellite cells improperly express myogenic factors and are arrested in their cell cycle progression, a consequence of the upregulation of p21, a FoxO downstream effector (Garry et al., 2000; Hawke et al., 2003). On the contrary, Kitamura et al. report Foxo1 cooperatively functions with Notch signaling to prevent myogenic differentiation (Kitamura et al., 2007). Furthermore, myostatin (GDF8), a negative regulator of myogenesis, inhibits myoblast proliferation and differentiation (Langley et al., 2002). Deletion of *myostatin* results in a dramatic increase in muscle mass (McPherron et al., 1997). Short-term blockade of myostatin enhances muscle regeneration and promotes satellite cell self-renewal through a Pax7-dependent mechanism (McFarlane et al., 2008).

1.14. Rationale and Hypothesis

Whereas Pax3 is critical for survival of muscle stem cells in the embryo, Pax7 is essential for the maintenance and survival of satellite cells, postnatally. Moreover, Pax3 does not play a critical role during muscle regeneration and cannot compensate for the loss of Pax7.

Because Paired-box proteins such as Pax3 and Pax7 are DNA binding transcription factors that regulate the expression of downstream genes, we rationed that the satellite cell-specific roles of Pax7 could be defined by its unique transcriptional outputs, or regulon. **More specifically, we hypothesized that Pax7 acts as an essential regulator of key gene networks required for the maintenance of satellite cells and that these networks could be defined by identification of its target genes**.

To define the unique signature of the Pax7 regulon, location analysis of Pax3 and Pax7 binding was performed by overexpressing tagged constructs in satellite cell-derived

primary myoblasts and generated binding maps by ChIP-Seq. The identification of unique binding sites and proximal genes combined with gene expression studies of Pax7 downstream genes allowed us to build a model transcriptional network for Pax7-specific regulation and illuminate its direct molecular functions. Additionally, contrasting the properties of genome-wide Pax3 and Pax7 binding provided a mechanistic explanation for differences in Pax3 and Pax7 function in satellite cells.

Chapter 2 – Transcriptional Dominance of Pax7 in Adult Myogenesis is Due to High-Affinity Recognition of Homeodomain Motifs Relative to Pax3

Transcriptional Dominance of Pax7 in Adult Myogenesis is Due to High-Affinity Recognition of Homeodomain Motifs Relative to Pax3

Vincent G. Punch[1,2], Vahab D. Soleimani[1,2], Yoichi Kawabe[1,2], Gareth A. Palidwor[1], Christopher J. Porter[1], Christel E.M. Kockx[3], Wilfred F.J. van IJcken[3], Frank Grosveld[3], Theodore J. Perkins[1,2], and Michael A. Rudnicki[1,2,4]

1. Sprott Center For Stem Cell Research
   Ottawa Hospital Research Institute
   Ottawa, ON
   Canada

2. University of Ottawa
   Department of Medicine
   Ottawa
   Canada

3. Erasmus University and Medical Center & Center for Biomedical Genetics
   3000 CA Rotterdam
   The Netherlands

4. Corresponding author: Michael A. Rudnicki
   Regenerative Medicine Program
   Ottawa Hospital Research Institute

Summary

Pax3 and Pax7 are closely related transcription factors that together regulate the migration, myogenic identity, and survival of satellite stem cells in adult skeletal muscle. However, an understanding of their unique roles has remained elusive. We ectopically expressed Pax3 and Pax7 in adult primary myoblasts and identified their binding sites by ChIP-Seq, and performed gene expression analysis to identify target genes. Strikingly, we found that Pax3 bound a small subset (6.4%) of Pax7 binding sites, which were biased towards the inclusion of paired domain motifs. By contrast, Pax7 exhibited high affinity binding to sites that preferentially contained homeodomain motifs. Almost 25% of Pax7-specific target genes were involved in regulating transcription underscoring a potentially major role played by Pax7 in defining myogenic identity. The reciprocal bias exhibited by Pax3 and Pax7 for binding paired versus homeodomain motifs mechanistically explains the distinctive attributes of Pax3 versus Pax7 function during myogenesis.

Introduction

The maintenance and repair of adult muscle tissue is directed by satellite cells, closely juxtaposed against the surface of muscle fibers underneath the basal lamina. These quiescent cells are activated by exercise or injury and reenter the cell cycle to produce progeny myogenic precursor cells that undergo many rounds of division before entering terminal differentiation and fusing to form multinucleated myofibers. Moreover, satellite cells exist as a population heterogeneous for *Myf5* expression, a feature that divides the satellite cell pool into a subpopulation of self-renewing stem cells (*Myf5*$^{-}$) and committed progenitors (*Myf5*$^{+}$) (Kuang et al., 2007).

The transcriptional heirarchies directing myogenic differentiation of satellite cell-derived myoblasts and embryonic myogenic cells are similar (Parker et al., 2003a). Progenitors cells express the paired box transcription factors Pax3 and Pax7 which lie genetically upstream of the myogenic regulatory factors (MRFs) MyoD and Myf5 (Bois et al., 2005; Punch et al., 2009; Relaix et al., 2005), that are responsible for myogenic specification and the initiation of the differentiation program. Satellite cells uniformly express Pax7, however the frequency of those co-expressing Pax3 varies among muscles; hindlimb cells rarely express Pax3, whereas Pax3 is expressed by half of forelimb and almost all diaphragm satellite cells (Relaix et al., 2006).

Pax7 plays a critical role in regulating the function of satellite cells (Kuang et al., 2006; Seale et al., 2000). Extensive analyses of *Pax7*$^{-/-}$ mice have confirmed the progressive loss of the satellite cell lineage in multiple muscle groups (Oustanina et al., 2004; Seale et al., 2000). Small numbers of Pax7-deficient cells do survive in the satellite cell position but these cells arrest and die upon entering mitosis. *Pax7*$^{-/-}$ muscles are reduced in size, the fibers

contain approximately 50% the normal number of nuclei, and fiber diameters are significantly reduced (Kuang et al., 2006). Pax7 recruits the Wdr5/Ash2L/MML2 histone methyltransferase (HMT) complex to target genes such as Myf5 leading to Histone 3 Lysine 4 trimethylation (H3K4me3) and subsequent gene activation (McKinnell et al., 2008). Together, these data support an essential role for Pax7 in regulating the myogenic potential of satellite cells.

The postnatal expression of Pax3 in quiescent satellite cells is limited to a small fraction of muscles, including the diaphragm (Kassar-Duchossoy et al., 2005; Relaix et al., 2004). This suggests that Pax3 is not required for adult satellite cell function. Developmentally, Pax3 has been shown to be involved in the delamination and migration of embryonic myoblasts towards developing limb buds (Bober et al., 1994). Pax3 can also directly activate Myf5 and MyoD in the embryo, thereby initiating myogenic differentiation (Bajard et al., 2006; Brunelli et al., 2007; Maroto et al., 1997; Tajbakhsh and Cossu, 1997).

The current understanding is that Pax3 and Pax7 play some overlapping but mostly non-redundant roles in establishing satellite cell lineage (Buckingham and Relaix, 2007). During early myogenesis, Pax3 and Pax7 expression defines a population of embryonic progenitors that have been suggested to later give rise to satellite cells. In the absence of both Pax3/7, these progenitors undergo apoptosis, or adopt alternative non-myogenic cell fates (Kassar-Duchossoy et al., 2005; Relaix et al., 2005). The inability of Pax3 and Pax7 to replace each other's function in knock-in or knock-out animals supports the notion that these proteins have distinct functions (Kuang et al., 2006; Relaix et al., 2006; Relaix et al., 2004). The requirement for Pax7 by satellite cells has been suggested to be dispensible following juvenile growth when satellite cells are transitioning to a quiescent state (Lepper et al., 2009). This raises important questions about how myogenic gene networks are regulated in

satellite cells postnatally. Collectively, these data together argue that Pax3 mediates the migratory phase of the lineage whereas Pax7 is required to achieve their myogenic potential.

The transcriptional network that controls satellite cell lineage progression is strikingly similar to that deployed during embryonic myogenesis (reviewed in (Parker et al., 2003b). The embryonic progenitors that give rise to satellite cells are characterized by Pax3/Pax7 expression and the lack of expression of other myogenic regulatory factors (MRFs: Myf5, MyoD, Mrf4 and Myogenin) (Kassar-Duchossoy et al., 2005; Relaix et al., 2005). Pax3/7^{+} cells enter the myogenic program by up-regulation of Myf5/MyoD, or give rise to satellite cells during late fetal myogenesis without up-regulating MRFs. Pax3/7^{+}MRF^{-} progenitors are first found to align with nascent myotubes at E15.5, then become satellite cells by taking a sublaminar position (Relaix et al., 2005). Intriguingly, Pax3/7^{+}MRF^{-} progenitors rapidly up-regulate Myf5 and down-regulate Pax3 expression upon assuming the satellite cell position (Kassar-Duchossoy et al., 2005). Notably, lineage tracing suggests that Pax3^{+} cells contribute to embryonic myoblasts and the endothelial lineage whereas Pax7^{+} cells contribute to fetal myoblasts supporting the notion that these represent distinct myogenic lineages (Hutcheson et al., 2009). Satellite stem cells (Pax7^{+}/Myf5^{-}) in adult muscle may represent a lineage continuum of the embryonic Pax3/7^{+}MRF^{-} progenitors (Kuang et al., 2007).

To investigate the mechanistic basis for the distinct features of Pax3 and Pax7 function, we ectopically expressed tagged versions Pax3 and Pax7 in satellite cell-derived primary myoblasts to define their binding sites. We combined microarray gene expression analysis with ultra-high throughput sequencing of DNA isolated from immunoprecipitated chromatin (ChIP-Seq) to derive functional associations between binding sites and Pax7-regulated genes. Our data indicate that Pax3 binds to a small subset of Pax7 sites that are

enriched in paired domain motifs whereas Pax7 preferentially binds sites enriched for homeodomain motifs. These experiments therefore elucidate how two highly related transcription factors Pax3 and Pax7 manifest developmental regulatory activities that are distinct from one another.

Results

Genome-Wide Mapping of Pax7 and Pax3 Binding

We examined Pax7 and Pax3 binding targets in adult satellite cell-derived primary myoblasts by ChIP-Seq. To facilitate a direct comparison between Pax3 and Pax7 binding maps, we generated stable cultures of primary myoblasts overexpressing Pax3 or Pax7 tagged with tandem affinity purification (TAP) tags by retroviral infection. We expressed the skeletal muscle splice isoforms Pax7d (Ziman and Kay, 1998) and Pax3a (Tsukamoto et al., 1994; Vogan et al., 1996). Pax7d is the predominantly expressed isoform in muscle and both Pax3a and Pax7d contain the inclusion of glutamine at the exon 2/3 boundary. Western blots demonstrated similar expression levels of Pax3-TAP and Pax7-TAP (Figure 1A). Intriguingly, endogenous Pax7 was efficiently repressed by proviral expression of either Pax3-TAP or Pax7-TAP. To estimate expression levels we compared proviral expression against wild type myoblasts. Pax7-TAP cells expressed approximately 6.0×10^5 molecules per cell, whereas we estimated that wild-type cells expressed 2.75×10^4 endogenous Pax7 molecules per cell (Figure S1).

The TAP tag was composed of 3xFLAG sequences and 6xHIS residues separated by a Tobacco Etch Virus (TEV) cleavage signal, fused in-frame to the 3' end of Pax7d or Pax3 (Figure S2A). Complexes were immunoprecipitated using anti-FLAG conjugated agarose

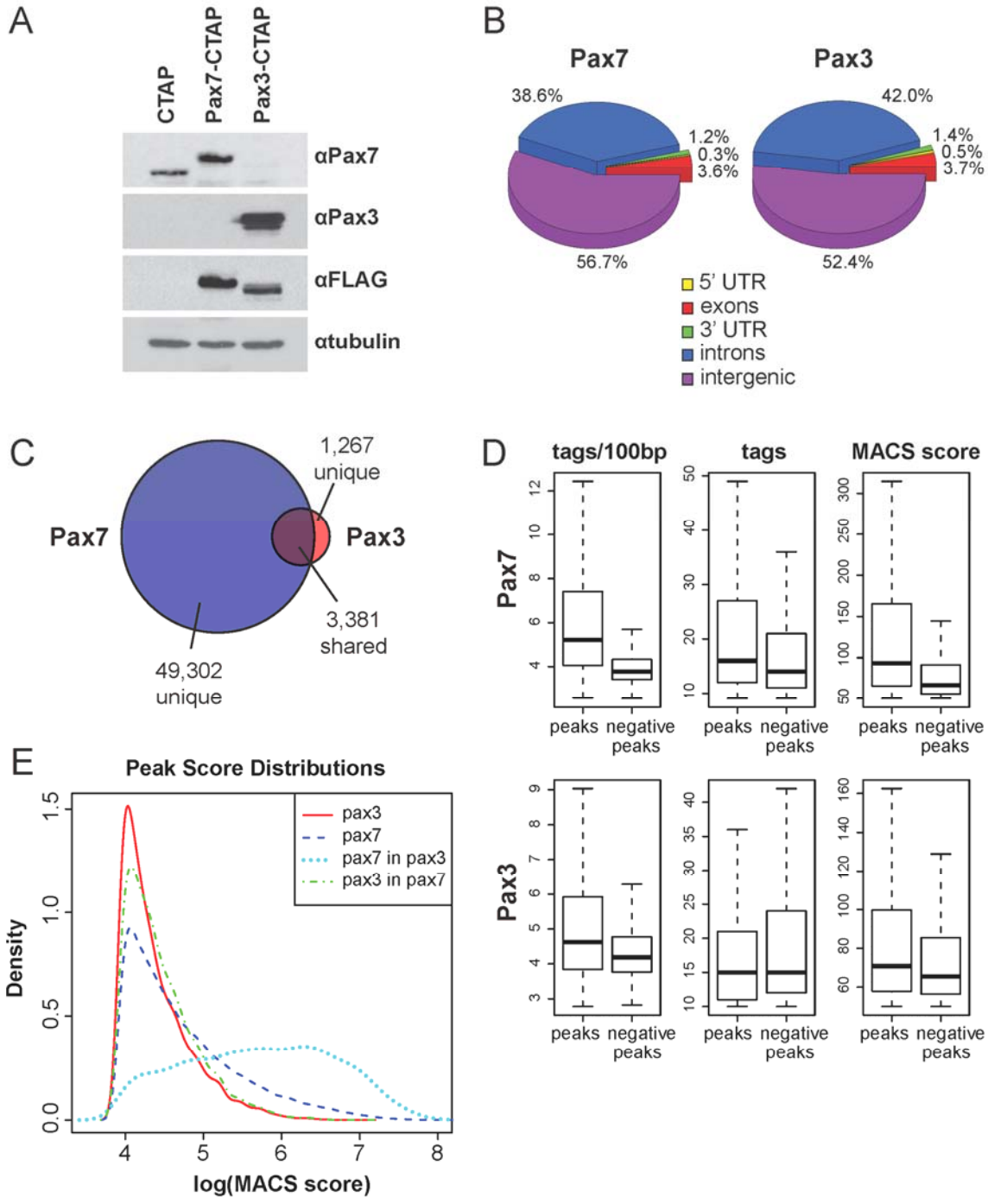

A
CTAP
Pax7-CTAP
Pax3-CTAP
αPax7
αPax3
αFLAG
αtubulin
B
Pax7
Pax3
38.6%
1.2%
0.3%
3.6%
56.7%
42.0%
1.4%
0.5%
3.7%
52.4%
5' UTR
exons
3' UTR
introns
intergenic
C
1,267
unique
Pax7
Pax3
3,381
shared
49,302
unique
D
tags/100bp
tags
MACS score
Pax7
Pax3
peaks
negative peaks
E
Peak Score Distributions
pax3
pax7
pax7 in pax3
pax3 in pax7
Density
log(MACS score)

Figure 1. Pax3 Binds a Subset of Pax7-Binding Sites

(A) Overexpression of Pax7- and Pax3-CTAP in primary myoblasts reveals a reduction in endogenous levels of Pax7. Pax3 is also not expressed in control cells.
(B) Analysis of peak locations relative to gene start sites indicates Pax3 and Pax7 binding is predominantly in intergenic or intronic regions.
(C) Overlap of Pax3 and Pax7 peak data reveals that Pax3 binds only 6.4% of Pax7 sites.
(D) Peak height (tags/100bp), tag number, and MACS score was compared between true peaks and negative peaks (≤800bp) for Pax3 and Pax7. Box boundaries denote the first and third quartile scores, black bars denote median scores.
(E) Distributions of MACS scores (a metric of overall peak quality) for peak subcategories. The area under the curve has been normalized across sample groups. These distributions demonstrate that the MACS scores of shared peaks are generally higher than Pax3 or Pax7 only peaks and may represent the highest-affinity sites.

and purified through enzymatic cleavage with TEV protease and subsequent purification on a Ni^{2+} column (Figure S2B, S2C). For the control, background enrichment was directly assessed using the same affinity reagents and by ChIP-Seq of TAP tag-expressing primary myoblasts in growth phase.

Immunoprecipitated fragments ranging from 100 bp to 300 bp were recovered and analyzed by high-throughput sequencing on an Illumina Genome Analyzer IIGx and the resulting 36-base pair reads mapped to the mouse NCBI m37 genome assembly using the Illumina Eland pipeline versions 1.5.1 (Pax7) and 1.4.0 (GM control and Pax3)[*]. Data sets of uniquely-mapping Solexa reads from Pax3, Pax7 and the control sample were generated, allowing for one or two mismatches. This resulted in 11.44 million uniquely mapping reads for the Pax7-TAP sample, and 12.43 million for the Pax3-TAP sample, and 8.54 million for the control sample.

Genomic locations with significant enrichment for Pax7 or Pax3 association (hereafter referred to as "peaks") were identified using MACS (Model-based Analysis of ChIP-Seq) analysis of each transcription factor data set using the growth phase control (Zhang et al., 2008). Peaks were filtered to remove those longer than 800bp, which eliminated approximately 1% of peaks and 13% of negative peaks (regions enriched in control ChIP relative to the treatment ChIP, derived as part of the MACS analysis). This filtering was done to further shape peak distributions, allowing us to compare peaks with similar properties and remove the rare broad binding regions, which are not expected to be biologically significant (Figure S3). We observed a total of 52,678 filtered peaks for Pax7 and 4,647 filtered peaks for Pax3. Negative peaks numbered 34,147 for Pax7 and 22,905 for Pax3.

[*] Pax7 raw data was obtained following a version update to Eland

Far fewer Pax3 peaks were observed than Pax7 peaks despite each transcription factor having a comparable numbers of reads and equivalent expression levels. Moreover, fewer tags in peaks suggest that Pax3 has a lower affinity for binding than Pax7. Taken together, these data argue that Pax3 and Pax7 binding dynamics are markedly different in primary adult myoblasts.

Pax3 Binds a Subset of Pax7 Sites

The vast majority of Pax3 and Pax7 binding sites were located within intergenic or intronic regions (Figure 1B). Very few binding sites were observed within 5' or 3' UTR regions. Notably, most Pax3 binding sites were found to overlap with Pax7 binding sites (Figure 1C). In many Pax7-specific binding sites, a number of Pax3 reads mapped to the region, indicating a measurable amount of Pax3 binding below the MACS peak calling threshold. To identify a set of Pax7 peaks with minimal evidence of Pax3 binding, peaks were annotated with the number of Pax7 and Pax3 reads contained within. A subset of 17,743 Pax7 peaks that contain fewer than three Pax3 mapped reads was identified as a set of Pax7-only peaks. The remaining 31,547 Pax7 peaks show some evidence of weak Pax3 binding.

In examining the quality of Pax3 and Pax7 peaks, we noted the disparity between Pax7 peaks and negative peaks was greater in terms of peak height (calculated as tags/100bp), total number of tags, and MACS score (defined as $-10 * \log_{10}$(pvalue) reported by MACS) than Pax3 peaks relative to negative peaks (Figure 1D). Pax7 peaks had significantly higher scores relative to negative peaks for peak height, tags, and MACS score, whereas Pax3 peaks had significantly higher peak height and MACS scores, but not tag numbers. Overall, this data suggested that Pax3-specific peaks were of generally lower quality peaks than Pax7 ones suggesting that Pax3 binds with lower affinity than Pax7.

To address this differential binding between Pax3 and Pax7, we assessed the distributions of peak quality in Pax3-only, Pax7-only, and common peak subsets (Figure 1E). Pax3-specific peaks averaged a MACS score of 65.3, while Pax3 peaks overlapping Pax7 peaks averaged a score of 98.8. By comparison, Pax7-only peaks averaged 107.2 while Pax7 peaks overlapping Pax3 peaks averaged 452.3. Because high-scoring Pax3 peaks are predominantly shared with Pax7, we conclude that Pax3 binding sites are a subset of Pax7 binding sites. Additionally, because Pax7 sites shared with Pax3 are highly scored, we propose that common targets represent the highest affinity sites for both Pax3 and Pax7.

Next, we examined the location of peaks relative to nearby genes. Peaks were annotated and with their location relative to gene start sites (taken from Ensembl release 55_37h). 5915 Pax7 peaks (11.2%) were found within 10kb of gene transcriptional start sites (TSS). Given that the DNA binding domains of Pax proteins are highly conserved across species, we anticipated finding a number of binding sites in evolutionarily conserved regions that might define functionally important elements (i.e. gene regulatory elements). Accordingly, we determined conservation scores for Pax3 and Pax7 peaks using the UCSC placental mammal conservation track. Effective conservation scores were calculated as the average conservation across entire peaks for Pax3, Pax7, and MACS negative peaks for both datasets. Pax7 peaks yielded higher conservation compared to negative peaks (mean scores of 0.176 and 0.150, respectively), as did Pax3 (mean scores of 0.185 and 0.160, respectively) (Figure 2A).

For further analysis, we considered Pax3 and Pax7 peaks conserved if their effective conservation was greater than the 90th percentile conservation score of negative peaks ≤800bp (Pax7 - 0.380; Pax3 – 0.402). For Pax7, this consisted of 7,570 peaks (14.4%) and

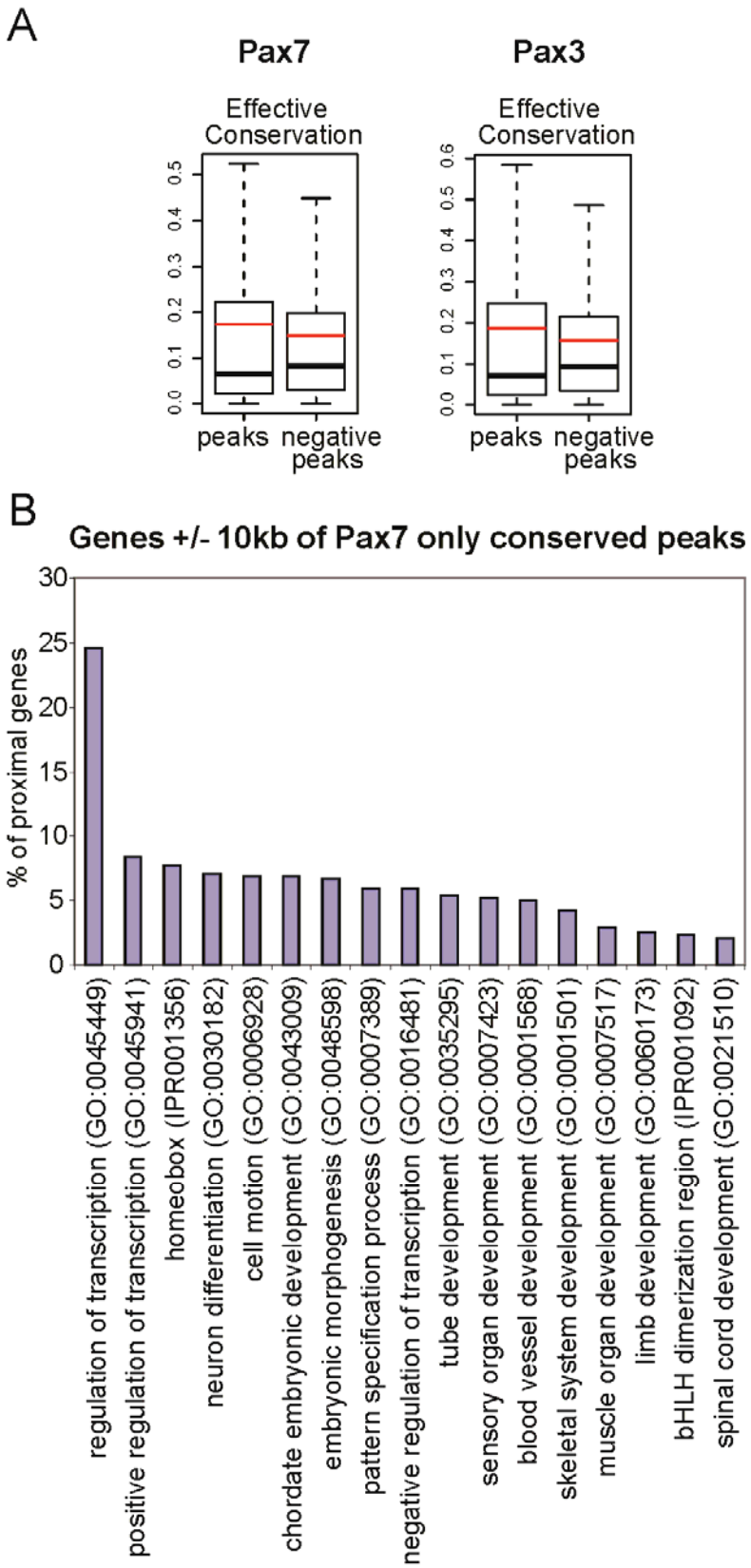

A
Pax7
Effective Conservation
0.0 0.1 0.2 0.3 0.4 0.5
peaks
negative peaks
Pax3
Effective Conservation
0.0 0.1 0.2 0.3 0.4 0.5 0.6
peaks
negative peaks
B
Genes +/- 10kb of Pax7 only conserved peaks
% of proximal genes
0 5 10 15 20 25 30
regulation of transcription (GO:0045449)
positive regulation of transcription (GO:0045941)
homeobox (IPR001356)
neuron differentiation (GO:0030182)
cell motion (GO:0006928)
chordate embryonic development (GO:0043009)
embryonic morphogenesis (GO:0048598)
pattern specification process (GO:0007389)
negative regulation of transcription (GO:0016481)
tube development (GO:0035295)
sensory organ development (GO:0007423)
blood vessel development (GO:0001568)
skeletal system development (GO:0001501)
muscle organ development (GO:0007517)
limb development (GO:0060173)
bHLH dimerization region (IPR001092)
spinal cord development (GO:0021510)

Figure 2. Pax7 Targets Genes Involved in Transcriptional Regulation

(A) Pax7 peaks averaged higher conservation scores than negative peaks (mean effective conservations of 0.176 and 0.150, respectively; red lines) as did Pax3 peaks (mean effective conservations of 0.185 and 0.160, respectively; red lines).
(B) GO analysis was carried out using genes with conserved peaks within 10kb of transcriptional start sites. Functional annotation clusters of GO terms over-represented in the gene lists were generated using DAVID. The most significant GO terms for each cluster are displayed. Conserved peaks were frequently observed near transcription factors, motility genes, and other genes involved in developmental processes.

for Pax3, 692 peaks (14.7%). The high conservation across these peaks justifies them as primary candidates for functional significance.

Pax7 Targets Discrete Gene Functional Groups

We sought to determine whether Pax7 binding sites were preferentially associated with functionally-related groups of genes. For each Pax7 peak, we identified genes with transcription start sites located within 10 kb upstream or downstream of the binding site. We separately analyzed the genes associated with Pax3/Pax7 and Pax7 unique binding sites for the most significantly enriched Gene Ontology (GO) terms using Functional Annotation Clustering tool provided by DAVID Bioinformatics Resources (2008) (Dennis et al., 2003). Affymetrix MoGene 1.0ST (v1) transcript cluster IDs were used as a background dataset.

Initially, we analyzed genes proximal to all peaks in each category regardless of conservation (Pax3/7 common: 412 genes; Pax7 unique: 3000 – as the maximum input for DAVID is 3000 IDs, the first 3000 genes were selected from a list of Pax7 peaks, sorted by MACS p-value). No GO terms were found to be significantly enriched for genes associated with common peaks (Benjamini < $5.0E^{-2}$). However, Pax7 unique peaks were associated with the significant enrichment of several functional gene groups, including cell and biological regulatory processes, developmental processes, cellular metabolism, and muscle development (Table S1).

We observed a more profound enrichment of functional categories when gene associations were restricted to conserved peaks (overlapping peaks = 55 genes; Pax7 unique peaks = 708 genes). While overlapping peaks were again not enriched for any GO terms, conserved Pax7 unique peaks were associated with strong enrichment of many functional groups (Benjamini < $1.0x10^{-5}$). Importantly, nearly one-fourth of these proximal genes were

annotated as transcriptional regulators. Other highly represented functional groups included genes associated with developmental processes, homeobox binding, neural development, and cell motility (Fig. 2b and Table S1). Taken together, these data underscore a major regulatory role played by Pax7 in defining myogenic cell identity and function.

Pax7 Binds or Supplants Known Pax3 Sites in Adult Myoblasts

The observation that the majority of sites are occupied uniquely by Pax7 or with low affinity by Pax3 raises the question of whether Pax7 unique sites reflect specific Pax7 functions that cannot be compensated for by Pax3 in adult muscle cells. Our ChIP-Seq data confirmed Pax7 binding to a number of previously-described binding sites targeted by Pax3 and describe an even greater number of new binding sites proximal to known regulatory targets of Pax3 and Pax7.

Many common peaks were observed within or near known Pax3 and Pax7 target genes, including *Myf5* (Bajard et al., 2006; McKinnell et al., 2008), *C-met* (Epstein et al., 1996), and *Cdh11* (McKinnell et al., 2008), while others were observed in and around genes linked to other muscle-related processes such as myogenic inhibition (i.e. *Mdfic;* (Ma et al., 2003)), myogenic signaling (i.e. *Stat1;* (Sun et al., 2007b)), and planar cell polarity (*Vangl2;* (Park and Moon, 2002)) (Figure S4 and data not shown). As previously described, *Myf5* has Pax3 and Pax7 binding sites located in the -57 kb enhancer (Bajard et al., 2006). Interestingly, we noted a larger peak located at -111 kb from the *Myf5* TSS in both datasets. Therefore, these data suggest the presence of a novel enhancer lying within the satellite-cell specific control region described previously (Zammit et al., 2004a).

Binding sites uniquely recognized by Pax7 were also found near key myogenic genes (Figure S5). Two prominent Pax7 binding sites were found in conserved intergenic regions

upstream of MyoD, and these sites were not bound by Pax3. In addition, Pax7 was found to bind the +22 kb Pax3 enhancer downstream from Fgfr4, but strikingly, Pax3 did not bind. This was particularly surprising, as Pax3 is known to bind the paired-containing motif during embryonic limb development (Bajard et al., 2006; Lagha et al., 2008). These data suggest a loss of Pax3 functional competency may occur in adult cells, and that binding might be supplanted by Pax7. Other Pax3 target genes were found to be adjacent to Pax7 uniquely-bound sites. A very strong peak was observed in the promoter of Lbx1, while binding sites were discovered within and downstream of Itm2a. These observations indicate that some enhancers that are targeted by Pax3 in the embryo become reoccupied by Pax7 in adult tissue. Alternatively, these sites may be common embryonic targets that lose their affinity for Pax3 postnatally.

Divergent Roles for Pax3 and Pax7 in Target Gene Regulation

The large number of Pax7-only sites and their discrete functional associations prompted us to investigate the specific gene regulatory activity of Pax3 and Pax7. Gene expression profiling was carried out on triplicate samples of Pax3-TAP, Pax7-TAP, and control myoblasts. Hybridizations were performed on Affymetrix Mouse 1.0 ST Gene arrays and normalized expression data was subjected to SAM analysis (Tusher et al., 2001) to generate rank ordered lists of differentially expressed genes.

We found 246 genes differentially expressed between Pax7 and Pax3 (rawp < 10^{-3}). Functional annotation clustering of related GO terms revealed signaling molecules and membrane receptors to be overrepresented within Pax7 upregulated genes (Table I). Strikingly, BMP4, BMP6 and Wnt7a were highly upregulated in Pax7-expressing cells, compared to both control and Pax3 cells, as well as Prickle2, a component of the PCP

signaling pathway. Conversely, Pax7 potently suppressed the expression of a large number of muscle-specific genes and cytoskeletal components, such as myogenin, Mef2c, and several myosin proteins. Caveolin-1 was strongly downregulated by Pax7, but virtually unchanged by Pax3 overexpression. A full gene list is presented in Appendix B.

Many genes proximal to common Pax3/7 peaks were also identified as common targets by gene overexpression. Cdh11, Mdfic, and Myf5 were among genes strongly induced by either factor, while Angpt1, Noggin, and Pbx1 were strongly inhibited. We observed a significant decrease in Notch signaling components including Hey1, Heyl, Notch3 and Dll1. A number of homeobox proteins including Dlx1, Pbx, Pitx2, and Six4 were also repressed by both Pax3 and Pax7, supporting the GO observation that Pax7 binding sites were adjacent to homeobox genes.

To form a comprehensive picture of Pax7's gene regulatory network, we furthered our expression study by eliminating Pax7 expression from wild-type myoblasts. RNA was isolated 48h following transfections of Pax7 or scramble siRNAs and expression analysis was carried out as above. A subset of genes was selected for validation by qPCR (Table S2). Pax7 knockdown resulted in the downregulation of 322 genes (implied as activated) and upregulation of 186 genes (implied as repressed). A full gene list is presented in Appendix C. Functional annotation clustering of each group suggested that activated genes are involved in cell growth and proliferation, while repressed genes are involved in differentiation, consistent with our overexpression data (Table I).

Taken together, our gene expression data suggest that a primary role of Pax7 is promoting myoblast growth and inhibiting myogenic differentiation. Furthermore, autocrine signaling mechanisms may be exploited by Pax7 to accomplish these roles.

Table I. Overrepresented GO terms associated with Pax7 regulation

Positively regulated genes (overexpression)

GO ID	Function Name	Fold Enrichment	Count	Size	Benjamini
PIR superfamily	glycoprotein	2.17	64	145	1.52E-08
PIR superfamily	signal	2.20	53	145	7.56E-07
PIR superfamily	membrane	1.88	85	145	2.70E-09
UniProt seq feature	topological domain:Extracellular	2.22	42	139	1.00E-04

Negatively regulated genes (overexpression)

GO ID	Function Name	Fold Enrichment	Count	Size	Benjamini
PIR superfamily	muscle protein	47.14	11	82	3.94E-12
GO:0043292	contractile fiber	21.85	11	65	3.19E-09
GO:0015629	actin cytoskeleton	11.15	12	65	1.67E-07
GO:0005509	calcium ion binding	4.33	18	64	5.89E-05
GO:0007517	muscle organ development	13.59	11	61	3.55E-06

Positively-regulated genes (siRNA)

GO ID	Function Name	Fold Enrichment	Count	Size	Benjamini
GO:0007049	cell cycle	7.92	89	250	1.01E-52
GO:0051301	cell division	11.61	60	250	6.85E-44
GO:0007067	mitosis	14.30	50	250	8.62E-41
GO:0000775	chromosome, centromeric region	18.69	35	211	2.21E-32
GO:0000793	condensed chromosome	17.17	31	211	3.19E-27
GO:0006260	DNA replication	10.73	30	250	1.50E-19

Negatively-regulated genes (siRNA)

GO ID	Function Name	Fold Enrichment	Count	Size	Benjamini
PIR superfamily	muscle protein	28.73	14	174	2.64E-13
GO:0007517	muscle organ development	10.94	18	127	5.23E-10
GO:0030016	myofibril	15.03	14	128	1.11E-09
GO:0016529	sarcoplasmic reticulum	28.73	10	128	1.57E-09
GO:0042692	muscle cell differentiation	8.23	9	127	1.89E-03

Gene Association

We presumed that Pax7 unique binding results in the unique regulation of many of its direct target genes. Therefore, we expected that Pax7 unique peaks could be physically associated with genes differentially regulated between Pax3 and Pax7, under the assumption that Pax3/7 transcriptional activity results in *cis*-regulation. Accordingly, we analyzed the location of Pax7 unique and Pax3/7 common peaks relative to TSS for up-, down-, and unregulated genes. For comparison, peaks were also plotted against Pax7-regulated genes observed from our siRNA experiments.

We considered peaks to be associated with genes if they were located within +/- 20kb of the TSS. While did not observe any relationship between fold change and the number of peaks associated with a particular gene, more peaks were found to be associated with upregulated genes (data not shown). Two-sided Kolmogorov-Smirnov (KS) tests revealed a significant difference ($p < 0.005$) in the distribution of Pax7 unique peaks associated with upregulated versus unregulated genes, supporting the hypothesis that Pax7 binding is responsible for the direct upregulation of these genes. No differences were observed in the distributions of common peaks.

In plotting the location of Pax7 MACS peaks relative to the TSS against expression fold changes under Pax7 siRNA treatment for peaks within +/- 10kb of TSS (Figure 3A), we also observed Pax7 peaks were more frequently proximal to activated genes than repressed genes, and noted that neither MACS nor conservation scores were correlated to fold changes in gene expression (Figure S6).

To test the effective range of Pax7 influence, we calculated the percentage of down and upregulated genes within 100 kb of Pax7 binding sites (Figure 3B). As a control

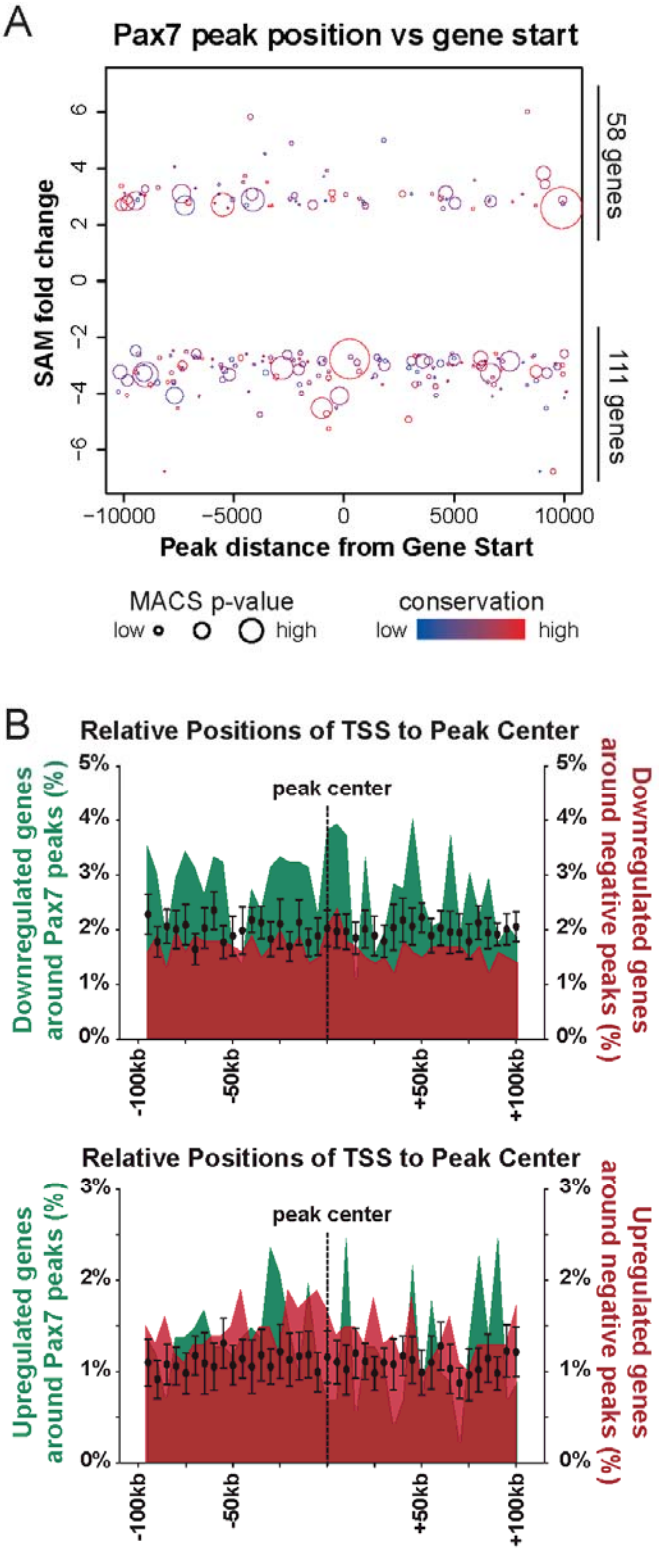

A
Pax7 peak position vs gene start
SAM fold change
6
4
2
0
−2
−4
−6
58 genes
111 genes
−10000
−5000
0
5000
10000
Peak distance from Gene Start
MACS p-value
low
high
conservation
low
high
B
Relative Positions of TSS to Peak Center
Downregulated genes around Pax7 peaks (%)
peak center
5%
4%
3%
2%
1%
0%
-100kb
-50kb
+50kb
+100kb
Downregulated genes around negative peaks (%)
Relative Positions of TSS to Peak Center
Upregulated genes around Pax7 peaks (%)
peak center
3%
2%
1%
0%
-100kb
-50kb
+50kb
+100kb
Upregulated genes around negative peaks (%)

Figure 3. Pax7 Acts as a Gene Activator

(A) Gene expression was compared between myoblasts treated with Pax7 and control siRNAs. The positions of Pax7 peaks relative to transcriptional start sites were plotted against expression fold differences for genes identified through SAM analysis as having significantly different expression between the two treatments. Pax7 peaks are more frequently found within 10kb of genes with lower expression under Pax7 siRNA treatment (activated by Pax7) than those with higher expression (repressed by Pax7). Association of peaks with genes is not influenced by peak quality (MACS p-value) or by the effective conservation scores of peaks.

(B) The effective range of Pax7 was calculated by determining the percentage of downregulated (activated, top panel) or upregulated (repressed, bottom panel) genes located +/- 100kb of Pax7 peaks (green). Enrichment of regulated genes was determined by comparing against the percentage of regulated genes around negative peaks (red) or twenty simulated peak sets of equal size (black dots) across a continuous window. We observed a significant enrichment of activated genes even at larger distances from Pax7 peaks, but the same effect was not observed with simulated peaks or negative peaks. Moreover, enrichment was not observed with repressed genes, supporting the notion that Pax7 acts primarily as a gene activator and can extend its influence to genes even at large distances.

comparison, we also performed the same analysis using an equal number of Pax7 negative peaks and with twenty sets of randomly simulated peak sets of equal size to assess statistical significance. We observed that genes activated by Pax7 were significantly enriched across a continuous window from -100kb to +100kb of Pax7 peaks, compared to negative peaks or simulated peak sets (Figure 3B, top). The same enrichment, however, was not observed with repressed genes (Figure 3B, bottom). Taken together, these data further support the conclusion that Pax7 is primarily a gene activator, and more importantly demonstrate that Pax7 can regulate gene expression at large distances.

Sequence Characteristics of Pax3/7 Binding Sites

Pax3 and Pax7 proteins contain paired and homeobox DNA-binding domains. Although Pax proteins are characterized by their paired domain, some family members contain partial homeobox domains (Pax2, Pax5, and Pax8), and others lack them entirely (Pax1 and Pax9). Therefore, the functional significance of homeobox-DNA interactions within Pax proteins remains unclear.

To understand the basis for Pax3 and Pax7 binding, we carried out *de novo* motif searches in Pax3 and Pax7 peaks using GLAM2 and MEME tools. GLAM2 analysis of Pax3 peaks yielded motifs composed of homeodomain (*hbox*) and paired (*prd*) domain motifs. *Hbox* motifs were structured as inverted repeats (i.e. TAAT and ATTA), separated by 2bp spacing elements that were commonly “TG”. Where *hbox* motifs were adjacent to *prd* motifs, we observed homeodomains in either orientation with respect to the paired domain. Overall, paired motifs formed a consensus that largely matched a sequence for Pax3 originally defined by Epstein *et al.* (GTCACG (C/G)TT(C/G)) (Epstein et al., 1996). In

addition, we identified a variant paired domain, *prdC*, containing an extra cytosine at position 6.

MEME and GLAM2 were similarly used to perform *de novo* motif discovery on Pax7 peak sequences (Figure 4A). The resulting motif was composed entirely of homeodomain motifs. Suspecting that a high abundance of homeodomain motifs might obscure the identification of other less abundant motifs, we masked *hbox* sequences in Pax7 data and repeated the motif search. MEME discovered additional motifs composed of both paired domains and a variant permutation of the *hbox* motif, termed *hbox3*, whereby three bases separated the inverted repeats. A juxtaposition of the *hbox3* and *prd* elements was observed in four cases (*h3prd*).

To accurately describe the paired motif in the Pax7 data, we used a regression-based approach to refine the paired sequence. The *prd* motif matrix derived from Pax3 peaks was used to seed a search within Pax7 data. Matches were iteratively used to generate refined motif matrices until the motif stabilized.

We used FIMO to search for instances of six motifs within Pax3 and Pax7 peak subsets. *h3prd*, *hbox*, *hbox3*, *prd*, and *prdC* motifs were generated from MEME and GLAM2 results. Additionally, we included the MEF3 motif (Six1/4), as it was previously reported that the Myf5 -57kb element requires both Pax3 and Six1/4 binding for the activation of *Myf5* (Bajard et al., 2006; Giordani et al., 2007). The peak subsets used were Pax7-only (fewer than three overlapping Pax3 reads), Pax7-only (subthreshold Pax3 binding), and overlapping Pax3/Pax7 peaks. Peak sequences matching these motifs were input into Weblogo to generate a representation of the motifs found in each peak subset (Figure 4A and data not shown).

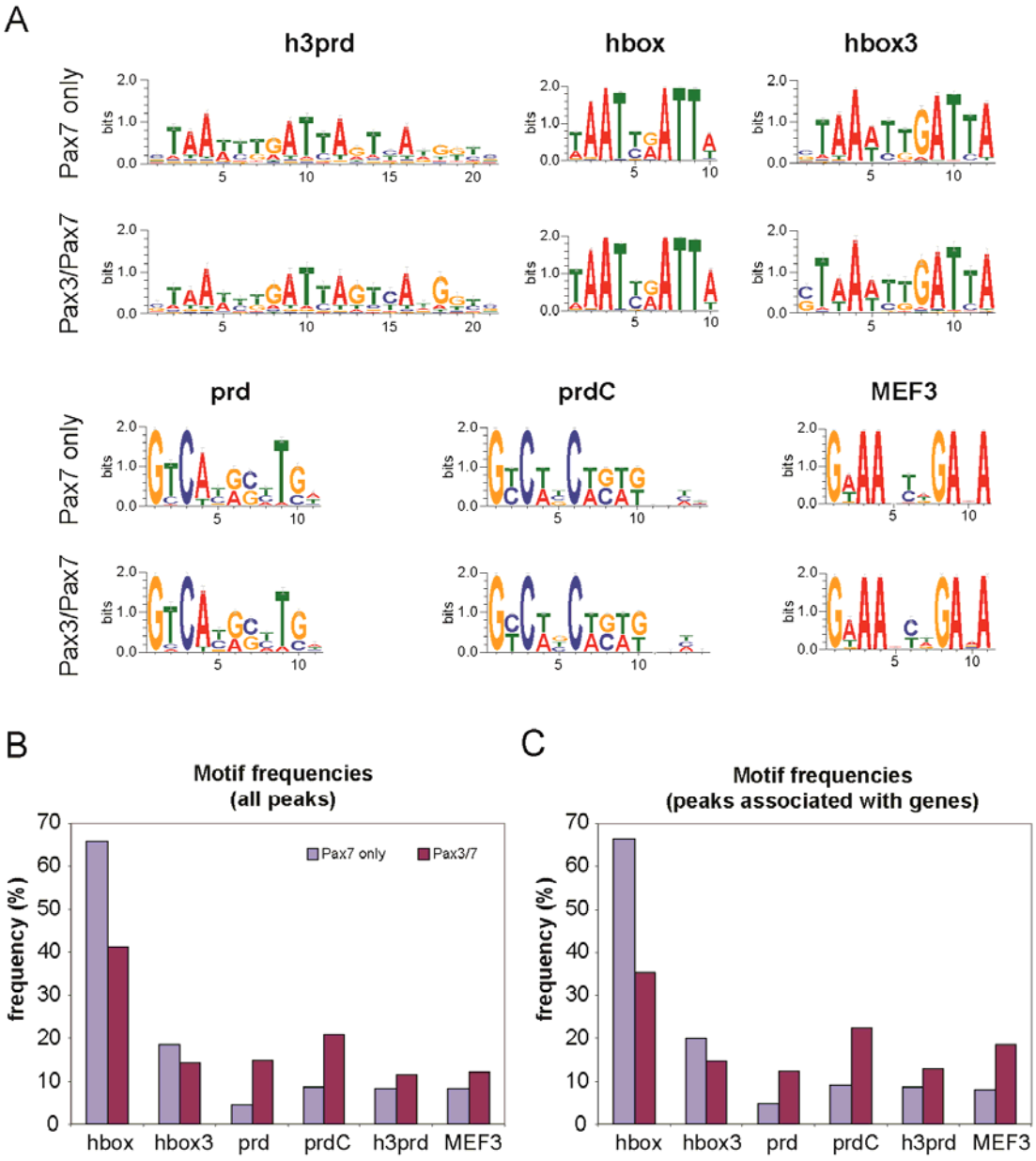

A
h3prd
hbox
hbox3
Pax7 only
Pax3/Pax7
prd
prdC
MEF3
bits
B
Motif frequencies
(all peaks)
Pax7 only
Pax3/7
frequency (%)
hbox
hbox3
prd
prdC
h3prd
MEF3
C
Motif frequencies
(peaks associated with genes)
frequency (%)
hbox
hbox3
prd
prdC
h3prd
MEF3

Figure 4. Pax3 and Pax7 Bind Common Motifs with Different Frequencies

(A) Top scoring peaks in either Pax3 or Pax7 were aligned by their peak center, truncated on either side, and searched *de novo* for motifs using MEME and GLAM2. Five motifs were found, including two variant homeobox motifs (*hbox*, *hbox3*), two paired box motifs (*prd*, *prdC*), and a single motif showing a juxtaposition of *hbox3* and *prd* motifs (*h3prd*). These resulting motif matrices were used to search Pax7 only and Pax7 peaks shared with Pax3 using FIMO and the resulting matches were assembled into consensus sequences. The seemingly identical composition of those recognized only by Pax7 and those also recognized by Pax3 suggests that differences in binding behavior are directed by sequence-independent mechanisms.
(B) The frequencies of each motif within peak sets were compared between Pax7 only and Pax3/7 shared peaks. Peaks unique to Pax7 contained a significantly higher incidence of homeobox motifs than those also recognized by Pax3. Conversely, peaks recognized by Pax3 contained a significantly higher incidence of paired motifs and MEF3 motifs.
(C) This observation is slightly mitigated in peaks +/- 10kb of gene transcriptional start sites. Together, these observations suggest Pax7 binding is biased towards homeodomain recognition, while Pax3 binding is biased towards paired domain recognition.

Motif frequencies were calculated for each peak subset as a percentage of total peaks (Figure 4B). Although we did not find any striking differences in motif composition though calculation of consensus sequences, peak groups differed strongly in motif frequencies. For Pax7-only peaks, 65.8% contained *hbox* motifs compared to 41.2% for Pax7 peaks overlapping with Pax3 peaks. Conversely, *prd* and *prdC* motifs were much more prevalent in overlapping peaks (14.8% and 20.8%, respectively) compared to single peaks (4.5% and 8.5%, respectively). *hbox3* motifs were slightly more common in Pax7-only peaks (18.4% versus 14.1%), while full *h3prd* motifs were slightly more common in overlapping peaks (8.3% versus 11.4%). Six1/4 motifs were also slightly more common in overlapping peaks (8.2% versus 12.1%) (Figure 4B). A Pearson Chi-square test on contingency table data indicated these differences were statistically significant (all $p < 1.0 \times 10^{-22}$).

We refined this analysis to examine only peaks proximal to Pax7-regulated genes to determine if motif frequency bias was a common effect across peaks in general or whether it was due to strong subsample bias. Peaks associated with genes (+/- 10 kb) displayed an increased contrast in their motif bias compared to all peaks (Figure 4C). Taken together, these data imply a strong bias of Pax7 towards homeodomain binding and Pax3 towards strong paired domain binding.

Differential Bias for Paired versus Homeodomain Binding by Pax3 and Pax7

The differences observed in motif binding frequency suggested that Pax3 and Pax7 have reciprocal affinities for binding paired motifs versus homeodomain motifs. Therefore, we examined the qualitative data for Pax3 and Pax7 peaks under the supposition that normalized peak height (calculated as the number of tags/100bp) is a reasonable metric for binding affinity. For each dataset, we divided the peaks into four subcategories; Pax3-only or Pax3-

shared, and Pax7-only and Pax7-shared binding sites. In addition, for each subset, we analyzed the distributions of binding sites containing only a homeodomain (*hbox* only), a paired domain (*prd* only), or both a paired domain and a homeodomain (*hbox+prd*).

Examination of sites that bound Pax3 and not Pax7, revealed 131 *prd*-only peaks, 36 *hbox*-only peaks, and 3 peaks containing both *hbox+prd*. However, the distributions of peak heights were not statistically different (*prd* vs *hbox,* $p = 0.39$; *prd* vs *hbox+prd,* $p = 0.94$; *hbox* vs *hbox+prd*, $p = 0.98$). Pax3 peaks overlapping with Pax7 sites contained 372 *prd*-only peaks, 1,272 *hbox*-only peaks, and 120 *hbox+prd* peaks, but these distributions were also not statistically different (*prd* vs *hbox*, $p = 0.23$; *prd* vs *hbox+prd*, $p = 0.04$; *hbox* vs *hbox+prd*, $p = 0.13$). These data indicate that while Pax3 unique sites more frequently contain paired motifs than homeobox motifs, the affinities for paired and homeoboxes are relatively equal (Figures 5A, 5B).

Conversely, sites uniquely bound by Pax7 contained 363 *prd*-only peaks, 11,224 *hbox*-only peaks, and 443 *hbox+prd* peaks. By KS test, *hbox*-only peaks had significantly higher peak height than *prd*-only peaks ($p < 2.2 \times 10^{-16}$), as did *hbox+prd* peaks ($p = 1.13 \times 10^{-12}$). However, *hbox* peaks were not significantly different in distribution of peak size than *hbox+prd* peaks ($p = 0.08$). Intriguingly, we saw a large majority of Pax7-only binding sites are characterized by exclusively homeodomain containing motifs.

Analysis of Pax7 binding sites that also bind Pax3, revealed 363 *prd*-only peaks, 1,259 *hbox*-only peaks, and 141 *hbox+prd* peaks. Notably, *hbox* and *hbox+prd* peaks had significantly higher peak scores than *prd* peaks ($p < 2.2 \times 10^{-16}$), but were not significantly different from each other in peak height distribution ($p = 0.12$). Of note, the majority of Pax7 sites that also bound Pax3 peaks contained homeodomain motifs. Overall, peak scores were significantly higher when the *hbox* motif was present. Collectively, these results suggest that

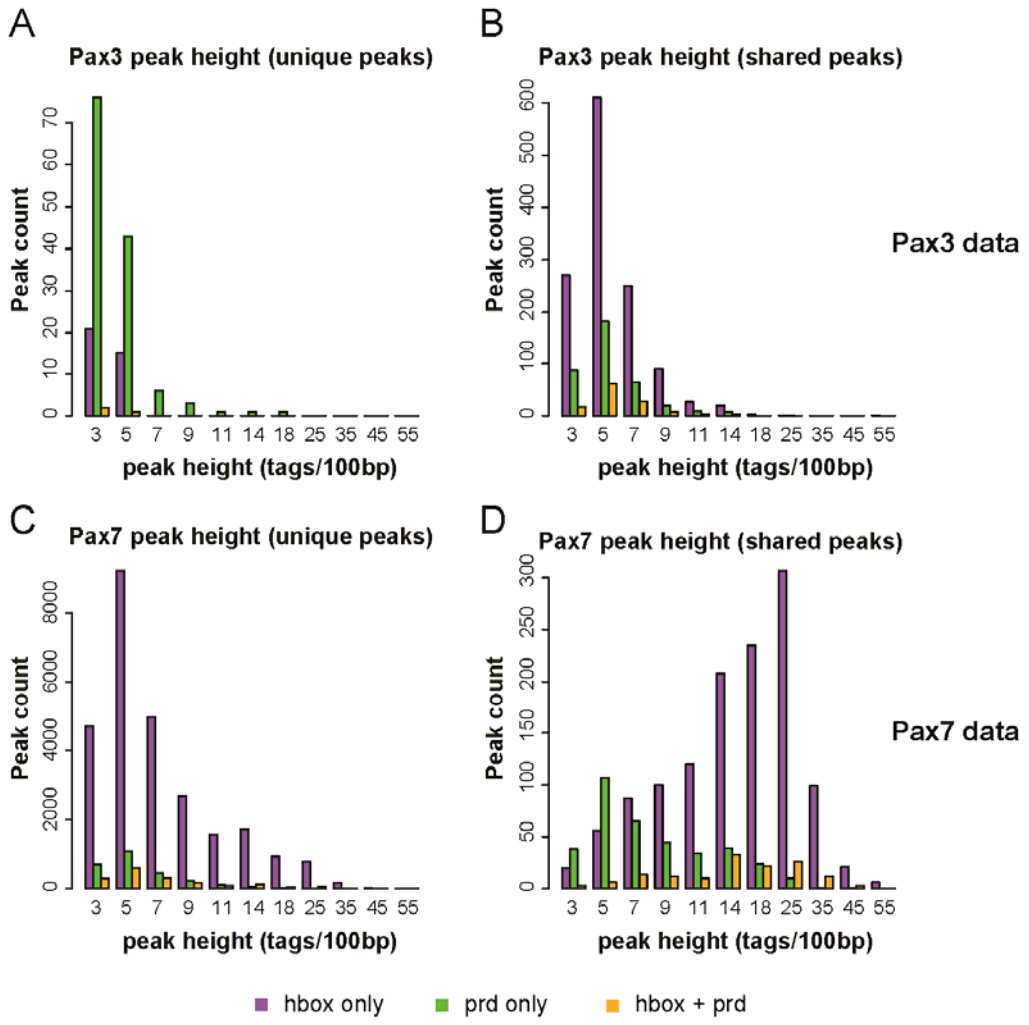


Figure 5. Pax3 and Pax7 Bind to Homeodomain and Paired Domain Motifs with Reciprocal Affinities

Pax3 and Pax7 peaks were classified by their inclusion of *prd* or *hbox* motifs and binned according to peak height, under the thesis that peak height (number of tags per 100bp) is a metric of binding affinity. Peak height categories are exclusive and category labels indicate the midpoint of the bin (i.e. Peak height "9" indicates the peak score is between 8 and 10). Pax3-only peaks contain a large proportion of *prd* motifs relative to *hbox* motifs. Conversely, Pax7-only peaks contain a large proportion of *hbox* motifs relative to *prd* motifs. In the Pax3 data, Pax3 peaks overlapping Pax7 peaks retained high counts of *prd* motifs, but displayed higher proportions of *hbox* motifs than Pax3 only peaks. In the Pax7 data, Pax7 peaks overlapping Pax3 peaks contained high numbers of both motifs, however higher affinity peaks were associated with *hbox*, while lower affinity peaks were associated with *prd*. Collectively, these data suggest that peak height correlates well with paired motifs in peaks specifically recognized by Pax3, but more so with homeobox motifs when recognized by Pax7.

Pax7 has a high binding affinity for homeodomains, and a relatively lower affinity for paired domains (Figure 5C, 5D). By contrast, Pax3 has a higher relative affinity for paired domains compared to Pax7, and thus more frequently binds sites containing paired domains.

We validated this hypothesis by EMSA using short oligonucleotides derived from a common Pax3/Pax7 binding site that contained both representative homeo and paired domains (Myf5-111; further described below). 30-bp oligomers were designed to span each binding motif and assayed for their affinity to bind purified Pax3-FLAG and Pax7-FLAG protein.

EMSA confirmed the different binding bias for homeobox versus paired motifs between Pax3 and Pax7. While Pax3 weakly bound to the homeobox probe, equivalent amounts of Pax7 protein binding resulted in a more robust shift of the probe (Figure 6A). Two species of shifted probe were identified, corresponding to expected monomer and dimer binding. This is consistent with previously published data that suggests homeobox repeats facilitate dimerization (Birrane et al., 2009). Conversely, Pax3 binding to the paired domain probe, was stronger than Pax7 (Figure 6B). These data support the different distributions of peak heights for Pax3 and Pax7 over homeo- and paired domain motifs observed in ChIP-seq data. Therefore, we conclude that Pax3 and Pax7 exhibit different motif affinities, with Pax3 biased towards paired domain binding, and Pax7 biased towards homeodomain binding.

Discussion

At the molecular level, our understanding of Pax7 activity is limited and it is unclear why the functions of Pax3 and Pax7 are divergent in adult muscle cells. Indeed, given the high amino acid sequence similarity between Pax3 and Pax7, and their ability to bind common regulatory sites in genes such as Myf5 (McKinnell et al., 2008), it has remained elusive as to

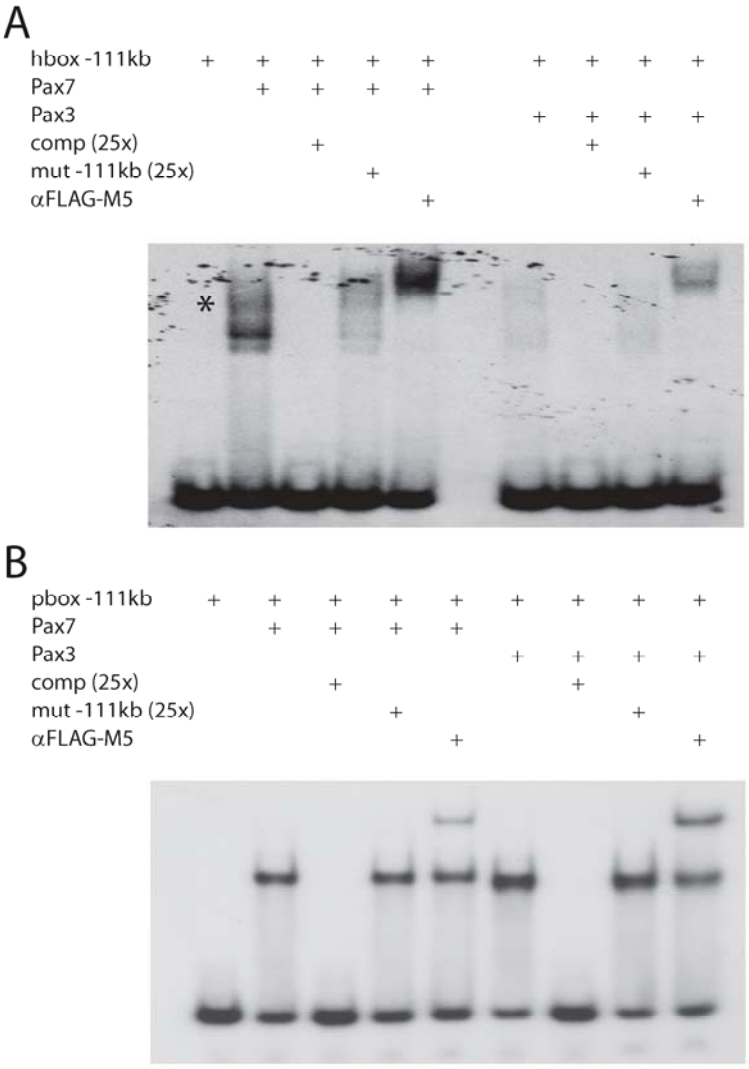


Figure 6. EMSA Analysis Indicates Pax7 is Biased Towards Homeodomains Whereas Pax3 is Biased Towards Paired Domains

Electrophoretic mobility shift assays (EMSA) were carried out on synthesized oligomers designed around the homeobox (*hbox*) and paired box (*pbox*) portions of the commonly bound Pax3/7 site -111 kb upstream of the *Myf5* gene.
(A) Binding at the *hbox* probe occurred as either a monomer or dimer (*) and a greater amount of probe was shifted in the presence of Pax7 compared to Pax3.
(B) Binding of the *pbox* probe yielded a single species that was more efficiently shifted in the presence of Pax3. Together, these data demonstrate that Pax3 and Pax7 have different binding affinities for *hbox* and *prd* motifs.

why Pax3 and Pax7 cannot functionally substitute for one another. However, whereas the transcriptional networks by which Pax3 regulates myogenic specification in embryonic tissues have been well studied, the transcriptional targets of Pax7 and how Pax7 specifies myogenic identity remains largely unknown (Punch et al., 2009). Although some candidate Pax7 target genes have been reported in the embryo (White and Ziman, 2008), a comprehensive analysis has not been carried out at the genome-wide level.

To investigate the basis for the genetic differences in Pax3 and Pax7 function, we ectopically expressed tagged versions Pax3 and Pax7 in satellite cell-derived primary myoblasts. We combined ultra-high throughput sequencing of ChIP DNA and gene expression analysis to identify genome-wide binding targets of Pax3 and Pax7 and to describe functional associations between binding sites and Pax7-regulated genes. Our data indicate that Pax3 binds to a small subset of Pax7 sites that more frequently contain paired domain motifs and that sites uniquely bound by Pax7 primarily contain homeodomain motifs.

Pax3 and Pax7 belong to a transcription factor family that has important and conserved roles in development. Gene replacement studies have demonstrated temporal and functional overlap Pax3 and Pax7 in both embryonic and adult tissue. However, both Pax3 and Pax7 have critical non-redundant roles in embryonic and adult muscle development, respectively (Kuang et al., 2006; Relaix et al., 2006; Relaix et al., 2004). Pax7 can competently replace Pax3 in neural crest cells, dorsal neural tube, and trunk muscles, but cannot compensate for Pax3 function in the delamination and migration of limb muscle progenitor cells (Relaix et al., 2004). Conversely, Pax7 alone is essential for satellite cells perinatally (Kuang et al., 2007; Oustanina et al., 2004; Relaix et al., 2006; Seale et al., 2000). Our results suggest that intrinsic differences between Pax3 and Pax7 affinities for paired-

versus homeodomain motifs are responsible for the different developmental roles played by Pax3 and Pax7

Additional mechanisms may differentially influence the DNA binding activity of Pax3 and Pax7. For example, we observed a number of examples where known embryonic Pax3 binding sites are only weakly associated with Pax3 in adult cells, but robustly bound by Pax7 (i.e. Fgfr4). DNA binding by Pax3 may require the presence of specific co-activators or post-translational modifications. Indeed, it has been suggested that Pax3 activity is dependent on phosphorylation, as Pax3 transcriptional activity is blocked by kinase inhibition (Amstutz et al., 2008; Miller et al., 2008). Pax3 is phosphorylated at Ser205, proximal to the octapeptide domain. Mutation of this site not only abolishes phosphorylation, but disrupts dimerization of Pax3 proteins. Homeodomain motifs are thought to facilitate dimerization through inverted repeats of "TAAT" and we see a substantial difference in affinity between Pax3 and Pax7 binding to homeodomain motifs.

Epigenetic modifications may also play a role in directing availability of binding sites to Pax3 or Pax7. For example, RARγ/RXRα heterodimers competently bind and induce retinoic acid response elements (RAREs) regulating Hoxa1 and Cyp26a1 in F9 teratocarcinoma stem cells, but not in Balb/c3T3 fibroblasts (Kashyap and Gudas, 2010). This is attributed to retinoic acid-induced reduction of polycomb protein Suz12 and associated H3K27 trimethylation of Hoxa1 and Cyp26a1 RAREs.

Our data support the notion that gene networks uniquely required by satellite cells and their daughter myoblasts are specifically regulated by Pax7. Pax7's unique ability to induce BMP4 and Wnt7a suggest a possible mechanism by which Pax7 functions in the survival of satellite cell populations. Wnt7a is known to drive the symmetric expansion of satellite cells through a non-canonical mechanism dependent on Fzd7 and PCP signaling (Le

Grand et al., 2009). Additionally, we saw an increase in the expression of PCP components Prickle2 and Vangl2. During embryogenesis, Wnt7a directly activates MyoD through a canonical Wnt signaling pathway (Tajbakhsh et al., 1998), yet we did not observe an increase in MyoD expression in either Pax3 or Pax7 cells. Bmp4 is well characterized in its ability to block the differentiation of myogenic cells (Dahlqvist et al., 2003; Katagiri et al., 1994). Therefore, it is unsurprising that Pax7-induced expression of Bmp4 correlates with the strong repression of myogenic differentiation genes such as myogenin and Mef2c. The observation that Cav1 is repressed by Pax7 is also curious. Cav1 is expressed in satellite cells, however, its overexpression delays muscle regeneration and impedes cellular proliferation (Volonte et al., 2005). The inhibition of Cav1 may indicate further aid Pax7 in promoting expansion of satellite cell populations.

Autocrine signaling may be a common mechanism exploited by Pax7 to maintain stem cell populations. Recently, it was shown that Angpt1 and its receptor, Tie2, function to inhibit proliferation and differentiation and promote a return to quiescence of satellite cells (Abou-Khalil et al., 2009). Angpt1 was strongly inhibited by overexpression of either Pax3 or Pax7, which may be reflected in the promotion of cell proliferation in cultured satellite cells. A strong induction of membrane-bound receptors and signaling molecules by Pax7 may indicate a role in other autocrine mechanisms that serve to maintain satellite cell reservoirs.

Our experiments indicate that in adult myoblasts, Pax7 is a more active transcription factor than Pax3. We propose that many Pax7-only binding sites are regulatory elements for genes essential to the normal function of adult myogenic cells, while their low affinity to Pax3 explains the inability of Pax3 to compensate for Pax7 in *Pax7*$^{-/-}$ muscles, such as the diaphragm. Ultimately, the unique characteristics of Pax3 and Pax7 binding provide a

framework for understanding the differences in transcriptional network organization between embryonic and adult myogenesis. Both factors may be sufficient to initiate myogenic programs, but different developmental contexts may utilize alternative networks to fulfill other essential roles.

Our data support the notion that gene networks uniquely required by satellite cells and their daughter myoblasts are specifically regulated by Pax7. By contrast, genetic analysis is consistent with the notion that Pax3 plays an essential role in embryonic myogenesis. Together, this raises the potential that cell-context is an important parameter in the analysis of Pax3 and Pax7 function and identification of targets. Future experiments investigating Pax3 binding in early somatic myocytes and progenitors would certainly clarify this issue.

Our affinity data provide a model by which paralogous proteins such as Pax3 and Pax7 provide specific responses through biased affinity towards common regulatory elements. Our comprehensive binding maps of Pax3 and Pax7 can facilitate the discovery of novel gene interactions and may provide novel approaches towards genetic programming of muscle stem cells for use in the amelioration of muscle diseases.

Experimental Procedures

Cell Culture and Retroviral Infection

Primary myoblasts were isolated from adult C57BL/6 mice as described previously (Huh et al., 2004) and cultured on collagen-coated plates (rat-tail collagen, Roche-Boehringer) in Ham's complete medium (Ham's F-10, 20% fetal calf serum, 1% penicillin/streptomycin, and 2.5ng/μl human recombinant bFGF (Invitrogen)). Expression plasmids (Pax3-TAP, Pax7-TAP, Pax7-FLAG, EGFP, and empty vectors) were constructed in a modified pHAN

backbone containing a SV40-driven puromycin resistance element. Pax7-FLAG was created by fusing mouse Pax7d in-frame with a carboxy-terminal 3x FLAG tag. Pax3- and Pax7-TAP were engineered with the in-frame addition of a c-terminal TAP tag, composed of 6 histidine residues and a 3x FLAG tag separated by a Tobacco Etch Virus (TEV) cleavage sequence. Retrovirus was produced in 293FT cells by transient transfection of expression constructs using Lipofectamine 2000 reagent (Invitrogen). Virus-containing supernatant was used to infect low-passage primary myoblasts supplemented with Ham's complete medium and 8ug/ml polybrene. Transgenic cells were obtained by antibiotic selection and maintained in Ham's complete medium containing 0.5ug/ml puromycin.

Chromatin Immunoprecipitation and Tandem Affinity ChIP (ChTAP)

ChIP was modified from previously described methods (Filippova et al., 2001). For ChTAP experiments, $1x10^7$ subconfluent primary myoblasts stably expressing Pax3-TAP, Pax7-TAP, or empty TAP vector were fixed with a 1% formaldehyde fixation solution for 10 min and cell pellets were flash frozen in liquid nitrogen. Cell lysates were diluted in ChIP lysis buffer (40 mM Tris-Cl pH 8; 1% Triton-X; 4 mM EDTA; 0.3 M NaCl) to 5 mg/ml of protein and sonicated for 45 minutes at 30s-ON/60s-OFF intervals in a Bioruptor (Diagenode) to obtain DNA fragments ranging from 100bp to 300bp. Lysates were further diluted to 15mL and TAP-tagged complexes were immunoprecipitated overnight with 200µl of monoclonal anti-FLAG M2-agarose (Sigma). Complexes were removed from M2-agarose beads by incubation in TEV cleavage buffer (0.5% NP40, 0.15mg/ml 3X FLAG peptide, 20U AcTEV protease in TBS). TEV eluate was purified on HIS-Select HF Nickel Affinity Gel (Sigma) and eluted with 400mM imidazole. Following reversal of crosslinks and proteinase K

treatment, ChIP DNA was purified by phenol:choloform extraction and resuspended in ddH_2O.

ChIP experiments with Pax7-FLAG, EGFP, and puro cell lines were fixed in 1% formaldehyde fixation buffer. Following immunoprecipitation with M2-agarose, wash conditions and DNA elution were performed according to the ChIP Assay Kit (Upstate). Enrichment of particular loci within ChIP DNA pools was determined by quantitative real-time PCR.

Gene Expression Analysis

Total RNA was purified from triplicate samples of Pax7 siRNA and control siRNA transfected cells or from stable Pax3- and Pax7-TAP cells using the RNEasy kit (Qiagen). Transfections were performed using Lipofectamine reagent (Invitrogen) using 30nm siRNA. ON-TARGET plus siRNAs (Dharmacon) used were Pax7 #2 (5' – CCAAGAUUCUGUGCCGAUA – 3') and #3 (5' – UGACCAAUGUACACCGAUU – 3') and Non-targeting siRNA #1 (D-001810-01-05).

Probes for the Mouse Gene Array ST 1.0 (Affymetrix) were prepared from 150ng RNA according to the manufacturer's instructions. cRNA probes were hybridized to Affymetrix Mouse Gene Array ST 1.0 chips and scanned using a GeneChip Scanner 3000 7G. Raw expression data was assembled using the Affymetrix GeneChip Command Console v1.1.

Selected targets from ChIP and microarray experiments were validated by PCR with a Mx3000P Real Time PCR System (Stratagene) using B-R SYBR Green Supermix for iQ (Quanta Biosciences). PCR primers for qPCR were designed using the online Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi; Whitehead Institute).

Primers for gene expression analysis were designed to span at least one intron where possible. Gene expression values were normalized to GAPDH. Primers for ChIP analysis were designed to generate products of 100-200bp. ChIP values were normalized to a control locus -3kb from GAPDH TSS (chosen from a panel of negative control loci) and expressed as fold enrichment over control IP through ΔCt calculation. Primer sequences are listed in Table S3.

For Western analysis, 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 (10-25 μg 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) with Biomax XAR film (Kodak). Primary antibodies used were: anti-FLAG (M2; 1:1000; 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 (1:1000; Sigma). Secondary antibodies were HRP-conjugated anti-mouse and anti-rabbit (1:4000; Biorad).

Association of Genes with Binding Sites

The list of TCIDs, ranked by raw p-value, was annotated based on the MoGene-1_0-st-v1.na30.1.mm9 transcript annotations provided by Affymetrix. TCIDs with zero or multiple ENSEMBL gene identifiers were stripped from the list, and the list was further filtered to limit it to known protein coding ENSEMBL genes present in the mouse ENSEMBL build 55_37h (Hubbard et al., 2009) based on NCBI m37(Church et al., 2009), with raw p-values <

0.01. The resultant significantly up and down regulated gene set was then mapped to associated ChIP-Seq peaks in the Pax7 data whose peak centers were from -10kb to +10kb base pairs relative to ENSEMBL gene start, which corresponds to the beginning of the earliest transcript associated with that gene. The fold change (Y-axis) was derived from the R fold change (R.fold) provided by SAM. Fold changes <1 (negative fold change) are notated as -1/R.fold. In Figure 3a, the size of the circle is proportional to the log of the MACS peak score, and the color is proportional to the log scaled value of the effective conservation score.

Motif Analysis

MEME and GLAM2 (Bailey et al., 2009) were used to discover over-represented motifs in subsets of the peak sequences identified. Peaks were filtered by width, MACS score (-10 * $\log_{10}$(pvalue)) and repeat content to select regions considered most likely to reflect specific high-affinity binding events, and least likely to identify spurious motifs within repetitive sequences. For motif identification, peaks were limited to 600bp width to eliminate adjacent overlapping binding sites, and repeat content was restricted to 50% (determined by RepeatMasker annotations obtained from UCSC). GLAM2 was used in addition to MEME as it allows the identification of gapped motifs, such as the predicted binding of homeo- and paired domains by a Pax family TF. The MEME *maxw* parameter was set at 24 or 30 in order to allow the identification of wider motifs; other parameters were set appropriately for the size of input sequence file. Other parameters were left at their default settings.

Peak Motif Analysis

A set of five motifs was assembled from MEME and GLAM2 runs, and full sequences for all peaks ≤800bp in width from both Pax3 and Pax7 were extracted. FIMO (Bailey et al., 2009)

was used to search for those motif instances in the peak sequences. A file of motif sequences was generated for each subset of peaks (Pax7 only, Pax3 only, Pax7 & Pax3) and WebLogo (Crooks et al., 2004) was used to generate a representation of the motifs found in each peak subset.

Electrophoretic Mobility Shift Assays (EMSA)

All probes were synthesized as full-length oligonucleotides, purified by HPLC, unless stated otherwise by Integrated DNA Technologies. 3µl of 10µM single stranded probes (+) were radiolabeled by incubation with T4 Kinase (Invitrogen) in the presence of 3µl γ-[^{32}P]-ATP for 1 hour at 37'C. Labeling was stopped by addition of STE buffer (40mM Tris-Cl pH 8.0, 40mM NaCl, 0.8mM EDTA). Double stranded probes were generated by hybridizing with 6µl of 10µM opposite strand probe (-), boiling for 2 minutes, and slowly cooled to room temperature over 2 hours. Unincorporated [^{32}P]-ATP was removed by purifying double stranded probes through Microspin G25 columns (GE Healthcare). Competitor probes were synthesized by combining 1µl of each single stranded probe (100µM) with 5µl STE buffer and 1µl water, boiling for 2 minutes and slowly cooling to room temperature. 0.5 µg of baculovirus-purified Pax3-FLAG or Pax7-FLAG was incubated with 1.25ng DNA probe and nonspecific carrier (poly dI-dC) in a binding buffer containing 75mM NaCl, 1mM EDTA, 1mM DTT, 10mM Tris pH 7.5, 6% glycerol, and 0.25% BSA at room temperature for 20 minutes. 25x cold probe was used for competitive binding assays. Supershifts were carried out with the addition of 1ug M2- or M5-FLAG antibody. Reactions were analyzed on a 5% nondenaturing polyacrylamide gel (0.5X TBE), dried onto 3MM Whatman paper and exposed to BioMax MS x-ray film (Kodak). Probe sequences are listed in Table S3.

See Supplemental Experimental Procedures for a more extensive description of procedures and computational methods.

Acknowledgements

M.A.R. holds the Canada Research Chair in Molecular Genetics and is an International Research Scholar of the Howard Hughes Medical Institute. This work was supported by grants to M.A.R. from the National Institutes of Health, the Howard Hughes Medical Institute, the Canadian Institutes of Health Research, and the Canada Research Chair Program. The authors declare no conflict of interest.

References

Abou-Khalil, R., Le Grand, F., Pallafacchina, G., Valable, S., Authier, F.J., Rudnicki, M.A., Gherardi, R.K., Germain, S., Chretien, F., Sotiropoulos, A., et al. (2009). Autocrine and paracrine angiopoietin 1/Tie-2 signaling promotes muscle satellite cell self-renewal. Cell Stem Cell 5, 298-309.

Amstutz, R., Wachtel, M., Troxler, H., Kleinert, P., Ebauer, M., Haneke, T., Oehler-Jänne, C., Fabbro, D., Niggli, F.K., and Schäfer, B.W. (2008). Phosphorylation Regulates Transcriptional Activity of PAX3/FKHR and Reveals Novel Therapeutic Possibilities, pp. 3767-3776.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37, W202-208.

Bajard, L., Relaix, F., Lagha, M., Rocancourt, D., Daubas, P., and Buckingham, M.E. (2006). A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. Genes Dev 20, 2450-2464.

Birrane, G., Soni, A., and Ladias, J.A.A. (2009). Structural Basis for DNA Recognition by the Human PAX3 Homeodomain. Biochemistry 48, 1148-1155.

Bober, E., Franz, T., Arnold, H.H., Gruss, P., and Tremblay, P. (1994). Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. Development 120, 603-612.

Bois, P.R., Izeradjene, K., Houghton, P.J., Cleveland, J.L., Houghton, J.A., and Grosveld, G.C. (2005). FOXO1a acts as a selective tumor suppressor in alveolar rhabdomyosarcoma. J Cell Biol 170, 903-912.

Brunelli, S., Relaix, F., Baesso, S., Buckingham, M., and Cossu, G. (2007). Beta catenin-independent activation of MyoD in presomitic mesoderm requires PKC and depends on Pax3 transcriptional activity. Dev Biol 304, 604-614.

Buckingham, M., and Relaix, F. (2007). The role of Pax genes in the development of tissues and organs: Pax3 and Pax7 regulate muscle progenitor cell functions. Annu Rev Cell Dev Biol 23, 645-673.

Church, D.M., Goodstadt, L., Hillier, L.W., Zody, M.C., Goldstein, S., She, X., Bult, C.J., Agarwala, R., Cherry, J.L., DiCuccio, M., et al. (2009). Lineage-Specific Biology Revealed by a Finished Genome Assembly of the Mouse. PLoS Biol 7, e1000112.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res 14, 1188-1190.

Dahlqvist, C., Blokzijl, A., Chapman, G., Falk, A., Dannaeus, K., Ibanez, C.F., and Lendahl, U. (2003). Functional Notch signaling is required for BMP4-induced inhibition of myogenic differentiation. Development 130, 6089-6099.

Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 4, P3.

Epstein, J.A., Shapiro, D.N., Cheng, J., Lam, P.Y., and Maas, R.L. (1996). Pax3 modulates expression of the c-Met receptor during limb muscle development. Proc Natl Acad Sci U S A 93, 4213-4218.

Filippova, G.N., Thienes, C.P., Penn, B.H., Cho, D.H., Hu, Y.J., Moore, J.M., Klesert, T.R., Lobanenkov, V.V., and Tapscott, S.J. (2001). CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. Nat Genet 28, 335-343.

Giordani, J., Bajard, L., Demignon, J., Daubas, P., Buckingham, M., and Maire, P. (2007). Six proteins regulate the activation of Myf5 expression in embryonic mouse limbs. Proc Natl Acad Sci U S A 104, 11310-11315.

Hubbard, T.J., Aken, B.L., Ayling, S., Ballester, B., Beal, K., Bragin, E., Brent, S., Chen, Y., Clapham, P., Clarke, L., et al. (2009). Ensembl 2009. Nucleic Acids Res 37, D690-697.

Huh, M.S., Parker, M.H., Scime, A., Parks, R., and Rudnicki, M.A. (2004). Rb is required for progression through myogenic differentiation but not maintenance of terminal differentiation. J Cell Biol 166, 865-876.

Hutcheson, D.A., Zhao, J., Merrell, A., Haldar, M., and Kardon, G. (2009). Embryonic and fetal limb myogenic cells are derived from developmentally distinct progenitors and have different requirements for beta-catenin. Genes Dev 23, 997-1013.

Kashyap, V., and Gudas, L.J. (2010). Epigenetic regulatory mechanisms distinguish retinoic acid-mediated transcriptional responses in stem cells and fibroblasts. J Biol Chem 285, 14534-14548.

Kassar-Duchossoy, L., Giacone, E., Gayraud-Morel, B., Jory, A., Gomes, D., and Tajbakhsh, S. (2005). Pax3/Pax7 mark a novel population of primitive myogenic cells during development. Genes Dev 19, 1426-1431.

Katagiri, T., Yamaguchi, A., Komaki, M., Abe, E., Takahashi, N., Ikeda, T., Rosen, V., Wozney, J.M., Fujisawa-Sehara, A., and Suda, T. (1994). Bone morphogenetic protein-2 converts the differentiation pathway of C2C12 myoblasts into the osteoblast lineage. J Cell Biol 127, 1755-1766.

Kuang, S., Charge, S.B., Seale, P., Huh, M., and Rudnicki, M.A. (2006). Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. J Cell Biol 172, 103-113.

Kuang, S., Kuroda, K., Le Grand, F., and Rudnicki, M.A. (2007). Asymmetric self-renewal and commitment of satellite stem cells in muscle. Cell 129, 999-1010.

Lagha, M., Kormish, J.D., Rocancourt, D., Manceau, M., Epstein, J.A., Zaret, K.S., Relaix, F., and Buckingham, M.E. (2008). Pax3 regulation of FGF signaling affects the progression of embryonic progenitor cells into the myogenic program. Genes Dev 22, 1828-1837.

Le Grand, F., Jones, A.E., Seale, V., Scime, A., and Rudnicki, M.A. (2009). Wnt7a activates the planar cell polarity pathway to drive the symmetric expansion of satellite stem cells. Cell Stem Cell 4, 535-547.

Lepper, C., Conway, S.J., and Fan, C.M. (2009). Adult satellite cells and embryonic muscle progenitors have distinct genetic requirements. Nature 460, 627-631.

Ma, R., Rundle, D., Jacks, J., Koch, M., Downs, T., and Tsiokas, L. (2003). Inhibitor of Myogenic Family, a Novel Suppressor of Store-operated Currents through an Interaction with TRPC1, pp. 52763-52772.

Maroto, M., Reshef, R., Munsterberg, A.E., Koester, S., Goulding, M., and Lassar, A.B. (1997). Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue. Cell 89, 139-148.

McKinnell, I.W., Ishibashi, J., Le Grand, F., Punch, V.G., Addicks, G.C., Greenblatt, J.F., Dilworth, F.J., and Rudnicki, M.A. (2008). Pax7 activates myogenic genes by recruitment of a histone methyltransferase complex. Nature cell biology 10, 77-84.

Miller, P.J., Dietz, K.N., and Hollenbach, A.D. (2008). Identification of serine 205 as a site of phosphorylation on Pax3 in proliferating but not differentiating primary myoblasts. Protein Sci 17, 1979-1986.

Oustanina, S., Hause, G., and Braun, T. (2004). Pax7 directs postnatal renewal and propagation of myogenic satellite cells but not their specification. Embo J 23, 3430-3439.

Park, M., and Moon, R.T. (2002). The planar cell-polarity gene stbm regulates cell behaviour and cell fate in vertebrate embryos. Nature cell biology 4, 20-25.

Parker, M.H., Seale, P., and Rudnicki, M.A. (2003a). Looking back to the embryo: defining transcriptional networks in adult myogenesis. Nat Rev Genet 4, 497-507.

Parker, M.H., Seale, P., and Rudnicki, M.A. (2003b). Looking Back to the Embryo: Transcriptional Networks in Adult Myogenesis. Nat Reviews Gen in press.

Punch, V.G., Jones, A.E., and Rudnicki, M.A. (2009). Transcriptional networks that regulate muscle stem cell function. Wiley Interdisciplinary Reviews: Systems Biology and Medicine 1, 128-140.

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

Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. (2004). Divergent functions of murine Pax3 and Pax7 in limb muscle development. Genes Dev 18, 1088-1105.

Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. (2005). A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. Nature 435, 948-953.

Seale, P., Sabourin, L.A., Girgis-Gabardo, A., Mansouri, A., Gruss, P., and Rudnicki, M.A. (2000). Pax7 is required for the specification of myogenic satellite cells. Cell 102, 777-786.

Sun, L., Ma, K., Wang, H., Xiao, F., Gao, Y., Zhang, W., Wang, K., Gao, X., Ip, N., and Wu, Z. (2007). JAK1-STAT1-STAT3, a key pathway promoting proliferation and preventing premature differentiation of myoblasts. J Cell Biol 179, 129-138.

Tajbakhsh, S., Borello, U., Vivarelli, E., Kelly, R., Papkoff, J., Duprez, D., Buckingham, M., and Cossu, G. (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.

Tajbakhsh, S., and Cossu, G. (1997). Establishing myogenic identity during somitogenesis. Current opinion in genetics & development 7, 634-641.

Tsukamoto, K., Nakamura, Y., and Niikawa, N. (1994). Isolation of two isoforms of the PAX3 gene transcripts and their tissue-specific alternative expression in human adult tissues. Hum Genet 93, 270-274.

Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 98, 5116-5121.

Vogan, K.J., Underhill, D.A., and Gros, P. (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, 6677-6686.

Volonte, D., Liu, Y., and Galbiati, F. (2005). The modulation of caveolin-1 expression controls satellite cell activation during muscle repair. Faseb J 19, 237-239.

White, R.B., and Ziman, M.R. (2008). Genome-wide discovery of Pax7 target genes during development. Physiol Genomics 33, 41-49.

Zammit, P.S., Carvajal, J.J., Golding, J.P., Morgan, J.E., Summerbell, D., Zolnerciks, J., Partridge, T.A., Rigby, P.W., and Beauchamp, J.R. (2004). Myf5 expression in satellite cells and spindles in adult muscle is controlled by separate genetic elements. Dev Biol 273, 454-465.

Zhang, Y., Liu, T., Meyer, C., Eeckhoute, J., Johnson, D., Bernstein, B., Nussbaum, C., Myers, R., Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biology 9, R137.

Ziman, M.R., and Kay, P.H. (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.

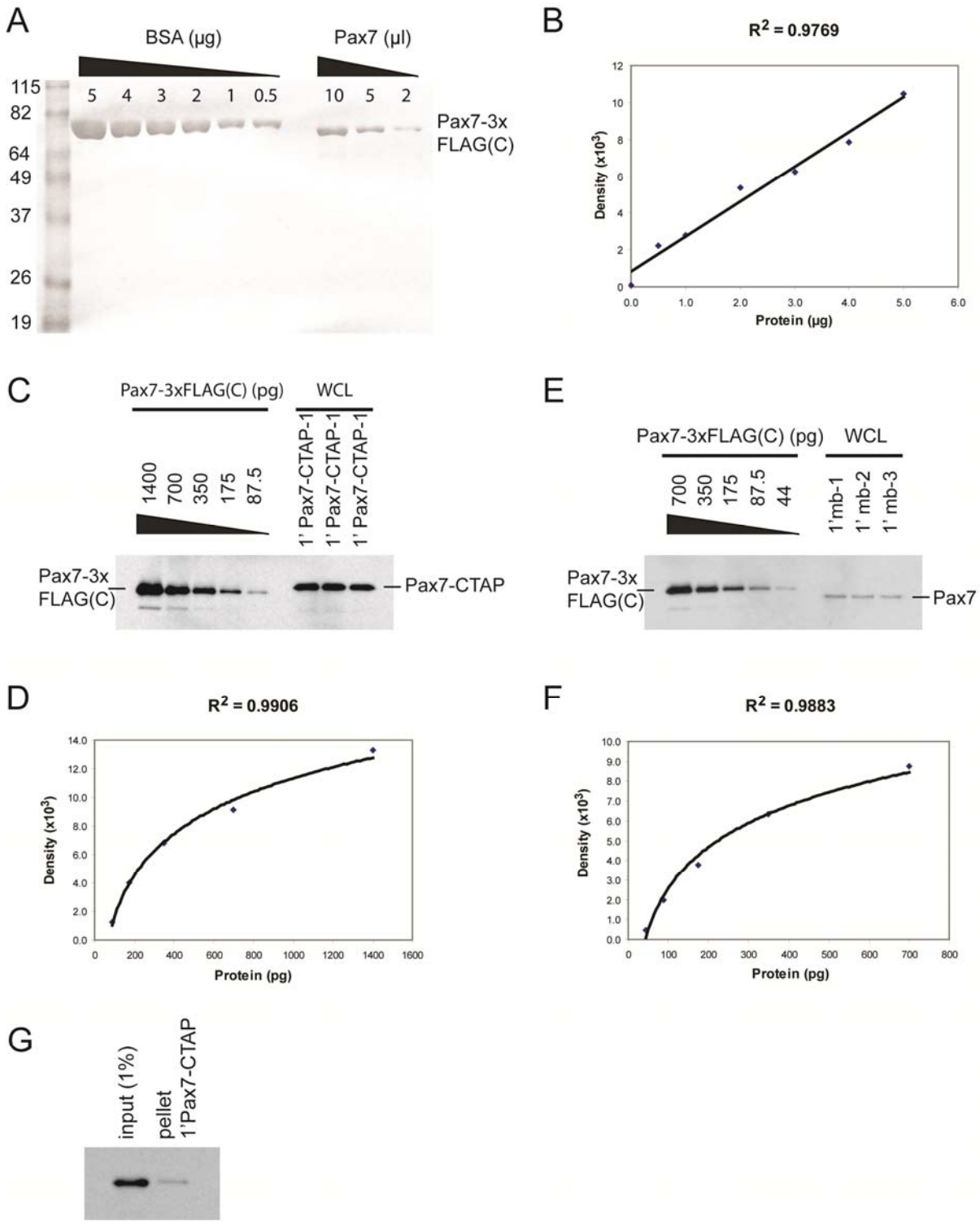

A
BSA (µg)
Pax7 (µl)
5 4 3 2 1 0.5
10 5 2
115
82
64
49
37
26
19
Pax7-3x FLAG(C)
B
R2 = 0.9769
Density (x10³)
Protein (µg)
C
Pax7-3xFLAG(C) (pg)
WCL
1400 700 350 175 87.5
1' Pax7-CTAP-1
1' Pax7-CTAP-1
1' Pax7-CTAP-1
Pax7-3x FLAG(C)
Pax7-CTAP
D
R2 = 0.9906
Density (x10³)
Protein (pg)
E
Pax7-3xFLAG(C) (pg)
WCL
700 350 175 87.5 44
1'mb-1
1' mb-2
1' mb-3
Pax7-3x FLAG(C)
Pax7
F
R2 = 0.9883
Density (x10³)
Protein (pg)
G
input (1%)
pellet 1'Pax7-CTAP

Figure S1. Quantification of Number of Molecules of Pax7 in Primary Myoblasts

(A) Coomassie Brilliant Blue staining of varying quantities of BSA loaded against an unknown concentration of purified Pax7-FLAG protein generated by baculoviral infection of Sf9 cells.
(B) Protein quantities were calculated by density scanning of exposed films using ImageJ software. Interpolation of the linear equation determined that purified Pax7-FLAG protein was 175 ng/µl.
(C) Protein extracts from 5,600 Pax7-TAP cells were loaded against known quantities of purified Pax7-FLAG protein.
(D) Interpolation of the logarithmic curve estimated Pax7-TAP at 7.0×10^5 molecules per cell.
(E) Protein extracts from 24,000 wild-type myoblasts were loaded against known quantities of Pax7-FLAG.
(F) Interpolation of the logarithmic curve estimated 2.75×10^4 molecules per cell of endogenous Pax7.
(G) Western blot depicting complete lysis of Pax7-CTAP cells used for quantification of Pax7 protein.

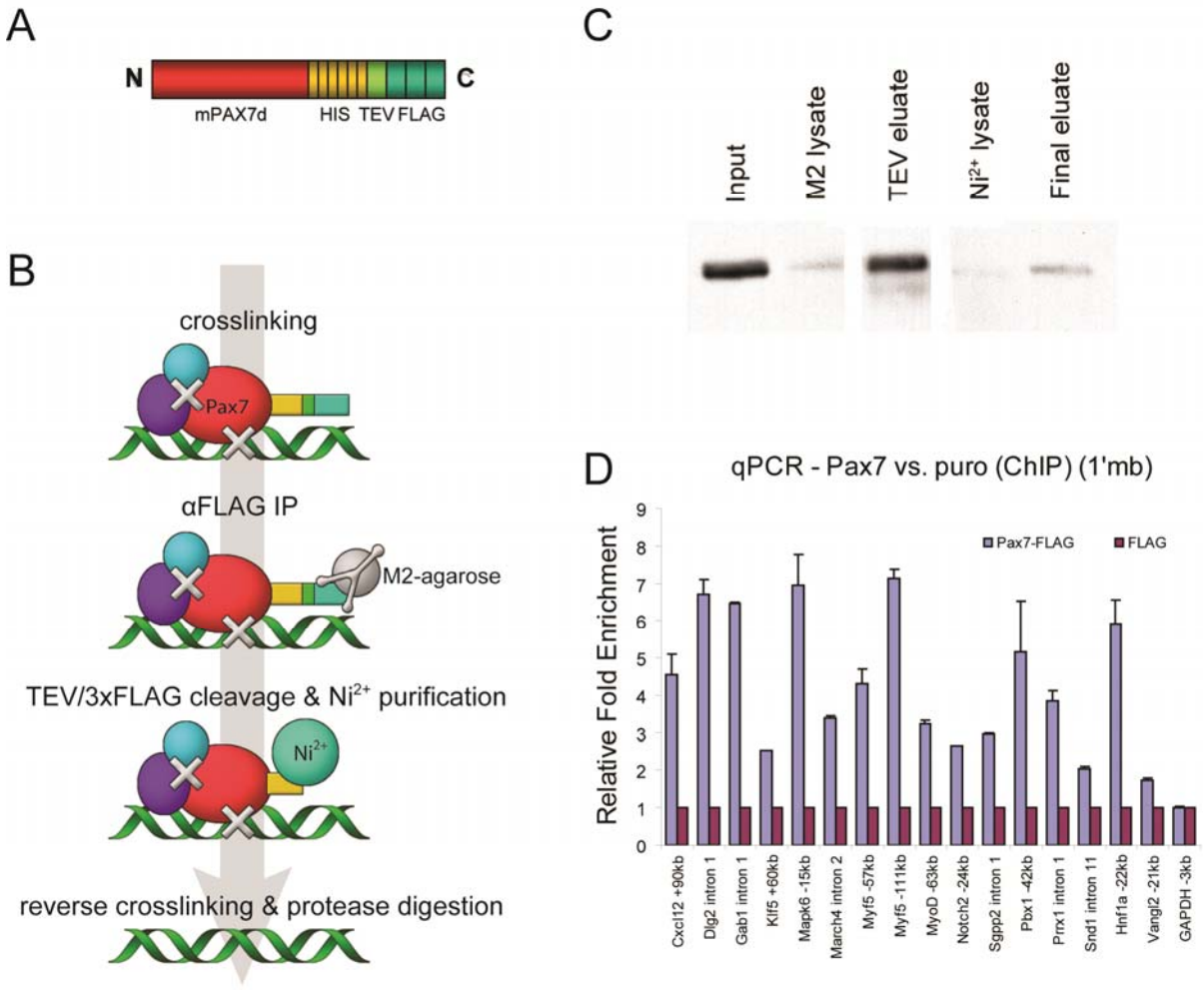


Figure S2. Tandem Affinity Purification of TAP-tagged Pax3 and Pax7 Chromatin

(A) Schematic for the TAP tag fused in-frame with Pax3 or Pax7. The TAP tag is comprised of six copies of HIS and three copies of FLAG epitopes, separated by a Tobacco Etch Virus (TEV) cleavage signal, fused to the C-terminal end of Pax3/7. Cell lines were generated by antibiotic selection of retrovirally-transduced primary myoblasts.
(B) Schematic of the ChIP-TAP protocol. Following immunoprecipitation with anti-FLAG (M2)-conjugated agarose, bound complexes were washed and enzymatically removed by incubation with TEV protease and 3xFLAG peptide. Subsequent purification was performed on a Ni2+ column, and associated DNA fragments were recovered by phenol:chloroform extraction following reversal of crosslinks and protease digestion.
(C) Western blot showing Pax7-TAP at various stages of TAP purification.
(D) Quantitative PCR validation of several Pax7-bound loci identified by ChIP-Seq. Independent ChIP was performed on primary myoblasts expressing Pax7-FLAG or FLAG alone according to protocols provided by Upstate. Fold enrichment was calculated by ΔCt, normalized to GAPDH -3 kb locus.

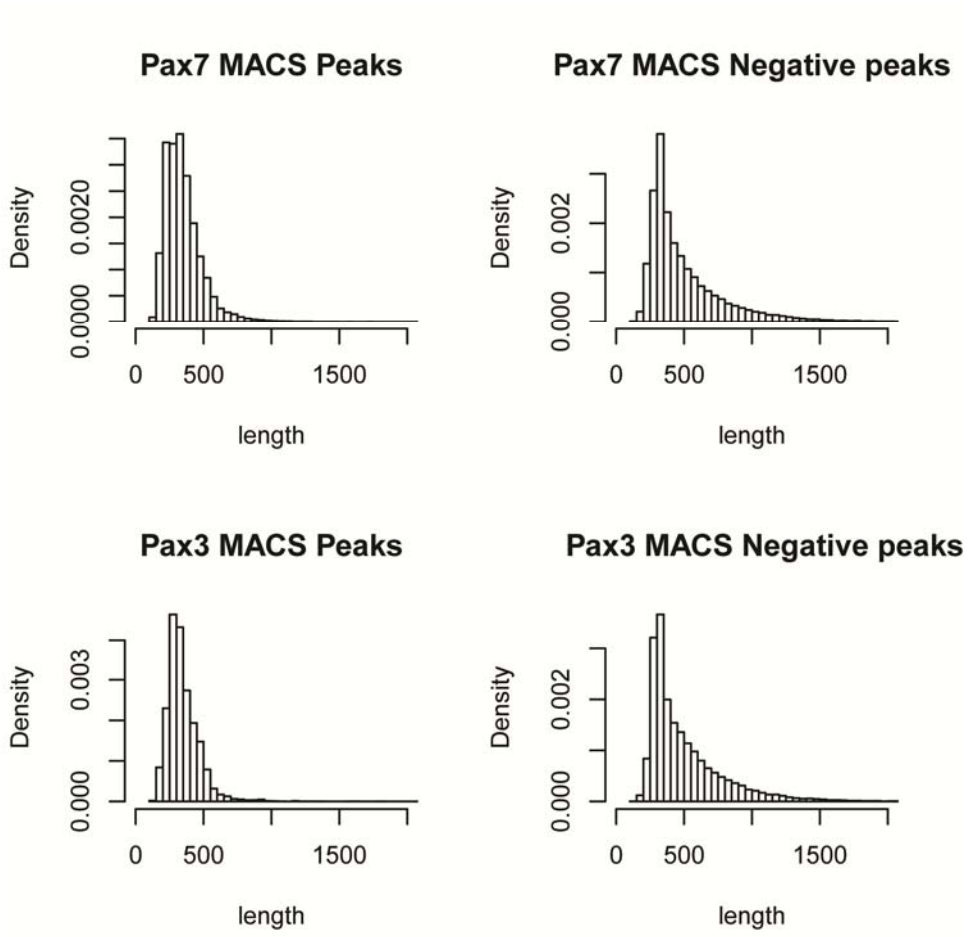


Figure S3. Peak Characterization and Normalization

To interpret differences in peak properties between datasets, we had to ensure peak distributions were comparable. Shown are normalized density plots for peak lengths. Negative peaks for both Pax3 and Pax7 tended to be broader, and an 800bp filter was selected to shape negative peak distributions to be similar to those for Pax3 and Pax7 peaks.

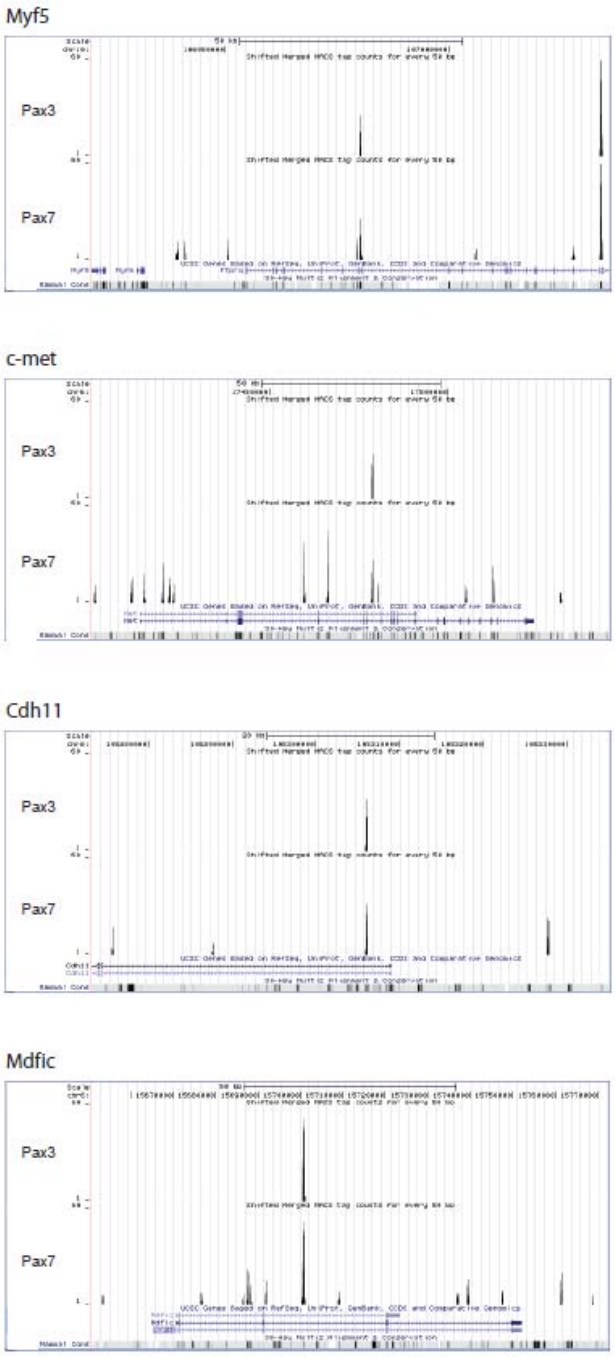


Figure S4. Binding of Pax3 and Pax7 to Common Sites in Myogenic Genes

Pax3/7 peaks were mapped to the UCSC genome browser, scaled to equivalent heights. Shown are examples of common binding around muscle-related genes. Pax7 appears to target multiple additional sites, however, common sites are frequently associated with the highest local MACS scores. Upstream of Myf5, both Pax3 and Pax7 are observed at the -57kb enhancer, utilized to direct Myf5 expression in limb buds. A larger, more prominent peak, however, is observed at -111kb, residing within a control region (Zammit et al., 2004) known to direct Myf5 expression in satellite cells.

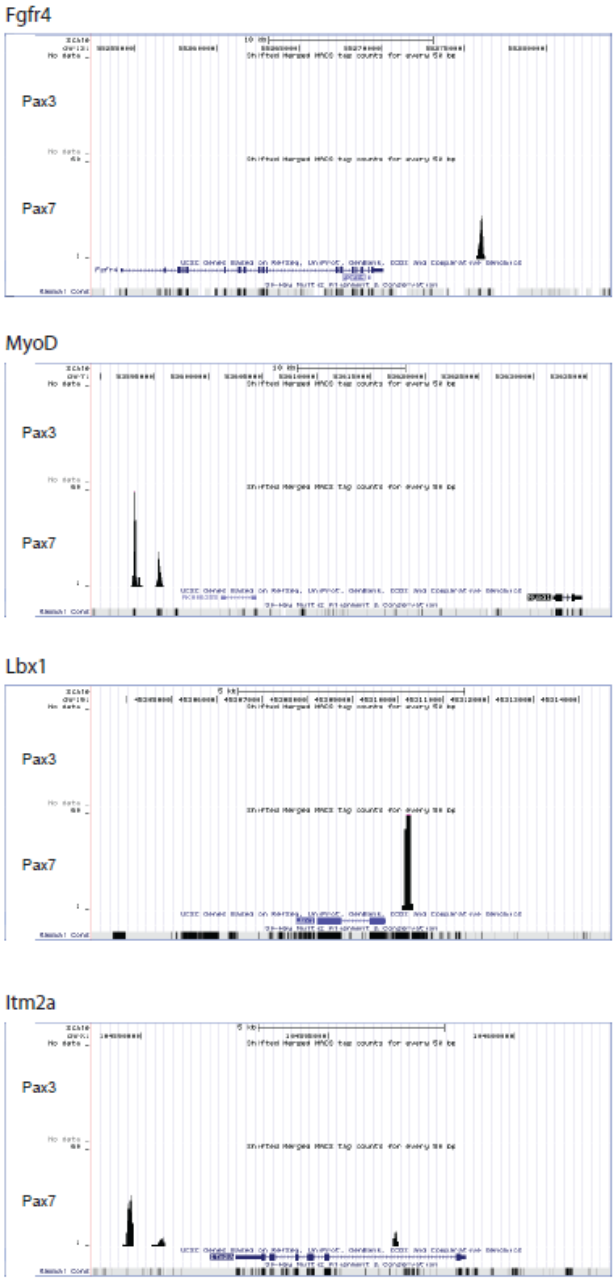


Figure S5. Pax7 Binds Uniquely to Sites that Bind Pax3 During Embryonic Myogenesis

Examples of various genes reported as downstream targets of Pax3 regulation in the embryo with no proximal Pax3 peaks, but strong proximal Pax7 peaks. In particular, whereas the Fgfr4 binding site at +22 kb is a direct target of Pax3 and directs Fgfr4 expression at myogenic sites in the embryo (Lagha et al., 2008), this site is occupied by Pax7 in adult-derived primary myoblasts.

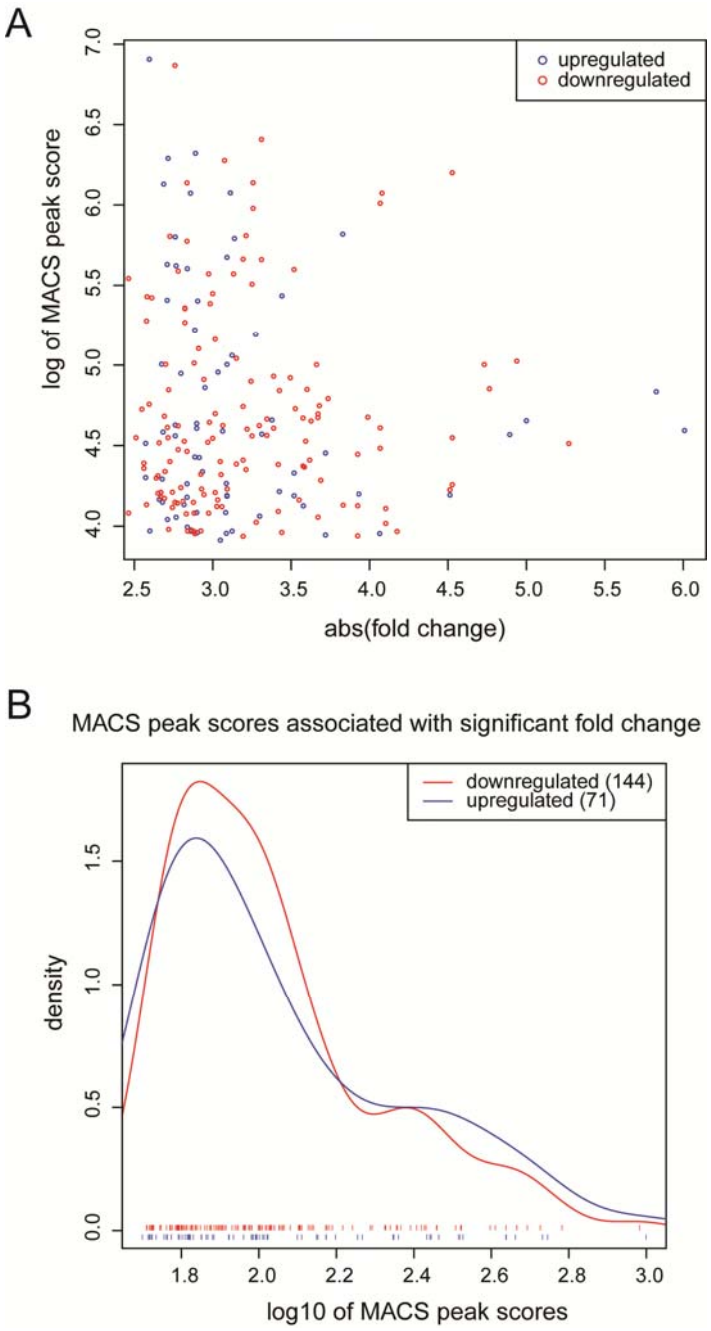


Figure S6. MACS Peak Score Does Not Predict Expression of Pax7-Target Genes

(A) The absolute fold changes of genes within 10 kb of Pax7 peaks were plotted against the MACS scores for those peaks. No significant correlation was noted between peak scores and fold changes.
(B) The distributions of MACS scores for peaks associated with significant fold changes were not different between genes upregulated or downregulated from knockdown of Pax7 siRNA.

Table S1. Enriched GO Terms Among Genes within 10kb of Pax7 Peaks

Pax7 unique peaks

Term	Count	%	P-value	Fold Enrich	Benjamini
GO:0044424~intracellular part	1170	51.34%	2.28E-07	1.09	1.75E-04
GO:0007275~multicellular organismal development	345	15.14%	3.35E-07	1.26	8.34E-04
GO:0065007~biological regulation	626	27.47%	5.51E-06	1.15	3.43E-03
GO:0044237~cellular metabolic process	943	41.38%	2.20E-06	1.10	3.64E-03
GO:0007517~muscle development	44	1.93%	1.33E-05	1.95	5.49E-03
GO:0044422~organelle part	373	16.37%	3.87E-04	1.16	3.65E-02

Pax7 unique, conserved peaks

Term	Count	%	P-value	Fold Enrich	Benjamini
GO:0007275~multicellular organismal development	134	24.54%	4.36E-18	2.07	2.17E-14
GO:0045449~regulation of transcription	130	23.81%	8.91E-18	2.09	2.22E-14
GO:0043565~sequence-specific DNA binding	45	8.24%	2.47E-12	3.36	1.56E-09
GO:0009887~organ morphogenesis	42	7.69%	3.49E-09	2.81	7.90E-07
GO:0006366~transcription from RNA polymerase II promoter	40	7.33%	4.65E-09	2.87	1.01E-06
GO:0007399~nervous system development	49	8.97%	8.40E-09	2.48	1.67E-06
IPR001356:Homeobox†	25	4.58%	3.93E-10	4.76	2.14E-06
GO:0030154~cell differentiation	95	17.40%	1.42E-08	1.77	2.63E-06
GO:0022008~neurogenesis	32	5.86%	8.95E-08	2.99	1.49E-05
GO:0007389~pattern specification process	23	4.21%	7.69E-07	3.45	1.16E-04
GO:0045941~positive regulation of transcription	29	5.31%	1.87E-06	2.77	2.58E-04
GO:0009790~embryonic development	33	6.04%	3.64E-06	2.48	4.54E-04
GO:0051674~localization of cell	28	5.13%	4.69E-06	2.71	5.30E-04
GO:0048523~negative regulation of cellular process	53	9.71%	1.69E-05	1.85	1.72E-03
GO:0001655~urogenital system development	11	2.01%	4.59E-05	5.12	4.15E-03
IPR007087:Zinc finger, C2H2-type†	30	5.49%	3.09E-06	2.66	4.22E-03
GO:0007423~sensory organ development	17	3.11%	7.56E-05	3.21	6.25E-03
GO:0048568~embryonic organ development	8	1.47%	1.09E-04	6.98	8.71E-03
GO:0008283~cell proliferation	31	5.68%	5.31E-04	1.96	3.38E-02

† INTERPRO terms

Table S2. Validation of Select Pax7 Regulated Genes

Gene Symbol	Gene	Fold change 1.0ST	Fold change qPCR
Cdh11	cadherin 11	-2.29	-1.15
Myf5	myogenic factor 5	-2.03	-2.94
Pax7	paired box 7	-1.92	-2.08
Syne2	synaptic nuclear envelope 2	-1.82	-6.25
Ezh2	enhancer of zeste homolog 2 (Drosophila)	-1.00	-1.16
Sox11	SRY-box containing gene 11	-1.00	-1.28
Wnt5a	wingless-related MMTV integration site 5A	-1.00	-1.39
Egfr	epidermal growth factor receptor	-1.00	-1.69
Hmgb2	high mobility group box 2	-1.00	-1.46
Fgfr4	fibroblast growth factor receptor 4	-1.00	-1.40
Grem1	gremlin 1	-1.42	-5.26
Fstl1	follistatin-like 1	-1.34	-2.38
Mdm2	transformed mouse 3T3 cell double minute 2	-1.32	-2.44
Lbh	limb-bud and heart	-1.27	-1.47
Prrx1	paired related homeobox 1	-1.26	-1.89
Ywhah	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide	-1.25	-2.50
Cxcl12	chemokine (C-X-C motif) ligand 12	-1.23	-2.08
Ddx39	DEAD (Asp-Glu-Ala-Asp) box polypeptide 39	-1.23	-1.56
Fhl2	four and a half LIM domains 2	-1.21	-1.52
Itga7	integrin alpha 7	1.45	1.00
Anxa	annexin A3	1.47	1.00
Srgap3	SLIT-ROBO Rho GTPase activating protein 3	1.47	1.00
Dvl1	dishevelled, dsh homolog 1 (Drosophila)	1.47	1.19
Acta1	actin, alpha 1, skeletal muscle	1.51	1.00
Il4ra	interleukin 4 receptor, alpha	1.54	1.00
Musk	muscle, skeletal, receptor tyrosine kinase	1.58	1.00
Myog	myogenin	1.59	1.47
Itgb4	integrin beta 4	1.59	1.00
Fgf7	fibroblast growth factor 7	2.40	1.00

Table S3. Primer and Probe Sequences

EMSA	
Myf5 -111kb Δprd Δhbox	TTCACAACGGCTAGCCCTCTTCGGACGGAAACCATCATTTCCTGATC TCAGCTACCCTAT
Myf5 -111kb PRD	TCATTTCCTGATTGTCATGCTTCTATCCTTGCACA
Myf5 -111kb HBOX	CTGGTTTTCACAATAATGCATTTTCTGTAACTAAC
qPCR	
Fgfr4-2F	CCCTGTTGAGCATCTTTCAG
Fgfr4-3R	CCCTCTTTGTACCAGTGACG
Hmgb2-2F	GAGGAGCACAAGAAGAAGCA
Hmgb2-3R	CCCTTTGGGAGGAACATAGT
Egfr-2F	GTCCTTGGGAACTTGGAAAT
Egfr-3R	AGGATGGCTAAGGCATAGGT
Wnt5a-3F	ATCAAGGAATGCCAGTACCA
Wnt5a-4R	CAGCCACAGGTAGACAGCTC
Ezh2-2F	CCAGACTGGGAAGAAATCTG
Ezh2-3R	TGATGTGCACAGGCTGTATC
Sox11-F	TTCAAGAACATCACCAAGCA
Sox11-R	CTGGGAGAAATTCAAGCTCA
Itga7-4F	AGTTTGGGTTCTGTCAGCAG
Itga7-5R	GGGTCTTGCTCCTTCTCG
Anxa3-3F	TTGTCAAGCAGTACCAAGCA
Anxa3-4R	GCAGTAACAAGAGCCACCAT
Srgap3-2F	GCAGGAAAGCTGAGATTGAG
Srgap3-3R	GCGAGAGGAGGTATTGATCC
Acta1-3F	CCTGCCATGTATGTGGCTAT
Acta1-4R	CGCTCAGTGAGGATTTTCAT
Il4ra-4F	CCCACCTGCTTCTCTGACTA
Il4ra-5R	GGCCTATTCATTTCCATGTG
Musk-3F	ATTCTGAGCGTGGAAGACAG
Musk-4R	ACACAGATGGTTTGGGGTTA
Itgb4-8F	AACAAACTGCAAGGAGAACG
Itgb4-9R	TCGTAGTGGAAGGCAGACTC
Fgf7-2F	CCAGTGGTACCTGAGGATTG
Fgf7-3R	TTCCTTGTTCATGGCAAGAT
Cdh11-F	TCCATGGACACCATGAGAAG
Cdh11-R	TGTCTTGGTGGCATGAATGT
Myf5-1F	TGAGGGAACAGGTGGAGAAC
Myf5-3R	AGCTGGACACGGAGCTTTTA
Pax7-F	CTGGATGAGGGCTCAGATGT
Pax7-R	GGTTAGCTCCTGCCTGCTTA
Syne2-6F	GTAGAAGGAAACCCGTCCAT
Syne2-7R	GGGAGGACTTGAGGTAGGAG
Grem1 1F	ACCTGGAGACCCAGAGTACC
Grem1 2R	TTGGGAACCTTTCTTTTTCC
Fstl1-F	CAGACAGAAGCCACAGGACA
Fstl1-R	CAGGAGCAGCTAAGGACCAC

Mdm2-4F	GAGCGCAAAACGACACTTAC
Mdm2-5R	AGAAACTCGGGACTCCAAAC
Lbh 2F	ATCTGAGATCGGCTGAGATG
Lbh 3R	TCAGTGGGTTCCACCACTAT
Prrx1 1F	CACAAGCAGACGAAAGTGTG
Prrx1 2R	GCGAGATCTTCTCGAACAAA
Ywhah 1F	GGAGCGCTACGACGATATG
Ywhah 2R	TCTCCCGGTAGGCTTTAACT
Cxcl12-F	ACAGCAGCTCAGAGAACCAT
Cxcl12-R	TGGAAGGACACAAGGAATGT
Ddx39 2F	AGTTGACTGTGGCTTTGAGC
Ddx39 3R	ATTGACAGGCTCAATCTGCT
Fhl2 3F	ACCGACTGCTATTCCAATGA
Fhl2 4R	CTGACAGGTGAAGCAGGTCT
Dvl1-F	CTTACCAGGACCCTGGCTTC
Dvl1-R	CCTGACTTCGAGGGCTACTG
Myog-F	GAAAGTGAATGAGGCCTTCG
Myog-R	ACGATGGACGTAAGGGAGTG
GAPDH-F	TGTCCGTCGTGGATCTGAC
GAPDH-R	GGTCCTCAGTGTAGCCCAAG
MRF4-1F	GGAGTGCCATCAGCTAACATT
MRF4-2R	GAATGATCCGAAACACTTGG
MyoD-F	TACCCAAGGTGGAGATCCTG
MyoD-R	CATCATGCCATCAGAGCAGT
Pax3-F	TATTCCACAAGCCGTGTCA
Pax3-R	GAGCTGCTGTCTGCGTTC

ChIP

Cxcl12 +90kb-F	ACAGCAGCTCAGAGAACCAT
Cxcl12 +90kb-R	TGGAAGGACACAAGGAATGT
Dlg2 intron 1-F	TTCAACATCGTGAAGCTGTG
Dlg2 intron 1-R	GGAATCCTCCTCCACGTAAT
Gab1 intron 1-F	CACATCCCAGGTCCTCTATG
Gab1 intron 1-R	CCAGTGAGAGCTGAGGGTAA
Hnf1a -22kb-F	AGTGATTTCAGGCTGATTGC
Hnf1a -22kb-R	TGTAGACCAGCAGGAAGGAG
Klf5 +60kb-F	CAAAGACAAAAGGCAAAACG
Klf5 +60kb-R	TGCACTCAGAGGTGAGTCAG
Mapk6 -15kb-F	CTTTCTTAGCGGCATCTGAA
Mapk6 -15kb-R	TGGAGAATGGGTTTGGTTTA
March4 intron 2-F	TCCATCCTCACTGTCTCGAT
March4 intron 2-R	AGAGAGCGAAAGAGAGCACA
Myf5 -111kb-F	CATCCCACATAATCCAATCAC
Myf5 -111kb-R	ACACAGATGGATGGGAAAGA
Myf5 -57kb-F	TGTGGCTCTCTCTCCGTATG
Myf5 -57kb-R	AATACAGACATGCAGGCTTCAC
MyoD -63kb-F	TCTCCTTGGTGTAGGCTCAG
MyoD -63kb-R	CCTGACCTTGAACGTGAATC
Notch2 -24kb-F	CCCTTCTCTCTACCCTCACC
Notch2 -24kb-R	AAAGGGAAGAAACAGGTGGT

Pbx1 -42kb-F	TGAGCTCAAGTGAAGGTTCC
Pbx1 -42kb-R	AAGTTGAAAAAGCCGATTCC
Prrx1 intron1-F	CGTTTACACCTCTGCCTTTG
Prrx1 intron1-R	CACGCACACATTCACCAG
Sgpp2 intron1-F	GTGACCAGCTGTTCTCCATC
Sgpp2 intron1-R	AATGTTGAGCAACCTTCAGC
Snd1 intron 11-F	CCTGTCTTCTAGGCCTGTCA
Snd1 intron 11-R	CATGATGAACAGCTGCAGAG
Vangl2 -21kb-F	GGAGTTAGCAGCAGAACAGC
Vangl2 -21kb-R	GCACTATGCCTTGTGCTCAG
GAPDH -3kb-F	TCCAGTGAGGACGGTATGAT
GAPDH -3kb-R	CATAAAGATGGGGCAAAATG

Supplementary Information

Experimental Procedures

ChIP-Seq Peak Calling

Solexa ChTAP-Seq reads from Pax7 (GSM615619), Pax3 (GSM615620), and control (GSM615621) samples were mapped to the mm9 mouse genome assembly (NCBI M37) using ELAND (part of the Illumina/Solexa GA pipeline) with default parameters. Aligned Pax7 and Pax3 tags with up to two mismatches were provided to MACS (Zhang et al., 2008) together with control tags. This resulted in negative (tag enriched control regions) and positive MACS peaks sets for both Pax7 and Pax3. Peaks ≥ 800 bp were excluded from the analysis, based on the size of excised ChIP fragments used for Solexa sequencing. ChIP-Seq peaks with p-values less than $1x10^{-5}$ were called significant by MACS. Negative peaks were used as negative controls in comparative conservation and motif analyses.

Filtering for Non-Conserved ChIP-Seq Peaks.

PhastCons conservation scores of placental mammals (30-way) were downloaded from UCSC Genomic Bioinformatics (http://hgdownload.cse.ucsc.edu/goldenPath/mm9/phastCons30way/placental/) (Rhead et al.). Average conservation scores for each MACS peak and negative peak were calculated as follows. Peaks were masked for the coding sequences of RefSeq genes (UCSC refGene table) and repetitive regions (UCSC rmsk table), and the average conservation score was calculated across the unmasked peak length. In addition, the longest region of contiguous bases with conservation score ≥0.4 was identified.

Average conservation scores were also calculated for Pax3 and Pax7 negative peaks and used to determine the 90th percentile negative peak conservation score for each factor. This score was used to filter the corresponding treatment peaks; a peak was accepted if its average conservation was greater than or equal to the 90^{th} percentile negative peak conservation, or if its longest region of contiguous bases with a conservation score ≥0.4 was 10 or more bp in length. The resultant sets of peaks were then used for subsequent analysis.

Our MACS called peaks have a notably high false discovery rate (FDR) for Pax3 (99.8%) and Pax7 (44.4%). Both transcription factors use the same control data, and the high FDR rates are the result of a large number of enriched regions in the control data. MACS was written to use controls derived from input DNA which provides a general background level across the genome. Our control data, however, are derived from ChIP from myoblasts expressing the TAP tag only. MACS estimates FDR by calculating "positive" peaks and the "negative" peaks with the treatment and control swapped. The p-values for the positive and negative peaks are ranked and the FDR is generated by comparing the two. If FDR rates are high, the enrichment in the control is comparable to that in the treatment. This is not unexpected for our data as mock ChIP protocols can generate localized enrichment. Rather than using FDR to derive cutoffs, we have taken a number of steps to direct our analysis towards the most biologically active binding sites. First, we have considered cross-species conservation at binding sites which can reflect the relative importance of non-coding genomic elements to essential processes. Secondly, we have combined ChIP-Seq data with gene expression data to define *cis* elements near Pax7-regulated genes and demonstrate that peaks are most frequently proximal to positively-regulated genes.

Peak List Comparison

Lists of peaks identified in the Pax3 and Pax7 experiments were compared using BEDTools (Quinlan and Hall. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics (2010) vol. 26 (6) pp. 841-2) to identify peaks found only in Pax3, Pax7, or both peak lists. We observed that peaks common to both TFs typically have a larger number of Pax7 than Pax3 reads. Many of the peaks identified as Pax7-specific also contain a smaller number of Pax3 reads, albeit at a level below the threshold for peak calling by MACS. In order to identify the peaks with greatest Pax7 specificity we developed a BioPerl script to count the number of Pax3 reads mapping to each of the 'Pax7-specific' peaks. Using these counts we generated a set of Pax7-specific peaks which contain three or fewer Pax3 reads; this provides a reasonably sized set of peaks for analysis, while avoiding peaks for which the Pax7-specificity is unclear.

Motif Iteration

To refine scoring matrices for motifs identified within sets of peaks, an iterative process was run to modify nucleotide weights based on occurrences within the peak set. The starting motif generated by MEME or GLAM2 was passed to FIMO to identify all instances of that motif found within the set of peak sequences. The sequences of these instances are aligned, and base frequencies counted, adding a pseudocount of 1 to each position. These frequencies are used to regenerate the scoring matrix, which is then passed to FIMO. This process is iterated until the number of motifs identified stabilizes.

Motif Counting

FIMO was used with the set of motif scoring matrices to identify motif instances in the sequences of all peaks ≤800bp wide. The number of peaks containing each motif in the

Pax3-specific, Pax7-specific, and overlapping peak sets were obtained, and compared on a pairwise basis using Fisher's Exact Test.

Protein Expression and Purification

C-terminally Flag-tagged Pax7 and Pax3 were expressed using baculovirus system. Recombinant baculoviruses are generated by using Bac-to-Bac system (Gibco-BRL) and infected to Sf9 cells. Infected Sf9 cells were incubated for 2 days and then collected by centrifugation. The cells were washed with cold calcium- and magnesium-free phosphate-buffered saline and stored at −80 °C until use. To purify Flag-tagged Pax7 and Pax3, cells were thawed, suspended in hypotonic buffer (10 mM Tris HCl, pH 7.6, 1% Nonidet P-40, 2mM MgCl2, 2 mM DTT, protease inhibitor cocktail (nacalai tesque) and incubated on ice for 10min. The nuclei were then collected by centrifugation, resuspended in High Salt buffer (25mM Tris HCl, pH 7.6, 0.1 % Triton X-100, 2mM MgCl2, 420mM NaCl, 2 mM DTT, protease inhibitor cocktail), and incubated on ice for 30min. After clarification by centrifugation, the nuclear extracts were incubated with protein G sepharose (GE healthcare life science) at 4 °C for 30 min to preclear the lysates and then incubated with anti-Flag M2-agarose beads (Sigma) at 4 °C for 3 hr. The resin was washed with wash buffer 1 (25 mM Tris HCl pH 7.6, 0.1 % Triton X-100, 2 mM MgCl2, 300mM NaCl, 2 mM DTT) for three times and then washed with wash buffer 2 (25 mM Tris HCl pH 7.6, 0.1 % Triton X-100, 2 mM MgCl2, 100mM NaCl) for twice. The proteins that remained bound to the resin were eluted by Flag peptide (Sigma) containing wash buffer 2 for twice. Eluted fractions were dialyzed with dialysis buffer (10 mM Tris HCl pH 7.6, 100mM NaCl, 2 mM DTT) and then frozen at −80 °C. The purity of the eluted fractions was verified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Microarray Analysis

Transcript cluster identifiers (TCID) of the Affymetrix MoGene 1.0 chipset were RMA (Irizarry et al., 2003) normalized using the xps (Stratowa, 2003) Bioconductor (Gentleman et al., 2004) package in R-2.10.0 (Team, 2010) with the Affymetrix provided MoGene-1_0-st-v1.r4 chip layout and scheme files. GSM616006, GSM616007, and GSM616008 were used as controls and GSM616009, GSM616010, and GSM616011 (Pax7 siRNA) as the treatment. The RMA normalized data was $\log_2$ transformed and analyzed using the Significance Analysis of Microarrays (Tusher et al., 2001) method as implemented in the bioconductor siggenes package (Schwender, 2009).

References

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., *et al.* (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol *5*, R80.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics (Oxford, England) *4*, 249-264.

Rhead, B., Karolchik, D., Kuhn, R.M., Hinrichs, A.S., Zweig, A.S., Fujita, P.A., Diekhans, M., Smith, K.E., Rosenbloom, K.R., Raney, B.J., *et al.* The UCSC Genome Browser database: update 2010. Nucleic Acids Res *38*, D613-619.

Schwender, H. (2009). siggenes: Multiple testing using SAM and Efron's empirical Bayes approaches. R package version 1200.

Stratowa, C. (2003). XPS, a Novel Framework for Distributed Storage and Analysis of Microarray Data in the Terabyte Range: An Alternative to BioConductor. Paper presented at: Proceedings of the 3rd International Workshop on Distributed Statistical Computing

Team, R.D.C. (2010). R: A language and environment for statistical computing (Vienna, Austria, R Foundation for Statistical Computing).

Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A *98*, 5116-5121.

Zhang, Y., Liu, T., Meyer, C., Eeckhoute, J., Johnson, D., Bernstein, B., Nussbaum, C., Myers, R., Brown, M., Li, W., *et al.* (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biology *9*, R137.

Chapter 3 – *Myf5* is a Direct Target of Pax7 Regulation in Adult Myoblasts

Myf5 is a Direct Target of Pax7 Regulation in Adult Myoblasts

Vincent G. Punch[1,2], Jaime J. Carvajal[3], Yoichi Kawabe[1,2], Peter W.J. Rigby[3], and Michael A. Rudnicki[1,2]*

1. Sprott Center For Stem Cell Research
   Ottawa Health Research Institute
   Regenerative Medicine Program

2. University of Ottawa
   Department of Medicine

3. Division of Gene Function and Regulation.
   The Institute of Cancer Research

Summary

Heterogeneity of Myf5 expression in satellite cells is a characteristic feature that divides them into subpopulations of uncommitted (stem) and committed myogenic progenitors. Control of Myf5 expression is highly complex, with distinct regulatory elements directing Myf5 expression in different anatomical locations. The regulatory control of Myf5 expression in quiescent satellite cells has been mapped to a region between -88kb and -140kb upstream of *Myf5*. We have investigated the role of Pax7 on Myf5 expression and observed that Myf5 is directly regulated by Pax7 in satellite cell-derived primary myoblasts. By ChIP-Seq, we have identified several novel binding sites for Pax7 and Pax3 upstream of *Myf5*. We show that a binding site at -111kb (ECR111) acts as a Pax3/7-regulated enhancer and is required for Myf5 expression in quiescent satellite cells, but not muscle spindles. We propose that the expression of Myf5 in quiescent satellite cells, therefore, is directly under the control of Pax7 acting through a satellite cell-specific enhancer at ECR111.

Introduction

Satellite cells are muscle stem cells juxtaposed against the surface of muscle fibers, beneath the basal lamina, and are responsible for the growth, maintenance, and repair of postnatal skeletal muscle. In their resting state, satellite cells are quiescent, however, their activation in response to stress or trauma causes them to re-enter the cell cycle, undergoing several rounds of division, producing myogenic precursor cells that differentiate and fuse to form multinucleated myotubes. Additionally, some progeny return to quiescence for the self-renewal of the satellite cell pool (Zammit et al., 2004a). The decision between self-renewal and myogenic differentiation appears to be highly influenced by the heterogeneity of Myf5 expression in satellite cell populations. Transplantation studies have shown that Pax7+/Myf5- cells contribute extensively to the expansion of the satellite cell pool, whereas Pax7+/Myf5+ cells tend towards precocious differentiation. Moreover, asymmetric divisions of Pax7+/Myf5- cells frequently give rise to Pax7+/Myf5+ cells, suggesting that Myf5 expression reflects a hierarchy within satellite cells comprised of stem cells and committed progenitors (Kuang et al., 2007).

Satellite cells, like embryonic muscle precursor cells, originate from progenitors present in the somite (Kassar-Duchossoy et al., 2005). Therefore, it is unsurprising that the regulatory networks directing myogenic specification and progression are analogous between embryo and adult. Initiation of the myogenic program is controlled by four related basic helix-loop-helix (bHLH) transcription factors that are known as the myogenic regulatory factors (MRFs); MyoD (Davis et al., 1987), Myf5 (Braun et al., 1989), myogenin (Myog (Edmondson and Olson, 1990; Wright et al., 1989)), and Myf6 (MRF4 (Braun et al., 1990a; Miner and Wold, 1990; Rhodes and Konieczny, 1989)). Myf5 and MyoD direct

commitment to the myogenic lineage (Rudnicki et al., 1993; Tajbakhsh et al., 1996), while myogenin and Myf6 are essential for terminal differentiaion and myoblast fusion (Hasty et al., 1993; Nabeshima et al., 1993; Rawls et al., 1998). In the embryo, Myf5 is expressed first, at 8.0dpc (Ott et al., 1991), however, MyoD and Myf5 are eventually both expressed in proliferating progenitor cells and can partially substitute for initiating the myogenic program. Mice lacking MyoD compensate by upregulating Myf5 in skeletal muscles (Rudnicki et al., 1992), although the initiation of differentiation in developing limbs is delayed. Conversely, mice lacking Myf5 are delayed in myotome formation (Braun et al., 1994; Braun et al., 1992; Kaul et al., 2000). The loss of both elements causes progenitor cells in the somite to develop alternative fates (Kablar and Rudnicki, 1999; Tajbakhsh et al., 1997).

Other key regulatory genes involved in myogenic control are Pax3 and Pax7. Both are transcription factors belonging to the paired box family. In Splotch (sp) mutant mice, mutations that disrupt Pax3 function impede the formation of limb muscles due to a failure of progenitor cells to migrate from the hypaxial somite to the limb buds (Franz et al., 1993; Goulding et al., 1994). In developing limb muscles, Pax3-expressing progenitor cells activate the expression of MyoD and Myf5 only following their arrival at limb buds, and it has been shown that Pax3 can directly activate Myf5 expression in this context (Bajard et al., 2006; Tajbakhsh and Buckingham, 1994). Furthermore, compound Pax3/Myf5 mutant embryos do not express MyoD and phenotypically resemble MyoD/Myf5 compound mutants, implying that Pax3 can also direct MyoD expression (Tajbakhsh et al., 1997).

Gene replacement studies have shown that Pax7 is functionally equivalent to Pax3 in somitogenesis and development of trunk musculature (Relaix et al., 2004), however, in adult satellite cells, Pax7 alone is required for cell survival and population expansion (Kuang et al., 2006; Oustanina et al., 2004; Seale et al., 2000). By contrast, the expression of Pax3 in

satellite cells is highly variable. Pax3-positive cells are readily detectible in diaphragm muscles and some positive satellite cells are found in forelimbs, however hindlimb muscles do not express Pax3. Furthermore, Pax3 fails to compensate for the loss of Pax7 in regenerative myogenesis (Kuang et al., 2006).

Previously, we have shown that Pax7 can influence the expression of Myf5 (McKinnell et al., 2008), however the genetic control of Myf5 is extremely complex. *Myf5* is linked with *Myf6* in vertebrates, separated by 8.7kb in the mouse. Substantial work has gone into mapping the regulatory elements that direct Myf5 and Myf6 expression, and we now know that discrete and dispersed elements drive expression in different populations of precursor cells (Carvajal et al., 2001; Summerbell et al., 2000). It has been suggested that this complexity allows the Myf5 to respond to a variety of inductive signals which may be present at different anatomical niches. In the current model proposed by Carvajal et al. (2008), the *Myf5/Myf6* locus is not regulated through straightforward promoter competition, but rather that promoters act in equilibrium with enhancers and other regulatory elements, such as cryptic promoters termed TRABS (transcription balancing sequences), to direct dynamic, specific expression of each gene (Carvajal et al., 2008). By using a series of BAC deletion constructs carrying AP and lacZ cloned at the transcriptional start sites of Myf6 and Myf5, respectively, we have shown that 140kb of upstream sequence is sufficient to recapitulate the complete embryonic expression patterns of both genes, demonstrating that regulatory elements for the locus are spread over a broad region (Carvajal et al., 2001). Other attention has been focused towards identifying the genetic elements responsible for Myf5 expression in adult muscles. We have also shown that Myf5 expression in muscle spindles is controlled by separate genetic elements than those directing Myf5 expression in quiescent satellite cells and their proliferating progeny (Zammit et al., 2004a). BAC clones

retaining up to 88kb of *Myf5* upstream sequence retained Myf5 expression in muscle spindles, but not quiescent satellite cells, indicating that the regulatory elements specific for satellite cell expression are located between -140kb and -88kb.

In order to understand the differences between Pax3 and Pax7 regulation in myogenic cells, we have recently performed a genome-wide analysis for Pax3 and Pax7 binding sites by ChIP-Seq. Here, we report that Pax3, but primarily Pax7 binding decorates the 140kb region upstream of *Myf5/Myf6* and show that Pax3/7 strongly bind a novel short conserved sequence at -111kb, within the satellite cell control region, which we hereafter refer to as ECR111 (ECR: Evolutionary Conserved Region). We show that a DNaseI-insensitive region of ECR111 contains both paired domain and homeodomain motifs and that these motifs are necessary for Pax7 binding. Furthermore, we demonstrate that ECR111 is an enhancer and is required for Myf5 expression in quiescent satellite cells but not muscle spindles. Taken together, this work demonstrates that Myf5 is directly regulated by Pax7 and provides novel insight to the regulatory mechanisms directing Myf5 expression in adult muscle.

Results

Previous work has demonstrated that forced expression of Pax7 can enhance Myf5 levels in C2C12 myoblasts and 10T1/2 fibroblasts (McKinnell et al., 2008). Furthermore, Myf5 has been shown to be directly regulated by Pax3 and Six1/4 in the developing limb (Bajard et al., 2006; Giordani et al., 2007). To determine the regulatory impact of Pax7 on myogenic genes in primary myoblasts, we transfected primary cultures with siRNA directed against Pax7. qPCR analysis of cDNA after 48 hours noted a decrease in MyoD, Pax3, and Myf6, but a

dramatic decrease in *Myf5* RNA (Figure. 1A). We also noted an increase in myogenin levels, suggesting the loss of Pax7 may initiate the differentiation program. Conversely, we overexpressed Pax7 in primary cells through proviral infection of Pax7-FLAG. Myf5 expression was significantly increased due to overexpression of Pax7 (Figure 1B). Western blot of whole cell lysates from siRNA-treated and overexpression myoblasts mirrored the impact on gene expression observed by qPCR (Figure 1C). These observations demonstrate that Myf5 is regulated in a linear fashion by Pax7.

Pax3/7 binding across the regulatory region controlling Myf5/Myf6 expression

Multiple enhancers scattered across a broad regulatory region (Figure 2A) function in equilibria with promoters and TRABS to direct Myf5 expression in different anatomical locations within the embryo. In the adult, satellite cell-specific expression of Myf5-lacZ is lost from BAC deletions of -88kb to -140kb, whereas expression in muscle spindles is retained even when only 59kb up of upstream sequence remains (Zammit et al., 2004a). ChIP-Seq analysis of Pax3 and Pax7 binding identified a number of peaks across the Myf5 regulatory region. Additionally, peaks were masked to filter out non-conserved sites (Punch et al, submitted), under the supposition that important regulatory elements would be retained through evolution. Within the 140kb upstream regulatory region, Pax3 peaks were observed at conserved loci at -57.5kb (ECR57), -111kb (ECR111), and -130kb (ECR130) (Figure 2B). Pax7 peaks were more numerous, but the largest peaks were found to overlap with Pax3 peaks (Figure 2C). Many peaks were found within introns of the *Ptprq* gene, and although *Ptprq* is not expressed in adult skeletal muscle, it is expressed in the CNS where it may be regulated by Pax genes.

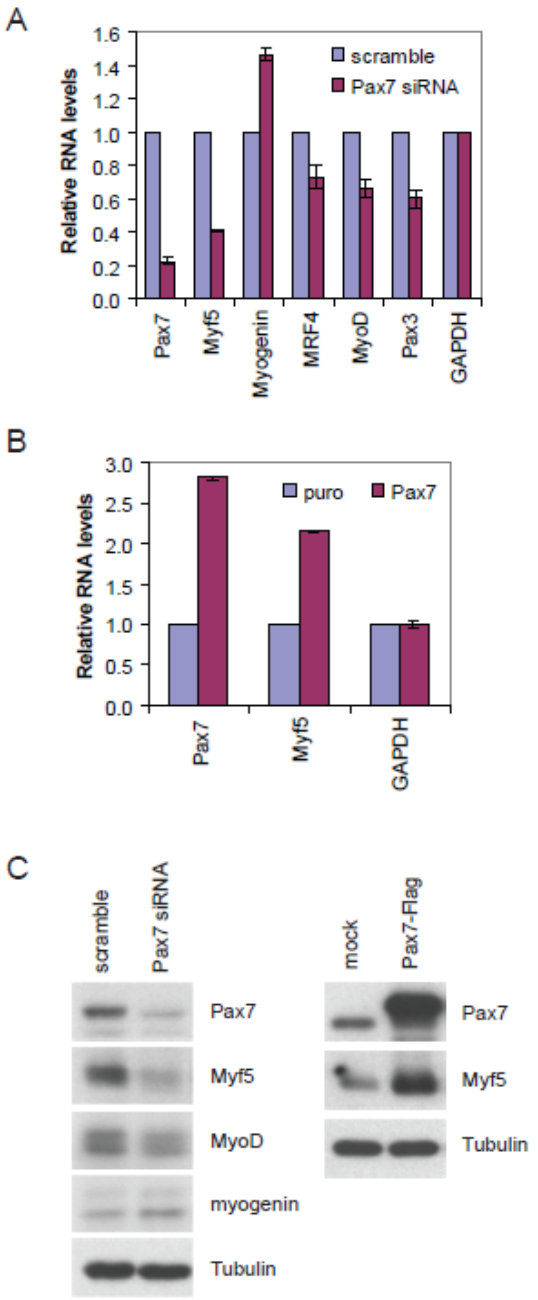


Figure 1. Pax7 drives the expression of Myf5

(A) qPCR analysis of RNA from primary myoblasts transfected with Pax7 or control siRNA. *Myf5* transcript levels are significantly reduced alongside Pax7, while other MRFs are influenced to a lesser extent.
(B) Stable overexpression of Pax7 in primary myoblasts by proviral infection causes an increase in *Myf5* transcripts compared to empty vector controls. Taken together, these data demonstrate a linear relationship between *Pax7* and *Myf5* expression.
(C) Western blot showing that alterations in Pax7-dependent changes in the expression of myogenic genes are congruent at the protein level.

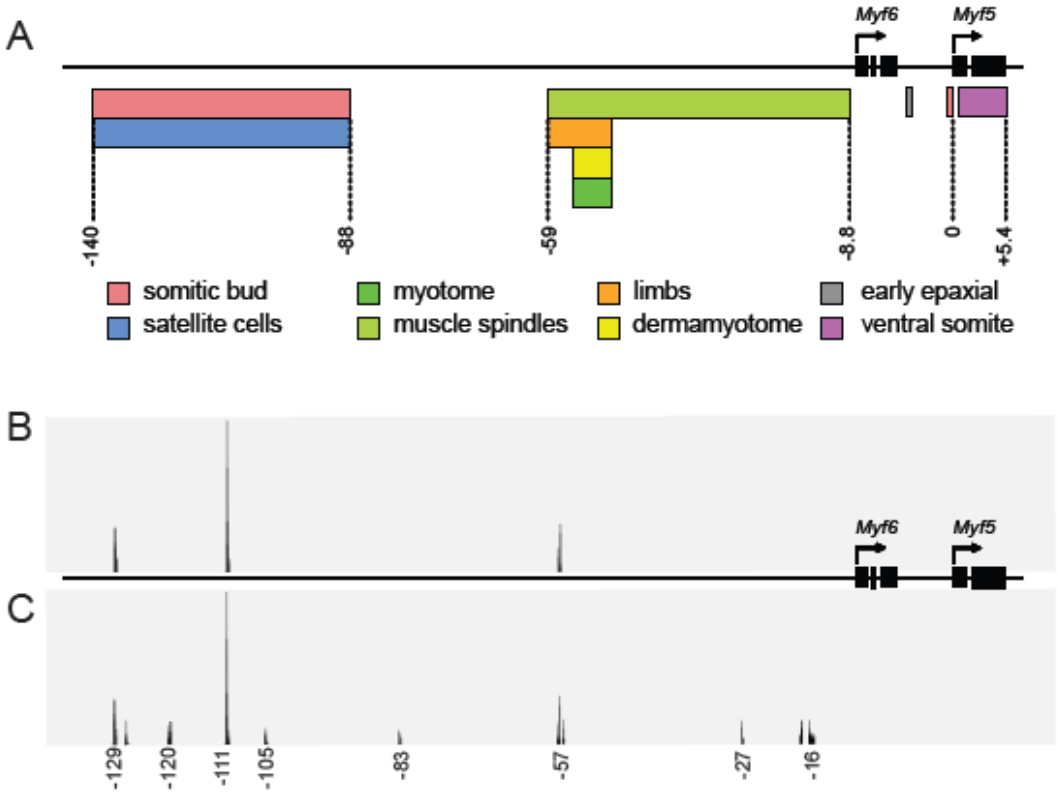


Figure 2. Pax3 and Pax7 DNA binding mapped around the *Myf5/Myf6* locus

(A) Schematic of the regulatory regions for *Myf5* depicting the multitude of interdigitated enhancers that direct Myf5 expression at different times and at different anatomical locations (Carvajal et al., 2008; Zammit et al., 2004a).
(B) Pax3 binding sites are found at ECR130, ECR111, and ECR57, as determined by ChIP-Seq analysis primary myoblasts stably expressing Pax3. The regulatory element at ECR57 directs *Myf5* expression in somites, trunk, and limb muscles (Bajard et al., 2006).
(C) Pax7 binding sites decorate the Myf5/6 regulatory region, and are found within the satellite cell region as well as regions directing expression in limbs and muscle spindles. Pax7 commonly binds all Pax3 sites, however the peak at ECR111 is significantly more enriched than other binding sites.

We have reported that Pax3 and Pax7 recognize identical motifs (Punch et al, submitted), therefore it is unsurprising that Pax3 binding sites aligned perfectly with some of our Pax7 sites. Binding within ECR57 overlaps with the 145bp regulatory element described by Bajard et al. that responds to Pax3 and Six1/4 in directing Myf5 expression in limb buds during embryogenesis (Bajard et al., 2006). Evidence from BAC work indicates this element alone would be insufficient for Myf5 expression in satellite cells, however, its location is concordant with the notion that this enhancer could be used by adult cells to drive expression of Myf5 in muscle spindles. The binding site within ECR111 was the largest peak, representing a high affinity binding site for both Pax3 and Pax7. Its location within the satellite cell control region primes it as a candidate satellite cell-specific enhancer for Myf5.

Binding of Pax7 ECR57 and ECR111 was validated by independent ChIP experiments. Primary myoblasts expressing Pax7-FLAG following retroviral infection were immunoprecipitated with anti-FLAG antibody. Quantitative PCR confirmed enrichment of ECR57 and ECR111 over a control locus (GAPDH -3kb) (Figure 3C).

Collectively, these data provide evidence for Pax3/7 binding at many sites within the regulatory region. That we see Pax7 binding at the known Pax3 enhancer at ECR57 and elsewhere suggests that Myf5 is a direct regulatory target of Pax7.

Identification of precise Pax3/Pax7 binding domains within conserved peaks

To precisely define the regions bound by Pax3/7 at ECR57 and ECR111, DNaseI footprinting assays were performed on probes spanning the regions identified by ChIP-Seq. FAM-labeled probes were generated by PCR and incubated with excess BSA or purified Pax3 or Pax7 produced through baculoviral infection of Sf9 cells (Figure S1A). Following brief treatment with a dilute concentration of DNaseI, fragment spectra were plotted on a

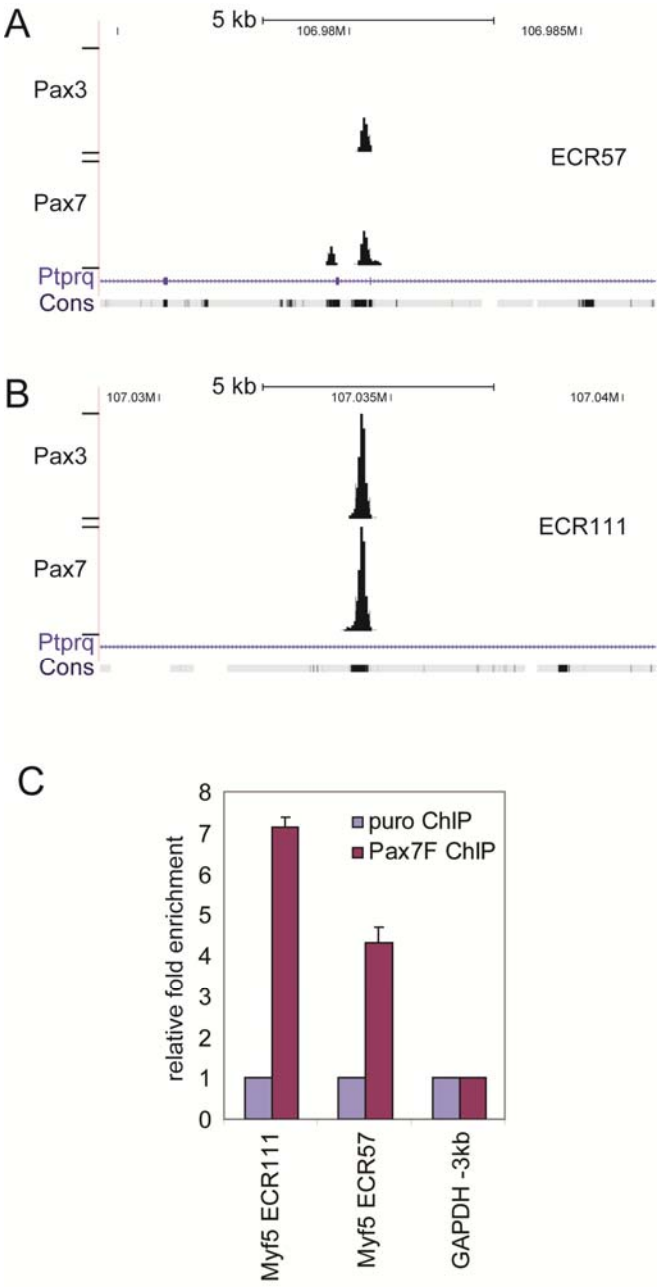


Figure 3. Pax3 and Pax7 binding at Myf5 ECR57 and ECR111 is highly conserved

(A) The binding site at ECR57 is located within an intron of *Ptprq* and is highly conserved and contains putative Pax3 and Six1/4 binding domains (Bajard et al., 2006; Giordani et al., 2007).
(B) The novel binding site at ECR111 also resides within an intron of *Ptprq* at a highly conserved locus. Enrichment of the locus is similar between Pax3 and Pax7.
(C) Primary myoblasts were stably infected with Pax7-FLAG or empty vector and chromatin immunoprecipitation with anti-FLAG antibody was carried out to validate binding to ECR57 and ECR111 sites observed by ChIP-Seq. qPCR analysis confirms enrichment of each loci, with stronger enrichment evident on ECR111.

DNA analyzer. With the ECR57 fragment we observed identical spectra patterns between Pax3 and Pax7 treatments, revealing a 29-bp protected region that was sensitive to digestion only in the presence of BSA (Figure 4A). We found the protected region was centered about the 10-bp paired-like motif reported by Bajard et al. to bind Pax3 and drive expression of Myf5 in developing limbs. With the ECR111 fragment, we observed a 54-base pair region that was insensitive to DNaseI digestion when incubated with Pax7 or Pax3, but not BSA (Figure 4B). Like the ECR57 probe, the spectra pattern were identical providing further evidence that Pax3 and Pax7 recognize identical motifs. The protected region within ECR111 contained candidate paired and homeodomain motifs, separated by 29bp.

The paired motif shares 90% identity with the motif reported by Epstein et al. (Epstein et al., 1996) , compared to the ECR57 paired motif which matches the Epstein motif by only 70%. Concurrent with our *de novo* motif data, these data support the notion that Pax3/7 proteins can recognize degenerate paired motifs. Conversely, the homeodomain motif was absent from the ECR57 sequence, however the ECR111 motif matches the consensus sequence reported by Wilson et al. (Wilson et al., 1993). In this case, the motif is an imperfect inverted repeat of the TAAT motif. Additionally, we performed binding on a control probe around the IgH locus and determined identical spectra between Pax3, Pax7 and BSA (Figure S1B). Taken together, we find that the ECR57 and ECR111 loci contain known Pax3/7 binding motifs and suggest that recognition of the ECR111 site may occur through paired and/or homeodomain recognition.

Binding at -111kb occurs through multiple recognition sites

The binding of many Pax proteins to their native recognition paired motif sequences is well studied. Monomeric binding via the paired domain results in protected regions of

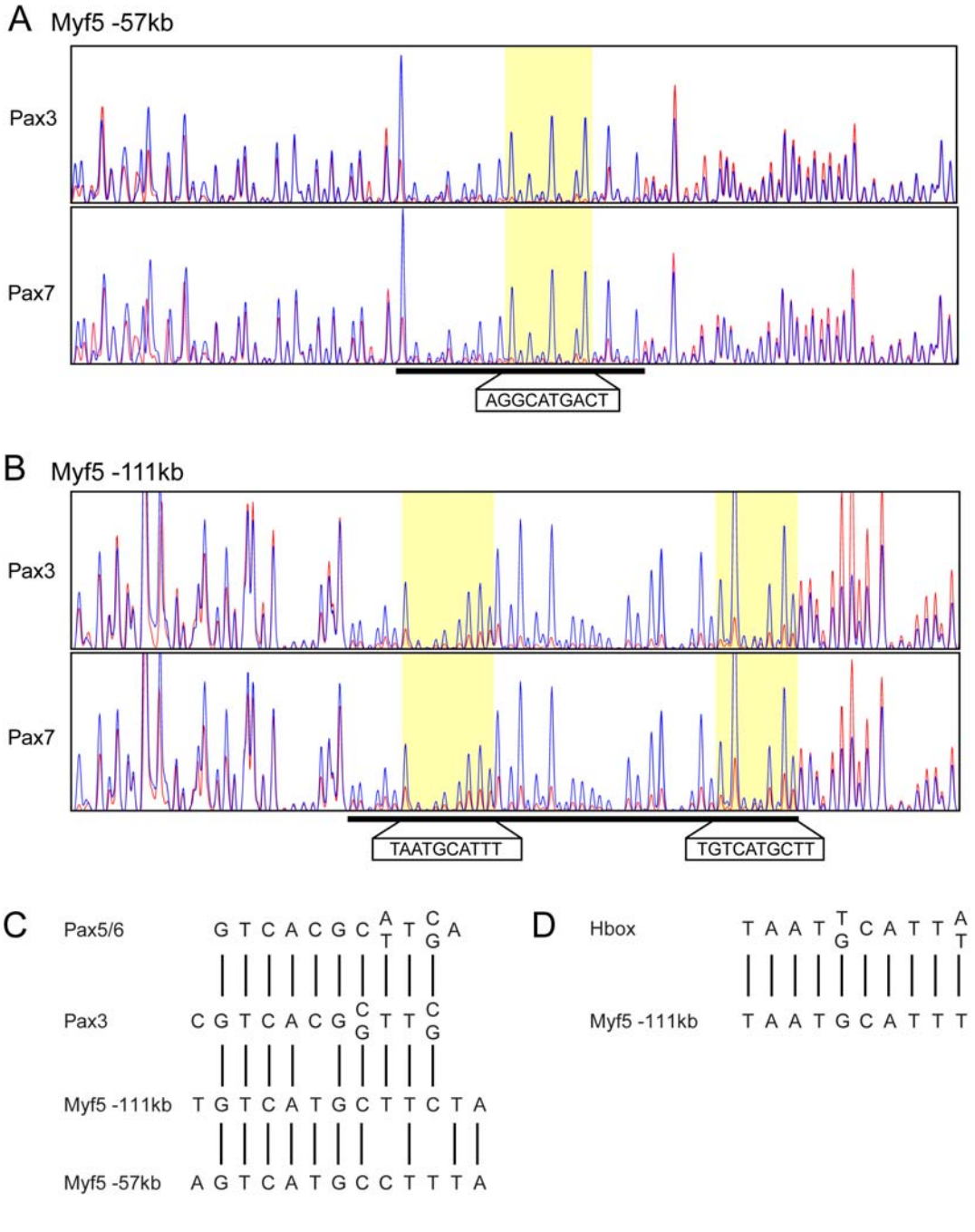

A
Myf5 -57kb
Pax3
Pax7
AGGCATGACT
B
Myf5 -111kb
Pax3
Pax7
TAATGCATTT
TGTCATGCTT
C
Pax5/6
G T C A C G C A/T T C/G A
Pax3
C G T C A C G C/G T T C/G
Myf5 -111kb
T G T C A T G C T T C T A
Myf5 -57kb
A G T C A T G C C T T T A
D
Hbox
T A A T T/G C A T T A/T
Myf5 -111kb
T A A T G C A T T T

Figure 4. DNaseI footprinting of the ECR57 and ECR111 loci defines the boundaries of Pax3 and Pax7 binding

(A) Chromatogram of *Myf5* ECR57 probe incubated with Pax3 or Pax7 (red) or BSA (blue) and digested with DNaseI. The 29-bp footprint (black bar) is identical between Pax3 and Pax7, centered around a paired domain motif.
(B) Chromatogram of *Myf5* ECR111 probe. Identical 54-bp protected regions (black bar) were observed in the presence of Pax3 or Pax7. The footprint spans a region containing a putative paired domain and a separate homeodomain motif(Wilson et al., 1993). The distance between each motif and the length of the footprint implies that the binding site might be simultaneously occupied by multiple Pax7 molecules.
(C) Consensus alignment of paired motifs within ECR57 and ECR111 to known sequences (Epstein et al., 1996; Xu et al., 1999). ECR57 bears 70% identity to the consensus described by Epstein et al., whereas ECR111 aligns with 90% identity.
(D) Consensus alignment of homeodomain motifs. While ECR57 site does not contain a full homeodomain motif, ECR111 site contains imperfect palindromic repeats, appropriately separated by a 2bp spacer sequence "GC" (Birrane et al., 2009; Wilson et al., 1993).

approximately 30bp or less (Czerny et al., 1999; Esposito et al., 1998). While the size of the ECR57 protected region is consistent with this, we saw a much larger protected region in the ECR111 probe. Because larger protected regions in Pax binding assays have been observed elsewhere when binding sites are adjacent (Czerny et al., 1999), we sought to determine if the ECR111 probe was bound by multiple molecules of Pax7.

We performed electrophoretic mobility shift assays on ECR111 and ECR57 probes. A 60-mer probe spanning both the paired and homeodomain sites of ECR111 and a 30-mer probe spanning the paired domain of ECR57 were shifted with purified Pax7-FLAG. A single species was observed for ECR57, however multiple species were detected in the longer ECR111 probe, suggesting that multiple molecules of Pax7 were recognizing the site (Figure 5A). Furthermore, these species were dependent on the presence of intact paired and homeodomain sequences. While excess cold probe was sufficient to compete for Pax7 binding, a mutant probe containing scrambled paired and homeodomain sequences was unable to bind Pax7 as a competitor. By contrast, a 30-mer ECR111 probe spanning only the paired motif bound Pax7 as a single species (Figure 5B). Taken together, these observations suggest that monomeric binding occurs at paired motifs at both ECR57 and ECR111 sites, whereas additional molecules are recruited to the homeodomain portion of ECR111.

Myf5 ECR111 acts as a Pax3/Pax7 enhancer

To determine whether ECR111 confers enhancer activity, we cloned the ECR111 element into a reporter vector driven by a minimal *Myf5* promoter. Three reporter constructs were engineered to test the independent potential of each binding site as well as the full site; M5-PRD was comprised of three copies of the paired-containing sequence, M5-HBOX of three copies of the homeodomain-containing sequence, and M5-PDHD of three copies of the full

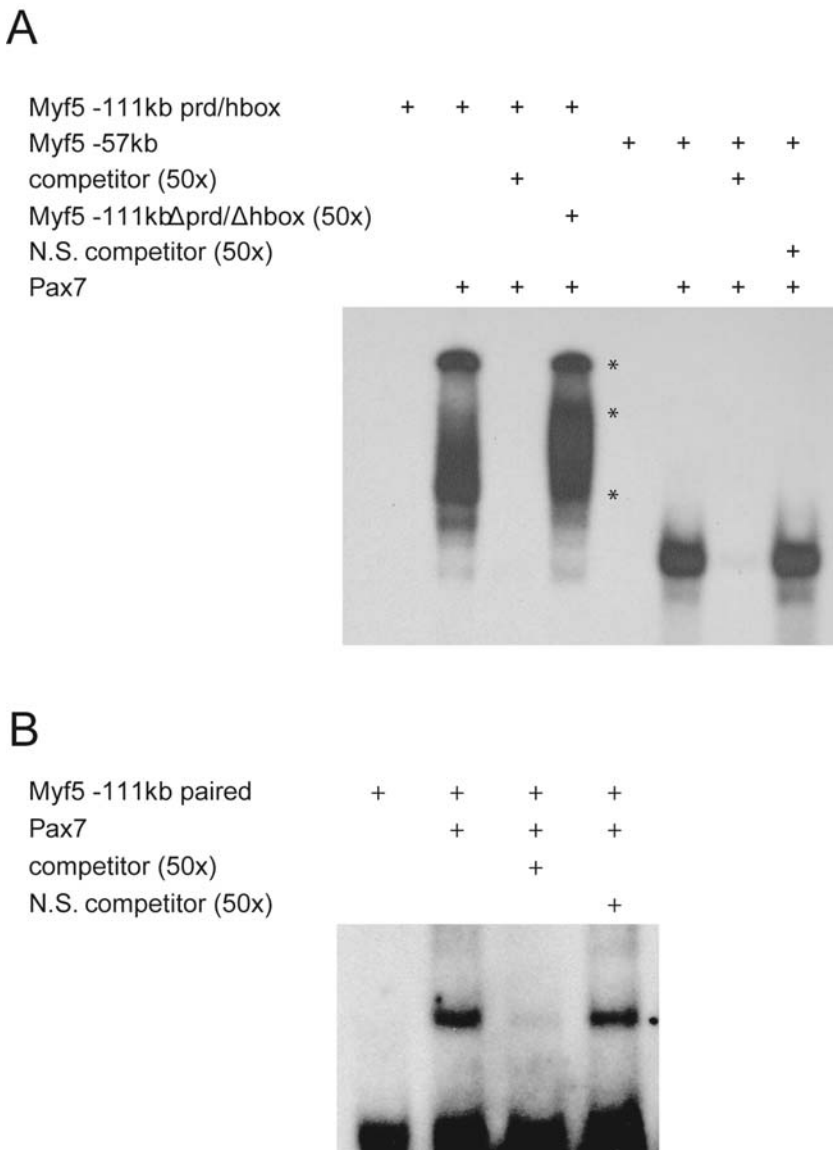


Figure 5. Mutagenesis of the paired and homeodomain motifs interferes with Pax7 binding at ECR111

(A) EMSA of ECR111 and ECR57 binding sites using specific oligonucleotides covering the footprinted regions. Purified Pax7 protein was incubated with labeled probes and complexes were resolved by 5% native PAGE in 0.5x TBE. Excess cold mutant probe with homeo and paired motifs scrambled was unable to compete for Pax7 binding. Several species of shifted bands were observed on the ECR111 fragment suggesting Pax7 may bind the probe in many configurations (i.e. multiple molecules, dimers, etc).
(B) EMSA of the paired-containing half of the ECR111 binding site indicates that Pax7 recognizes this motif as a monomer species.

sequence. Reporters were cotransfected along with Pax3, Pax7, Pax3-VP16, Pax7-VP16, or empty vector, into 10T1/2 fibroblasts. Reporter activity was measured by dual luciferase assay and normalized against *Renilla* luciferase. Because Pax3 and Pax7 are poor transactivators on their own (Corry and Underhill, 2005), VP16 fusions were constructed to fuse the transactivation domain from the herpes simplex virus VP16 protein in-frame with the 3' end of each.

Pax3-VP16 and Pax3, but neither Pax7-VP16 nor Pax7 were able to induce mild transactivation of M5-PRD (Figure 6A), consistent with our previous observations that Pax3 has a higher relative affinity for the paired motif than Pax7 (Punch et al., submitted). Conversely, Pax7-VP16 was able to induce a much higher transactivation of M5-HBOX relative to Pax3-VP16. Pax3 and Pax7 did not significantly induce transactivation at all compared to controls (Figure 6B). Strikingly, Pax3- and Pax7-VP16 competently induced reporter activity from the M5-PDHD construct, demonstrating a hugely synergistic effect when both sites are juxtaposed (Figure 6C). Mild transactivation of this construct was observed with Pax3 and Pax7. Taken together, these data support that ECR111 can act as enhancer of *Myf5* when recognized by Pax3 or Pax7 and is concordant with the idea this element is biologically significant as evidenced by its high degree of conservation.

ECR111 is an enhancer required for Myf5 expression in satellite cells.

We have previously shown that a BAC construct carrying 195kb upstream of the Myf5 transcriptional start site is sufficient to recapitulate the expression pattern of Myf5 and Mrf4 in the embryo (Carvajal et al., 2001) and in the adult (Zammit et al., 2004a). It has also been demonstrated that the -140kb to -88kb interval is necessary for the expression of Myf5 in

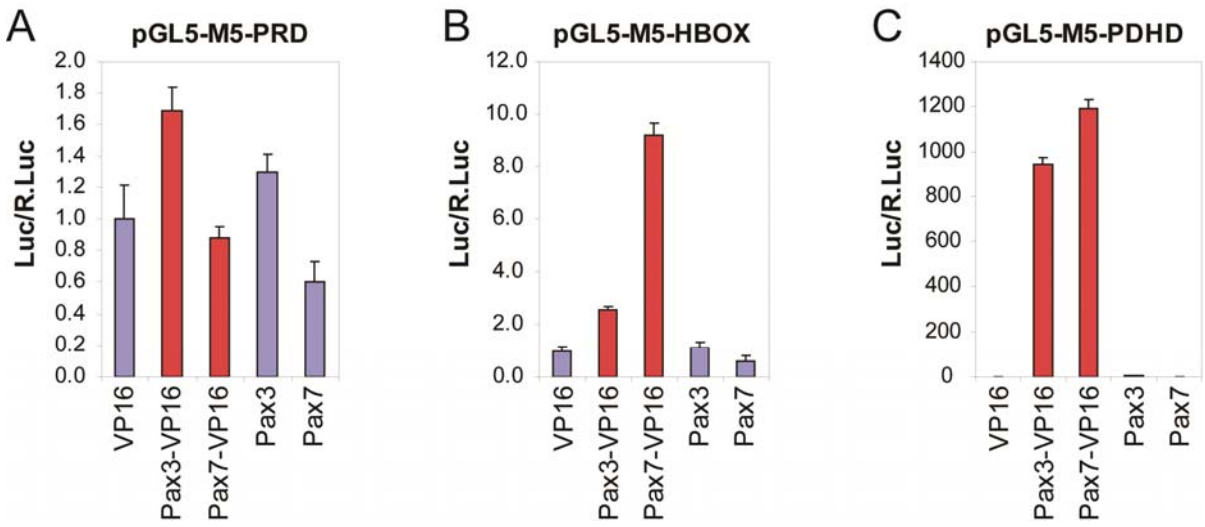


Figure 6. *Myf5* ECR111 is a novel enhancer regulated by Pax3 or Pax7

(A) Three copies of the prd-containing half of the ECR111 locus were directionally cloned into a pGL4.10 reporter vector driven by a *Myf5* minimal promoter (Summerbell et al., 2000). 10T1/2 fibroblasts were cotransfected with Pax3, Pax7, or fusion constructs consisting of Pax3 or 7 fused in-frame with a VP16 transcriptional activation domain (Lin et al., 1991) at the C-terminus. Expression of *luciferase* was measured and normalized to *renilla* expression. Relative luciferase activity was normalized against a control (VP16) reporter construct. Cotransfection with Pax3 or Pax3-VP16 produced a slight increase in relative luciferase expression compared to controls, whereas Pax7 constructs did not yield any increase.
(B) Three copies of the hbox-containing half of the ECR111 locus were concatamerized and cloned as above. Pax7-VP16 produced a significant increase in luciferase expression relative to controls and Pax3-VP16. Pax3-VP16 produced only a minor increase in luciferase activity relative to controls.
(C) Three copies of ECR111 containing both paired and homeodomain motifs were cloned as described. Both Pax3- and Pax7-VP16 drove luciferase expression to high levels, demonstrating that ECR111 has enhancer activity and suggests that a combination of paired and homeodomain motifs produce synergistic effects on enhancer activity.

quiescent satellite cells, while another element(s) controlling Myf5 expression in activated satellite cells is located in the -59kb to +40.6kb interval (Zammit et al., 2004a). Within the 52kb upstream interval harboring the satellite cell enhancer we have identified four evolutionary conserved regions, of which ECR111 is conserved among all vertebrates. Deletion of the ECR111 sequence abolishes transgene expression in satellite cells in the adult, while the expression of Myf5-nLacZ in muscle spindle cells is not affected, confirming that expression at these two sites in adult skeletal muscles is under the control of separate enhancers (Figure 7). In addition, activated satellite cells from cultured EDL single-fibers also express Myf5-nLacZ in the absence of ECR111 (data not shown), further indication that sequences within this enhancer are only required for expression in quiescent satellite cells.

Discussion

Regulation of *Myf5* is a complex situation whereby a large number of enhancers activate *Myf5* at different times and in different locations during development. It has been proposed that Myf5 might respond to a diverse panel of inductive signals though a number of different enhancers in order to function as a common entry point to skeletal muscle commitment (Carvajal et al., 2008).

By ChIP-Seq analysis of Pax3 and Pax7 binding, we have identified a number of putative Pax3/7 enhancers at ECRs. These observations suggest that either factor may utilize several different binding sites in order to confer Myf5 expression in different contexts. This notion is supported by the fact that expression of Myf5 can be driven by Pax3 or Pax7 in both embryonic and skeletal muscle and that adult Myf5 enhancers operate independently of other embryonic Myf5 enhancers. It is known that embryonic regulation of Myf5 can occur

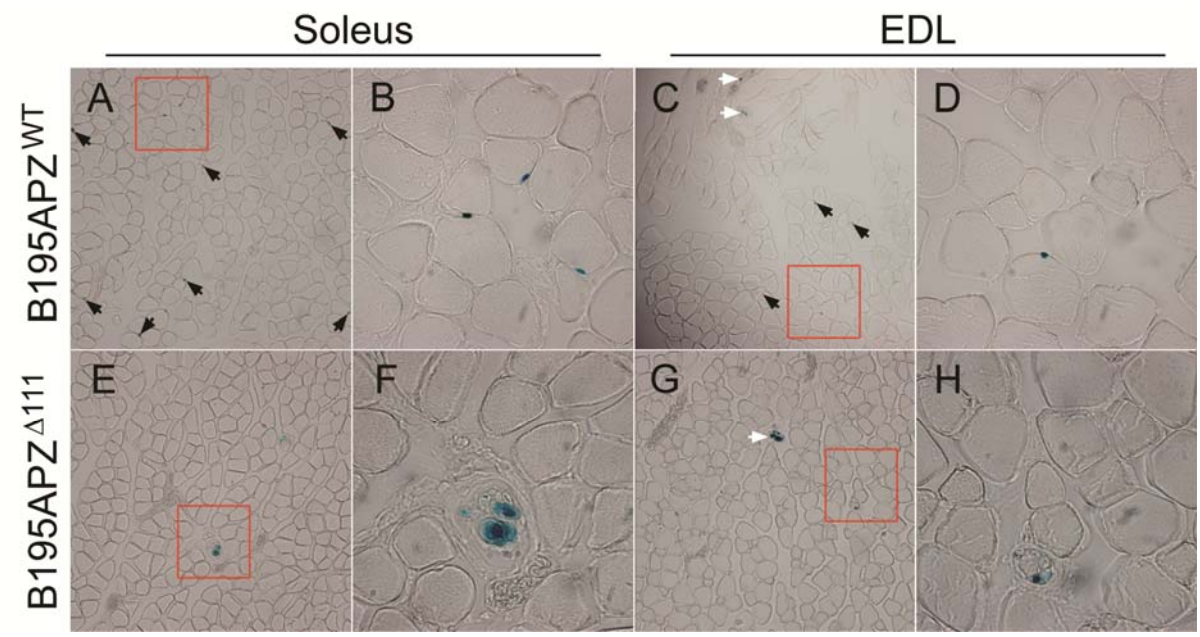


Figure 7. Deletion of the ECR111 enhancer abolishes Myf5-nLacZ expression in quiescent satellite cells

Cryosections of adult Soleus (A,B,E,F) and EDL (B,C,G,H) muscles from B195APZWT (A-D) and B195APZΔ111 (E-H) transgenic lines show that in the absence of the ECR111 enhancer harbouring the Pax7 binding site expression in the satellite cells is abolished. Note that in transgenic animals carrying the deleted BAC construct expression in muscle spindle cells, under the transcriptional control of a separate enhancer, is not affected. Red boxes in A, C, E and G indicate the close up regions shown in B, D, E and H, respectively. Additional satellite and muscle spindle cells not included within the close-up figures are also indicated (black and white arrows, respectively).

through the ECR57, driven by Pax3. Deletions within Myf5-LacZ BAC systems support the idea that the ECR111 element lies within the -140kb and -88kb region critical for the expression of Myf5 in quiescent satellite cells (Zammit et al., 2004a). Our work confirms that ECR111 is a Pax7-dependent enhancer that is evolutionarily conserved and directs Myf5 expression in this context. That we see continued expression of Myf5-nLacZ in the ECR111 mutant in muscle spindles reinforces the idea that separate genetic elements control Myf5 expression between quiescent satellite cells and muscle spindles (Zammit et al., 2004a).

These findings provide an understanding of adult Myf5 regulation that was not made clear by the studies of Chang et al. in their work describing the adult role of the A17 enhancer (Chang et al., 2004). The A17 element, located 17kb upstream of Myf5 directs Myf6 expression in early embryonic skeletal muscle of the myotome and Myf5 expression in later fetal skeletal muscle. In the adult, an A17-Myf5-nLacZ transgene is expressed by most satellite cells in the diaphragm and intercostal muscles, but by few satellite cells in soleus or EDL muscles (Chang et al., 2007). The limited extent to which A17 can drive Myf5 expression in adult cells argues that other regulatory elements are required for full Myf5 expression in satellite cells. Moreover, the A17 sequence contains a critical E-box which binds USF, but does not contain any Pax binding sites. Its lack of Pax binding sites implies that it operates in a Pax3/7-independent manner, and it has been demonstrated that Myf5 can regulate myogenesis independent of Pax gene expression (Relaix et al., 2006). Regardless, these data strongly indicate that Pax3/7 influences Myf5 expression by other means.

Because Pax7 null mice show no skeletal muscle defects through embryogenesis, it is suggested that Pax3 and Pax7 are redundant for initiating myogenesis in the embryo (Mansouri et al., 1996). Pax7 null animals rapidly lose their satellite cell pools following birth, demonstrating that Pax7 plays an essential role in maintenance of satellite cells by

driving self-renewal and propagation (Oustanina et al., 2004). Moreover, satellite cells are lost postnatally through apoptosis in Pax7 null mice (Relaix et al., 2006) regardless of Pax3 expression, indicating that Pax3 is not capable of replacing key functions of Pax7 in the adult. Consistent with this, our recent ChIP-Seq analysis of global Pax3 and Pax7 binding indicates that in adult cells, Pax3 recognizes only a small subset of Pax7 binding sites (Punch et al, submitted). Furthermore, Pax7 can prevent precocious differentiation of satellite cells by inhibiting the expression of MyoD and myogenenin (Olguin and Olwin, 2004; Relaix et al., 2006). Accordingly, we also observed that siRNA knockdown of Pax7 results in an upregulation of myogenin and other markers of differentiation.

Satellite cells exhibit heterogeneity for Myf5 expression despite their unform expression of Pax7. By using Myf5-Cre and Rosa26-YFP Cre-reporter alleles, Kuang et al. showed that roughly 10% of satellite cells have never expressed Myf5 (Kuang et al., 2007). Importantly, whereas Myf5+ cells were predisposed to myogenic differentiation, Myf5- cells contributed extensively to replenishing the satellite cell pool following intramuscular injection into *mdx* mice. Myf5- satellite cells predominantly divided in the apical-basal plane to give rise to Myf5- and Myf5+ daughter cells, whereas cells undergoing planar divisions typically generated identical daughter cells, demonstrating the surrounding niche plays an important role in the identity of newly formed daughter cells. It has also been suggested that Myf5 has a role linked to cell cycle dynamics, based on temporal expression data of MRFs (Day et al., 2009) and the observation that MyoD and Myf5 levels oscillate differently throughout the cell cycle (Kitzmann et al., 1998; Kitzmann and Fernandez, 2001). Furthermore, blocking of Myf5 degradation perturbs normal cell division (Doucet et al., 2005; Lindon et al., 2000) and *Myf5* mutant myoblasts are perturbed in their proliferation rate (Ustanina et al., 2007). Pax3 and Pax7 have also been shown to influence the kinetics of

cell cycle progression. We casually observe increased cell proliferation in Pax3/7 overexpressing myoblasts consistent with existing reports that have demonstrated the same in C2C12 and satellite cells, while showing that the expression of dominant negative forms greatly slow down proliferation (Collins et al., 2009; Olguin and Olwin, 2004). Our data demonstrate a direct regulatory relationship between Pax3/7 and Myf5 in adult cells, and taken with these reports, suggest that Pax7 might influence the proliferation of myogenic cells through modulation of Myf5. This is consistent with other studies which suggest a primary role of Pax7 is to maintain the proliferation of progenitor cells and prevent precocious differentiation (Olguin and Olwin, 2004; Zammit et al., 2006).

Collectively, these studies form a model for satellite cell behavior whereby Pax7 expression is required for the propagation and survival of satellite cells and the expression of Myf5. Signals emanating from the niche may play a large role in determining permissive conditions for the activation of Myf5 when Pax7 is expressed, but these relationships remain to be fully understood. Moreover, it was recently suggested that adult satellite cells undergo age-dependent changes in their requirements for Pax7 (Lepper et al., 2009). Conditional inactivation of Pax7 by tamoxifen-induced Cre recombination in satellite cells between postnatal days 60-90 did not impair regeneration and produced Pax7 mutant descendent cells that occupy the satellite cell niche and contribute to regenerating fibers. This raises important questions about alternative mechanisms of Myf5 regulation in adult cells.

Work in *Drosophila* suggests that Pax proteins recognize different target genes through various combinations of their DNA binding domains and maintains that synergistic binding of paired and homeodomains is required for the expression of *even-skipped* (Jun and Desplan, 1996). With our reporter constructs, we show that the presence of both paired and homeodomains drives luciferase expression in a highly synergistic fashion. Although the

mechanism by which synergy is achieved is not well defined, evidence of multiple species of shifted bands in our EMSA experiments with full length ECR111 probe suggests that the combination of sites facilitates binding of multiple complex configurations or multiple molecules. In support of this, while in the case of Myf5 ECR57, the size of the protected region containing the paired motif corresponds with the sizes footprinted for other Pax proteins (Chalepakis et al., 1991; Czerny et al., 1993; Epstein et al., 1994; Xu et al., 1999), the protected region for ECR111 is nearly twice that of ECR57, with the paired and homeodomain motifs separated by more than 25bp.

In summary, we have shown that Pax3 and Pax7 decorate the regulatory region for Myf5/Myf6. In addition to binding the known enhancer at ECR57, we determined that Pax3 and Pax7 recognize a novel conserved enhancer at ECR111 which contains both homeodomain and paired domain binding motifs, both of which are required for full enhancer activity. Significantly, the ECR111 enhancer drives the expression of Myf5 in quiescent satellite cells, but not muscle spindles.

Experimental Procedures

Homologous Recombination in E. coli

We generated a deletion cassette using the following oligonucleotides for the 5'homology arm: 5HA-F, 5' TCT AAA TTA TAT TGC ACA GAG TTC 3' and 5HA-R, 5' GAA CCA CCT TAA GAA CCA ATA C 3'; for the 3' homology arm: 3HA-F 5' tgg ttc tta agg tgg ttc TTT TCT TTT TAT GTG ACT ACA GTA A 3' and 3HA-R, 5' ATA TAA AAT GTG GAA AGA TCA GTC 3'. The lowercase in the 3HA-F oligonucleotide is complementary to the 5HA-R. After amplification individual homology arms were joined by PCR to generate the deletion cassette. The cassette was then used to delete ECR111 from the BAC construct

B195APZ (Carvajal *et al.*, 2001) to generate B195APZΔ111 by linear recombination as previously described (Carvajal *et al.*, 2008).

Production of transgenic mice and muscle sections

All *in vivo* experimentation was performed according to United Kingdom Home Office Regulations. BAC DNA was prepared using the QIAgen Maxiprep kit (Qiagen) following our own modification of the protocol (Carvajal *et al.* 2001) and used to inject fertilized mouse eggs from CBA/Ca × C57Bl/6 crosses, as previously described (Yee and Rigby 1993). Specific muscles where isolated from 20week old transgenic mice, fixed for 5min in 4% (v/v) PFA in PBS and stained for ß-galactosidase activity as previously described (Zammit et al., 2004). Muscle biopsies where embedded in Lamb/OCT (Thermo Fisher Scientific) following manufacturer's recomendations and 10µm sections obtained using a Leica CM1950 cryostat.

Luciferase Assays

10T1/2 fibroblasts were seeded at 1.0 x 10^4 cells/well in 24-well plates and supplemented with DMEM containing 10% FBS and no antibiotics. Cells were cotransfected with combinations of Pax3 and Pax7 expression vectors and luciferase reporter vectors using Lipofectamine 2000 reagent (Invitrogen). Myf5 ECR111 reporter vectors were constructed as follows. For the PRD and HBOX reporters, three copies of 30bp sequences covering the prd and hbox motifs were directionally concatamerized and cloned upstream of a minimal Myf5 promoter driving the *luc* gene of pGL4.10 (Promega). A 70bp sequence covering the entire ECR111 region, including both the prd and hbox regions was also concatamerized and cloned as above to generate the PDHD reporter. Reporter vectors were cotransfected with a

renilla vector standard and either empty vector (VP16), VP16-Pax7, VP16-Pax3, Pax7 or Pax3 for 17 hours. Luciferase assays were carried out using the Dual-Reporter Luciferase Assay System (Promega) and analysed on a Lumistar Optima (BMG Labtech) fluorescent plate reader.

Quantitative PCR

Selected targets from ChIP and microarray experiments were validated by PCR with a Mx3000P Real Time PCR System (Stratagene) using B-R SYBR Green Supermix for iQ (Quanta Biosciences). PCR primers for qPCR were designed using the online Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi; Whitehead Institute). Primers for gene expression analysis were designed to span at least one intron where possible. Gene expression values were normalized to GAPDH. Primers for ChIP analysis were designed to generate products of 100-200bp. ChIP values were normalized to a control locus -3kb from GAPDH TSS (chosen from a panel of negative control loci) and expressed as fold enrichment over control IP through ΔCt calculation. Primer sequences are listed in Table S1.

Protein expression and purification

C-terminally Flag-tagged Pax7 and Pax3 were expressed using baculovirus system. Recombinant baculoviruses are generated by using Bac-to-Bac system (Gibco-BRL) and infected to Sf9 cells. Infected Sf9 cells were incubated for 2 days and then collected by centrifugation. The cells were washed with cold calcium- and magnesium-free phosphate-buffered saline and stored at −80 ◦C until use. To purify Flag-tagged Pax7 and Pax3, cells were thawed, suspended in hypotonic buffer (10 mM Tris HCl, pH 7.6, 1% Nonidet P-40,

2mM MgCl2, 2 mM DTT, protease inhibitor cocktail (nacalai tesque) and incubated on ice for 10min. The nuclei were then collected by centrifugation, resuspended in High Salt buffer (25mM Tris HCl, pH 7.6, 0.1 % Triton X-100, 2mM MgCl2, 420mM NaCl, 2 mM DTT, protease inhibitor cocktail), and incubated on ice for 30min. After clarification by centrifugation, the nuclear extracts were incubated with protein G sepharose (GE healthcare life science) at 4 ◦C for 30 min to preclear the lysates and then incubated with anti-Flag M2-agarose beads (Sigma) at 4 ◦C for 3 hr. The resin was washed with wash buffer 1 (25 mM Tris HCl pH 7.6, 0.1 % Triton X-100, 2 mM MgCl2, 300mM NaCl, 2 mM DTT) for three times and then washed with wash buffer 2 (25 mM Tris HCl pH 7.6, 0.1 % Triton X-100, 2 mM MgCl2, 100mM NaCl) for twice. The proteins that remained bound to the resin were eluted by Flag peptide (Sigma) containing wash buffer 2 for twice. Eluted fractions were dialyzed with dialysis buffer (10 mM Tris HCl pH 7.6, 100mM NaCl, 2 mM DTT) and then freezed at −80 ◦C. The purity of the eluted fractions was verified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Electrophoretic Mobility Shift Assays (EMSA)

All probes were synthesized as full-length oligonucleotides, purified by HPLC, unless stated otherwise by Integrated DNA Technologies. 3µl of 10µM single stranded probes (+) were radiolabeled by incubation with T4 Kinase (Invitrogen) in the presence of 3µl γ-[^{32}P]-ATP for 1 hour at 37'C. Labeling was stopped by addition of STE buffer (40mM Tris-Cl pH 8.0, 40mM NaCl, 0.8mM EDTA). Double stranded probes were generated by addition 6µl of 10µM opposite strand probe (-) and mixtures were boiled for 2 minutes and slowly cooled to room temperature over 2 hours. Unincorporated [^{32}P]-ATP was removed by purifying double stranded probes through Microspin G25 columns (GE Healthcare). Probes were then

diluted 1:5 in water to yield a final concentration of ~125nM. Competitor probes were synthesized by combining 1µl of each single stranded probe (100µM) with 5µl STE buffer and 1µl water, boiling for 2 minutes and slowly cooling to room temperature. Probe concentrations were 12.5µM. Approximately 10,000 cpm of probe was used for each binding reaction. Reactions were analyzed on a 5% nondenaturing polyacrylamide gel (0.5X TBE), dried onto 3MM Whatman paper and exposed to BioMax MS x-ray film (Kodak). Probe sequences are listed in Table S1.

DNase I Footprinting

Genomic DNA from Balb/c mice was used as a PCR template for all probes. The Myf5 ECR111 probe was generated using the primers Myf5 -111 F FAM and Myf5 -111 R. Myf5 ECR57 probe was generated using the primers Myf5 -57 F FAM and Myf5 -57 R. Oligos were synthesized by Integrated DNA Technologies (IDT) and sequences are presented in Table S1. Triplicate PCR reactions were perfomed for 35 cycles under the conditions (95'C – 30sec; 58'C – 30sec; 72'C-45sec) and pooled. PCR reactions were cleaned up using the Purelink PCR purification kit (Invitrogen). 0.5-1.0ug of purified Pax7 or Pax3 protein or 20ug of BSA (NEB) was equilibrated for 10 minutes at room temperature in the following buffer: (150mM KCl, 5mM MgCl2, 0.1mM EDTA, 8% glycerol, 30mM Tris-Cl pH 8.0, 1mM DTT) along with 1ug of poly dI-dC (Sigma). 150 ug of each probe was then added for 20 minutes at room temperature. Probes were then digested with $7.5x10^{-3}$u of DNase I (Worthington Biochemicals) for 15 minutes at room temperature. Digests were stopped with the addition of 166mM EDTA and immediately purified using MinElute enzymatic reaction cleanup (Qiagen) and eluted into 10µl water. Control digestions were performed on an IgH probe. 5µl of digested DNA was added to 4.9µl of HiDi formamide (Applied Biosystems)

and 0.1µl GeneScan 500 LIZ size standards (Applied Biosystems). Mixtures were then heat denatured for 5' at 95'C, immediately cooled on ice, and analyzed with a 3730 DNA Analyzer (ABI) running a G5 dye set.

Western Blotting

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 (10-25 µg 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) with Biomax XAR film (Kodak). Primary antibodies used were: anti-FLAG (M2; 1:1000; 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 (1:1000; Sigma). Secondary antibodies were HRP-conjugated anti-mouse and anti-rabbit (1:4000; Biorad).

Acknowledgements

We thank Ricardo Ribas for providing the B195APZΔECR111 construct. M.A.R. holds the Canada Research Chair in Molecular Genetics and is an International Research Scholar of the Howard Hughes Medical Institute. This work was supported by grants to M.A.R. from the National Institutes of Health, the Howard Hughes Medical Institute, the Canadian Institutes of Health Research, and the Canada Research Chair Program. The authors declare no conflict of interest.

References

Bajard, L., Relaix, F., Lagha, M., Rocancourt, D., Daubas, P., and Buckingham, M.E. (2006). A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. Genes Dev 20, 2450-2464.

Birrane, G., Soni, A., and Ladias, J.A.A. (2009). Structural Basis for DNA Recognition by the Human PAX3 Homeodomain. Biochemistry 48, 1148-1155.

Braun, T., Bober, E., Rudnicki, M.A., Jaenisch, R., and Arnold, H.H. (1994). MyoD expression marks the onset of skeletal myogenesis in Myf-5 mutant mice. Development 120, 3083-3092.

Braun, T., Bober, E., Winter, B., Rosenthal, N., and Arnold, H.H. (1990). Myf-6, a new member of the human gene family of myogenic determination factors: evidence for a gene cluster on chromosome 12. Embo J 9, 821-831.

Braun, T., Buschhausen-Denker, G., Bober, E., Tannich, E., and Arnold, H.H. (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., Rudnicki, M.A., Arnold, H.H., and Jaenisch, R. (1992). Targeted inactivation of the muscle regulatory gene Myf-5 results in abnormal rib development and perinatal death. Cell 71, 369-382.

Carvajal, J.J., Cox, D., Summerbell, D., and Rigby, P.W. (2001). A BAC transgenic analysis of the Mrf4/Myf5 locus reveals interdigitated elements that control activation and maintenance of gene expression during muscle development. Development 128, 1857-1868.

Carvajal, J.J., Keith, A., and Rigby, P.W. (2008). Global transcriptional regulation of the locus encoding the skeletal muscle determination genes Mrf4 and Myf5. Genes Dev 22, 265-276.

Chalepakis, G., Fritsch, R., Fickenscher, H., Deutsch, U., Goulding, M., and Gruss, P. (1991). The molecular basis of the undulated/Pax-1 mutation. Cell 66, 873-884.

Chang, T.H., Primig, M., Hadchouel, J., Tajbakhsh, S., Rocancourt, D., Fernandez, A., Kappler, R., Scherthan, H., and Buckingham, M. (2004). An enhancer directs differential expression of the linked Mrf4 and Myf5 myogenic regulatory genes in the mouse. Dev Biol 269, 595-608.

Chang, T.H., Vincent, S.D., Buckingham, M.E., and Zammit, P.S. (2007). The A17 enhancer directs expression of Myf5 to muscle satellite cells but Mrf4 to myonuclei. Dev Dyn 236, 3419-3426.

Collins, C.A., Gnocchi, V.F., White, R.B., Boldrin, L., Perez-Ruiz, A., Relaix, F., Morgan, J.E., and Zammit, P.S. (2009). Integrated functions of Pax3 and Pax7 in the regulation of proliferation, cell size and myogenic differentiation. PloS one 4, e4475.

Corry, G.N., and Underhill, D.A. (2005). Pax3 target gene recognition occurs through distinct modes that are differentially affected by disease-associated mutations. Pigment cell research / sponsored by the European Society for Pigment Cell Research and the International Pigment Cell Society 18, 427-438.

Czerny, T., Halder, G., Kloter, U., Souabni, A., Gehring, W.J., and Busslinger, M. (1999). twin of eyeless, a Second Pax-6 Gene of Drosophila, Acts Upstream of eyeless in the Control of Eye Development. Molecular Cell 3, 297-307.

Czerny, T., Schaffner, G., and Busslinger, M. (1993). DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site, pp. 2048-2061.

Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell 51, 987-1000.

Day, K., Paterson, B., and Yablonka-Reuveni, Z. (2009). A distinct profile of myogenic regulatory factor detection within Pax7+ cells at S phase supports a unique role of Myf5 during posthatch chicken myogenesis. Dev Dyn 238, 1001-1009.

Doucet, C., Gutierrez, G.J., Lindon, C., Lorca, T., Lledo, G., Pinset, C., and Coux, O. (2005). Multiple phosphorylation events control mitotic degradation of the muscle transcription factor Myf5. BMC biochemistry 6, 27.

Edmondson, D.G., and Olson, E.N. (1990). A gene with homology to the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. Genes Dev 4, 1450.

Epstein, J.A., Glaser, T., Cai, J., Jepeal, L., Walton, D.S., and Maas, R.L. (1994). Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing, pp. 2022-2034.

Epstein, J.A., Shapiro, D.N., Cheng, J., Lam, P.Y., and Maas, R.L. (1996). Pax3 modulates expression of the c-Met receptor during limb muscle development. Proc Natl Acad Sci U S A 93, 4213-4218.

Esposito, C., Miccadei, S., Saiardi, A., and Civitareale, D. (1998). PAX 8 activates the enhancer of the human thyroperoxidase gene. The Biochemical journal 331 (Pt 1), 37-40.

Franz, T., Kothary, R., Surani, M.A., Halata, Z., and Grim, M. (1993). The Splotch mutation interferes with muscle development in the limbs. Anat Embryol (Berl) 187, 153-160.

Giordani, J., Bajard, L., Demignon, J., Daubas, P., Buckingham, M., and Maire, P. (2007). Six proteins regulate the activation of Myf5 expression in embryonic mouse limbs. Proc Natl Acad Sci U S A 104, 11310-11315.

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

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

Jun, S., and Desplan, C. (1996). Cooperative interactions between paired domain and homeodomain. Development 122, 2639-2650.

Kablar, B., and Rudnicki, M.A. (1999). Development in the absence of skeletal muscle results in the sequential ablation of motor neurons from the spinal cord to the brain. Dev Biol 208, 93-109.

Kassar-Duchossoy, L., Giacone, E., Gayraud-Morel, B., Jory, A., Gomes, D., and Tajbakhsh, S. (2005). Pax3/Pax7 mark a novel population of primitive myogenic cells during development. Genes Dev 19, 1426-1431.

Kaul, A., Koster, M., Neuhaus, H., and Braun, T. (2000). Myf-5 revisited: loss of early myotome formation does not lead to a rib phenotype in homozygous Myf-5 mutant mice. Cell 102, 17-19.

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

Kitzmann, M., and Fernandez, A. (2001). Crosstalk between cell cycle regulators and the myogenic factor MyoD in skeletal myoblasts. Cell Mol Life Sci 58, 571-579.

Kuang, S., Charge, S.B., Seale, P., Huh, M., and Rudnicki, M.A. (2006). Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. J Cell Biol 172, 103-113.

Kuang, S., Kuroda, K., Le Grand, F., and Rudnicki, M.A. (2007). Asymmetric self-renewal and commitment of satellite stem cells in muscle. Cell 129, 999-1010.

Lepper, C., Conway, S.J., and Fan, C.M. (2009). Adult satellite cells and embryonic muscle progenitors have distinct genetic requirements. Nature 460, 627-631.

Lin, Y.S., Ha, I., Maldonado, E., Reinberg, D., and Green, M.R. (1991). Binding of general transcription factor TFIIB to an acidic activating region. Nature 353, 569-571.

Lindon, C., Albagli, O., Domeyne, P., Montarras, D., and Pinset, C. (2000). Constitutive instability of muscle regulatory factor Myf5 is distinct from its mitosis-specific disappearance, which requires a D-box-like motif overlapping the basic domain. Mol Cell Biol 20, 8923-8932.

Mansouri, A., Stoykova, A., Torres, M., and Gruss, P. (1996). Dysgenesis of cephalic neural crest derivatives in Pax7-/- mutant mice. Development 122, 831-838.

McKinnell, I.W., Ishibashi, J., Le Grand, F., Punch, V.G., Addicks, G.C., Greenblatt, J.F., Dilworth, F.J., and Rudnicki, M.A. (2008). Pax7 activates myogenic genes by recruitment of a histone methyltransferase complex. Nature cell biology 10, 77-84.

Miner, J.H., and Wold, B. (1990). Herculin, a fourth member of the MyoD family of myogenic regulatory genes. Proc Natl Acad Sci U S A 87, 1089-1093.

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

Olguin, H.C., and Olwin, B.B. (2004). Pax-7 up-regulation inhibits myogenesis and cell cycle progression in satellite cells: a potential mechanism for self-renewal. Dev Biol 275, 375-388.

Ott, M.O., Bober, E., Lyons, G., Arnold, H., and Buckingham, M. (1991). Early expression of the myogenic regulatory gene, myf-5, in precursor cells of skeletal muscle in the mouse embryo. Development 111, 1097-1107.

Oustanina, S., Hause, G., and Braun, T. (2004). Pax7 directs postnatal renewal and propagation of myogenic satellite cells but not their specification. Embo J 23, 3430-3439.

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

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

Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. (2004). Divergent functions of murine Pax3 and Pax7 in limb muscle development. Genes Dev 18, 1088-1105.

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

Rudnicki, M.A., Braun, T., Hinuma, S., and 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.

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

Seale, P., Sabourin, L.A., Girgis-Gabardo, A., Mansouri, A., Gruss, P., and Rudnicki, M.A. (2000). Pax7 is required for the specification of myogenic satellite cells. Cell 102, 777-786.

Summerbell, D., Ashby, P.R., Coutelle, O., Cox, D., Yee, S., and Rigby, P.W. (2000). The expression of Myf5 in the developing mouse embryo is controlled by discrete and dispersed enhancers specific for particular populations of skeletal muscle precursors. Development 127, 3745-3757.

Tajbakhsh, S., and Buckingham, M.E. (1994). Mouse limb muscle is determined in the absence of the earliest myogenic factor myf-5. Proc Natl Acad Sci U S A 91, 747-751.

Tajbakhsh, S., Rocancourt, D., and Buckingham, M. (1996). Muscle progenitor cells failing to respond to positional cues adopt non-myogenic fates in myf-5 null mice. Nature 384, 266-270.

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

Ustanina, S., Carvajal, J., Rigby, P., and Braun, T. (2007). The myogenic factor Myf5 supports efficient skeletal muscle regeneration by enabling transient myoblast amplification. Stem Cells 25, 2006-2016.

Wilson, D., Sheng, G., Lecuit, T., Dostatni, N., and Desplan, C. (1993). Cooperative dimerization of paired class homeo domains on DNA. Genes Dev 7, 2120-2134.

Wright, W.E., Sassoon, D.A., and Lin, V.K. (1989). Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. Cell 56, 607-617.

Xu, H.E., Rould, M.A., Xu, W., Epstein, J.A., Maas, R.L., and Pabo, C.O. (1999). Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. Genes Dev 13, 1263-1275.

Zammit, P.S., Carvajal, J.J., Golding, J.P., Morgan, J.E., Summerbell, D., Zolnerciks, J., Partridge, T.A., Rigby, P.W., and Beauchamp, J.R. (2004). Myf5 expression in satellite cells and spindles in adult muscle is controlled by separate genetic elements. Dev Biol 273, 454-465.

Zammit, P.S., Relaix, F., Nagata, Y., Ruiz, A.P., Collins, C.A., Partridge, T.A., and Beauchamp, J.R. (2006). Pax7 and myogenic progression in skeletal muscle satellite cells. J Cell Sci 119, 1824-1832.

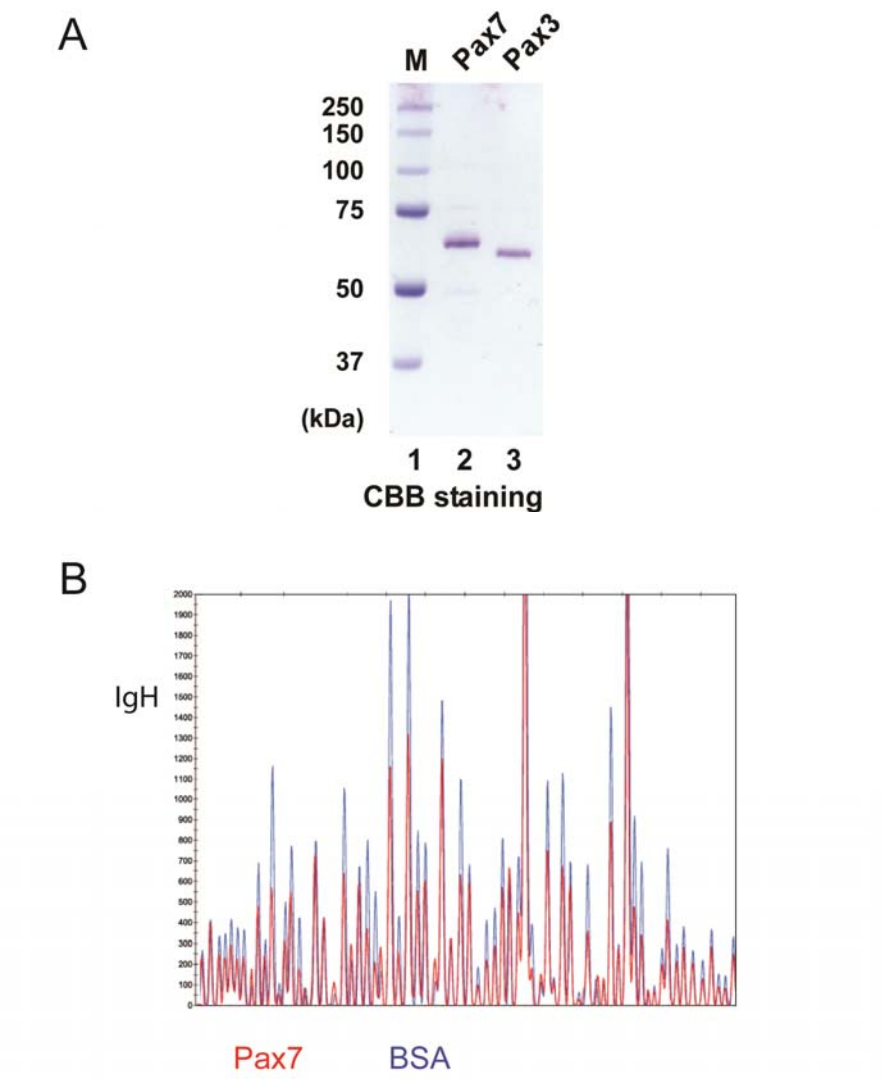


Figure S1. Pax7 does not Bind the IgH Control Locus

(A) Coomassie Brilliant Blue staining of Pax3-FLAG and Pax7-FLAG protein, purified from Sf9 cells infected by recombinant baculovirus. Purified protein was used for all footprinting and EMSA applications.
(B) The IgH locus is not hypersensitive to DNaseI digestion when incubated with Pax7 (red). Control incubation with BSA is in blue.

Table S1. Primer and Probe Sequences

EMSA	
Myf5 -57kb	CCATGCAATTAGTCATGCCTTTATGATT
Myf5 -111kb paired	CCATCATTTCCTGATTGTCATGCTTCTATCCTTGCACAGG
Myf5 -111kb	TTCACAATAATGCATTTTCTGTAACTAACAACCATCATTTCCTGATTGTCATG CTTCTAT
Myf5 -111kb ∆prd ∆hbox	TTCACAACGGCTAGCCCTCTTCGGACGGAAACCATCATTTCCTGATCTCAG CTACCCTAT
Luciferase	
Myf5 -111kb PD F BamHI	GATCTTCATTTCCTGATTGTCATGCTTCTATCCTTGCACAG
Myf5 -111kb PD R BglII	GATCCTGTGCAAGGATAGAAGCATGACAATCAGGAAATGAA
Myf5 -111kb HD F BamHI	GATCTCTGGTTTTCACAATAATGCATTTTCTGTAACTAACG
Myf5 -111kb HD R BglII	GATCCGTTAGTTACAGAAAATGCATTATTGTGAAAACCAGA
Myf5 -111kb HDPD F BamHI	GATCCTGTGCAAGGATAGAAGCATGACAATCAGGAAATGATGGTTGTTA GTTACAGAAAATGCATTATTGTGAAAACCAGA
Myf5 -111kb HDPD R BglII	GATCTCTGGTTTTCACAATAATGCATTTTCTGTAACTAACAACCATCATTTCC TGATTGTCATGCTTCTATCCTTGCACAG
Footprinting	
Myf5 -57 F FAM	(6FAM) - CTGTGCCTCTTACTGTGTGG
Myf5 -57 R	CTTGGGGCTTAAAAGCCATA
Myf5 -111 F FAM	(6FAM) - CATCCCACATAATCCAATCAC
Myf5 -111 R	TGCTAGGCAAAATGATGGTC
IgH	(6FAM) – CCGATCAGAACCAGAACACCTGCAGCAGCTGGCAGGAA GCAGGTCATGTGGCAAGGCTATTTGGGGAAGGGAAAATAAAACCACTAGG TAAACTTGTAGCTGTGGTTTGAAGAAGTGGTTTTGAAACACTCTGTCCAGC CCCACCA
qPCR	
Myf5-1F	TGAGGGAACAGGTGGAGAAC
Myf5-3R	AGCTGGACACGGAGCTTTTA
Pax7-F	CTGGATGAGGGCTCAGATGT
Pax7-R	GGTTAGCTCCTGCCTGCTTA
Myog-F	GAAAGTGAATGAGGCCTTCG
Myog-R	ACGATGGACGTAAGGGAGTG
GAPDH-F	TGTCCGTCGTGGATCTGAC
GAPDH-R	GGTCCTCAGTGTAGCCCAAG
MRF4-1F	GGAGTGCCATCAGCTAACATT
MRF4-2R	GAATGATCCGAAACACTTGG
MyoD-F	TACCCAAGGTGGAGATCCTG
MyoD-R	CATCATGCCATCAGAGCAGT
Pax3-F	TATTCCACAAGCCGTGTCA
Pax3-R	GAGCTGCTGTCTGCGTTC

ChIP	
GAPDH-3kb-F	TCCAGTGAGGACGGTATGAT
GAPDH-3kb-R	CATAAAGATGGGGCAAAATG
Myf5-111-F	CATCCCACATAATCCAATCAC
Myf5-111-R	ACACAGATGGATGGGAAAGA
Myf5-57F	TGTGGCTCTCTCTCCGTATG
Myf5-57R	AATACAGACATGCAGGCTTCAC

Chapter 4 – General Discussion

Discussion

4.1. Overview

Initially, this research project set out to identify DNA targets of Pax7, as it had been identified as a critical factor for satellite cell survival in adult mice (Seale et al., 2000). While Pax3 had been extensively studied as a regulator of embryonic myogenesis and progenitor cell migration in limb buds, relatively little molecular information has emerged about Pax7 molecular functions. Based on the known roles of Pax3 as a regulator of *MyoD* and *Myf5* expression in the embryo, we anticipated that Pax7 might directly regulate MRF family members in adult satellite cells, and that its molecular targets might explain why Pax7 is required for satellite cell survival and function.

Early technical approaches combined ChIP with representational difference analysis (RDA) and shotgun cloning strategies to sequence DNA targets. As RDA is best suited to background subtraction of cDNA libraries, and not genomic DNA, this approach was later replaced with the construction of SAGE-style sequence tag vector libraries. However, both methods failed to acknowledge the abundance of background ChIP fragments and the propensity for repetitive DNA elements to contribute to a high false positive rate. Accordingly, these libraries could not be sequenced to saturating points where enriched sequences could be distinguished from background noise. The emergence of ultra-high throughput sequencing (i.e. Solexa) provided the necessary technical platform for transcription factor analysis, as millions of sequence tags provided a saturating density map of ChIP fragments. Through this approach we identified tens of thousands of putative Pax7 binding sites. A highly enriched binding site discovered in the regulatory region of *Myf5* suggested that Pax7 might directly regulate Myf5 expression in satellite cells.

Mass spectrometry analysis of Pax7 interacting proteins indicated that Pax7 associates with Ash2L and Wdr5, members of the MLL2 histone methyltransferase complex (Appendix A) (McKinnell et al., 2008). Concurrently, the *Myf5* -57.5kb enhancer had emerged as a target of Pax3 and Six1/4 binding in the embryo (Bajard et al., 2006). ChIP-Seq of Pax7 had also identified the ECR57 enhancer, although ChIP-PCR validation experiments routinely suggested that the ECR111 putative enhancer was a stronger, more robust binding site. Additionally, it was known that the satellite cell-specific element for *Myf5* expression resided between -88kb and -140kb upsteam (Zammit et al., 2004a). However, both sites were used in the initial analysis of Pax7's role in chromatin remodeling. Modulation of Pax7 levels by siRNA transfection or proviral overexpression led to a concomitant change in the enrichment of Ash2L occupancy at ECR111 (not shown) and ECR57 sites and a matching change in the level of H3K4 trimethylation along the *Myf5* locus downstream of either site. Accordingly, we concluded that Pax7 regulates the expression of target genes through chromatin remodeling events.

Analysis of the ECR111 putative enhancer entailed DNaseI footprinting to precisely define the binding site. Moreover, our observations with ECR57 enrichment and known Pax3 binding to the ECR57 site led us to hypothesize that Pax3 and Pax7 might recognize identical binding sites. Because Pax3 is not expressed by most satellite cells and because of Pax7's role in chromatin modeling, it was not known whether the two highly homologous proteins would function equivalently at the molecular level. Footprinting and EMSA experiments confirmed that both Pax3 and Pax7 bound to identical sites within the ECR57 and ECR111. Furthermore, we identified both paired and homeodomain-like motifs in the ECR111 ChIP site, confirming the relevance of this site as a direct target of Pax3/7 binding. Followup work on the Myf5 ECR111 locus provided strong evidence of binding by both

Pax3 and Pax7 *in vivo* (ChIP). Luciferase reporter experiments validated the locus as an enhancer for *Myf5* and suggested that the combination of homeo and paired motifs synergistically drives expression. By use of an ECR111-deleted BAC, it was shown that this element is required for *Myf5* expression in satellite cells.

These observations prompted us to perform ChIP-Seq on Pax3, and these experiments were carried out concurrently with a repetition of Pax7 and empty vector controls to facilitate direct comparison. The use of TAP tagged constructs was ideal for generating comparable cell lines; because Pax3 is not expressed by most satellite cells, achieving comparable expression in primary myoblasts required low level proviral expression of both Pax3-TAP and Pax7-TAP. Expression levels of these genes were high enough to out-compete endogenous Pax7 binding and assert that stoichiometry generally favored homodimers of tagged proteins (rather than heterodimers with endogenous Pax7). Moreover, this system benefitted from a surprising observervation that endogenous Pax7 expression was suppressed by either construct.

Comparison of Pax3 and Pax7 ChIP-Seq data revealed a striking difference in the abundance of enriched regions between each sample. Pax3 enriched sites comprised a small fraction (6.4%) of Pax7 sites. The majority of high-scoring Pax3 sites, as derived from MACS p-values, were found to overlap with Pax7 sites. Conversely, Pax7 sites overlapping with Pax3 were among the top scoring Pax7 sites. Collectively, these data argue that Pax3 recognizes high affinity Pax7 sites.

Consensus sequences derived from *de novo* motif searching of top scoring Pax3 and Pax7 peaks uncovered only paired and homeodomain-like motifs. We anticipated possibly finding MEF3 motifs, as work on the ECR57 element suggested that Pax3 might function in conjunction with Six1/4 for *Myf5* regulation (Giordani et al., 2007). However, MEF3 motifs

were not uncovered *de novo*; FIMO identified degenerate sites in approximately 10% of peaks and this motif was only slightly more common in sites that overlapped with Pax3 compared to Pax7. While this does not preclude cooperation of Six proteins and Pax3/7 for the regulation of target genes, we did not find compelling evidence that this mechanism is significant.

Strikingly, we observed that Pax7 sites and Pax3 sites exhibited a bias in their relative abundance of paired and homeodomain motifs. Pax7 unique sites more frequently contained homeodomain motifs, while Pax3 sites more frequently contained paired domain motifs. In comparing the sequences of motifs between the two sets of binding sites, we did not observe any differences in composition nor were degenerate bases significantly different in their ratios. These data suggested that Pax3 and Pax7 were reciprocally biased in their binding affinity to each motif, a result validated by EMSA and luciferase reporter experiments.

Taken together, these data not only provided maps of Pax3 and Pax7 transcriptional output, but also provided a novel molecular mechanism to explain biological differences in Pax3 and Pax7 behavior; although Pax3 and Pax7 can recognize common binding sites, their relative affinities for binding individual sites are likely dependent on the presence and/or abundance of homeo and paired motifs (Figure 1). However, this alone may not explain the degree of dominance by Pax7 over Pax3 in adult cells; cellular context may be important for binding sites that are strongly targeted by Pax3 in embryonic cells but not in adult cells.

4.2. Future Directions

That these paralogous proteins have differential genomic binding behaviors despite having seemingly identical binding motifs raises important questions about other sequence-independent mechanisms that might further govern binding affinity. Moreover, because

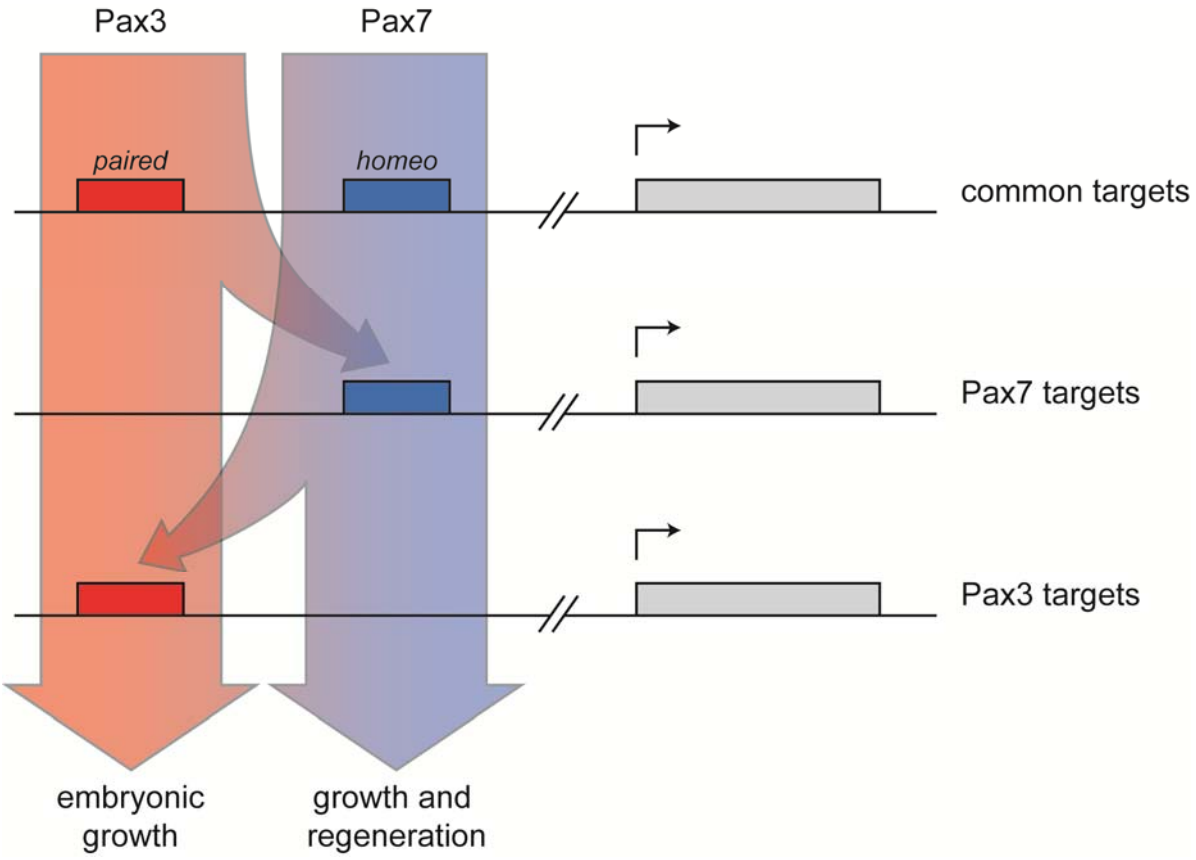


Figure 1. A model for how biased affinities of Pax3 and Pax7 result in the unique activation of genetic programs.

Pax3 and Pax7 recognize both common and unique gene targets. Although both factors can bind to identical motifs, affinity biases drive dominant associations between each factor and its unique gene targets. The resulting transcriptional signatures reflect unique the biological functions of each factor. In adult satellite cells, progenitor cells responsible for postnatal growth and regeneration, Pax7 uniquely activates genetic programs that promote satellite cell proliferation and survival. Conversely, in embryonic muscle progenitors, Pax3 drives the myogenic program, functions in the survival of progenitor cells, and facilitates migration of precursor cells to sites of future muscle development.

binding affinity appears to be dynamic at some sites (i.e. Fgfr4, where Pax7 binding supplants Pax3 binding in adult cells), we are curious what mechanisms are dependent on cellular context for modulating Pax3 or Pax7 binding affinity, resulting in distinct biological roles for each factor.

The argument for sequence-independent mechanisms is motivated by the fact that *de novo* searching of Pax3 and Pax7 peaks identified identical motif sequences. A caveat of this observation may be that our seed motifs derived from GLAM2 and MEME searches created a self-fulfilling prophecy when used as input for the FIMO searches; in a worst-case scenario, this could have resulted in the generation of identically-weighted consensus motifs merely because the FIMO search was directed by a seed sequence. However, this is unlikely because a high degree of degeneracy was allowed in the FIMO matching algorithm, and consensus motifs were assembled from an array of loosely matching hits. Therefore, it is more likely that sequence-independent mechanisms are responsible for differences between Pax3 and Pax7 binding.

A significant amount of work has demonstrated the importance of Pax3 phosphorylation for its normal function; phosphorylation of Pax3 is crucial for efficient DNA binding and subsequent transcriptional activity (Amstutz et al., 2008). Certain cofactors may be required for efficient Pax3 binding, but not for Pax7. Our mass spectrometry analysis identified a cohort of Pax7 interacting proteins that may be relevant to its transcriptional activity. Furthermore, Pax3 appears to be phosphorylated during proliferation, but this is rapidly lost upon the induction of differentiation, a time when transcriptional programs are undergoing dynamic shifts (Dietz et al., 2009; Miller et al., 2008). Therefore, it is conceivable that phosphorylation alters the biological activity of Pax3 and/or Pax7. Future studies should be undertaken to explore and compare post-translational

modifications of Pax3 and Pax7 in quiescent and activated satellite cells, to determine their biological significance.

Specific functions for *Pax3* and *Pax7* isoforms have remained unclear, however the relative abundance of isoforms are shown to differ among Pax7-expressing tissues (Ziman and Kay, 1998). The insertion of glutamine in Q+ isoforms alters the spacing between the helix-turn-helix motifs within the paired domain, which may modify binding site recognition. If the splice modifications result in variability in binding affinity or specificity, a ChIP-Seq study comparing the binding maps of all Pax7 isoforms may provide valuable insight into their relative roles.

Forced expression of Pax7-TAP and Pax3-TAP resulted in the unexpected downregulation of endogenous Pax7 protein. This observation suggests that auto- and cross-regulatory feedback mechanisms exist to tightly control Pax7 expression, however, the regulation of Pax7 remains poorly understood. Only recently have studies began to uncover mechanisms that operate through Pax7 to regulate satellite cell self-renewal, proliferation, and differentiation. Myostatin, a member of the TGF-β superfamily, negatively regulates myogenesis by impeding satellite cell activation and self-renewal (McCroskery et al., 2003). It is now further evident that myostatin negatively regulates Pax7 expression in myoblasts through a mechanism dependent on ERK1/2 MAPK signaling (McFarlane et al., 2008). TGF-β is also known to inhibit myogenesis through an ERK-dependent pathway (Pena et al., 2000), therefore it is conceivable that other superfamily members may regulate Pax7 in a similar fashion. MicroRNAs have also been shown to regulate the expression of genes and signaling pathways important for myogenesis (van Rooij et al., 2008). miR-1 and miR-133 have been shown to play central roles in modulating myoblast proliferation and differentiation (Chen et al., 2006). More recently, miR-1 and miR-206 were demonstrated to

facilitate the differentiation of satellite cells by restricting their proliferative potential through the direct regulation of Pax7 (Chen et al., 2010). Future studies aimed at dissecting regulatory inputs for Pax7 may be critically important for understanding or manipulating satellite cell self-renewal, proliferation, or differentiation.

This study was aimed at determining and comparing the direct targets of Pax3 and Pax7 in primary myoblasts. In order to generate comparable datasets in our ChIP-Seq experiments, both Pax3 and Pax7 were tagged and overexpressed in our cell lines. Pax7 is expressed at higher levels in satellite cells than activated satellite cells (Fukada et al., 2007), however the role of Pax7 in quiescent versus activated cells is not well understood. Pax7 dosage may be an important factor for the maintenance of a quiescent state or resistance to apoptosis. The gene signature of quiescent and activated cells reveals intriguing differences, but whether any differences can be attributed to Pax7 levels remains to be examined. Similarly, Pax3 expression is absent from most satellite cells and is not expressed in our myoblast preparations. Diaphragm muscle, on the other hand, is rich in Pax3-expressing satellite cells. Although Pax3 cannot rescue the loss of Pax7, Pax3 may have specific functions in these cells that are not evident in our Pax3 exogenously-expressing cells. Ideally, future experiments may investigate the binding behaviors of endogenous Pax3 and Pax7 proteins, however this will require the development of specific antibodies suitable for ChIP that are not currently commercially available.

4.3. Significance

Pax7 is a crucial factor for the postnatal survival of myogenic satellite cells and regulates the expression of myogenic determinants such as *Myf5*. Pax3, a paralog of Pax7, also regulates the expression of key myogenic proteins and is required for the migration of

muscle precursor cells into developing limb buds, however is not expressed by most satellite cells nor is it sufficient to compensate for the loss of Pax7.

This thesis provides novel insight into the mechanisms behind their divergent function. By ChIP-Seq, we observed that Pax7, in contrast to Pax3, is a dominant transcription factor in adult primary myoblasts. Although both proteins are able to recognize identical sequence motifs, Pax3 only binds a small subset of Pax7 sites. By comparison, we see that Pax3 and Pax7 exhibit reciprocal affinities for paired and homeodomain motifs. Accordingly, the abundance of homeodomain motifs in Pax3/7 binding sites could be a major reason for transcriptional dominance of Pax7.

We find that transcription factors comprise a major functional group within genes proximal to Pax7 binding sites. Cross-analysis of binding data with gene expression data indicates that Pax7 functions mainly as a gene activator and that its target genes correspond with a proliferative cellular state while repressed genes are associated with myogenic differentiation. Taken together, these data suggest that Pax7 functions by activating the transcriptional programs that promote the proliferation of myogenic precursor cells. By contrast, Pax3's relatively underwhelming affinity for binding sites likely reflects its inability to drive these transcriptional programs, providing insight into why it cannot compensate for the loss of Pax7.

Finally, we have confirmed that *Myf5* is a direct target of Pax7 regulation in satellite cells. We have identified a novel Pax3/7-dependent enhancer that is required for the expression of *Myf5* specifically in satellite cells, but not muscle spindles. Importantly, the location of this enhancer is consistent with deletion constructs that have delineated satellite cell-specific requirements for *Myf5* expression.

4.4. Biomedical Implications

A large amount of current research is directed at exploiting the therapeutic potential of muscle stem cells for the treatment of degenerative diseases such as muscular dystrophies and age-related muscle wasting (sarcopenia). Work by Kuang et al. has demonstrated the potential for satellite cell transplantation as a viable method for repopulation of the satellite cell compartment, however, it is clear that recapitulating signals from the niche are critical for effective regeneration (Kuang et al., 2007). Accordingly, it is evident that cell-based strategies require a strong understanding of the pathways that regulate differentiation of stem cell populations in order to effectively manipulate them.

Dissecting the molecular networks regulated by Pax7 is a critical step towards manipulating the availability and efficacy of muscle stem cells for therapeutic purposes. The transplantation of myofiber-associated satellite cells produces more effective engraftment of satellite cells than cultured cells. When coupled with injury, myofiber-associated cells yield long-term enrichment in muscle mass, stem cell numbers, and muscular force due to increased regenerative capacity of donor satellite cells (Hall et al., 2010). Some myogenic cells isolated from Pax7-/- muscle can undergo long-term expansion, however these cells lack the myogenic potential to regenerate muscle. Rescuing Pax7 expression in these cells allows them to regenerate dystrophin-expressing fibers in *mdx* mice (Lu et al., 2008). Similarly, CD45+/Sca1+ adult stem cells from injured Pax7-/- muscle can regain their myogenic potential by forced induction of Pax7 (Seale et al., 2004). Our own data suggests that Pax7 stimulates growth rates of satellite cells and enhances the expression of cell surface receptors and signaling molecules. Thus, we may now be able to interpret how signals from

the niche increase the efficacy of transplanted satellite cells, and by pharmacological or genetic means, improve the efficacy of cultured cells.

Rhabdomyosarcomas are common soft tissue carcinomas that occur in children and are related to the skeletal muscle lineage. Fusions between the potent transactivation domain of FoxO1 (FKHR) and the DNA-binding domains of either Pax3 or Pax7 are generally observed in alveolar RMS, while Pax7 expression has been associated with embryonal RMS (Tiffin et al., 2003). The transcriptional and cell survival functions of Pax3/7 may prove valuable to the understanding and treatment of this disease, as misregulation of Pax3/7 target genes by Pax3-FKHR and Pax7-FKHR likely contributes to disease development and progression.

4.5. Conclusion

In contrast to Pax3, Pax7 regulates a large number of transcriptional networks crucial for the proliferation and survival of satellite cells. In particular, it targets a large number of transcription factors which may promote cell growth and inhibit myogenic differentiation. Moreover, Pax7 regulates the myogenic program directly through the activation of *Myf5* via a satellite cell-specific enhancer.

References

Abou-Khalil, R., Le Grand, F., Pallafacchina, G., Valable, S., Authier, F.J., Rudnicki, M.A., Gherardi, R.K., Germain, S., Chretien, F., Sotiropoulos, A., *et al.* (2009). Autocrine and paracrine angiopoietin 1/Tie-2 signaling promotes muscle satellite cell self-renewal. Cell Stem Cell *5*, 298-309.

Amstutz, R., Wachtel, M., Troxler, H., Kleinert, P., Ebauer, M., Haneke, T., Oehler-Jänne, C., Fabbro, D., Niggli, F.K., and Schäfer, B.W. (2008). Phosphorylation Regulates Transcriptional Activity of PAX3/FKHR and Reveals Novel Therapeutic Possibilities, pp. 3767-3776.

Amthor, H., Christ, B., Weil, M., and Patel, K. (1998). The importance of timing differentiation during limb muscle development. Curr Biol *8*, 642-652.

Anderson, J., and Pilipowicz, O. (2002). Activation of muscle satellite cells in single-fiber cultures. Nitric Oxide *7*, 36-41.

Asakura, A., and Rudnicki, M.A. (2002). Cellular and molecular mechanisms regulating skeletal muscle development. In Mouse Development, J. Rossant, and P.P. Tam, eds. (New York, Academic Press), pp. 253-278.

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

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res *37*, W202-208.

Bajard, L., Relaix, F., Lagha, M., Rocancourt, D., Daubas, P., and Buckingham, M.E. (2006). A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. Genes Dev *20*, 2450-2464.

Barr, F.G. (2001). Gene fusions involving PAX and FOX family members in alveolar rhabdomyosarcoma. Oncogene *20*, 5736-5746.

Bennicelli, J.L., Edwards, R.H., and Barr, F.G. (1996). Mechanism for transcriptional gain of function resulting from chromosomal translocation in alveolar rhabdomyosarcoma. Proc Natl Acad Sci U S A *93*, 5455-5459.

Birrane, G., Soni, A., and Ladias, J.A.A. (2009). Structural Basis for DNA Recognition by the Human PAX3 Homeodomain. Biochemistry *48*, 1148-1155.

Bischoff, R. (1994). The satellite cell and muscle regneration. In Myology, A.G. Engel, and C. Franzini-Armstrong, eds. (New York, McGraw-Hill), pp. 97-118.

Bittner, R.E., Schofer, C., Weipoltshammer, K., Ivanova, S., Streubel, B., Hauser, E., Freilinger, M., Hoger, H., Elbe-Burger, A., and Wachtler, F. (1999). Recruitment of bone-marrow-derived cells by skeletal and cardiac muscle in adult dystrophic mdx mice. Anat Embryol (Berl) *199*, 391-396.

Black, B.L., Martin, J.F., and Olson, E.N. (1995). The mouse MRF4 promoter is trans-activated directly and indirectly by muscle-specific transcription factors. J Biol Chem *270*, 2889-2892.

Blais, A., Tsikitis, M., Acosta-Alvear, D., Sharan, R., Kluger, Y., and Dynlacht, B.D. (2005). An initial blueprint for myogenic differentiation. Genes Dev *19*, 553-569.

Bober, E., Franz, T., Arnold, H.H., Gruss, P., and Tremblay, P. (1994). Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. Development *120*, 603-612.

Bober, E., Lyons, G.E., Braun, T., Cossu, G., Buckingham, M., and Arnold, H.H. (1991). The muscle regulatory gene, Myf-6, has a biphasic pattern of expression during early mouse development. J Cell Biol *113*, 1255-1265.

Bois, P.R., Izeradjene, K., Houghton, P.J., Cleveland, J.L., Houghton, J.A., and Grosveld, G.C. (2005). FOXO1a acts as a selective tumor suppressor in alveolar rhabdomyosarcoma. J Cell Biol *170*, 903-912.

Borycki, A.G., Brunk, B., Tajbakhsh, S., Buckingham, M., Chiang, C., and Emerson, C.P., Jr. (1999). Sonic hedgehog controls epaxial muscle determination through Myf5 activation. Development *126*, 4053-4063.

Brack, A.S., Conboy, I.M., Conboy, M.J., Shen, J., and Rando, T.A. (2008). A temporal switch from Notch to Wnt signaling in muscle stem cells is necessary for normal adult myogenesis. Cell Stem Cell, 50-59.

Brack, A.S., Conboy, M.J., Roy, S., Lee, M., Kuo, C.J., Keller, C., and Rando, T.A. (2007). Increased Wnt signaling during aging alters muscle stem cell fate and increases fibrosis. Science *317*, 807-810.

Braun, T., Bober, E., Rudnicki, M.A., Jaenisch, R., and Arnold, H.H. (1994). MyoD expression marks the onset of skeletal myogenesis in Myf-5 mutant mice. Development *120*, 3083-3092.

Braun, T., Bober, E., Winter, B., Rosenthal, N., and Arnold, H.H. (1990a). Myf-6, a new member of the human gene family of myogenic determination factors: evidence for a gene cluster on chromosome 12. Embo J *9*, 821-831.

Braun, T., Buschhausen-Denker, G., Bober, E., Tannich, E., and Arnold, H.H. (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., Rudnicki, M.A., Arnold, H.H., and Jaenisch, R. (1992). Targeted inactivation of the muscle regulatory gene Myf-5 results in abnormal rib development and perinatal death. Cell *71*, 369-382.

Braun, T., Winter, B., Bober, E., and Arnold, H.H. (1990b). Transcriptional activation domain of the muscle-specific gene-regulatory protein myf5. Nature *346*, 663-665.

Brohmann, H., Jagla, K., and Birchmeier, C. (2000). The role of Lbx1 in migration of muscle precursor cells. Development *127*, 437-445.

Brunelli, S., Relaix, F., Baesso, S., Buckingham, M., and Cossu, G. (2007). Beta catenin-independent activation of MyoD in presomitic mesoderm requires PKC and depends on Pax3 transcriptional activity. Dev Biol *304*, 604-614.

Buchberger, A., Ragge, K., and Arnold, H.H. (1994). The myogenin gene is activated during myocyte differentiation by pre-existing, not newly synthesized transcription factor MEF-2. J Biol Chem *269*, 17289-17296.

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

Buckingham, M., and Relaix, F. (2007). The role of Pax genes in the development of tissues and organs: Pax3 and Pax7 regulate muscle progenitor cell functions. Annu Rev Cell Dev Biol *23*, 645-673.

Cao, L., Yu, Y., Bilke, S., Walker, R.L., Mayeenuddin, L.H., Azorsa, D.O., Yang, F., Pineda, M., Helman, L.J., and Meltzer, P.S. (2010). Genome-wide identification of PAX3-FKHR binding sites in rhabdomyosarcoma reveals candidate target genes important for development and cancer. Cancer research *70*, 6497-6508.

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

Carvajal, J.J., Cox, D., Summerbell, D., and Rigby, P.W. (2001). A BAC transgenic analysis of the Mrf4/Myf5 locus reveals interdigitated elements that control activation and maintenance of gene expression during muscle development. Development *128*, 1857-1868.

Carvajal, J.J., Keith, A., and Rigby, P.W. (2008). Global transcriptional regulation of the locus encoding the skeletal muscle determination genes Mrf4 and Myf5. Genes Dev *22*, 265-276.

Chalepakis, G., Fritsch, R., Fickenscher, H., Deutsch, U., Goulding, M., and Gruss, P. (1991). The molecular basis of the undulated/Pax-1 mutation. Cell *66*, 873-884.

Chang, H.-Y., Kao, M.-C., Way, T.-D., Ho, C.-T., and Fu, E. (2011). Diosgenin Suppresses Hepatocyte Growth Factor (HGF)-Induced Epithelial–Mesenchymal Transition by Down-regulation of Mdm2 and Vimentin. Journal of Agricultural and Food Chemistry *59*, 5357-5363.

Chang, T.H., Primig, M., Hadchouel, J., Tajbakhsh, S., Rocancourt, D., Fernandez, A., Kappler, R., Scherthan, H., and Buckingham, M. (2004). An enhancer directs differential expression of the linked Mrf4 and Myf5 myogenic regulatory genes in the mouse. Dev Biol *269*, 595-608.

Chang, T.H., Vincent, S.D., Buckingham, M.E., and Zammit, P.S. (2007). The A17 enhancer directs expression of Myf5 to muscle satellite cells but Mrf4 to myonuclei. Dev Dyn *236*, 3419-3426.

Charge, S.B., and Rudnicki, M.A. (2004). Cellular and molecular regulation of muscle regeneration. Physiol Rev *84*, 209-238.

Chen, J.F., Mandel, E.M., Thomson, J.M., Wu, Q., Callis, T.E., Hammond, S.M., Conlon, F.L., and Wang, D.Z. (2006). The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nat Genet *38*, 228-233.

Chen, J.F., Tao, Y., Li, J., Deng, Z., Yan, Z., Xiao, X., and Wang, D.Z. (2010). microRNA-1 and microRNA-206 regulate skeletal muscle satellite cell proliferation and differentiation by repressing Pax7. J Cell Biol *190*, 867-879.

Choi, J., Costa, M.L., Mermelstein, C.S., Chagas, C., Holtzer, S., and Holtzer, H. (1990). MyoD converts primary dermal fibroblasts, chondroblasts, smooth muscle, and retinal pigmented epithelial cells into striated mononucleated myoblasts and multinucleated myotubes. Proc Natl Acad Sci U S A *87*, 7988-7992.

Church, D.M., Goodstadt, L., Hillier, L.W., Zody, M.C., Goldstein, S., She, X., Bult, C.J., Agarwala, R., Cherry, J.L., DiCuccio, M., *et al.* (2009). Lineage-Specific Biology Revealed by a Finished Genome Assembly of the Mouse. PLoS Biol *7*, e1000112.

Collins, C.A., Gnocchi, V.F., White, R.B., Boldrin, L., Perez-Ruiz, A., Relaix, F., Morgan, J.E., and Zammit, P.S. (2009). Integrated functions of Pax3 and Pax7 in the regulation of proliferation, cell size and myogenic differentiation. PloS one *4*, e4475.

Collins, C.A., and Partridge, T.A. (2005). Self-renewal of the adult skeletal muscle satellite cell. Cell cycle (Georgetown, Tex *4*, 1338-1341.

Conboy, I.M., Conboy, M.J., Smythe, G.M., and Rando, T.A. (2003). Notch-mediated restoration of regenerative potential to aged muscle. Science *302*, 1575-1577.

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

Conboy, I.M., and Rando, T.A. (2005). Aging, stem cells and tissue regeneration: lessons from muscle. Cell cycle (Georgetown, Tex *4*, 407-410.

Cornelison, D.D., Filla, M.S., Stanley, H.M., Rapraeger, A.C., and 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.

Cornelison, D.D., Wilcox-Adelman, S.A., Goetinck, P.F., Rauvala, H., Rapraeger, A.C., and Olwin, B.B. (2004). Essential and separable roles for Syndecan-3 and Syndecan-4 in skeletal muscle development and regeneration. Genes Dev *18*, 2231-2236.

Corry, G.N., and Underhill, D.A. (2005). Pax3 target gene recognition occurs through distinct modes that are differentially affected by disease-associated mutations. Pigment cell research / sponsored by the European Society for Pigment Cell Research and the International Pigment Cell Society *18*, 427-438.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res *14*, 1188-1190.

Czerny, T., Halder, G., Kloter, U., Souabni, A., Gehring, W.J., and Busslinger, M. (1999). twin of eyeless, a Second Pax-6 Gene of Drosophila, Acts Upstream of eyeless in the Control of Eye Development. Molecular Cell *3*, 297-307.

Czerny, T., Schaffner, G., and Busslinger, M. (1993). DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site, pp. 2048-2061.

Dahlqvist, C., Blokzijl, A., Chapman, G., Falk, A., Dannaeus, K., Ibanez, C.F., and Lendahl, U. (2003). Functional Notch signaling is required for BMP4-induced inhibition of myogenic differentiation. Development *130*, 6089-6099.

Daston, G., Lamar, E., Olivier, M., and Goulding, M. (1996). Pax-3 is necessary for migration but not differentiation of limb muscle precursors in the mouse. Development *122*, 1017-1027.

Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell *51*, 987-1000.

Day, K., Paterson, B., and Yablonka-Reuveni, Z. (2009). A distinct profile of myogenic regulatory factor detection within Pax7+ cells at S phase supports a unique role of Myf5 during posthatch chicken myogenesis. Dev Dyn *238*, 1001-1009.

De Angelis, L., Berghella, L., Coletta, M., Lattanzi, L., Zanchi, M., Cusella-De Angelis, M.G., Ponzetto, C., and Cossu, G. (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*, 869-878.

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

Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol *4*, P3.

Devlin, R.B., and Emerson, C.P., Jr. (1978). Coordinate regulation of contractile protein synthesis during myoblast differentiation. Cell *13*, 599-611.

Dietrich, S., Abou-Rebyeh, F., Brohmann, H., Bladt, F., Sonnenberg-Riethmacher, E., Yamaai, T., Lumsden, A., Brand-Saberi, B., and Birchmeier, C. (1999). The role of SF/HGF and c-Met in the development of skeletal muscle. Development *126*, 1621-1629.

Dietz, K.N., Miller, P.J., and Hollenbach, A.D. (2009). Phosphorylation of serine 205 by the protein kinase CK2 persists on Pax3-FOXO1, but not Pax3, throughout early myogenic differentiation. Biochemistry *48*, 11786-11795.

Doucet, C., Gutierrez, G.J., Lindon, C., Lorca, T., Lledo, G., Pinset, C., and Coux, O. (2005). Multiple phosphorylation events control mitotic degradation of the muscle transcription factor Myf5. BMC biochemistry *6*, 27.

Edmondson, D.G., and Olson, E.N. (1989). A gene with homology to the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. Genes Dev *3*, 628-640.

Edmondson, D.G., and Olson, E.N. (1990). A gene with homology to the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. Genes Dev *4*, 1450.

Epstein, J.A., Glaser, T., Cai, J., Jepeal, L., Walton, D.S., and Maas, R.L. (1994). Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing, pp. 2022-2034.

Epstein, J.A., Shapiro, D.N., Cheng, J., Lam, P.Y., and Maas, R.L. (1996). Pax3 modulates expression of the c-Met receptor during limb muscle development. Proc Natl Acad Sci U S A *93*, 4213-4218.

Esner, M., Meilhac, S.M., Relaix, F., Nicolas, J.F., Cossu, G., and Buckingham, M.E. (2006). Smooth muscle of the dorsal aorta shares a common clonal origin with skeletal muscle of the myotome. Development *133*, 737-749.

Esposito, C., Miccadei, S., Saiardi, A., and Civitareale, D. (1998). PAX 8 activates the enhancer of the human thyroperoxidase gene. The Biochemical journal *331 (Pt 1)*, 37-40.

Ferrari, G., Cussela-De Angelis, G., Coletta, M., Paolucci, E., Stornaiuolo, A., Cossu, G., and Mavilio, F. (1998). Muscle regeneration by bone marrow-derived myogenic progenitors. Science *279*, 1528-1530.

Filippova, G.N., Thienes, C.P., Penn, B.H., Cho, D.H., Hu, Y.J., Moore, J.M., Klesert, T.R., Lobanenkov, V.V., and Tapscott, S.J. (2001). CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. Nat Genet *28*, 335-343.

Flanagan-Steet, H., Hannon, K., McAvoy, M.J., Hullinger, R., and Olwin, B.B. (2000). Loss of FGF receptor 1 signaling reduces skeletal muscle mass and disrupts myofiber organization in the developing limb. Dev Biol *218*, 21-37.

Franz, T., Kothary, R., Surani, M.A., Halata, Z., and Grim, M. (1993). The Splotch mutation interferes with muscle development in the limbs. Anat Embryol (Berl) *187*, 153-160.

Fukada, S., Uezumi, A., Ikemoto, M., Masuda, S., Segawa, M., Tanimura, N., Yamamoto, H., Miyagoe-Suzuki, Y., and Takeda, S. (2007). Molecular signature of quiescent satellite cells in adult skeletal muscle. Stem Cells *25*, 2448-2459.

Garry, D.J., Meeson, A., Elterman, J., Zhao, Y., Yang, P., Bassel-Duby, R., and Williams, R.S. (2000). Myogenic stem cell function is impaired in mice lacking the forkhead/winged helix protein MNF. Proc Natl Acad Sci U S A *97*, 5416-5421.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J.*, et al.* (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol *5*, R80.

Gibson, M.C., and Schultz, E. (1983). Age-related differences in absolute numbers of skeletal muscle satellite cells. Muscle Nerve *6*, 574-580.

Giordani, J., Bajard, L., Demignon, J., Daubas, P., Buckingham, M., and Maire, P. (2007). Six proteins regulate the activation of Myf5 expression in embryonic mouse limbs. Proc Natl Acad Sci U S A *104*, 11310-11315.

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

Grifone, R., Demignon, J., Giordani, J., Niro, C., Souil, E., Bertin, F., Laclef, C., Xu, P.X., and Maire, P. (2007). Eya1 and Eya2 proteins are required for hypaxial somitic myogenesis in the mouse embryo. Dev Biol *302*, 602-616.

Gross, M.K., Moran-Rivard, L., Velasquez, T., Nakatsu, M.N., Jagla, K., and Goulding, M. (2000). Lbx1 is required for muscle precursor migration along a lateral pathway into the limb. Development *127*, 413-424.

Gussoni, E., Soneoka, Y., Strickland, C.D., Buzney, E.A., Khan, M.K., Flint, A.F., Kunkel, L.M., and Mulligan, R.C. (1999). Dystrophin expression in the mdx mouse restored by stem cell transplantation. Nature *401*, 390-394.

Gustafsson, M.K., Pan, H., Pinney, D.F., Liu, Y., Lewandowski, A., Epstein, D.J., and Emerson, C.P., Jr. (2002). Myf5 is a direct target of long-range Shh signaling and Gli regulation for muscle specification. Genes Dev *16*, 114-126.

Hall, J.K., Banks, G.B., Chamberlain, J.S., and Olwin, B.B. (2010). Prevention of Muscle Aging by Myofiber-Associated Satellite Cell Transplantation. Science Translational Medicine *2*, 57ra83.

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

Hawke, T.J., Jiang, N., and Garry, D.J. (2003). Absence of p21CIP rescues myogenic progenitor cell proliferative and regenerative capacity in Foxk1 null mice. J Biol Chem *278*, 4015-4020.

Heanue, T.A., Reshef, R., Davis, R.J., Mardon, G., Oliver, G., Tomarev, S., Lassar, A.B., and Tabin, C.J. (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.

Hirsinger, E., Duprez, D., Jouve, C., Malapert, P., Cooke, J., and Pourquie, O. (1997). Noggin acts downstream of Wnt and Sonic Hedgehog to antagonize BMP4 in avian somite patterning. Development *124*, 4605-4614.

Holowacz, T., Zeng, L., and Lassar, A.B. (2006). Asymmetric localization of numb in the chick somite and the influence of myogenic signals. Dev Dyn *235*, 633-645.

Holterman, C.E., Le Grand, F., Kuang, S., Seale, P., and Rudnicki, M.A. (2007). Megf10 regulates the progression of the satellite cell myogenic program. J Cell Biol *179*, 911-922.

Horst, D., Ustanina, S., Sergi, C., Mikuz, G., Juergens, H., Braun, T., and Vorobyov, E. (2006). Comparative expression analysis of Pax3 and Pax7 during mouse myogenesis. The International journal of developmental biology *50*, 47-54.

Hu, J.S., Olson, E.N., and Kingston, R.E. (1992). HEB, a helix-loop-helix protein related to E2A and ITF2 that can modulate the DNA-binding ability of myogenic regulatory factors. Mol Cell Biol *12*, 1031-1042.

Huang, H., Kao, M.C., Zhou, X., Liu, J.S., and Wong, W.H. (2004). Determination of local statistical significance of patterns in Markov sequences with application to promoter element identification. J Comput Biol *11*, 1-14.

Hubbard, T.J., Aken, B.L., Ayling, S., Ballester, B., Beal, K., Bragin, E., Brent, S., Chen, Y., Clapham, P., Clarke, L., *et al.* (2009). Ensembl 2009. Nucleic Acids Res *37*, D690-697.

Huh, M.S., Parker, M.H., Scime, A., Parks, R., and Rudnicki, M.A. (2004). Rb is required for progression through myogenic differentiation but not maintenance of terminal differentiation. J Cell Biol *166*, 865-876.

Hutcheson, D.A., Zhao, J., Merrell, A., Haldar, M., and Kardon, G. (2009). Embryonic and fetal limb myogenic cells are derived from developmentally distinct progenitors and have different requirements for beta-catenin. Genes Dev *23*, 997-1013.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics (Oxford, England) *4*, 249-264.

Ishibashi, J. (2006). Transcription Factors Providing Myogenic Identity to Adult Muscle Satellite Cells. In Cellular and Molecular Medicine (Ottawa, University of Ottawa), pp. 233.

Jones, N.C., Tyner, K.J., Nibarger, L., Stanley, H.M., Cornelison, D.D., Fedorov, Y.V., and Olwin, B.B. (2005). The p38alpha/beta MAPK functions as a molecular switch to activate the quiescent satellite cell. J Cell Biol *169*, 105-116.

Jun, S., and Desplan, C. (1996). Cooperative interactions between paired domain and homeodomain. Development *122*, 2639-2650.

Kablar, B., Krastel, K., Ying, C., Asakura, A., Tapscott, S.J., and Rudnicki, M.A. (1997). MyoD and Myf-5 differentially regulate the development of limb versus trunk skeletal muscle. Development *124*, 4729-4738.

Kablar, B., Krastel, K., Ying, C., Tapscott, S.J., Goldhamer, D.J., and Rudnicki, M.A. (1999). Myogenic determination occurs independently in somites and limb buds. Dev Biol *206*, 219-231.

Kablar, B., and Rudnicki, M.A. (1999). Development in the absence of skeletal muscle results in the sequential ablation of motor neurons from the spinal cord to the brain. Dev Biol *208*, 93-109.

Kashyap, V., and Gudas, L.J. (2010). Epigenetic regulatory mechanisms distinguish retinoic acid-mediated transcriptional responses in stem cells and fibroblasts. J Biol Chem *285*, 14534-14548.

Kassar-Duchossoy, L., Giacone, E., Gayraud-Morel, B., Jory, A., Gomes, D., and Tajbakhsh, S. (2005). Pax3/Pax7 mark a novel population of primitive myogenic cells during development. Genes Dev *19*, 1426-1431.

Katagiri, T., Yamaguchi, A., Komaki, M., Abe, E., Takahashi, N., Ikeda, T., Rosen, V., Wozney, J.M., Fujisawa-Sehara, A., and Suda, T. (1994). Bone morphogenetic protein-2 converts the differentiation pathway of C2C12 myoblasts into the osteoblast lineage. J Cell Biol *127*, 1755-1766.

Kaul, A., Koster, M., Neuhaus, H., and Braun, T. (2000). Myf-5 revisited: loss of early myotome formation does not lead to a rib phenotype in homozygous Myf-5 mutant mice. Cell *102*, 17-19.

Keller, C., Hansen, M.S., Coffin, C.M., and Capecchi, M.R. (2004). Pax3:Fkhr interferes with embryonic Pax3 and Pax7 function: implications for alveolar rhabdomyosarcoma cell of origin. Genes Dev *18*, 2608-2613.

Kitamura, T., Kitamura, Y.I., Funahashi, Y., Shawber, C.J., Castrillon, D.H., Kollipara, R., DePinho, R.A., Kitajewski, J., and Accili, D. (2007). A Foxo/Notch pathway controls myogenic differentiation and fiber type specification. J Clin Invest *117*, 2477-2485.

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

Kitzmann, M., and Fernandez, A. (2001). Crosstalk between cell cycle regulators and the myogenic factor MyoD in skeletal myoblasts. Cell Mol Life Sci *58*, 571-579.

Konieczny, S.F., and Emerson, C.P., Jr. (1984). 5-Azacytidine induction of stable mesodermal stem cell lineages from 10T1/2 cells: evidence for regulatory genes controlling determination. Cell *38*, 791-800.

Konigsberg, I.R. (1963). Clonal analysis of myogenesis. Science *140*, 1273-1284.

Kuang, S., Charge, S.B., Seale, P., Huh, M., and Rudnicki, M.A. (2006). Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. J Cell Biol *172*, 103-113.

Kuang, S., Gillespie, M.A., and Rudnicki, M.A. (2008). Niche regulation of muscle satellite cell self-renewal and differentiation. Cell Stem Cell, 22-31.

Kuang, S., Kuroda, K., Le Grand, F., and Rudnicki, M.A. (2007). Asymmetric self-renewal and commitment of satellite stem cells in muscle. Cell *129*, 999-1010.

Laclef, C., Hamard, G., Demignon, J., Souil, E., Houbron, C., and Maire, P. (2003). Altered myogenesis in Six1-deficient mice. Development *130*, 2239-2252.

Lagha, M., Kormish, J.D., Rocancourt, D., Manceau, M., Epstein, J.A., Zaret, K.S., Relaix, F., and Buckingham, M.E. (2008). Pax3 regulation of FGF signaling affects the progression of embryonic progenitor cells into the myogenic program. Genes Dev *22*, 1828-1837.

Langley, B., Thomas, M., Bishop, A., Sharma, M., Gilmour, S., and Kambadur, R. (2002). Myostatin inhibits myoblast differentiation by down-regulating MyoD expression. J Biol Chem *277*, 49831-49840.

Le Grand, F., Auda-Boucher, G., Levitsky, D., Rouaud, T., Fontaine-Perus, J., and Gardahaut, M.F. (2004). Endothelial cells within embryonic skeletal muscles: a potential source of myogenic progenitors. Exp Cell Res *301*, 232-241.

Le Grand, F., Jones, A.E., Seale, V., Scime, A., and Rudnicki, M.A. (2009). Wnt7a activates the planar cell polarity pathway to drive the symmetric expansion of satellite stem cells. Cell Stem Cell *4*, 535-547.

Lepper, C., Conway, S.J., and Fan, C.M. (2009). Adult satellite cells and embryonic muscle progenitors have distinct genetic requirements. Nature *460*, 627-631.

Lepper, C., and Fan, C.M. (2010). Inducible lineage tracing of Pax7-descendant cells reveals embryonic origin of adult satellite cells. Genesis *48*, 424-436.

Leroy, P., and Mostov, K.E. (2007). Slug is required for cell survival during partial epithelial-mesenchymal transition of HGF-induced tubulogenesis. Mol Biol Cell *18*, 1943-1952.

Lin, Y.S., Ha, I., Maldonado, E., Reinberg, D., and Green, M.R. (1991). Binding of general transcription factor TFIIB to an acidic activating region. Nature *353*, 569-571.

Lindon, C., Albagli, O., Domeyne, P., Montarras, D., and Pinset, C. (2000). Constitutive instability of muscle regulatory factor Myf5 is distinct from its mitosis-specific disappearance, which requires a D-box-like motif overlapping the basic domain. Mol Cell Biol *20*, 8923-8932.

Liu, Y., Liu, X.S., Wei, L., Altman, R.B., and Batzoglou, S. (2004). Eukaryotic regulatory element conservation analysis and identification using comparative genomics. Genome Res *14*, 451-458.

Lu, A., Cummins, J.H., Pollett, J.B., Cao, B., Sun, B., Rudnicki, M.A., and Huard, J. (2008). Isolation of myogenic progenitor populations from Pax7-deficient skeletal muscle based on adhesion characteristics. Gene Ther *15*, 1116-1125.

Ma, R., Rundle, D., Jacks, J., Koch, M., Downs, T., and Tsiokas, L. (2003). Inhibitor of Myogenic Family, a Novel Suppressor of Store-operated Currents through an Interaction with TRPC1, pp. 52763-52772.

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

Maroto, M., Reshef, R., Munsterberg, A.E., Koester, S., Goulding, M., and Lassar, A.B. (1997). Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue. Cell *89*, 139-148.

Mauro, A. (1961). Satellite cells of skeletal muscle fibres. J Biophy Biochem Cytol *9*, 493-495.

McCroskery, S., Thomas, M., Maxwell, L., Sharma, M., and Kambadur, R. (2003). Myostatin negatively regulates satellite cell activation and self-renewal. J Cell Biol *162*, 1135-1147.

McDermott, A., Gustafsson, M., Elsam, T., Hui, C.C., Emerson, C.P., Jr., and Borycki, A.G. (2005). Gli2 and Gli3 have redundant and context-dependent function in skeletal muscle formation. Development *132*, 345-357.

McFarlane, C., Hennebry, A., Thomas, M., Plummer, E., Ling, N., Sharma, M., and Kambadur, R. (2008). Myostatin signals through Pax7 to regulate satellite cell self-renewal. Exp Cell Res *314*, 317-329.

McKinnell, I.W., Ishibashi, J., Le Grand, F., Punch, V.G., Addicks, G.C., Greenblatt, J.F., Dilworth, F.J., and Rudnicki, M.A. (2008). Pax7 activates myogenic genes by recruitment of a histone methyltransferase complex. Nature cell biology *10*, 77-84.

McPherron, A.C., Lawler, A.M., and Lee, S.J. (1997). Regulation of skeletal muscle mass in mice by a new TGF-beta superfamily member. Nature *387*, 83-90.

Megeney, L.A., and Rudnicki, M.A. (1995). Determination versus differentiation and the MyoD family of transcription factors. Biochem Cell Biol *73*, 723-732.

Miller, P.J., Dietz, K.N., and Hollenbach, A.D. (2008). Identification of serine 205 as a site of phosphorylation on Pax3 in proliferating but not differentiating primary myoblasts. Protein Sci *17*, 1979-1986.

Minasi, M.G., Riminucci, M., De Angelis, L., Borello, U., Berarducci, B., Innocenzi, A., Caprioli, A., Sirabella, D., Baiocchi, M., De Maria, R., *et al.* (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.

Miner, J.H., and Wold, B. (1990). Herculin, a fourth member of the MyoD family of myogenic regulatory genes. Proc Natl Acad Sci U S A *87*, 1089-1093.

Molkentin, J.D., and Olson, E.N. (1996). Combinatorial control of muscle development by basic helix-loop-helix and MADS-box transcription factors. Proc Natl Acad Sci U S A *93*, 9366-9373.

Montarras, D., Morgan, J., Collins, C., Relaix, F., Zaffran, S., Cumano, A., Partridge, T., and Buckingham, M. (2005). Direct isolation of satellite cells for skeletal muscle regeneration. Science *309*, 2064-2067.

Munsterberg, A.E., Kitajewski, J., Bumcrot, D.A., McMahon, A.P., and Lassar, A.B. (1995). Combinatorial signaling by Sonic hedgehog and Wnt family members induces myogenic bHLH gene expression in the somite. Genes Dev *9*, 2911-2922.

Murre, C., McCaw, P.S., Vaessin, H., Caudy, M., Jan, L.Y., Jan, Y.N., Cabrera, C.V., Buskin, J.N., Hauschka, S.D., Lassar, A.B., *et al.* (1989). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell *58*, 537-544.

Nabeshima, Y., Hanaoka, K., Hayasaka, M., Esumi, E., Li, S., Nonaka, I., and Nabeshima, Y. (1993). Myogenin gene disruption results in perinatal lethality because of severe muscle defect. Nature *364*, 532-535.

Naya, F.J., Wu, C., Richardson, J.A., Overbeek, P., and Olson, E.N. (1999). Transcriptional activity of MEF2 during mouse embryogenesis monitored with a MEF2-dependent transgene. Development *126*, 2045-2052.

Noden, D.M. (1991). Cell movements and control of patterned tissue assembly during craniofacial development. J Craniofac Genet Dev Biol *11*, 192-213.

Olguin, H.C., and Olwin, B.B. (2004). Pax-7 up-regulation inhibits myogenesis and cell cycle progression in satellite cells: a potential mechanism for self-renewal. Dev Biol *275*, 375-388.

Ott, M.O., Bober, E., Lyons, G., Arnold, H., and Buckingham, M. (1991). Early expression of the myogenic regulatory gene, myf-5, in precursor cells of skeletal muscle in the mouse embryo. Development *111*, 1097-1107.

Oustanina, S., Hause, G., and Braun, T. (2004). Pax7 directs postnatal renewal and propagation of myogenic satellite cells but not their specification. Embo J *23*, 3430-3439.

Park, M., and Moon, R.T. (2002). The planar cell-polarity gene stbm regulates cell behaviour and cell fate in vertebrate embryos. Nature cell biology *4*, 20-25.

Parker, M.H., Perry, R.L., Fauteux, M.C., Berkes, C.A., and Rudnicki, M.A. (2006). MyoD synergizes with the E-protein HEB beta to induce myogenic differentiation. Mol Cell Biol *26*, 5771-5783.

Parker, M.H., Seale, P., and Rudnicki, M.A. (2003a). Looking back to the embryo: defining transcriptional networks in adult myogenesis. Nat Rev Genet *4*, 497-507.

Parker, M.H., Seale, P., and Rudnicki, M.A. (2003b). Looking Back to the Embryo: Transcriptional Networks in Adult Myogenesis. Nat Reviews Gen *in press*.

Pena, T.L., Chen, S.H., Konieczny, S.F., and Rane, S.G. (2000). Ras/MEK/ERK Up-regulation of the fibroblast KCa channel FIK is a common mechanism for basic

fibroblast growth factor and transforming growth factor-beta suppression of myogenesis. J Biol Chem *275*, 13677-13682.

Petropoulos, H., and Skerjanc, I.S. (2002). Beta-catenin is essential and sufficient for skeletal myogenesis in P19 cells. J Biol Chem *277*, 15393-15399.

Polesskaya, A., Seale, P., and Rudnicki, M.A. (2003). Wnt signaling induces the myogenic specification of resident CD45+ adult stem cells during muscle regeneration. Cell *113*, 841-852.

Punch, V.G., Jones, A.E., and Rudnicki, M.A. (2009). Transcriptional networks that regulate muscle stem cell function. Wiley Interdisciplinary Reviews: Systems Biology and Medicine *1*, 128-140.

Puri, P.L., Wu, Z., Zhang, P., Wood, L.D., Bhakta, K.S., Han, J., Feramisco, J.R., Karin, M., and Wang, J.Y. (2000). Induction of terminal differentiation by constitutive activation of p38 MAP kinase in human rhabdomyosarcoma cells. Genes Dev *14*, 574-584.

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

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

Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. (2004). Divergent functions of murine Pax3 and Pax7 in limb muscle development. Genes Dev *18*, 1088-1105.

Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. (2005). A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. Nature *435*, 948-953.

Reshef, R., Maroto, M., and Lassar, A.B. (1998). Regulation of dorsal somitic cell fates: BMPs and Noggin control the timing and pattern of myogenic regulator expression. Genes Dev *12*, 290-303.

Rhead, B., Karolchik, D., Kuhn, R.M., Hinrichs, A.S., Zweig, A.S., Fujita, P.A., Diekhans, M., Smith, K.E., Rosenbloom, K.R., Raney, B.J., *et al.* The UCSC Genome Browser database: update 2010. Nucleic Acids Res *38*, D613-619.

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

Ridgeway, A.G., Petropoulos, H., Wilton, S., and Skerjanc, I.S. (2000). Wnt signaling regulates the function of MyoD and myogenin. J Biol Chem *275*, 32398-32405.

Ridgeway, A.G., and Skerjanc, I.S. (2001). Pax3 is essential for skeletal myogenesis and the expression of Six1 and Eya2. J Biol Chem *276*, 19033-19039.

Rudnicki, M.A., Braun, T., Hinuma, S., and 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.

Rudnicki, M.A., and Jaenisch, R. (1995). The MyoD family of transcription factors and skeletal myogenesis. Bioessays *17*, 203-209.

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

Sandmann, T., Jensen, L.J., Jakobsen, J.S., Karzynski, M.M., Eichenlaub, M.P., Bork, P., and Furlong, E.E. (2006). A temporal map of transcription factor activity: mef2

directly regulates target genes at all stages of muscle development. Dev Cell *10*, 797-807.

Sassoon, D., Lyons, G., Wright, W.E., Lin, V., Lassar, A., Weintraub, H., and Buckingham, M. (1989). Expression of two myogenic regulatory factors myogenin and MyoD1 during mouse embryogenesis. Nature *341*, 303-307.

Schafer, K., and Braun, T. (1999). Early specification of limb muscle precursor cells by the homeobox gene Lbx1h. Nat Genet *23*, 213-216.

Schienda, J., Engleka, K.A., Jun, S., Hansen, M.S., Epstein, J.A., Tabin, C.J., Kunkel, L.M., and Kardon, G. (2006). Somitic origin of limb muscle satellite and side population cells. Proc Natl Acad Sci U S A *103*, 945-950.

Schultz, E. (1976). Fine structure of satellite cells in growing skeletal muscle. Am J Anat *147*, 49-70.

Schuster-Gossler, K., Cordes, R., and Gossler, A. (2007). Premature myogenic differentiation and depletion of progenitor cells cause severe muscle hypotrophy in Delta1 mutants. Proc Natl Acad Sci U S A *104*, 537-542.

Schwender, H. (2009). siggenes: Multiple testing using SAM and Efron's empirical Bayes approaches. R package version 1200.

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

Seale, P., Polesskaya, A., and Rudnicki, M.A. (2003). Adult stem cell specification by Wnt signaling in muscle regeneration. Cell cycle (Georgetown, Tex *2*, 418-419.

Seale, P., and 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., and Rudnicki, M.A. (2000). Pax7 is required for the specification of myogenic satellite cells. Cell *102*, 777-786.

Semenov, M.V., Habas, R., Macdonald, B.T., and He, X. (2007). SnapShot: Noncanonical Wnt Signaling Pathways. Cell *131*, 1378.

Shinin, V., Gayraud-Morel, B., Gomes, D., and Tajbakhsh, S. (2006). Asymmetric division and cosegregation of template DNA strands in adult muscle satellite cells. Nature cell biology *8*, 677-687.

Snow, M.H. (1977). The effects of aging on satellite cells in skeletal muscles of mice and rats. Cell Tissue Res *185*, 399-408.

Stratowa, C. (2003). XPS, a Novel Framework for Distributed Storage and Analysis of Microarray Data in the Terabyte Range: An Alternative to BioConductor. Paper presented at: Proceedings of the 3rd International Workshop on Distributed Statistical Computing

Summerbell, D., Ashby, P.R., Coutelle, O., Cox, D., Yee, S., and Rigby, P.W. (2000). The expression of Myf5 in the developing mouse embryo is controlled by discrete and dispersed enhancers specific for particular populations of skeletal muscle precursors. Development *127*, 3745-3757.

Sun, H., Li, L., Vercherat, C., Gulbagci, N.T., Acharjee, S., Li, J., Chung, T.K., Thin, T.H., and Taneja, R. (2007a). Stra13 regulates satellite cell activation by antagonizing Notch signaling. J Cell Biol *177*, 647-657.

Sun, L., Ma, K., Wang, H., Xiao, F., Gao, Y., Zhang, W., Wang, K., Gao, X., Ip, N., and Wu, Z. (2007b). JAK1-STAT1-STAT3, a key pathway promoting proliferation and preventing premature differentiation of myoblasts. J Cell Biol *179*, 129-138.

Tajbakhsh, S., Borello, U., Vivarelli, E., Kelly, R., Papkoff, J., Duprez, D., Buckingham, M., and Cossu, G. (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.

Tajbakhsh, S., and Buckingham, M. (2000). The birth of muscle progenitor cells in the mouse: spatiotemporal considerations. Current topics in developmental biology *48*, 225-268.

Tajbakhsh, S., and Buckingham, M.E. (1994). Mouse limb muscle is determined in the absence of the earliest myogenic factor myf-5. Proc Natl Acad Sci U S A *91*, 747-751.

Tajbakhsh, S., and Cossu, G. (1997). Establishing myogenic identity during somitogenesis. Current opinion in genetics & development *7*, 634-641.

Tajbakhsh, S., Rocancourt, D., and Buckingham, M. (1996). Muscle progenitor cells failing to respond to positional cues adopt non-myogenic fates in myf-5 null mice. Nature *384*, 266-270.

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

Tatsumi, R., Anderson, J.E., Nevoret, C.J., Halevy, O., and Allen, R.E. (1998). HGF/SF is present in normal adult skeletal muscle and is capable of activating satellite cells. Dev Biol *194*, 114-128.

Tatsumi, R., Hattori, A., Ikeuchi, Y., Anderson, J.E., and Allen, R.E. (2002). Release of hepatocyte growth factor from mechanically stretched skeletal muscle satellite cells and role of pH and nitric oxide. Mol Biol Cell *13*, 2909-2918.

Tatsumi, R., Liu, X., Pulido, A., Morales, M., Sakata, T., Dial, S., Hattori, A., Ikeuchi, Y., and Allen, R.E. (2006). Satellite cell activation in stretched skeletal muscle and the role of nitric oxide and hepatocyte growth factor. Am J Physiol Cell Physiol *290*, C1487-1494.

Team, R.D.C. (2010). R: A language and environment for statistical computing (Vienna, Austria, R Foundation for Statistical Computing).

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

Tiffin, N., Williams, R.D., Shipley, J., and Pritchard-Jones, K. (2003). PAX7 expression in embryonal rhabdomyosarcoma suggests an origin in muscle satellite cells. Br J Cancer *89*, 327-332.

Torrente, Y., Belicchi, M., Sampaolesi, M., Pisati, F., Meregalli, M., D'Antona, G., Tonlorenzi, R., Porretti, L., Gavina, M., Mamchaoui, K., *et al.* (2004). Human circulating AC133(+) stem cells restore dystrophin expression and ameliorate function in dystrophic skeletal muscle. J Clin Invest *114*, 182-195.

Trainor, P.A., Tan, S.S., and Tam, P.P. (1994). Cranial paraxial mesoderm: regionalisation of cell fate and impact on craniofacial development in mouse embryos. Development *120*, 2397-2408.

Tsukamoto, K., Nakamura, Y., and Niikawa, N. (1994). Isolation of two isoforms of the PAX3 gene transcripts and their tissue-specific alternative expression in human adult tissues. Hum Genet *93*, 270-274.

Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A *98*, 5116-5121.

Ustanina, S., Carvajal, J., Rigby, P., and Braun, T. (2007). The myogenic factor Myf5 supports efficient skeletal muscle regeneration by enabling transient myoblast amplification. Stem Cells *25*, 2006-2016.

van Rooij, E., Liu, N., and Olson, E.N. (2008). MicroRNAs flex their muscles. Trends Genet *24*, 159-166.

Vasyutina, E., Lenhard, D.C., Wende, H., Erdmann, B., Epstein, J.A., and Birchmeier, C. (2007). RBP-J (Rbpsuh) is essential to maintain muscle progenitor cells and to generate satellite cells. Proc Natl Acad Sci U S A *104*, 4443-4448.

Venters, S.J., and Ordahl, C.P. (2005). Asymmetric cell divisions are concentrated in the dermomyotome dorsomedial lip during epaxial primary myotome morphogenesis. Anat Embryol (Berl) *209*, 449-460.

Vogan, K.J., Underhill, D.A., and Gros, P. (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*, 6677-6686.

Volonte, D., Liu, Y., and Galbiati, F. (2005). The modulation of caveolin-1 expression controls satellite cell activation during muscle repair. Faseb J *19*, 237-239.

Wang, Q., Fang, W.H., Krupinski, J., Kumar, S., Slevin, M., and Kumar, P. (2008). Pax genes in embryogenesis and oncogenesis. J Cell Mol Med *12*, 2281-2294.

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

Wexler, L.H., and Helman, L.J. (1994). Pediatric soft tissue sarcomas. CA: a cancer journal for clinicians *44*, 211-247.

White, R.B., and Ziman, M.R. (2008). Genome-wide discovery of Pax7 target genes during development. Physiol Genomics *33*, 41-49.

Wilson, D., Sheng, G., Lecuit, T., Dostatni, N., and Desplan, C. (1993). Cooperative dimerization of paired class homeo domains on DNA. Genes Dev *7*, 2120-2134.

Wozniak, A.C., and Anderson, J.E. (2007). Nitric oxide-dependence of satellite stem cell activation and quiescence on normal skeletal muscle fibers. Dev Dyn *236*, 240-250.

Wright, W.E., Sassoon, D.A., and Lin, V.K. (1989). Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. Cell *56*, 607-617.

Wu, Z., Woodring, P.J., Bhakta, K.S., Tamura, K., Wen, F., Feramisco, J.R., Karin, M., Wang, J.Y., and Puri, P.L. (2000). p38 and extracellular signal-regulated kinases regulate the myogenic program at multiple steps. Mol Cell Biol *20*, 3951-3964.

Xu, H.E., Rould, M.A., Xu, W., Epstein, J.A., Maas, R.L., and Pabo, C.O. (1999). Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. Genes Dev *13*, 1263-1275.

Yaffe, D. (1968). Retention of differentiation potentialities during prolonged cultivation of myogenic cells. Proc Natl Acad Sci U S A *61*, 477-483.

Yu, M., Smolen, G.A., Zhang, J., Wittner, B., Schott, B.J., Brachtel, E., Ramaswamy, S., Maheswaran, S., and Haber, D.A. (2009). A developmentally regulated inducer of EMT, LBX1, contributes to breast cancer progression. Genes Dev *23*, 1737-1742.

Zammit, P.S., Carvajal, J.J., Golding, J.P., Morgan, J.E., Summerbell, D., Zolnerciks, J., Partridge, T.A., Rigby, P.W., and Beauchamp, J.R. (2004a). Myf5 expression in satellite cells and spindles in adult muscle is controlled by separate genetic elements. Dev Biol *273*, 454-465.

Zammit, P.S., Golding, J.P., Nagata, Y., Hudon, V., Partridge, T.A., and Beauchamp, J.R. (2004b). Muscle satellite cells adopt divergent fates: a mechanism for self-renewal? J Cell Biol *166*, 347-357.

Zammit, P.S., Relaix, F., Nagata, Y., Ruiz, A.P., Collins, C.A., Partridge, T.A., and Beauchamp, J.R. (2006). Pax7 and myogenic progression in skeletal muscle satellite cells. J Cell Sci *119*, 1824-1832.

Zetser, A., Gredinger, E., and Bengal, E. (1999). p38 mitogen-activated protein kinase pathway promotes skeletal muscle differentiation. Participation of the Mef2c transcription factor. J Biol Chem *274*, 5193-5200.

Zhang, W., Behringer, R.R., and Olson, E.N. (1995). Inactivation of the myogenic bHLH gene MRF4 results in up-regulation of myogenin and rib anomalies. Genes Dev *9*, 1388-1399.

Zhang, Y., Liu, T., Meyer, C., Eeckhoute, J., Johnson, D., Bernstein, B., Nussbaum, C., Myers, R., Brown, M., Li, W.*, et al.* (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biology *9*, R137.

Zhao, P., and Hoffman, E.P. (2004). Embryonic myogenesis pathways in muscle regeneration. Dev Dyn *229*, 380-392.

Ziman, M.R., and Kay, P.H. (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

Published in final edited form as:
Nat Cell Biol. 2008 January ; 10(1): 77–84. doi:10.1038/ncb1671.

Pax7 activates myogenic genes by recruitment of a histone methyltransferase complex

Iain W. McKinnell[1,*], **Jeff Ishibashi**[1,*], **Fabien Le Grand**[1], **Vincent J. G. Punch**[1], **Gregory C. Addicks**[1], **Jack F. Greenblatt**[2], **F. Jeffrey Dilworth**[1], and **Michael A. Rudnicki**[1,3]

[1] Sprott Centre for Stem Cell Research, Ottawa Health Research Institute, 501 Smyth Road, Ottawa, Ontario, Canada K1H 8L6

[2] Banting and Best Department of Medical Research, University of Toronto, 112 College Street, Ontario, Canada M5G 1L6

Abstract

Satellite cells purified from adult skeletal muscle can participate extensively in muscle regeneration and can also re-populate the satellite cell pool, suggesting that they have direct therapeutic potential for treating degenerative muscle diseases[1,2]. The paired-box transcription factor Pax7 is required for satellite cells to generate committed myogenic progenitors[3]. In this study we undertook a multi-level approach to define the role of Pax7 in satellite cell function. Using comparative microarray analysis, we identified several novel and strongly regulated targets; in particular, we identified *Myf5* as a gene whose expression was regulated by Pax7. Using siRNA, fluorescence-activated cell sorting (FACS) and chromatin immunoprecipitation (ChIP) studies we confirmed that *Myf5* is directly regulated by Pax7 in myoblasts derived from satellite cells. Tandem affinity purification (TAP) and mass spectrometry were used to purify Pax7 together with its co-factors. This revealed that Pax7 associates with the Wdr5–Ash2L–MLL2 histone methyltransferase (HMT) complex that directs methylation of histone H3 lysine 4 (H3K4, refs 4–10). Binding of the Pax7–HMT complex to Myf5 resulted in H3K4 tri-methylation of surrounding chromatin. Thus, Pax7 induces chromatin modifications that stimulate transcriptional activation of target genes to regulate entry into the myogenic developmental programme.

Satellite cells arise from a population of muscle progenitor cells that originate in the central domain of the dermomyotome. These progenitors express the paired-box transcription factors Pax3 and Pax7 (refs 11,12), and although neither their emergence nor their maintenance requires Pax3 function[13], recent studies have demonstrated that Pax7 is uniquely indispensable for these cells 14. In the absence of Pax7, satellite cells die and thus fail to re-populate their niche[11,14,15]. Pax7, is therefore essential for the formation and maintenance of a population of functional satellite cells. Analysis of the physiological functions of Pax7 has been hindered by relatively weak *trans*-activation properties resulting from *cis*-repression[16]. Consequently, the mechanisms by which Pax7 activates downstream target genes remain unclear. To address this problem, we used a comparative microarray approach to globally identify Pax7 myogenic targets and we propose that the regulation of those genes by Pax7 is intimately related to the protein complexes with which it interacts.

Pools of C2C12 myoblasts were transfected with retrovirus expressing either mouse Pax7–FLAG, or a control virus expressing only the puromycin resistance gene (Puro). As expected, persistent Pax7 expression resulted in the maintenance of a proliferative phenotype and

[3]Correspondence should be addressed to M.A.R. (e-mail: mrudnicki@ohri.ca).
[*]These authors contributed equally to this work.

inhibition of differentiation17 (Supplementary Information, S1). Total RNA was harvested to generate probes for hybridization to Affymetrix GeneChip microarrays; 43 genes were up-regulated by Pax7–FLAG, including several that exhibited striking changes in expression (Table 1). Real-time PCR analysis confirmed that the candidate transcripts were increased in response to Pax7 (Fig. 1a). Pax7 and Pax3 are very closely related proteins; however, Pax7 and Pax3 differed significantly in their abilities to regulate the target genes that were identified. When expressed in C2C12 myoblasts, the levels of each target were substantially increased by Pax7–FLAG whereas Pax3–FLAG had either no effect (*PlagL1; Syne2*) or only modest effects (*Cipar1; Lix1; Mest; Trim54*) (Fig. 1a). To further assess the specificity and responsiveness of these genes, 10T1/2 fibroblast cells were transfected with Pax7–FLAG, Pax3–FLAG or Puro control Despite the non-muscle environment of these cells, the levels of all the candidate genes tested, with the exception of *Mest*, were increased by Pax7 and they were more responsive to Pax7 than to Pax3 (Fig. 1b), confirming the observations made in C2C12 myoblasts. This is consistent with the phenotypic observations that Pax7 and Pax3 have divergent, non-redundant functions during both embryonic14 and post-natal15 myogenesis.

The most significant target identified in our microarray study was the gene for the myogenic regulatory factor Myf5, one of two well-established primary master regulators of skeletal muscle commitment (Myf5) and differentiation (MyoD) in satellite cells3,18,19. This suggests a direct link between Pax7 and myogenic commitment of adult myoblasts. *Myf5* expression was increased by Pax7–FLAG, as shown by microarray analysis (2.2-fold; Table 1) and real-time PCR (~3-fold; data not shown). Importantly, changes in Myf5 protein levels were consistent with the observed increases in mRNA levels. Expression of Pax7–FLAG in C2C12 cells resulted in a dramatic increase in Myf5 protein compared with the Puro control (Fig. 1c). An exaggerated change in Myf5 protein versus mRNA levels has been reported previously in *MyoD*$^{-/-}$ myoblasts20. As shown for previously described targets, Pax3-induced levels of Myf5 protein were substantially lower than those induced by Pax7 in C2C12 myoblasts; in contrast, *MyoD* levels were unaffected by Pax 3 or Pax 7 (Fig. 1c). 10T1/2 fibroblasts are capable of myogenic differentiation in the presence of exogenous Myf5 or MyoD. Remarkably, Pax7 expression in 10T1/2 cells resulted in the induction of Myf5 at both RNA and protein levels (Fig. 1d), whereas Pax3 produced no detectable change. This indicates that Pax7 readily stimulates Myf5 transcription in both muscle and non-muscle cell lines and provides a molecular connection between Pax7 and myogenic commitment.

To test our hypothesis that Pax7 activates target genes by recruiting regulatory complexes to these loci, proteins interacting with Pax7 were purified using a 6x-histidine–TEV (tobacco etch virus)cleavage–3xFLAG tandem affinity purification (TAP) tag fused to the carboxy (C)-terminus of full-length mouse Pax7 (Fig. 2a). This construct (Pax7–CTAP), or the tag alone (HisFLAG–tag) control, was expressed in C2C12 myoblasts. The Pax7–CTAP protein migrated at the predicted size; showed no sign of degradation; was amenable to cleavage by TEV protease and was detected by western-blot analysis of the final eluate (Fig. 2b).

We performed large-scale TAPs to identify interacting proteins via MALDI–TOF mass spectrometry of silver-stained SDS–PAGE bands (Fig. 2c). A complex mixture of largely uncharacterized, interacting proteins was identified, together with the target protein Pax7. Of particular interest, one of the Pax7-associated proteins identified was the WD40-domain-containing protein Wdr5. The observation that Pax7 was co-isolated with Wdr5 is intriguing given that Wdr5, the trithorax-group protein Ash2L, and RbBP5 are the three common components of histone methyltransferase (HMT) complexes that methylate histone H3 at lysine 4 (H3K4, refs4–10,21). Extensive biochemical and structural studies have revealed that Wdr5 is a crucial component of HMT because it associates directly with the amino acid tail of histone H3 (refs 4,22–25), whereas Ash2L, RbBP5 and Wdr5 together create the core structural platform for subsequent protein complex assembly4. This suggests that Pax7 interacts with a

core component of the complex, namely Wdr5, to recruit HMT complexes to target promoters for inscription of epigenetic modifications.

We used nuclear extracts from satellite cell-derived primary myoblasts to validate the Wdr5 interaction in an endogenous setting and to determine whether native Pax7 associates with the other core components of HMT. Using antibodies to the endogenous Pax7, Wdr5 and Ash2L proteins, reciprocal co-immunoprecipitation analysis revealed that antibodies reactive with Pax7 specifically co-precipitated Wdr5 and Ash2L (versus IgG controls). In addition, antibodies reactive with Wdr5 co-precipitated Pax7 and Ash2L, and antibodies reactive with Ash2L co-precipitated Pax7 and Wdr5 (Fig. 2d). These results support the assertion that the three proteins are found in the same complex. We also performed control western blots with α-Erk1/2, a nuclear protein present in abundance in many protein complexes. The failure to detect Erk1/2 in Pax7-immunoprecipitates demonstrated the specificity of the observed interactions. Together, these data confirm that in primary myoblasts Wdr5 and Ash2L interact with Pax7, and support an association of Pax7 with a HMT complex. To identify the methyltransferase responsible for this activity, we performed co-immunoprecipitation with Pax7 and probed for the members of the MLL family of HMTs. The MLL proteins are obvious candidates for this activity because of their known association with the Wdr5–Ash2L–RbBP5 core complex[21]. Indeed, Pax7 interacted with MLL2 (Fig. 2e) but not MLL1 (data not shown), indicating that MLL2 is a specific HMT recruited to this complex.

To further investigate these interactions we produced Pax7–FLAG constructs that contained amino (N)- and C-termini truncations, and specific deletions of the Pax7 paired-domain or homeodomain (Fig. 2f). Deletion of the paired-domain of Pax7, whether following truncation of a significant portion of the N-terminus or a more subtle deletion of only the paired-domain, resulted in near complete abolition of Wdr5 binding to Pax7 (a moderate decrease in binding was observed with Ash2L). In contrast, the effect of deleting the homeodomain appeared to be negligible (Fig. 2g). To ensure that these effects did not result from a failure of the constructs to enter the correct cellular compartment, we performed immunocytochemistry and noted that all of the Pax7–FLAG deletion constructs were localized to the nucleus (Fig. 2h). Given that the paired-domain is involved in the DNA-binding activity of Pax7, this result suggests that DNA binding may be required for the Pax7–HMT complex to form and/or remain bound.

To directly confirm that the Pax7–Wdr5–Ash2L–MLL2 complex possesses methyltransferase activity, we incubated Pax7 immunocomplexes (from primary myoblast nuclear extracts) with core histones in the presence of the tritiated methyl-donor S-adenosyl-methionine (SAM). Histones incubated with the Pax7–immunoprecipitated complex showed an elevated level of tritium incorporation (>1,000 c.p.m.) compared with the IgG control, demonstrating that the Pax7–immunocomplex indeed possesses HMT activity. Histones incubated with immunoprecipitates from primary myotubes (which do not express Pax7) showed no incorporation of SAM above baseline levels of activity (a representative experiment is shown in Fig. 3a).

Modifications of the histone tail significantly influence the nature of gene expression[26]. Methylation of H3K4 marks chromatin in a conformation permissive for transcription, while trimethylated H3K4 is restricted to the 53 promoter and coding regions and is considered to be a definitive marker of active genes[27–30]. To test the prediction that the Pax7-Wdr5–Ash2L–MLL2 complex specifically directs methylation of H3K4, we repeated the previous experiment and performed SDS–PAGE on the samples. Coomassie staining and fluorography indicated that methyltransferase activity is indeed directed against H3 (Fig. 3b). Using H3K4 modification-specific antibodies, we observed that the Pax7 complex mediated trimethylation and (to a markedly lesser extent) dimethylation of H3K4 (Fig. 3c). These modifications are indicative of active or transcriptionally permissive chromatin[27]. In contrast, western blots with

antibodies specific for dimethylated H3K9 (a marker of repressed chromatin31) showed no increases over background levels in the presence of the Pax7 complex (data not shown).

These data demonstrate that Pax7 recruits a protein complex capable of methylating H3K4. As such, we predicted that Pax7 physically associates with sites of open chromatin that are methylated on H3K4. To test this, we performed immunoprecipitations using methylation-specific antibodies and probed western blots for the presence of Pax7. Pax7 was observed to associate strongly with regions of dimethylated H3K4, and, to a lesser extent, with trimethylated H3K4 (Fig. 3d). No association of Pax7 with dimethylated H3K9 was observed (Fig. 3d).

Intriguingly, certain structural data show that Wdr5 binds directly to the H3 tail with an affinity not affected by the state of lysine-4 methylation22,23, although other data indicate a preference for dimethylated H3K424,25. In neither case is the question addressed of how the HMT complex is recruited to specific target genes. This function must be conferred by the unique component(s) of the overall complex, such as a site-specific transcriptional activator, rather than any of the three core elements21,22,25. Our data indicate that Pax7 provides the missing selective function that targets HMT complexes to specific genes while remaining consistent with either model of Wdr5 function.

To directly examine the effect of Pax7 on myogenic commitment, we focused on the Pax7 regulation of *Myf5* in satellite cells. We proposed that *Myf5* would be regulated by Pax7 in satellite cells, would be bound by Pax7 and HMT–complex proteins and would exhibit an altered methylation status. Expression of Pax7–FLAG in satellite cell-derived myoblasts produced a large increase in *Myf5* RNA expression, with an exaggerated change at the protein level (Fig. 4a), consistent with our observations in C2C12 myoblasts and those of other groups20. Accordingly, when we knocked down *Pax7* expression by 50% using siRNA, *Myf5* expression was reduced to approximately 40% of that observed in the scrambled siRNA control cells (Fig. 4b), demonstrating that *Myf5* expression is responsive to Pax7 in satellite cell-derived myoblasts. To test this *in vivo,* we sorted satellite cells from *Pax7*-null mice. These mice possess reduced numbers of satellite cells (3.7% −/− versus 7.6% +/− Fig. 4c), and were previously shown to have a severely impaired ability to regenerate damaged muscle and maintain viable satellite cell populations3,11–13,15. However, we were able to obtain sufficient numbers of satellite cells from four-week old pups to examine the expression level of *Myf5*. Initially we examined the levels of *LacZ*, and two key markers of satellite cells, myocyte nuclear factor and integrin-α7. Expression was equivalent between the heterozygous and homozygous-null populations of cells, indicating that we were indeed sorting satellite cells from the mutant. Strikingly, *Myf5* was absent from Pax7-null satellite cells (Fig. 4d). Immunocytochemistry of the sorted cells revealed that the Pax7$^{-/-}$ cells maintained a rounded, non-activated, non-myogenic phenotype (Fig. 4e). Thus, Pax7 is required for the expression of *Myf5* and the commitment of satellite cells to the myogenic lineage.

A Pax3-paired domain binding site (~57.5 kb upstream of the transcriptional start) that activates the *Myf5* gene in primary myoblasts was recently defined32,33. Chromatin immunoprecipitations (performed from Pax7–FLAG and EGFP control satellite cell-derived myoblasts using α-FLAG) revealed that this site is bound by Pax7 (Fig. 4f). To test our model of Pax7 recruiting a HMT complex, we performed ChIP using anti?]-Ash2L antibodies from three satellite cell-derived myoblast populations: Pax7–FLAG, EGFP control, and siRNA–*Pax7* knockdown cells. This revealed that, in the presence of high Pax7 (Pax7–FLAG), Ash2L levels on the −57.7 Myf5 promoter were increased by more than 125% compared with controls (EGFP); however, reduced Pax7 levels in myoblasts (siRNA-*Pax7*) decreased Ash2L binding to this region to 40% of that seen in the controls (Fig. 4g), demonstrating a Pax7-dependent recruitment of the HMT complex to this region.

In keeping with the hypothesis that the recruitment of this complex results in downstream trimethylation of the coding region and gene activation, we performed ChIP experiments using an anti-H3K4 trimethylation antibody. Accordingly, and despite the fact that Myf5 is already activated in myoblasts, H3K4 trimethylation in the immediate 53 and coding regions of *Myf5* in myoblasts with ectopic Pax7 expression (Pax7–FLAG) was increased by 250–300% above that seen in EGFP cells, consistent with a propagation of the methylation signal by the HMT complex downstream from the Pax7 binding site. Significantly, in myoblasts where Pax7 expression (and Myf5, Fig. 4b) was reduced (siRNA–*Pax7*), the level of H3K4 trimethylation was markedly diminished (Fig. 4h).

Taken together, our data support a model whereby specific binding of Pax7 to dimethylated-H3K4 regulatory elements in target genes leads to the recruitment of HMT core complexes, strong H3K4 trimethylation of their regulatory elements, and ensuing activation of gene expression. This is consistent with structural data showing an interaction between the Wdr5–Ash2L core complex and histone H3 (refs21–25), and answers the question of how this generic complex is recruited to specific genes. It also corroborates other genomic data26,27–29 indicating that recruitment of this complex and subsequent trimethylation of the 5′ promoter and coding regions result in target gene activation. Finally, it defines a specific role and molecular mechanism for Pax7 in myogenic stem cells1–3,11–16 in regulating *Myf5* and other target genes. Together, these experiments indicate that Pax7 enforces satellite cell commitment by recruiting a HMT complex to*Myf5*, resulting in transcriptional activation. Notably, Pax-family genes are essential for the embryonic specification of diverse tissues; thus, Pax recruitment of HMT complexes could be a conserved mechanism for seeding lineage-specific gene-expression programmes during development.

METHODS

Generation of cell lines

C2C12 myoblasts and 10T1/2 fibroblasts were cultured in DMEM with 10% fetal calf serum and 1% penicillin/streptomycin. Primary myoblasts were isolated from adult C57BL/6 mice and cultured using two previously described methods1,20. Pax7– or Pax3–FLAG were cloned onto the pHAN backbone (with puromycin resistance driven from a distinct SV40 promoter). Control virus expressed puromycin-resistance alone. Additionally, full-length mouse Pax7d was fused in-frame to a C-terminal TAP tag and cloned into the pBRIT retroviral plasmid to make pBRIT–Pax7–CTAP (Fig. 2a). A plasmid comprising only the TAP tag was used as a control. Retrovirus was produced by transient co-transfection as described19; or using Lipofectamine 2000 (Invitrogen) into Phoenix-Eco cells (a gift from the Nolan lab, Stanford University Medical Center). Viral supernatant containing 8 μg ml^{-1} polybrene was used to infect C2C12 cells. Stable lines were created by antibiotic selection (1.0–1.5 μg ml^{-1} puromycin; Sigma).

Microarray analysis

Triplicate pools of C2C12 cells stably infected with retro-virus 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. Samples were hybridized to MOE430A GeneChips (Affymetrix,) at the Ottawa Genome Centre (Ontario, Canada) and analysed as previously described19. Criteria used to derive Tables 1 and S1 included: log-fold change of greater than one (that is, 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 the StemBase, Ontario Genomics Innovation Centre (http://www.scgp.ca:8080/StemBase/) as experiment E102 (samples S137, S310) and from the

Gene Expression Omnibus, National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/geo/) under series accession no. GSE3224 (GSM72628, -30, -32, -34…6).

Primers

PCR primers for all RT- and real-time PCR were designed using the online Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). All primer sequence specifications are provided in Supplementary Information, Table S2.

Real-time RT-PCR

Total RNA was isolated as described for microarray samples and used as a template for first-strand reverse transcription (RT) using the RNA PCR Core kit (Perkin Elmer) with random hexamer primers. SYBR Green real-time PCR reactions were performed in triplicate using an MX4000 PCR machine (Stratagene) with fold-change normalized against GAPDH β-Actin, or tubulin. Primer specificity was validated by denaturation curve analysis (55–94 °C) and direct sequencing of the PCR products. Amplification-curve plotting and calculation of C_t values were performed using the MX4000 software (v4.20; Stratagene), with further calculations performed using Microsoft Excel. Each experiment was performed independently on at least three occasions.

Tandem affinity purification

Whole cell extracts were prepared from Pax7–CTAP and HisFLAG–tag C2C12 cells, corresponding to approximately 200 × 100 mm plates. Cells were homogenized in 10 mM Tris-Cl pH 7.9, 0.1 M NaCl, 1.5 mM $MgCl_2$, 0.18% NP-40, on ice. One volume of 50 mM Tris-Cl pH 7.9, 0.6 M NaCl, 1.5 mM $MgCl_2$, 25% glycerol was then added, and the mixture was homogenized again before being treated with benzonase nuclease (Novagen) for 30 min at 4 °C. The lysate was centrifuged, dialysed against 10 mM Tris-Cl pH 7.9, 0.1 M NaCl, 0.1 mM EDTA, 10% glycerol and cleared by centrifugation. Pax7–CTAP or HisFLAG-tag complexes were immunoprecipitated with M2-agarose (Sigma) at 4 °C then eluted with TEV protease (Invitrogen) and FLAG peptide (Sigma) in 20 mM Tris-Cl pH 7.9, 0.1 M NaCl, 0.1 mM EDTA, 0.1% NP40. The eluate was incubated with ProBond resin (Invitrogen) and complexes eluted with 100 mM EDTA, dialysed against 10 mM Tris-Cl pH 7.9, 0.1 M NaCl, 0.1 mM EDTA, 10% glycerol and concentrated on a Centricon column (Amicon).

Mass spectrometry

Samples processed for MALDI-TOF were subjected to SDS–PAGE and silver-stained as described previously34. Bands were excised directly from the gel and analysed at the Ottawa Genome Centre. MALDI-MS/MS spectra were acquired using a QSTAR XL tandem mass spectrometer (ABI/Sciex) with a MALDI-2 source, α-Cyano-4-hydroxycinnamic acid (Agilent, Palo Alto, CA) matrix and Analyst QS version 1.1, build 9865. Spectra were searched against the NCBI database using Mascot daemon version 2.0.5 on an in-house Mascot server version 2.0.04.

Co-immunoprecipitation analysis

Mouse primary myoblasts were used for immunoprecipitation analysis. For differentiation protocols, cells were maintained for 7 days in DMEM + 2% horse serum. Nuclear extracts were produced as follows: cells were re-suspended in hypotonic lysis buffer (10 mM Hepes pH 7.6, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT and protease inhibitors), lysed using a 27-gauge needle and the nuclei collected by centrifugation. Nuclei were re-suspended in 20 mM Hepes pH 7.6, 1.5 mM $MgCl_2$, 420 mM NaCl, 2 mM DTT, 0.2 mM EDTA, 25% glycerol, plus protease inhibitors, and lysed using a 27-gauge needle. The nuclear suspension was agitated

for 60 min at 4 °C, cleared by centrifugation, and adjusted to 20 mM Hepes pH 7.6, 1.5 mM $MgCl_2$, 150 mM NaCl, 2 mM DTT, 0.2 mM EDTA, 20% glycerol, plus protease inhibitors, for subsequent use. Primary antibodies (concentrated α-Pax7 hybridoma supernatant, DSHB and α-Wdr5, kindly donated by Winship Herr; α-Ash2L, kindly donated by Marjorie Brand; control normal IgG, (Upstate) were either incubated at a concentration of 1–2 μg 500 μg^{-1} nuclear extract overnight at 4 °C, before being collected with Protein G/A; or alternatively, were cross-linked to Protein G/A via DMP before incubation with nuclear extract. Immuno-complexes were washed and subjected to SDS–PAGE and western blotting with the antibodies listed above and α-Erk1/2 (Chemicon).

HMT activity assay

Mouse primary myoblast nuclear extracts were immuno-precipitated with α-Pax7 or control IgG as described above and incubated with 2 μg core histones (Upstate) in the presence of ^{3}H-adenosyl-L-methionine (PerkinElmer), for 60 mins at 30 °C, spotted onto P81 phosphocellulose squares (Upstate), and read in a Beckman LS 6500 scintillation counter. Counts per minute for the α-Pax7 immunoprecipitations were normalized against the background counts observed with the IgG immunoprecipitations. Three independent experiments were performed and each sample was read in triplicate. In complementary experiments the HMT assay samples were subjected to SDS–PAGE and treated in one of two ways: (1) Gels were stained in 0.25% Coomassie Brilliant Blue R250, 50% methanol, 10% acetic acid; de-stained in 50% methanol, 10% acetic acid, amplified with Enlightning (NEN Life Sciences), dried and exposed for fluorography; (2) Gels were transferred to nitrocellulose membrane and subjected to western blotting with antibodies raised against dimethylated H3K4, trimethylated H3K4 and dimethylated H3K9 (Upstate).

ChIP analysis

Pax7–FLAG, Ash2L and H3K4 trimethyl protein-DNA complexes were crosslinked with 1% formaldehyde and sonicated. Processing of samples was performed according to the ChIP kit manufacturer's protocols (Upstate). DNA was recovered using PCR purification columns (Invitrogen). The immunoprecipitated DNA was subjected to real-time PCR assays with each primer pair, and normalized against primers for GAPDH as a control locus representing a control fragment that was immunoprecipitated non-specifically to produce robust and reproducible results. This strategy is validated by noting that, for example, Pax7–FLAG samples showed little or no enrichment at flanking loci.

Each DNA sample was analysed by real-time PCR in triplicate, in at least two independent experiments. PCR specificity was validated by denaturation curve analysis and direct sequencing of the products.

Additional methods

Information on western blots, Pax-7 domain deletion construction, histone pull-down assays, FACS of satellite cells, siRNA transfection and immunocytochemistry is available in Supplementary Information S5.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are indebted to Doug Borris and Murray Smith for mass spectrometry analysis; Jeff Baker and Joyce Li for technical assistance; and to Marjorie Brand, Jennifer McCann, Mark Gillespie and Dave Picketts for critical input.

This work was supported by grants to M.A.R. from the National Institutes of Health, the HHMI, the Canadian Institutes of Health Research, the Muscular Dystrophy Association and the CRC Program.

References

1. Kuang S, Kuroda K, Le Grand F, Rudnicki MA. Asymmetric self-renewal and commitment of satellite stem cells in muscle. Cell 2007;129:999–1010. [PubMed: 17540178]
2. Montarras D, et al. Direct isolation of satellite cells for skeletal muscle regeneration. Science 2005;309:2064–2067. [PubMed: 16141372]
3. Seale P, et al. Pax7 is required for the specification of myogenic satellite cells. Cell 2000;102:777–786. [PubMed: 11030621]
4. Steward MM, et al. Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. Nature Struct Mol Biol 2006;13:852–854. [PubMed: 16892064]
5. Wysocka J, Myers MP, Laherty CD, Eisenman RN, Herr W. Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. Genes Dev 2003;17:896–911. [PubMed: 12670868]
6. Roguev A, et al. The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. Embo J 2001;20:7137–7148. [PubMed: 11742990]
7. Yokoyama A, et al. Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. Mol Cell Biol 2004;24:5639–5649. [PubMed: 15199122]
8. Milne TA, et al. MLL associates specifically with a subset of transcriptionally active target genes. Proc Natl Acad Sci USA 2005;102:14765–14770. [PubMed: 16199523]
9. Milne TA, et al. Menin and MLL cooperatively regulate expression of cyclin-dependent kinase inhibitors. Proc Natl Acad Sci USA 2005;102:749–754. [PubMed: 15640349]
10. Hughes CM, et al. Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. Mol Cell 2004;13:587–597. [PubMed: 14992727]
11. Relaix F, Rocancourt D, Mansouri A, Buckingham M. A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. Nature 2005;435:948–953. [PubMed: 15843801]
12. Gros J, Manceau M, Thome V, Marcelle C. A common somitic origin for embryonic muscle progenitors and satellite cells. Nature 2005;435:954–958. [PubMed: 15843802]
13. Kassar-Duchossoy L, et al. Pax3/Pax7 mark a novel population of primitive myogenic cells during development. Genes Dev 2005;19:1426–1431. [PubMed: 15964993]
14. Relaix F, et al. Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. J Cell Biol 2006;172:91–102. [PubMed: 16380438]
15. Kuang S, Charge SB, Seale P, Huh M, Rudnicki MA. Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. J Cell Biol 2006;172:103–113. [PubMed: 16391000]
16. Bennicelli JL, Advani S, Schafer BW, Barr FG. PAX3 and PAX7 exhibit conserved cis-acting transcription repression domains and utilize a common gain of function mechanism in alveolar rhabdomyosarcoma. Oncogene 1999;18:4348–4356. [PubMed: 10439042]
17. Zammit PS, et al. Muscle satellite cells adopt divergent fates: a mechanism for self-renewal? J Cell Biol 2004;166:347–357. [PubMed: 15277541]
18. Berkes CA, Tapscott SJ. MyoD and the transcriptional control of myogenesis. Semin Cell Dev Biol 2005;16:585–595. [PubMed: 16099183]
19. Ishibashi J, Perry RL, Asakura A, Rudnicki MA. MyoD induces myogenic differentiation through cooperation of its NH2- and COOH-terminal regions. J Cell Biol 2005;171:471–482. [PubMed: 16275751]
20. Sabourin LA, Girgis-Gabardo A, Seale P, Asakura A, Rudnicki MA. Reduced differentiation potential of primary MyoD−/− myogenic cells derived from adult skeletal muscle. J Cell Biol 1999;144:631–643. [PubMed: 10037786]
21. Dou Y, et al. Regulation of MLL1 H3K4 methyltransferase activity by its core components. Nature Struct Mol Biol 2006;13:713–719. [PubMed: 16878130]
22. Ruthenburg AJ, et al. Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex. Nature Struct Mol Biol 2006;13:704–712. [PubMed: 16829959]

23. Couture JF, Collazo E, Trievel RC. Molecular recognition of histone H3 by the WD40 protein WDR5. Nature Struct Mol Biol 2006;13:698–703. [PubMed: 16829960]

24. Han Z, et al. Structural basis for the specific recognition of methylated histone H3 lysine 4 by the WD-40 protein WDR5. Mol Cell 2006;22:137–144. [PubMed: 16600877]

25. Wysocka J, et al. WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell 2005;121:859–872. [PubMed: 15960974]

26. Martin C, Zhang Y. The diverse functions of histone lysine methylation. Nature Rev Mol Cell Biol 2005;6:838–849. [PubMed: 16261189]

27. Bernstein BE, et al. Methylation of histone H3 Lys 4 in coding regions of active genes. Proc Natl Acad Sci USA 2002;99:8695–700. [PubMed: 12060701]

28. Liang G, et al. Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. Proc Natl Acad Sci USA 2004;101:7357–7362. [PubMed: 15123803]

29. Santos-Rosa H, et al. Active genes are tri-methylated at K4 of histone H3. Nature 2002;419:407–411. [PubMed: 12353038]

30. Bernstein BE, et al. Genomic maps and comparative analysis of histone modifications in human and mouse. Cell 2005;120:169–181. [PubMed: 15680324]

31. Rea S, et al. Regulation of chromatin structure by site-specific histone H3 methyl-transferases. Nature 2000;406:593–599. [PubMed: 10949293]

32. Bajard L, et al. A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. Genes Dev 2006;20:2450–2464. [PubMed: 16951257]

33. Buchberger A, Freitag D, Arnold HH. A homeo-paired domain-binding motif directs Myf5 expression in progenitor cells of limb muscle. Development 2007;134:1171–1180. [PubMed: 17301086]

34. Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem 1996;68:850–858. [PubMed: 8779443]

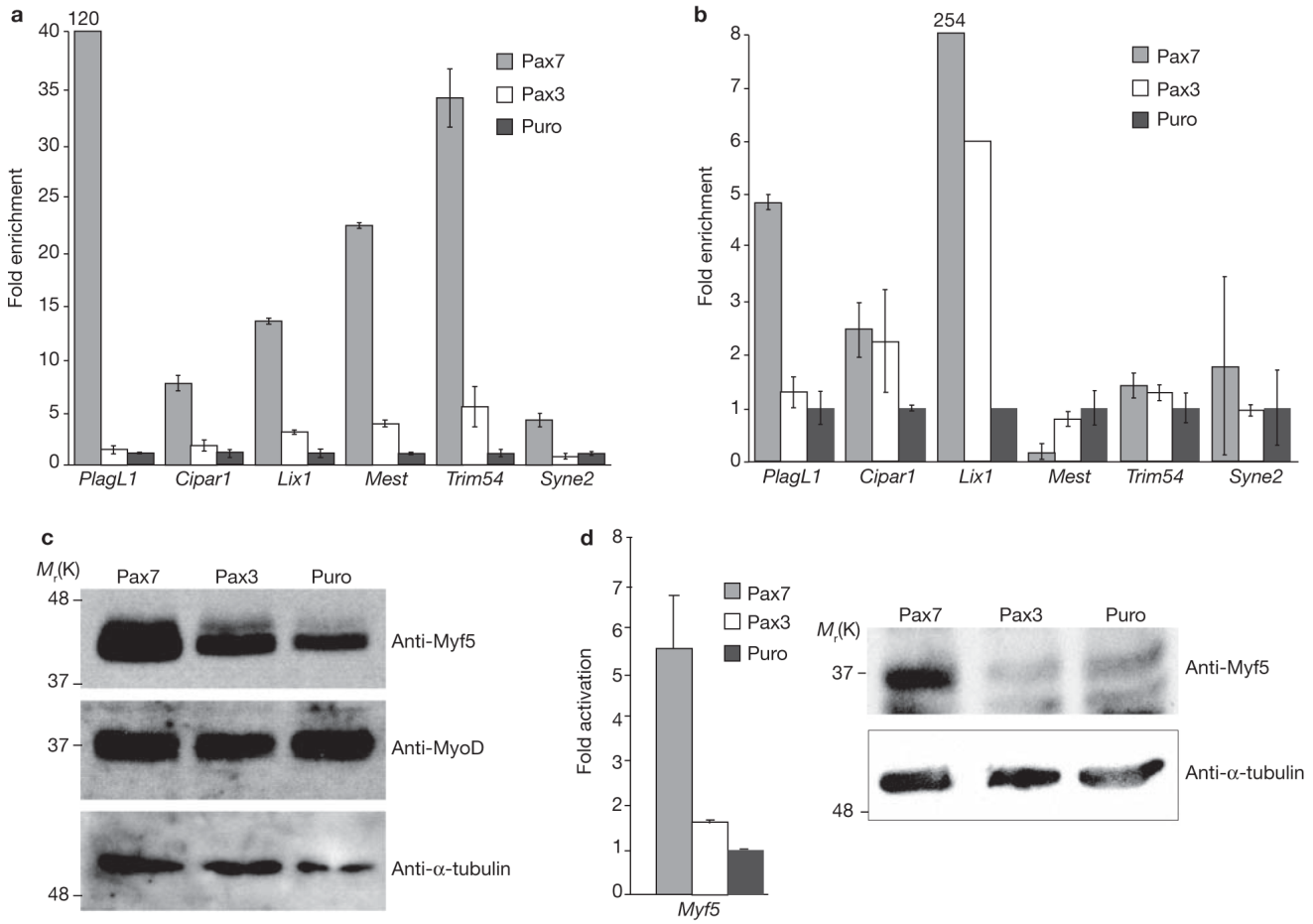


Figure 1.

Candidate target genes are specifically activated by Pax7 in C2C12 myoblasts and are only weakly responsive to Pax3. (**a**) Pax7 strongly activated *PlagL1*, *Cipar1*, *Lix1*, *Mest* and *Trim54* in C2C12 myoblasts whereas Pax3 was far less effective. *Syne2* was only weakly increased by Pax7. Expression was normalized to *GAPDH* transcript levels and is shown relative to the empty-vector (Puro) controls. (**b**) *PlagL1*, and particularly *Lix1*, were both strongly induced by Pax7 in non-muscle 10T1/2 fibroblasts but were weakly induced or unresponsive to Pax3. *Cipar, Trim54* and *Syne2* transcripts are also increased by Pax7d in 10T1/2 cells. (**c**) Myf5 protein levels were also increased by Pax7–FLAG and Pax3–FLAG expression in C2C12 myoblasts whereas MyoD protein levels remained unaffected;α-tubulin was included as a control. (**d**) Real-time PCR and western blot analysis demonstrated that Myf5 RNA and protein, respectively, were increased in 10T1/2 cells expressing Pax7-FLAG (error bars are standard error (s.e.m.)). Full-length scans of western blot data can be found in Supplementary Information S4.

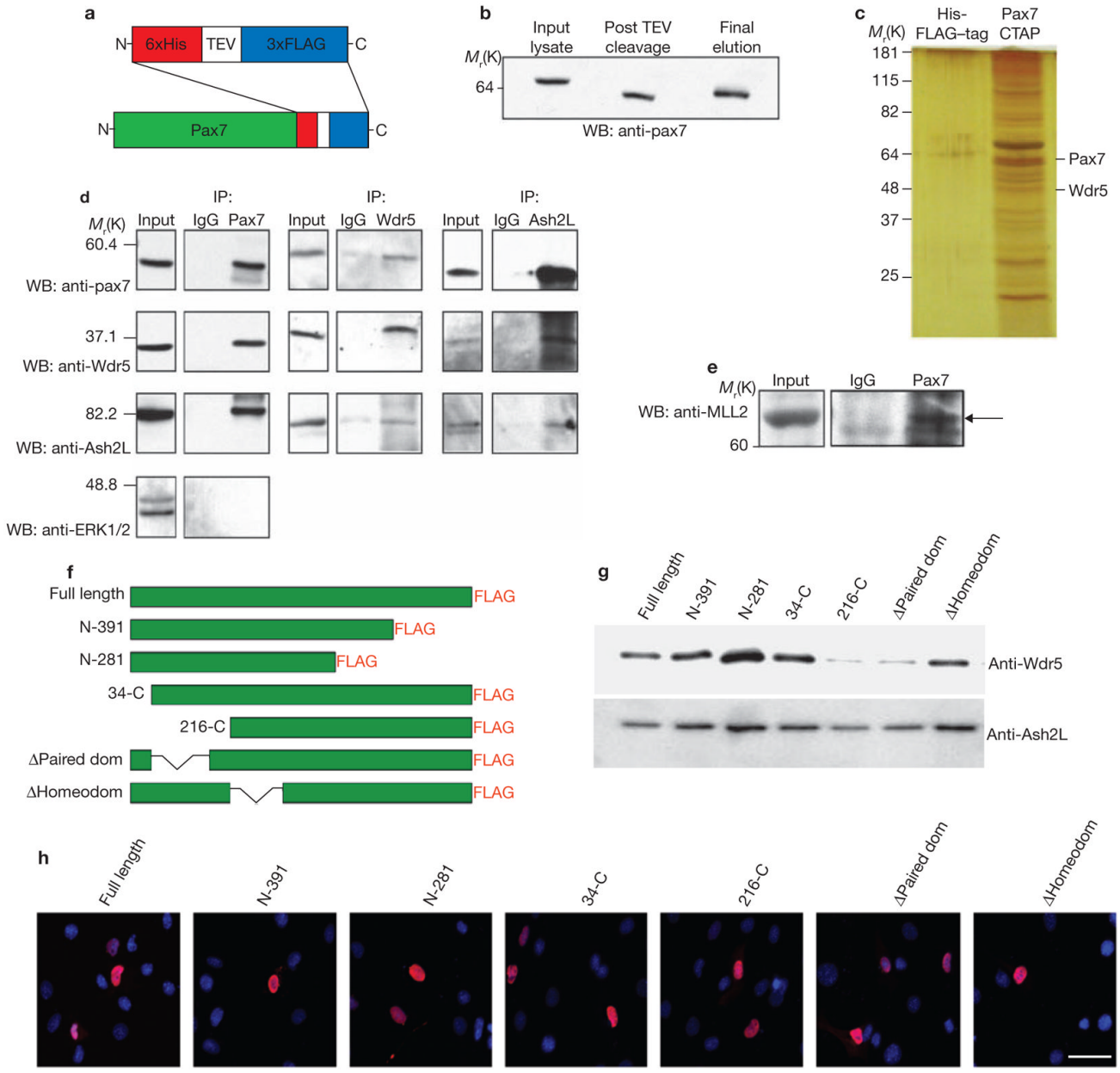


Figure 2.

Identification of Pax7-interacting co-factors. (**a**) A TAP tag, consisting of six histidine and three FLAG epitopes, separated by a TEV cleavage site, was fused to the C-terminus of Pax7 to create a Pax7–CTAP (His–TEV–FLAG) construct. A construct expressing only the tag (referred to as HisFLAG-tag) was used as a negative control. (**b**) High yield purification of the Pax7–CTAP protein as shown by detection of the fusion protein in the initial cell lysate, in the eluate following TEV cleavage, and in the final elution from Ni^{+}-resin. (**c,**) TAP of Pax7–CTAP- compared with HisFLAG-tag-associated proteins from C2C12 cells. Protein matches were generated from multiple data sets and were identified following MALDI–TOF (analysed via Mascot). Pax7–CTAP-interacting proteins were compared with those identified in HisFLAG-tag purifications to unequivocally identify those that were Pax7-specific (versus those that represented contaminants). (**d**) Pax7 co-immunoprecipitated with Wdr5 and Ash2L,

two conserved units of a HMT complex. As a negative control, the Pax7-immunoprecipitate was also probed with an antibody to a non-native-Pax7 interacting nuclear protein, ERK1/2. (**e**) Co-purification of Pax7 and MLL2 immunoprecipitated from primary myoblast nuclear extract. (**f**) Pax7-deletion constructs used to map Pax7 binding to the HMT complex. (**g**) Loss of the paired domain was observed to almost entirely abolish binding between Pax7 and Wdr5. (**h**) Pax7–FLAG truncations where observed to localize to the nucleus (DAPI-blue, Pax7–FLAG deletion constructs-red, scale bar = 20 μm). Full length scans of western blot data can be found in Supplementary Information S4.

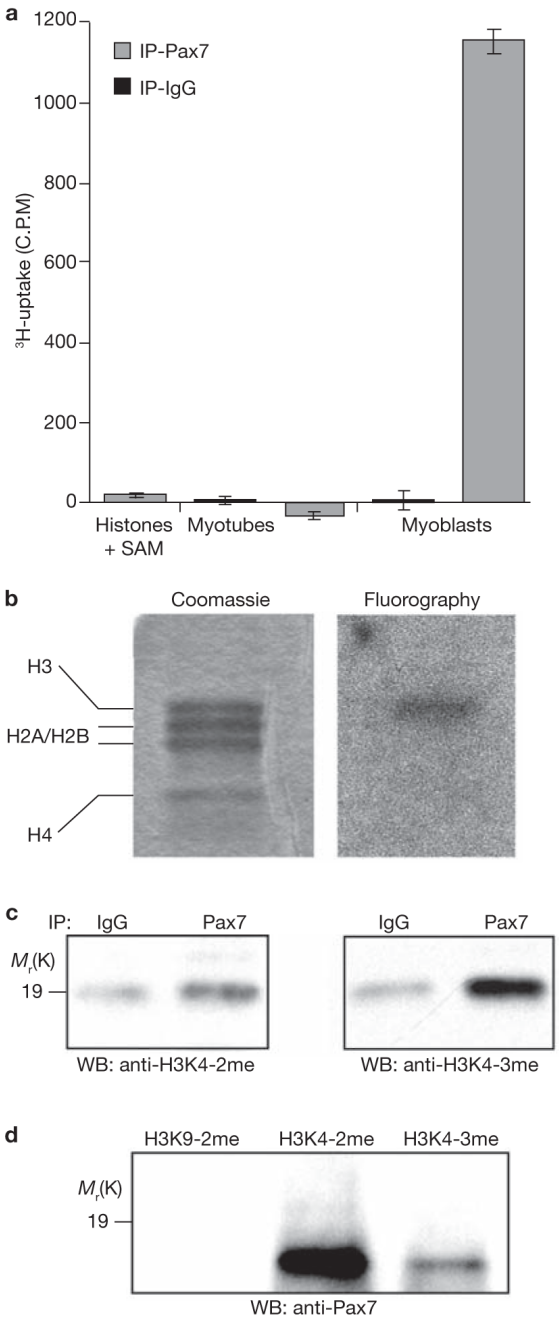


Figure 3.

Pax7 immunocomplex has HMT activity and associates with sites of H3K4 methylation. (**a**) Pax7 or IgG immunoprecipitates from differentiated myotubes (Pax7 negative) or primary myoblasts were incubated with core histones and ^{3}H-SAM. Pax7 immunocomplexes from primary myoblasts were consistently observed to induce a markedly higher level of histone methylation than IgG controls (error bars represent s.e.m.). (**b**) Fluorography of histones incubated with Pax7 immunocomplexes from primary myoblasts indicated that the methyltransferase activity associated with the Pax7 immunocomplex were directed at histone H3. (**c**) Western blot analysis with antisera directed against dimethylated H3K4 (H3K42me) and trimethylated H3K4 (H3K43me) indicate that histones incubated with the Pax7

immunocomplexes show a markedly increased level of H3K43me, with a concurrent but lesser increase in H3K42me. (**d**) IP-western blot analysis revealed that Pax7 was highly associated with chromatin isolated from primary myoblasts using antisera reactive with H3K42me and H3K43me, but not at all with chromatin isolated with antisera reactive with dimethylated H3K9. Full-length scans of western blot data can be found in Supplementary Information S4.

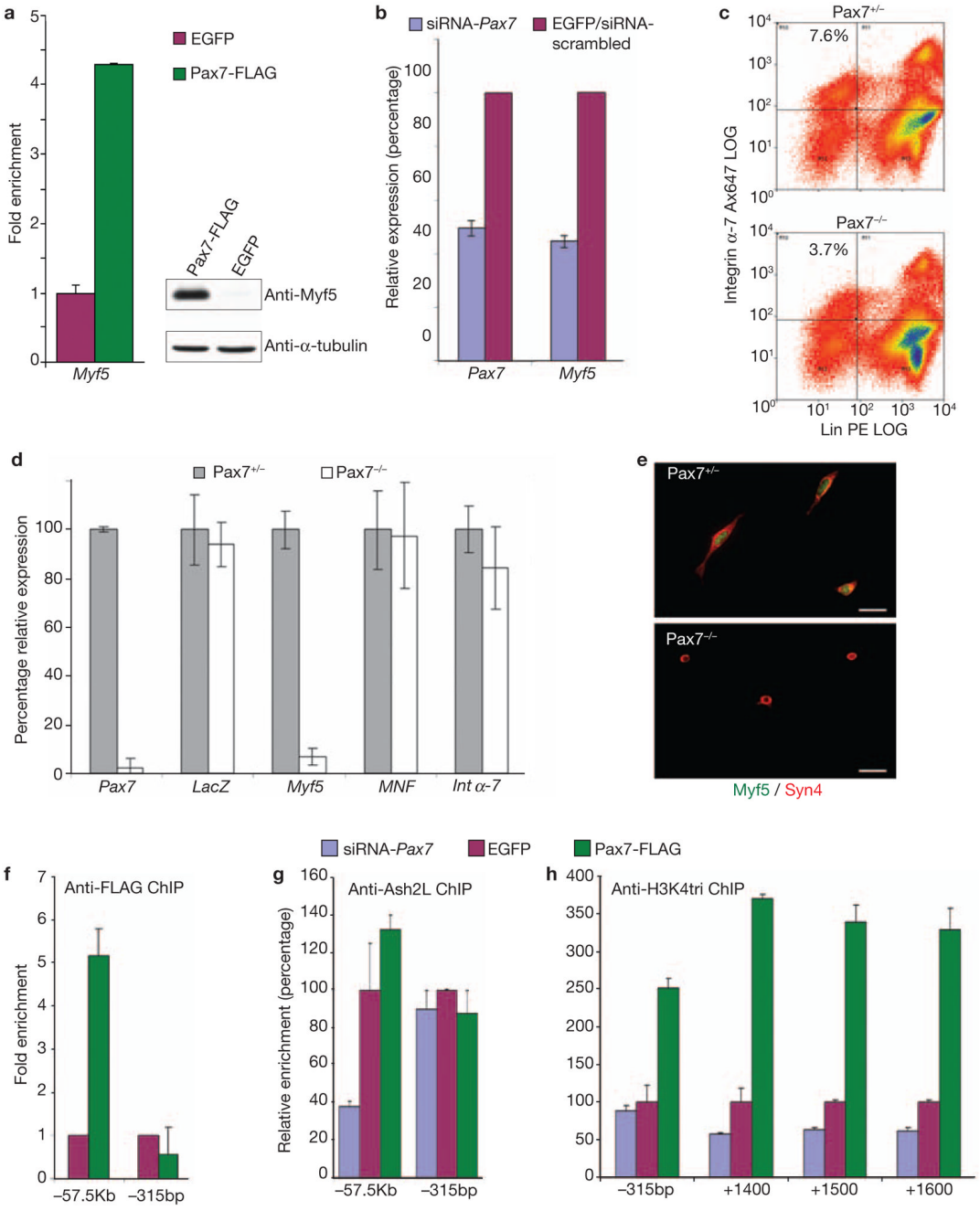


Figure 4.

Pax7 regulates *Myf5* expression directly in satellite cell-derived myoblasts. (**a**) Over-expression of Pax7 resulted in elevated levels of *Myf5* RNA and protein. (**b**) siRNA knockdown of *Pax7* in satellite cell-derived myoblasts results in concurrent knockdown of *Myf5* expression. (**c**) FACS sorted satellite cells (Integrin α-7+; Lin: CD31−, CD45−, CD11−, Sca1−) from $Pax7^{-/-}$ mice show complete absence of Myf5 expression via both (**d**) RNA (n = 3 animals per group, P =: 2.9×10^{-5} for Pax7 and 1.7×10^{-5}, for Myf5, * = $P < 0.001$), and (**e**) immunofluorescence (Syndecan 4-red, (Syn4) Myf5-green (Myf5), scale bar = 20 μm). Expression of *LacZ* was assayed as the *Pax7* knockout was created via the insertion of a *LacZ* sequence into the Pax7 locus and as such all satellite cells should be expressing LacZ in

place of Pax7. (**f**, **g**) ChIP demonstrated that Pax7 was bound directly to the −57.5 kb region of Myf5 in complex with Ash2L (**f**) and in a site-specific manner (**g**). The downstream −315 bp loci, which does not contain a binding site and hence did not show any enrichment, is shown for comparison purposes. (**h**) Antisera directed against H3K4 trimethylation ChIP demonstrated that, accordingly, the coding region of *Myf5* showed a marked increase in activation as measured by trimethylation of H3K4 (error bars = s.e.m.). Full-length scans of western blot data can be found in Supplementary Information S4.

Table 1

Pax7-induced increases in expression in C2C12 myoblasts

Symbol	GID	Fold	Name	RLT-PCR
Plagl1	**NM_009538**	**385.0**	[1]**Pleiomorphic adenoma gene-like 1**	**135.8**
Lix1	**NM_025681**	**12.3**	**Limb expression 1 homolog (chicken)**	**34.8**
Syne2	**BF582734**	**11.3**	**Synaptic nuclear envelope 2**	**2.2**
Cipar1	**AK008716**	**8.1**	**Castration-induced prostatic apoptosis-related 1**	**166.6**
Trim54	**NM_021447**	**7.2**	**Tripartite motif-containing 54**	**4.1**
Il13ra1	NM_133990	5.4	[2] Interleukin 13 receptor, alpha 1	–
Mest	**NM_008590**	**5.0**	**Mesoderm specific transcript**	**17.9**
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 tumour 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**	**3.0**
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	–

Symbol	GID	Fold	Name	RLT-PCR
–	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	–

Pax7–FLAG samples were stringently compared with Puro samples to derive sets of candidate Pax7-regulated genes. *PlagL1* was not detected in Puro samples but was highly expressed in Pax7–FLAG samples. Others that were also considerably increased included *Lix1* (12-fold); *Syne2* (11-fold); *Cipar1* (8-fold); *Trim54* (7-fold); and *Mest* (5-fold).

[1]Signal was very low and called Absent in control.

[2]Mean fold change for two distinct probesets directed at the same transcript.

[3]Mean fold change for 3 distinct probesets directed at the same transcript. Bold lines indicate those used for real-time PCR validation

Appendix B

Table I. Differential Expression of Pax3 and Pax7 target genes

Gene Symbol	Fold Change (Pax3/ctrl)	rawp	Fold Change (Pax7/ctrl)	rawp	Fold Change (Pax7/Pax3)	rawp
Pax7	0.1993	4.52E-04	3.4048	1.44E-03	17.0875	7.80E-06
Bmp4	3.2879	3.19E-03	42.1306	2.42E-04	12.8138	1.95E-05
Slc7a8	1.1017	1.78E-01	8.7044	9.98E-04	7.9006	1.13E-04
Rspo1	1.5148	1.01E-02	11.8921	5.07E-04	7.8507	1.56E-05
Slc2a3	2.6107	1.81E-03	19.9963	5.07E-05	7.6594	2.73E-05
Gpm6a	0.9763	7.11E-01	7.4238	7.25E-04	7.6041	1.17E-05
Gpr39	1.2789	1.61E-02	8.7898	7.41E-05	6.8732	2.34E-05
Klk14	2.983	5.84E-03	19.3109	1.52E-03	6.4736	3.51E-05
Bmp6	2.219	1.15E-03	14.2552	2.69E-04	6.4241	4.29E-05
Entpd3	1.5988	6.04E-03	9.14	6.04E-04	5.7167	5.07E-05
Ecel1	1.7717	4.32E-03	10.013	5.19E-04	5.6516	7.80E-05
Dnase1l3	1.8007	6.71E-03	9.3783	3.74E-04	5.2083	8.97E-05
Wnt7a	NA	NA	4.3893	1.89E-03	4.4173	1.17E-04
Perp	1.5159	1.10E-02	6.5933	9.67E-04	4.3495	5.85E-05
Cpa6	1.7792	1.26E-02	7.7117	2.25E-03	4.3344	2.53E-04
Isg20	0.4901	3.77E-03	2.0936	4.70E-03	4.2714	5.46E-05
Gas7	1.8789	4.93E-03	8.0063	7.64E-04	4.2611	6.63E-05
Mal	1.566	4.36E-03	6.5931	2.38E-04	4.2101	3.12E-05
Pgr15l	1.6482	7.11E-03	6.9299	1.01E-03	4.2045	8.58E-05
Nell2	0.8848	4.12E-02	3.537	8.03E-04	3.9973	8.19E-05
Ramp2	1.9331	5.22E-03	7.1384	8.89E-04	3.6927	6.24E-05
Scd1	2.9505	2.15E-03	10.8798	2.03E-04	3.6874	1.21E-04
Cd38	4.1046	3.00E-03	15.0647	8.19E-04	3.6703	3.00E-04
Kcnk1	1.1941	3.42E-02	4.3766	1.86E-03	3.6651	1.33E-04
Fxyd5	1.2884	2.04E-02	4.6517	2.01E-03	3.6104	1.01E-04
Cyp4a12b	1.7059	1.00E-02	6.0514	1.19E-03	3.5473	1.40E-04
Bace2	1.4947	7.76E-03	5.2469	5.15E-04	3.5104	4.68E-05
Prickle2	2.8532	7.37E-04	9.922	8.97E-05	3.4775	3.90E-05
Krt18	1.576	1.27E-02	5.4346	3.39E-03	3.4483	1.87E-04
Cck	1.1793	3.11E-02	4.0601	5.32E-03	3.4428	4.41E-04
Prokr1	1.7655	3.82E-03	5.9708	2.13E-03	3.3818	3.08E-04
Ly6a	1.7048	4.79E-03	5.7606	1.51E-03	3.379	1.83E-04
Mertk	6.2438	1.32E-03	20.5053	6.79E-04	3.2841	1.25E-04
Crym	1.6335	9.60E-03	5.311	2.07E-03	3.2513	1.09E-04
Clic5	1.1611	4.03E-02	3.7191	2.73E-03	3.2032	1.29E-04
Nox3	1.4224	8.65E-03	4.5164	7.76E-04	3.1752	9.75E-05
Tmprss11f	6.113	6.20E-04	18.7747	4.25E-04	3.0713	3.74E-04
Acan	1.3981	1.29E-02	4.2543	1.94E-03	3.0428	1.72E-04
Dhrs9	4.3785	2.08E-03	13.1444	1.12E-03	3.0021	2.11E-04
Rbm35b	1.8217	7.38E-03	5.3929	2.80E-03	2.9603	2.77E-04
Samd5	1.2725	5.23E-02	3.7315	9.42E-03	2.9324	1.60E-04
Sdpr	0.118	4.83E-04	0.3455	4.02E-03	2.9265	1.79E-04
Atp1b1	0.5847	7.50E-03	1.6982	4.15E-03	2.9044	2.69E-04
C030014K22Rik	2.6759	2.51E-03	7.6878	5.38E-04	2.8729	3.24E-04
Cadps	2.5979	9.65E-03	7.3402	4.53E-03	2.8254	5.93E-04
Slc25a13	0.6737	1.35E-02	1.8889	3.75E-03	2.8037	4.52E-04

Mdga2	1.342	1.07E-02	3.7473	4.32E-03	2.7923	5.81E-04
Adra2a	1.4669	1.09E-02	4.0945	1.70E-03	2.7912	2.73E-04
Cpe	0.9006	1.41E-01	2.5108	1.87E-03	2.7879	4.87E-04
Prkcq	1.1623	5.28E-02	3.2251	5.66E-03	2.7747	6.40E-04
Nppb	1.8258	9.82E-03	5.0642	4.61E-03	2.7737	6.51E-04
Irgm1	1.1933	2.76E-02	3.2881	2.61E-03	2.7554	2.03E-04
Pcdh10	1.9911	5.28E-03	5.4515	1.13E-03	2.7379	4.09E-04
Zfp185	6.8913	9.39E-04	18.6945	1.01E-04	2.7128	6.43E-04
Plekhg1	1.3515	1.04E-02	3.666	2.36E-03	2.7125	2.57E-04
F2rl1	3.2579	2.84E-03	8.7855	1.15E-03	2.6967	3.28E-04
Slc16a12	4.4823	5.07E-04	12.0642	2.30E-04	2.6915	1.75E-04
Elovl7	8.1744	4.68E-04	21.9051	5.46E-05	2.6797	4.37E-04
1810011O10Rik	2.0731	6.20E-03	5.5302	3.31E-04	2.6676	6.32E-04
Ntf3	1.0788	2.88E-01	2.8745	3.33E-03	2.6644	6.47E-04
Lypd6b	0.97	5.77E-01	2.5033	4.03E-03	2.5808	3.70E-04
Klf4	1.2333	1.47E-02	3.1741	2.09E-03	2.5736	2.30E-04
Prkch	0.3861	3.17E-03	1.0316	4.57E-01	2.5596	5.62E-04
Gper	1.3217	1.60E-02	3.3801	2.14E-03	2.5573	2.42E-04
Mmp12	1.4209	1.68E-02	3.6221	4.98E-03	2.5492	5.26E-04
Tbc1d30	1.3621	1.12E-02	3.3953	1.03E-03	2.4927	1.68E-04
Lix1	0.7649	2.09E-02	1.9002	4.86E-03	2.4842	5.03E-04
Jup	2.2506	3.20E-03	5.5538	1.08E-03	2.4677	2.07E-04
Plce1	0.5189	2.62E-03	1.2695	2.20E-02	2.4465	2.14E-04
Cd53	5.677	3.66E-04	13.8852	3.55E-04	2.4459	4.06E-04
Adora1	0.262	1.12E-03	0.6403	1.26E-02	2.4437	2.26E-04
Nfatc2	0.7926	1.92E-02	1.9268	5.48E-03	2.4309	4.68E-04
Pawr	1.3664	9.68E-03	3.3111	2.02E-03	2.4233	3.47E-04
Tspan18	2.6912	2.39E-03	6.4294	8.85E-04	2.3891	1.52E-04
Foxc2	0.4854	6.93E-03	1.1586	2.74E-02	2.3868	9.51E-04
Itgb3	1.2728	1.78E-02	3.0361	3.93E-03	2.3854	6.08E-04
Slc35f3	1.0622	3.73E-01	2.508	5.30E-03	2.3613	7.21E-04
AI593442	NA	NA	2.3141	7.89E-03	2.3254	7.80E-04
Pfkp	1.1564	5.06E-02	2.6759	3.16E-03	2.3141	8.34E-04
Blnk	NA	NA	2.2788	7.32E-03	2.3005	3.31E-04
Hpse	2.8055	4.21E-03	6.4263	2.29E-03	2.2906	5.73E-04
Fam59a	0.4562	5.85E-03	1.0347	4.34E-01	2.268	9.87E-04
Tuft1	3.2447	9.75E-04	7.3546	3.35E-04	2.2666	4.25E-04
Dmrta1	2.723	4.32E-03	6.1702	1.89E-03	2.266	7.06E-04
Il1rl1	1.7851	4.45E-03	3.9915	4.99E-04	2.236	4.13E-04
Gata4	1.6456	8.97E-03	3.6309	3.05E-03	2.2064	2.22E-04
Ptgr1	2.2689	1.29E-03	5.0016	2.50E-04	2.2044	2.34E-04
Xpnpep2	1.134	3.59E-02	2.4731	1.85E-03	2.1809	3.35E-04
Tinagl1	2.0576	2.79E-03	4.4789	4.91E-04	2.1768	2.61E-04
Thbd	2.9492	2.03E-03	6.4105	3.00E-04	2.1736	6.59E-04
Slc7a11	1.364	1.16E-02	2.9244	1.18E-03	2.1439	3.86E-04
Apbb1ip	2.7176	6.47E-03	5.7341	4.54E-03	2.11	7.84E-04
Sobp	1.4592	8.31E-03	3.0733	2.24E-03	2.1061	3.90E-04
Cltb	0.5962	9.21E-03	1.2556	2.70E-02	2.1061	7.68E-04
Inhbb	2.6139	3.66E-03	5.4035	2.30E-03	2.0672	3.04E-04
Lama3	1.5116	4.55E-03	3.1127	2.22E-03	2.0592	5.19E-04
Hs6st3	1.5087	5.98E-03	3.0798	2.47E-03	2.0413	6.12E-04

Tgfbr3	1.4408	7.95E-03	2.9345	1.22E-03	2.0368	1.48E-04
Tshr	1.8262	6.30E-03	3.7137	2.31E-03	2.0336	7.88E-04
Irx3	1.3624	1.57E-02	2.7553	4.49E-03	2.0225	2.85E-04
Fgfbp1	1.5088	1.23E-02	3.0332	4.67E-03	2.0103	4.17E-04
Abhd6	1.1516	2.51E-02	2.3148	2.65E-03	2.01	4.72E-04
S100a7a	3.2869	1.95E-03	6.5872	1.30E-03	2.0041	4.64E-04
Fam70a	1.4886	1.50E-02	2.977	6.43E-03	1.9999	7.37E-04
Rnf43	1.8679	7.44E-03	3.7188	3.61E-03	1.9908	1.95E-04
Nup210	1.587	7.74E-03	3.1565	3.05E-03	1.9889	6.01E-04
Epha4	0.4815	3.93E-03	0.9528	1.82E-01	1.9787	7.99E-04
Ehd3	2.4954	2.89E-03	4.9358	9.83E-04	1.978	6.94E-04
Cyba	1.4865	8.66E-03	2.9322	3.11E-03	1.9725	5.30E-04
Spint1	1.7388	2.59E-03	3.3918	1.46E-03	1.9507	5.23E-04
C1ql1	0.9241	1.24E-01	1.7895	8.28E-03	1.9365	6.63E-04
5730507A09Rik	0.6912	1.11E-02	1.3378	2.30E-02	1.9354	8.15E-04
Car14	1.1389	5.52E-02	2.2032	8.19E-03	1.9344	6.90E-04
Ahrr	1.7261	6.53E-03	3.3172	3.58E-03	1.9218	6.75E-04
Fzd1	0.6756	9.24E-03	1.2969	1.77E-02	1.9196	4.95E-04
Dock5	4.1753	6.67E-04	7.9909	3.90E-04	1.9139	4.29E-04
Maf	1.3453	1.25E-02	2.5489	2.43E-03	1.8947	6.98E-04
6230419C23Rik	1.1627	2.81E-02	2.1988	2.33E-03	1.891	5.65E-04
Osbpl6	2.4926	2.43E-03	4.7029	1.26E-03	1.8868	4.21E-04
Ptprk	1.8812	4.03E-03	3.5134	1.40E-04	1.8676	7.49E-04
Bcam	1.0985	7.37E-02	2.0506	4.79E-03	1.8668	7.17E-04
2200002K05Rik	1.3039	2.93E-02	2.4231	1.08E-02	1.8584	3.55E-04
Pir	NA	NA	1.8329	4.02E-03	1.856	7.56E-04
Adi1	0.3804	2.05E-03	0.706	1.11E-02	1.856	8.11E-04
Arrb1	NA	NA	1.8031	1.01E-02	1.8444	5.15E-04
Mgat5	1.2363	2.40E-02	2.2753	5.01E-03	1.8405	1.00E-03
Scn3b	8.3051	5.26E-04	15.2423	4.64E-04	1.8353	7.64E-04
Pmp22	0.7657	1.29E-02	1.3892	1.08E-02	1.8143	4.99E-04
Myo1d	2.831	8.26E-04	4.9352	4.02E-04	1.8136	3.67E-04
Pcdh9	1.4674	9.64E-03	2.5175	3.26E-03	1.8114	5.89E-04
Ablim1	0.2578	7.68E-04	0.465	4.12E-03	1.8037	8.19E-04
Tspan11	1.2403	1.64E-02	2.2201	5.26E-03	1.7899	9.94E-04
Osbpl10	1.1517	2.75E-02	2.0508	4.75E-03	1.7806	8.73E-04
Nqo1	2.4587	2.56E-03	4.2841	1.84E-03	1.7424	9.63E-04
Gpr125	0.9195	8.18E-02	1.5941	4.79E-03	1.7338	9.44E-04
Ppm1h	1.9632	3.82E-03	3.3658	2.21E-03	1.7145	7.33E-04
Itgb4	3.7291	9.90E-04	6.3639	7.80E-04	1.7066	3.59E-04
Podxl	2.5944	1.58E-03	4.4253	7.45E-04	1.7057	6.71E-04
Rap1gap	1.4227	7.26E-03	2.4172	3.34E-03	1.6989	9.83E-04
Mfge8	1.5858	6.96E-03	2.6584	2.58E-03	1.6764	9.20E-04
BC023744	1.4803	6.77E-03	2.4599	2.32E-03	1.6618	8.54E-04
Fmo1	1.8828	3.11E-03	3.1271	1.54E-03	1.6609	9.32E-04
Pla2g7	3.5535	1.74E-03	5.8613	1.79E-03	1.6495	8.93E-04
Hunk	1.3532	8.78E-03	2.2123	1.61E-03	1.6349	7.14E-04
Lrrc8c	1.4023	1.47E-02	2.2671	6.50E-03	1.6167	9.24E-04
Cd68	2.5615	2.43E-03	4.1117	2.02E-03	1.6052	8.31E-04
Npdc1	1.8243	2.16E-03	2.8768	8.46E-04	1.577	9.36E-04
Ezr	1.3187	1.26E-02	2.0612	3.51E-03	1.563	8.27E-04

Syt7	1.6447	9.95E-03	2.5693	5.96E-03	1.5622	8.89E-04
Atp8a1	9.0049	2.07E-04	13.9667	2.57E-04	1.551	6.36E-04
Prl2c5	80.9581	2.38E-04	125.1743	5.30E-04	1.5462	9.48E-04
Gcat	1.3476	1.35E-02	2.0805	5.44E-03	1.5439	8.46E-04
Prkcb	4.299	1.36E-03	6.5353	1.59E-03	1.5202	8.66E-04
Klf3	1.3502	7.61E-03	2.0106	1.82E-03	1.4891	9.71E-04
Rell1	1.7125	2.74E-03	2.5398	1.22E-03	1.4831	9.09E-04
Met	0.6861	6.21E-03	0.4501	2.01E-03	0.656	5.97E-04
Col3a1	0.8375	2.35E-02	0.5432	6.10E-03	0.6486	9.40E-04
4631416L12Rik	0.2777	7.21E-04	0.1758	5.03E-04	0.6332	7.72E-04
Acta2	0.8398	1.97E-02	0.5312	4.71E-03	0.6325	6.82E-04
St8sia2	0.2038	2.11E-04	0.128	1.95E-05	0.6279	9.98E-04
Cacna1s	0.3023	1.79E-03	0.1897	1.86E-03	0.6276	8.38E-04
Cdh11	6.0506	3.59E-04	3.7897	1.25E-03	0.6263	9.59E-04
Reep1	0.8577	2.91E-02	0.5364	6.78E-03	0.6253	7.95E-04
Srl	0.4195	2.11E-03	0.2603	1.59E-03	0.6205	8.58E-04
Otub2	0.7838	1.30E-02	0.4809	2.85E-03	0.6136	7.45E-04
Prkag3	0.4266	2.21E-03	0.261	9.63E-04	0.6118	8.97E-04
Acsl3	0.4922	2.57E-03	0.3005	1.69E-03	0.6105	8.85E-04
Gpr124	2.4696	7.64E-04	1.4951	7.00E-03	0.6054	4.76E-04
Dtx3l	0.4741	2.79E-03	0.2853	1.60E-03	0.6018	5.07E-04
Mtap4	0.5155	4.56E-03	0.3032	3.01E-03	0.5882	8.62E-04
Dapk2	0.2427	1.60E-04	0.1416	1.05E-04	0.5835	5.54E-04
Slc43a3	1.7969	3.58E-03	1.0308	5.16E-01	0.5736	5.77E-04
Stk17b	0.8723	4.58E-02	0.4995	7.82E-03	0.5726	7.53E-04
Fndc5	0.506	2.86E-03	0.2893	2.17E-03	0.5716	9.05E-04
Des	0.6982	7.17E-03	0.397	1.69E-03	0.5687	4.56E-04
Gal3st3	0.5641	3.12E-03	0.3196	1.49E-03	0.5666	5.58E-04
Alox5	0.6279	4.96E-03	0.3541	2.30E-03	0.5639	9.67E-04
Mypn	0.3041	2.66E-03	0.1708	2.41E-03	0.5615	2.38E-04
Mtap1b	0.8161	1.60E-02	0.4574	2.12E-03	0.5604	5.50E-04
Slc29a1	0.4254	1.12E-03	0.2383	6.98E-04	0.5603	7.76E-04
Il1r1	0.9085	4.77E-02	0.5072	5.56E-03	0.5583	8.07E-04
Aplp1	1.447	5.84E-03	0.7949	1.63E-02	0.5493	3.78E-04
Sema3a	0.7097	1.17E-02	0.3863	3.74E-03	0.5443	8.42E-04
Nfia	0.556	9.52E-03	0.3006	1.63E-03	0.5438	9.28E-04
Trim55	0.5139	3.16E-03	0.2784	1.19E-03	0.5416	7.92E-04
Sorcs2	0.4088	1.89E-03	0.221	1.39E-03	0.5407	7.02E-04
Ifi27l1	0.8426	1.81E-02	0.4534	2.92E-03	0.538	4.45E-04
Smyd1	0.2488	6.16E-04	0.1335	5.93E-04	0.5366	5.46E-04
Dusp27	0.5449	2.44E-03	0.2923	8.93E-04	0.5364	3.94E-04
Lrp4	0.8605	2.34E-02	0.4599	2.47E-03	0.5344	2.96E-04
Mfsd2	0.2569	3.55E-04	0.1371	1.79E-04	0.5337	4.80E-04
Cyp2d22	2.8603	2.19E-03	1.5116	1.23E-02	0.5285	8.23E-04
Adamts14	3.1732	1.24E-03	1.6733	1.05E-02	0.5273	2.18E-04
Cd200	0.1976	2.77E-04	0.1036	2.89E-04	0.5241	4.91E-04
Grb10	10.4235	3.90E-04	5.456	1.12E-03	0.5234	9.75E-04
Cxcl12	0.6858	7.69E-03	0.3561	1.72E-03	0.5192	2.89E-04
Mmp2	0.8915	3.57E-02	0.4621	3.00E-03	0.5183	4.84E-04
Loxl1	0.712	6.10E-03	0.3633	1.29E-03	0.5102	3.51E-04
Fggy	0.6125	8.13E-03	0.3516	1.62E-03	0.5082	3.39E-04

Mylpf	0.2777	9.82E-04	0.1407	1.45E-03	0.5068	9.79E-04
Dpep1	0.9502	5.47E-01	0.4785	9.40E-03	0.5035	8.70E-04
Igfbp7	1.997	2.38E-03	NA	NA	0.5028	6.55E-04
Has2	0.4441	2.17E-03	0.223	1.05E-03	0.502	7.29E-04
Tns1	1.7608	4.11E-03	0.756	2.17E-02	0.5001	3.43E-04
Tnnt3	0.4718	4.09E-03	0.2334	2.55E-03	0.4947	5.69E-04
Zfhx4	0.7521	1.59E-02	0.3715	2.47E-03	0.494	5.85E-04
Tsen15	0.7589	3.00E-02	0.3684	1.05E-02	0.4855	7.25E-04
Sgce	0.8473	2.95E-02	0.4098	4.09E-03	0.4836	3.63E-04
Actn3	0.3753	5.65E-04	0.1808	9.36E-05	0.4819	1.44E-04
Ccdc141	0.5949	6.22E-03	0.2852	3.38E-03	0.4795	6.67E-04
Actn2	0.3995	4.19E-03	0.1898	2.57E-03	0.4749	2.92E-04
Ryr1	0.4322	1.33E-03	0.2017	1.74E-03	0.4666	7.10E-04
Arpp21	0.3385	1.82E-03	0.1577	2.35E-03	0.4659	9.90E-04
Pstpip2	0.4339	2.70E-03	0.2015	1.38E-03	0.4644	6.16E-04
Lsp1	0.4696	7.52E-04	0.2172	3.16E-04	0.4626	1.99E-04
Gnao1	0.7849	1.87E-02	0.3619	1.81E-03	0.4611	6.20E-04
Mybph	0.3708	2.23E-03	0.1695	1.74E-03	0.4573	5.11E-04
Tnnt2	0.7719	2.09E-02	0.3527	4.50E-03	0.4569	8.77E-04
Myl9	0.6299	1.15E-02	0.2878	3.82E-03	0.4568	4.02E-04
Myog	0.4074	1.90E-03	0.1851	1.48E-03	0.4543	6.24E-04
Neb	0.2082	3.78E-04	0.0938	3.39E-04	0.4505	6.28E-04
Adamts6	1.0564	4.38E-01	0.4735	6.06E-03	0.4483	8.50E-04
Pygm	1.0862	1.19E-01	0.4863	3.92E-03	0.4477	5.38E-04
Cd59a	0.9443	3.25E-01	0.4219	2.93E-03	0.4468	3.98E-04
Sntb1	0.6388	8.57E-03	0.2846	2.98E-03	0.4456	4.33E-04
Gng2	0.8319	2.06E-02	0.3686	2.70E-03	0.4431	2.50E-04
Gpr97	4.523	5.42E-04	1.9993	2.40E-03	0.442	3.82E-04
Mef2c	0.4082	4.33E-03	0.1797	2.66E-03	0.4404	6.86E-04
Mustn1	0.596	8.26E-03	0.2603	2.43E-03	0.4367	3.20E-04
Fbln2	5.2477	2.85E-04	2.2839	3.13E-03	0.4352	1.56E-04
Kcnq4	0.8043	1.40E-02	0.3463	2.38E-03	0.4305	2.65E-04
Ttn	0.3922	2.32E-03	0.1687	1.75E-03	0.4301	9.01E-04
C1qtnf3	0.5079	2.38E-03	0.2118	6.86E-04	0.417	2.46E-04
Ankrd2	0.6119	9.98E-03	0.2519	4.31E-03	0.4116	5.34E-04
Cav1	0.9554	2.85E-01	0.3868	5.07E-03	0.4048	4.48E-04
Myl4	NA	NA	0.3942	9.62E-03	0.3902	8.03E-04
Extl1	3.6208	4.60E-04	1.4065	1.14E-02	0.3884	7.41E-05
Casq2	0.3778	3.10E-03	0.142	2.58E-03	0.3758	6.04E-04
Smpx	0.4293	5.78E-03	0.1609	3.60E-03	0.3747	8.81E-04
Srpx2	0.411	1.48E-03	0.1473	8.38E-04	0.3584	2.81E-04
Tagln	0.3291	4.87E-04	0.109	4.29E-04	0.3312	1.64E-04
Iqsec3	0.2936	1.56E-04	0.0942	3.94E-04	0.3209	1.91E-04
Myl1	0.6187	1.01E-02	0.1315	2.37E-03	0.2126	3.12E-04
Dio2	0.4365	1.13E-03	0.0914	4.09E-04	0.2093	9.36E-05
Car3	0.0765	5.07E-05	0.0127	2.07E-04	0.1664	1.36E-04
Pax3	37.2378	2.34E-05	1.9857	4.91E-03	0.0533	3.90E-06

* Entries marked "NA" were omitted by siggenes, and are not significantly differentially expressed

Table II. Differential gene expression between Pax7 siRNA and control myoblasts

Ensembl ID	Gene Symbol	Fold Change	rawp	d.value	stdev
ENSMUSG00000019817	Plagl1	0.362	3.90E-06	-4.77	0.0729
NA	NA	0.362	6.23E-04	-2.29	0.3619
ENSMUSG00000036768	Kif15	0.395	1.48E-04	-3.04	0.2066
ENSMUSG00000067578	Cbln4	0.407	7.79E-05	-3.35	0.1525
NA	Gas5	0.417	7.05E-04	-2.23	0.3316
ENSMUSG00000023919	Cenpq	0.417	7.21E-04	-2.21	0.3357
ENSMUSG00000020330	Hmmr	0.419	1.79E-04	-2.95	0.1901
ENSMUSG00000075316	Scn9a	0.426	6.23E-05	-3.42	0.1259
ENSMUSG00000031673	Cdh11	0.434	3.08E-04	-2.65	0.2187
ENSMUSG00000027379	Bub1	0.444	1.52E-04	-3.03	0.1511
ENSMUSG00000032332	Col12a1	0.445	7.01E-05	-3.39	0.1095
ENSMUSG00000051855	Mest	0.446	3.12E-05	-3.7	0.0795
ENSMUSG00000047534	C79407	0.45	4.64E-04	-2.43	0.2399
ENSMUSG00000021614	Vcan	0.453	1.40E-04	-3.09	0.1352
ENSMUSG00000015880	Ncapg	0.454	1.71E-04	-2.99	0.1463
ENSMUSG00000033031	C330027C09Rik	0.458	2.34E-04	-2.77	0.1725
ENSMUSG00000045328	Cenpe	0.459	2.06E-04	-2.89	0.1535
ENSMUSG00000012443	Kif11	0.459	5.69E-04	-2.34	0.246
ENSMUSG00000039385	Cdh6	0.459	6.12E-04	-2.31	0.2518
ENSMUSG00000027326	Casc5	0.46	1.17E-04	-3.17	0.1191
ENSMUSG00000024795	Kif20b	0.464	1.60E-04	-3.02	0.131
ENSMUSG00000025395	Prim1	0.464	2.84E-04	-2.69	0.1772
NA	Olfr1372-ps1	0.464	3.31E-04	-2.6	0.1915
ENSMUSG00000020914	Top2a	0.465	1.01E-04	-3.24	0.1062
ENSMUSG00000022322	Shcbp1	0.468	1.21E-04	-3.12	0.1165
ENSMUSG00000019773	Fbxo5	0.475	3.86E-04	-2.54	0.1881
ENSMUSG00000040084	Bub1b	0.476	1.09E-04	-3.19	0.1005
NA	Mki67	0.476	1.29E-04	-3.11	0.1103
ENSMUSG00000038379	Ttk	0.477	4.05E-04	-2.49	0.1936
ENSMUSG00000030641	4632434I11Rik	0.48	4.99E-04	-2.39	0.2086
ENSMUSG00000024989	Cep55	0.485	4.79E-04	-2.41	0.199
ENSMUSG00000034311	Kif4	0.489	4.25E-04	-2.47	0.1827
ENSMUSG00000000435	Myf5	0.491	1.87E-04	-2.93	0.1153
ENSMUSG00000075316	Scn9a	0.491	2.03E-04	-2.9	0.1188
ENSMUSG00000026068	Il18rap	0.493	1.32E-04	-3.1	0.0943
ENSMUSG00000023015	Racgap1	0.494	1.36E-04	-3.1	0.0936
ENSMUSG00000025758	Plk4	0.494	1.91E-04	-2.91	0.1145
ENSMUSG00000060832\\ ENSMUSG00000069274	Hist1h4f	0.498	1.15E-03	-2	0.2683
ENSMUSG00000025001	Hells	0.499	3.23E-04	-2.6	0.1502
ENSMUSG00000026683	Nuf2	0.5	6.90E-04	-2.24	0.2113
ENSMUSG00000045273	Cenph	0.5	8.34E-04	-2.15	0.2301
ENSMUSG00000051235	Gen1	0.501	5.34E-04	-2.37	0.186
ENSMUSG00000002297	Dbf4	0.502	7.91E-04	-2.17	0.2245
ENSMUSG00000001228	Uhrf1	0.503	3.70E-04	-2.56	0.1529
ENSMUSG00000034459	Ifit1	0.506	2.49E-04	-2.75	0.1226
NA	NA	0.506	1.48E-03	-1.83	0.302

ENSMUSG00000050711	Scg2	0.507	3.78E-04	-2.54	0.15
ENSMUSG00000032254	Kif23	0.509	1.95E-04	-2.91	0.1003
NA	Snhg1	0.509	2.69E-04	-2.73	0.1228
NA	NA	0.51	1.78E-03	-1.74	0.322
ENSMUSG00000021714	Cenpk	0.512	3.74E-04	-2.55	0.1436
ENSMUSG00000042029	Ncapg2	0.513	4.21E-04	-2.47	0.1549
ENSMUSG00000032400	Zwilch	0.513	6.27E-04	-2.29	0.1858
NA	Snord87	0.515	3.47E-04	-2.58	0.1361
ENSMUSG00000034349	Smc4	0.516	6.78E-04	-2.26	0.1879
ENSMUSG00000028736	Pax7	0.517	1.13E-04	-3.17	0.0652
NA	Gcnt4	0.518	5.06E-05	-3.46	0.0398
ENSMUSG00000032113	Chek1	0.519	5.10E-04	-2.39	0.1622
ENSMUSG00000026779	Mastl	0.521	7.44E-04	-2.2	0.194
NA	NA	0.522	7.56E-04	-2.19	0.1945
ENSMUSG00000028312	Smc2	0.524	7.09E-04	-2.22	0.1853
ENSMUSG00000037725	Ckap2	0.524	7.99E-04	-2.16	0.1963
ENSMUSG00000047757	Fancb	0.526	4.75E-04	-2.41	0.1496
ENSMUSG00000038047	Haus6	0.529	5.38E-04	-2.37	0.1535
NA	D17H6S56E-5	0.53	2.61E-04	-2.73	0.1005
ENSMUSG00000028678	Kif2c	0.531	4.09E-04	-2.49	0.1325
ENSMUSG00000056313	1810011O10Rik	0.531	5.49E-04	-2.36	0.1521
ENSMUSG00000029414	Kntc1	0.531	6.82E-04	-2.26	0.1702
ENSMUSG00000031262	Cenpi	0.532	1.36E-03	-1.87	0.2517
ENSMUSG00000005320	Fgfr4	0.533	1.83E-04	-2.93	0.0742
ENSMUSG00000058773	Hist1h1b	0.533	2.96E-04	-2.67	0.1057
ENSMUSG00000022360	Atad2	0.533	3.90E-04	-2.54	0.1225
ENSMUSG00000025912	Mybl1	0.534	1.21E-03	-1.96	0.2271
ENSMUSG00000039396	Neil3	0.535	7.01E-04	-2.23	0.1706
ENSMUSG00000028480	Glipr2	0.538	9.74E-05	-3.28	0.0379
ENSMUSG00000027074	Slc43a3	0.538	4.83E-04	-2.41	0.1369
ENSMUSG00000036777	Anln	0.538	5.26E-04	-2.37	0.142
ENSMUSG00000028718	Stil	0.538	6.93E-04	-2.24	0.1647
NA	NA	0.538	1.87E-03	-1.72	0.2843
ENSMUSG00000026605	Cenpf	0.539	7.40E-05	-3.37	0.0293
ENSMUSG00000028702	Rad54l	0.541	4.13E-04	-2.49	0.1212
ENSMUSG00000021377	Dek	0.541	4.95E-04	-2.39	0.1353
ENSMUSG00000031245	Nsbp1	0.542	8.53E-04	-2.14	0.1774
ENSMUSG00000024691	Fam111a	0.542	1.11E-03	-2.02	0.2028
ENSMUSG00000022881	Rfc4	0.543	1.08E-03	-2.05	0.1953
ENSMUSG00000040204	2810417H13Rik	0.544	4.48E-04	-2.44	0.1246
ENSMUSG00000027469	Tpx2	0.547	3.51E-04	-2.58	0.1033
ENSMUSG00000027820	Mme	0.547	6.08E-04	-2.32	0.1405
NA	NA	0.547	2.76E-03	-1.54	0.3308
NA	NA	0.549	3.05E-03	-1.49	0.3455
ENSMUSG00000032218	Ccnb2	0.55	7.48E-04	-2.19	0.1581
ENSMUSG00000033952	Aspm	0.55	8.45E-04	-2.14	0.1676
ENSMUSG00000005233	Spc25	0.551	8.49E-04	-2.14	0.1664
NA	NA	0.551	9.39E-04	-2.11	0.1732
ENSMUSG00000002055	Spag5	0.553	1.12E-03	-2.02	0.1892
ENSMUSG00000029816	Gpnmb	0.554	2.18E-04	-2.8	0.0697
ENSMUSG00000063450	Syne2	0.554	2.42E-04	-2.76	0.0737

ENSMUSG00000035683	Melk	0.555	2.73E-04	-2.72	0.0775
ENSMUSG00000063450	Syne2	0.559	3.62E-04	-2.56	0.093
ENSMUSG00000044201	Cdc25c	0.559	7.29E-04	-2.2	0.1466
ENSMUSG00000017146	Brca1	0.56	1.18E-03	-1.97	0.1899
ENSMUSG00000037544	Dlgap5	0.561	6.47E-04	-2.28	0.1306
ENSMUSG00000027699	Ect2	0.561	8.14E-04	-2.16	0.1508
ENSMUSG00000027115	Kif18a	0.561	1.13E-03	-2.01	0.179
ENSMUSG00000037474	Dtl	0.561	1.69E-03	-1.77	0.2371
ENSMUSG00000006678	Pola1	0.563	2.26E-04	-2.79	0.0621
ENSMUSG00000038943	Prc1	0.563	3.27E-04	-2.6	0.084
ENSMUSG00000049281	Scn3b	0.563	4.29E-04	-2.46	0.1011
ENSMUSG00000051378	Kif18b	0.564	2.92E-04	-2.67	0.0745
NA	NA	0.564	1.72E-03	-1.76	0.234
NA	NA	0.566	5.57E-04	-2.36	0.1127
ENSMUSG00000020415	Pttg1	0.567	3.04E-04	-2.66	0.0729
ENSMUSG00000035455	Fignl1	0.568	6.62E-04	-2.27	0.1252
NA	NA	0.568	1.25E-03	-1.95	0.1839
ENSMUSG00000028412	Slc44a1	0.568	1.83E-03	-1.74	0.2349
ENSMUSG00000027496	Aurka	0.57	4.32E-04	-2.46	0.0948
ENSMUSG00000019961	Tmpo	0.57	5.03E-04	-2.39	0.1041
ENSMUSG00000022945	Chaf1b	0.57	7.36E-04	-2.2	0.1345
ENSMUSG00000061615\\ ENSMUSG00000071516	Hist1h2ab	0.57	2.85E-03	-1.52	0.2981
NA	Gm8681	0.571	9.43E-04	-2.11	0.1484
ENSMUSG00000030671	Pde3b	0.572	1.08E-03	-2.05	0.1583
ENSMUSG00000025574	Tk1	0.572	1.09E-03	-2.05	0.1593
ENSMUSG00000027715	Ccna2	0.573	2.88E-04	-2.68	0.0655
ENSMUSG00000024074	Crim1	0.574	2.65E-04	-2.73	0.0585
ENSMUSG00000055612	Cdca7	0.574	3.58E-04	-2.57	0.0759
ENSMUSG00000036533	Cdc42ep3	0.574	6.16E-04	-2.3	0.1124
ENSMUSG00000023940	Sgol1	0.574	1.31E-03	-1.92	0.1819
ENSMUSG00000034023	Fancd2	0.575	7.25E-04	-2.21	0.127
NA	NA	0.575	1.68E-03	-1.78	0.214
ENSMUSG00000020649	Rrm2	0.576	2.45E-04	-2.76	0.0534
ENSMUSG00000041219	Arhgap11a	0.577	3.66E-04	-2.56	0.0744
ENSMUSG00000030677	Kif22	0.577	1.16E-03	-1.99	0.1636
ENSMUSG00000050953	Gja1	0.578	3.19E-04	-2.61	0.0686
ENSMUSG00000069793	Slfn9	0.578	1.57E-03	-1.81	0.2025
ENSMUSG00000021728	Emb	0.579	9.19E-04	-2.12	0.1378
ENSMUSG00000027160	Ccdc34	0.58	2.81E-04	-2.7	0.0567
ENSMUSG00000022070	6720463M24Rik	0.58	1.32E-03	-1.92	0.175
ENSMUSG00000034906	Ncaph	0.581	6.35E-04	-2.29	0.1074
ENSMUSG00000048327	Ckap2l	0.581	6.43E-04	-2.28	0.1085
ENSMUSG00000029910	Mad2l1	0.581	7.67E-04	-2.18	0.1253
ENSMUSG00000038112	AW551984	0.581	7.75E-04	-2.17	0.1264
ENSMUSG00000028364	Tnc	0.584	1.25E-04	-3.11	0.0147
ENSMUSG00000027239	Mdk	0.585	2.53E-04	-2.75	0.0464
ENSMUSG00000051220	Ercc6l	0.586	9.27E-04	-2.11	0.1304
ENSMUSG00000035365	4930547N16Rik	0.586	1.04E-03	-2.07	0.1377
ENSMUSG00000036202	Rif1	0.586	1.22E-03	-1.95	0.1607
ENSMUSG00000068039	Tcp1	0.586	1.76E-03	-1.75	0.2047

ENSMUSG00000069662	Marcks	0.587	1.38E-03	-1.87	0.1769
ENSMUSG00000005470	Asf1b	0.588	3.35E-04	-2.59	0.0608
ENSMUSG00000026039	Sgol2	0.588	9.66E-04	-2.09	0.1309
NA	Snhg1	0.588	1.06E-03	-2.06	0.1367
ENSMUSG00000027331	D2Ertd750e	0.589	1.59E-03	-1.8	0.1894
ENSMUSG00000003779	Kif20a	0.591	5.77E-04	-2.33	0.0908
ENSMUSG00000042489	Clspn	0.592	1.12E-03	-2.02	0.1396
ENSMUSG00000029790	Tsga14	0.593	1.40E-03	-1.86	0.1707
ENSMUSG00000030867	Plk1	0.594	5.42E-04	-2.37	0.0827
ENSMUSG00000021569	Trip13	0.594	5.96E-04	-2.32	0.0891
ENSMUSG00000034317	Trim59	0.594	9.58E-04	-2.1	0.1239
NA	NA	0.595	2.92E-03	-1.51	0.2606
ENSMUSG00000026355	Mcm6	0.596	3.97E-04	-2.51	0.0624
NA	NA	0.596	4.87E-04	-2.4	0.0756
ENSMUSG00000041859	Mcm3	0.596	7.60E-04	-2.18	0.1068
NA	NA	0.596	1.22E-03	-1.96	0.1459
ENSMUSG00000034520	Gjc1	0.597	7.64E-04	-2.18	0.1058
ENSMUSG00000022021	Diap3	0.597	1.04E-03	-2.07	0.1256
ENSMUSG00000053164	Gpr21	0.597	2.00E-03	-1.69	0.2054
NA	NA	0.598	4.60E-04	-2.43	0.0702
ENSMUSG00000045751	F730047E07Rik	0.598	8.18E-04	-2.16	0.1083
ENSMUSG00000059676	Tdpoz1	0.598	1.97E-03	-1.7	0.2025
NA	Eid3	0.599	1.65E-03	-1.78	0.1797
ENSMUSG00000030528	Blm	0.601	1.49E-03	-1.83	0.1656
ENSMUSG00000033904	6330503K22Rik	0.601	1.91E-03	-1.71	0.1947
ENSMUSG00000037313	Tacc3	0.602	7.13E-04	-2.22	0.0955
ENSMUSG00000003038	Gm16494	0.603	1.41E-03	-1.86	0.1582
ENSMUSG00000026196	Bard1	0.604	1.01E-03	-2.08	0.1156
ENSMUSG00000071266	1300003B13Rik	0.605	1.43E-03	-1.85	0.1562
ENSMUSG00000026678	Rgs5	0.606	1.18E-03	-1.98	0.1305
ENSMUSG00000027306	Nusap1	0.607	1.14E-03	-2	0.1253
ENSMUSG00000026234	Ncl	0.608	2.40E-03	-1.6	0.2143
ENSMUSG00000072082	Ccnf	0.61	6.31E-04	-2.29	0.0767
ENSMUSG00000023104	Rfc2	0.61	1.95E-03	-1.7	0.1841
ENSMUSG00000044906	4930503L19Rik	0.611	1.71E-03	-1.77	0.1673
ENSMUSG00000031939	Taf1d	0.611	2.90E-03	-1.51	0.2336
ENSMUSG00000029177	Cenpa	0.612	1.00E-03	-2.08	0.1058
ENSMUSG00000030924	2610020H08Rik	0.614	1.62E-03	-1.79	0.1578
ENSMUSG00000049539	Hist1h1a	0.615	2.38E-04	-2.76	0.0189
NA	NA	0.616	6.58E-04	-2.27	0.0728
ENSMUSG00000046186	Cd109	0.616	1.87E-03	-1.73	0.1708
NA	Rnu73b	0.616	2.45E-03	-1.59	0.2041
ENSMUSG00000032057	4833427G06Rik	0.616	3.02E-03	-1.49	0.2331
ENSMUSG00000020122	Egfr	0.617	4.36E-04	-2.46	0.0485
ENSMUSG00000022385	Gtse1	0.617	6.00E-04	-2.32	0.0662
ENSMUSG00000032477	Cdc25a	0.617	6.04E-04	-2.32	0.0661
ENSMUSG00000025925	Terf1	0.617	1.19E-03	-1.97	0.1189
ENSMUSG00000051278	4930422G04Rik	0.617	1.58E-03	-1.8	0.1525
ENSMUSG00000030607	Acan	0.617	2.93E-03	-1.51	0.2266
ENSMUSG00000025289	Prdx4	0.619	3.16E-04	-2.62	0.0294
ENSMUSG00000024660	Incenp	0.619	6.39E-04	-2.29	0.0683

ENSMUSG00000020427	Igfbp3	0.62	6.70E-04	-2.26	0.0698
ENSMUSG00000017716	Birc5	0.62	7.52E-04	-2.19	0.0808
NA	Gm13487	0.62	1.03E-03	-2.07	0.0987
ENSMUSG00000027641	Rbl1	0.621	1.27E-03	-1.94	0.12
ENSMUSG00000058799	Nap1l1	0.621	2.32E-03	-1.62	0.1909
ENSMUSG00000020493	Prr11	0.622	2.72E-03	-1.54	0.21
ENSMUSG00000034610	Zcchc11	0.623	6.86E-04	-2.25	0.0685
ENSMUSG00000007682	Dio2	0.623	1.38E-03	-1.87	0.1304
ENSMUSG00000050410	Tcf19	0.624	1.31E-03	-1.92	0.1196
ENSMUSG00000024085	Man2a1	0.625	1.47E-03	-1.83	0.1349
ENSMUSG00000021965	F630043A04Rik	0.625	1.60E-03	-1.8	0.1428
ENSMUSG00000063011	Msln	0.626	9.82E-04	-2.09	0.0885
ENSMUSG00000054889	Dsp	0.628	4.71E-04	-2.42	0.0427
ENSMUSG00000027323	Rad51	0.628	9.54E-04	-2.1	0.085
ENSMUSG00000027200	Sema6d	0.628	1.33E-03	-1.9	0.1186
ENSMUSG00000055240	Zfp101	0.628	2.36E-03	-1.61	0.1818
ENSMUSG00000007080	Pole	0.629	1.53E-03	-1.82	0.1318
ENSMUSG00000061533	4930534B04Rik	0.629	2.12E-03	-1.67	0.1669
ENSMUSG00000016984	Etaa1	0.629	3.02E-03	-1.49	0.2125
ENSMUSG00000058189	Hist1h2bm	0.63	1.29E-03	-1.93	0.1114
ENSMUSG00000034206	Polq	0.631	1.10E-03	-2.04	0.091
ENSMUSG00000033149	Phldb2	0.631	1.22E-03	-1.95	0.1057
NA	NA	0.632	1.27E-03	-1.93	0.108
ENSMUSG00000062949	Atp11c	0.632	2.64E-03	-1.56	0.1909
ENSMUSG00000026669	Mcm10	0.633	7.79E-04	-2.17	0.0687
ENSMUSG00000019942	Cdc2a	0.633	1.01E-03	-2.08	0.0823
ENSMUSG00000036875	Dna2	0.633	1.50E-03	-1.83	0.1258
ENSMUSG00000022978	2610039C10Rik	0.634	1.61E-03	-1.79	0.1316
ENSMUSG00000028693	Nasp	0.634	2.14E-03	-1.66	0.161
ENSMUSG00000020224	Llph	0.635	2.71E-03	-1.54	0.1903
ENSMUSG00000024501	Dpysl3	0.636	8.84E-04	-2.13	0.0718
ENSMUSG00000072980	Oip5	0.636	1.13E-03	-2.01	0.0896
ENSMUSG00000038252	Ncapd2	0.637	9.23E-04	-2.11	0.0735
ENSMUSG00000035293	G2e3	0.637	1.61E-03	-1.8	0.1276
ENSMUSG00000018217	Pmp22	0.637	1.75E-03	-1.75	0.1364
ENSMUSG00000024590	Lmnb1	0.639	9.90E-04	-2.09	0.0749
ENSMUSG00000028873	Cdca8	0.639	1.09E-03	-2.05	0.0811
ENSMUSG00000031756	Cenpn	0.639	1.78E-03	-1.74	0.136
ENSMUSG00000029730	Mcm7	0.64	1.11E-03	-2.04	0.0816
ENSMUSG00000015568	Lpl	0.64	1.68E-03	-1.77	0.1284
ENSMUSG00000056531	Ccdc18	0.641	2.77E-03	-1.54	0.1824
ENSMUSG00000027133	Nop10	0.642	9.08E-04	-2.12	0.0662
ENSMUSG00000001918	Slc1a5	0.642	1.15E-03	-1.99	0.0867
ENSMUSG00000029516	Cit	0.643	1.23E-03	-1.95	0.0916
ENSMUSG00000078453	3110003A17Rik	0.643	1.24E-03	-1.95	0.0917
ENSMUSG00000022676	Snai2	0.643	1.88E-03	-1.72	0.1354
ENSMUSG00000031939	Taf1d	0.645	1.14E-03	-2	0.0813
ENSMUSG00000038732	Mboat1	0.645	2.69E-03	-1.55	0.1736
NA	B430203M17Rik	0.645	2.87E-03	-1.52	0.1818
ENSMUSG00000029672	Fam3c	0.646	1.16E-03	-1.99	0.0819
ENSMUSG00000027454	Gins1	0.649	9.00E-04	-2.12	0.0593

ENSMUSG00000015354	Pcolce2	0.65	9.51E-04	-2.1	0.0604
ENSMUSG00000020900	Myh10	0.651	1.25E-03	-1.94	0.0846
ENSMUSG00000037572	Wdhd1	0.653	7.17E-04	-2.21	0.0425
ENSMUSG00000020262	Adarb1	0.654	5.06E-04	-2.39	0.0215
ENSMUSG00000063632	Sox11	0.654	6.51E-04	-2.28	0.0346
ENSMUSG00000058385	Hist2h2bb	0.655	1.84E-03	-1.73	0.118
NA	NA	0.655	3.05E-03	-1.49	0.1745
ENSMUSG00000006398	Cdc20	0.656	1.02E-03	-2.08	0.0581
ENSMUSG00000031877	2210023G05Rik	0.656	1.28E-03	-1.93	0.0802
ENSMUSG00000048922	Cdca2	0.656	1.37E-03	-1.87	0.091
ENSMUSG00000053469	Tg	0.656	1.88E-03	-1.72	0.1177
ENSMUSG00000004099	Dnmt1	0.657	7.87E-04	-2.17	0.0442
ENSMUSG00000021994	Wnt5a	0.657	1.54E-03	-1.82	0.0986
NA	NA	0.657	2.02E-03	-1.68	0.1247
ENSMUSG00000019979	Apaf1	0.659	8.77E-04	-2.14	0.047
ENSMUSG00000020897	Aurkb	0.659	1.01E-03	-2.08	0.054
ENSMUSG00000043850	Clrn1	0.659	1.37E-03	-1.87	0.0871
ENSMUSG00000024165	Hn1l	0.659	1.40E-03	-1.86	0.0876
ENSMUSG00000074994	Qser1	0.659	2.49E-03	-1.58	0.1452
ENSMUSG00000005410	Mcm5	0.66	5.84E-04	-2.33	0.0232
ENSMUSG00000034329	Brip1	0.66	2.61E-03	-1.56	0.1486
NA	NA	0.662	1.58E-03	-1.8	0.0957
ENSMUSG00000039748	Exo1	0.662	1.91E-03	-1.71	0.113
ENSMUSG00000022673	Mcm4	0.662	2.48E-03	-1.58	0.1403
ENSMUSG00000078773	Rad54b	0.664	1.62E-03	-1.79	0.0959
ENSMUSG00000046591	5730590G19Rik	0.665	1.55E-03	-1.81	0.0898
ENSMUSG00000074994	Qser1	0.665	2.25E-03	-1.63	0.1251
ENSMUSG00000020086	H2afy2	0.665	2.84E-03	-1.53	0.1508
ENSMUSG00000055044	Pdlim1	0.665	2.94E-03	-1.51	0.1566
ENSMUSG00000020808	6720460F02Rik	0.666	1.34E-03	-1.89	0.0751
ENSMUSG00000026134	Prim2	0.666	1.97E-03	-1.7	0.1108
ENSMUSG00000026547	Tagln2	0.667	2.08E-03	-1.67	0.1137
ENSMUSG00000022422	Dscc1	0.667	3.10E-03	-1.49	0.1584
ENSMUSG00000063021	Hist1h2ak	0.668	1.34E-03	-1.89	0.073
ENSMUSG00000021811	Dnajc9	0.668	1.89E-03	-1.72	0.1031
ENSMUSG00000031684	Slc10a7	0.668	1.94E-03	-1.7	0.1069
NA	NA	0.669	1.11E-03	-2.03	0.0512
ENSMUSG00000024070	Prkd3	0.669	2.19E-03	-1.65	0.1173
ENSMUSG00000008658	A2bp1	0.67	1.20E-03	-1.97	0.0586
ENSMUSG00000047259	Mc4r	0.671	1.75E-03	-1.76	0.0925
ENSMUSG00000033906	Zdhhc15	0.671	1.94E-03	-1.7	0.1025
ENSMUSG00000010830	Kdelr3	0.671	2.22E-03	-1.64	0.1152
ENSMUSG00000033721	Vav3	0.672	1.98E-03	-1.69	0.1036
NA	NA	0.673	1.73E-03	-1.76	0.0904
ENSMUSG00000055980	Irs1	0.674	1.90E-03	-1.72	0.0971
ENSMUSG00000048284	Tas2r126	0.674	2.05E-03	-1.68	0.1048
ENSMUSG00000028044	Cks1b	0.675	1.99E-03	-1.69	0.1001
ENSMUSG00000000028	Cdc45l	0.676	2.29E-03	-1.63	0.1121
NA	NA	0.677	2.00E-03	-1.69	0.0991
ENSMUSG00000031740	Mmp2	0.677	2.23E-03	-1.64	0.1085
ENSMUSG00000078453	3110003A17Rik	0.678	9.12E-04	-2.12	0.0298

ENSMUSG00000039982	Dtx4	0.678	1.31E-03	-1.92	0.0569
ENSMUSG00000002068	Ccne1	0.678	1.47E-03	-1.83	0.0714
NA	Tbrg3	0.678	2.27E-03	-1.63	0.109
ENSMUSG00000037286	Stag1	0.678	3.01E-03	-1.5	0.1396
ENSMUSG00000028885	Smpdl3b	0.679	1.26E-03	-1.94	0.0532
ENSMUSG00000024054	Smchd1	0.679	2.28E-03	-1.63	0.1083
ENSMUSG00000029638	Glcci1	0.679	2.93E-03	-1.51	0.1358
ENSMUSG00000018417	Myo1b	0.68	1.49E-03	-1.83	0.0687
ENSMUSG00000046808	Atp10d	0.68	2.24E-03	-1.63	0.1052
ENSMUSG00000031939	Taf1d	0.682	1.66E-03	-1.78	0.0748
ENSMUSG00000060090	Rp2h	0.682	2.06E-03	-1.68	0.0941
ENSMUSG00000046169	Adamts6	0.683	1.48E-03	-1.83	0.0654
ENSMUSG00000029788	Cpa5	0.683	2.48E-03	-1.58	0.1129
NA	NA	0.685	2.72E-03	-1.54	0.1187
NA	Snord82	0.685	2.74E-03	-1.54	0.1193
ENSMUSG00000015937	H2afy	0.686	1.47E-03	-1.83	0.0617
ENSMUSG00000030978	Rrm1	0.686	1.85E-03	-1.73	0.0796
ENSMUSG00000032235	Narg2	0.687	2.94E-03	-1.51	0.125
ENSMUSG00000034981	9130213B05Rik	0.689	2.07E-03	-1.68	0.086
ENSMUSG00000020534	Shmt1	0.69	1.42E-03	-1.86	0.0536
ENSMUSG00000051517	E130306D19Rik	0.69	2.92E-03	-1.51	0.119
ENSMUSG00000020492	Fam33a	0.691	1.96E-03	-1.7	0.0785
ENSMUSG00000049313	Sorl1	0.691	2.12E-03	-1.67	0.0853
NA	NA	0.691	2.58E-03	-1.57	0.1043
ENSMUSG00000047945	Marcksl1	0.691	3.00E-03	-1.5	0.1214
ENSMUSG00000029782	Tmem209	0.693	2.09E-03	-1.67	0.0822
ENSMUSG00000024665	Fads2	0.693	2.53E-03	-1.58	0.1016
ENSMUSG00000052798	Nup107	0.694	2.30E-03	-1.62	0.0906
NA	NA	0.694	2.81E-03	-1.53	0.1098
ENSMUSG00000031107	Rbmx2	0.694	3.07E-03	-1.49	0.1184
ENSMUSG00000051278	4930422G04Rik	0.695	1.92E-03	-1.71	0.0722
ENSMUSG00000027239	Mdk	0.698	2.78E-03	-1.53	0.1032
ENSMUSG00000029366	Dck	0.698	3.18E-03	-1.48	0.1158
ENSMUSG00000074934	Grem1	0.699	2.41E-03	-1.6	0.0874
ENSMUSG00000038070	Cntln	0.699	2.89E-03	-1.52	0.1064
ENSMUSG00000056394	Lig1	0.7	1.70E-03	-1.77	0.0564
NA	6330407A03Rik	0.701	2.57E-03	-1.57	0.0908
ENSMUSG00000006800	Sulf2	0.702	1.82E-03	-1.74	0.0591
ENSMUSG00000044783	Hjurp	0.702	2.26E-03	-1.63	0.0783
ENSMUSG00000075378	Olfr361	0.703	1.51E-03	-1.83	0.0439
ENSMUSG00000053141	Ptprt	0.703	1.85E-03	-1.73	0.0595
NA	2810026P18Rik	0.704	1.74E-03	-1.76	0.0534
ENSMUSG00000027752	Exosc8	0.704	2.54E-03	-1.57	0.0869
ENSMUSG00000027204	Fbn1	0.705	3.14E-03	-1.48	0.1062
ENSMUSG00000029687	Ezh2	0.706	1.54E-03	-1.82	0.0418
ENSMUSG00000029859	Epha1	0.706	1.64E-03	-1.78	0.0472
ENSMUSG00000040225	Bat2d	0.706	1.71E-03	-1.77	0.0501
NA	NA	0.706	1.89E-03	-1.72	0.0567
ENSMUSG00000002870	Mcm2	0.708	1.78E-03	-1.75	0.0502
ENSMUSG00000046818	Ddit4l	0.708	2.03E-03	-1.68	0.0613
ENSMUSG00000001517	Foxm1	0.709	1.60E-03	-1.8	0.0417

NA	NA	0.709	2.57E-03	-1.57	0.0811
ENSMUSG00000044934	Zfp367	0.709	2.82E-03	-1.53	0.09
ENSMUSG00000034855	Cxcl10	0.709	2.98E-03	-1.5	0.0955
ENSMUSG00000049932	H2afx	0.71	2.80E-03	-1.53	0.0883
ENSMUSG00000036036	Zfp57	0.711	2.34E-03	-1.61	0.071
ENSMUSG00000001986	Gria3	0.711	2.37E-03	-1.61	0.0713
ENSMUSG00000007653	Gabrb2	0.711	2.68E-03	-1.55	0.0823
ENSMUSG00000005397	Nid1	0.711	2.89E-03	-1.52	0.0904
ENSMUSG00000045414	1190002N15Rik	0.714	2.61E-03	-1.57	0.0762
ENSMUSG00000021699	Pde4d	0.715	1.43E-03	-1.85	0.0264
ENSMUSG00000026249	Serpine2	0.717	2.09E-03	-1.67	0.0519
ENSMUSG00000069312	Gm12260	0.72	1.89E-03	-1.72	0.0409
NA	Gm8956	0.72	2.22E-03	-1.64	0.0545
ENSMUSG00000021115	Vrk1	0.72	2.54E-03	-1.57	0.0657
ENSMUSG00000030265	Kras	0.72	2.99E-03	-1.5	0.0825
ENSMUSG00000058013	Sept11	0.722	2.80E-03	-1.53	0.0719
ENSMUSG00000074476	Spc24	0.723	3.16E-03	-1.48	0.0807
ENSMUSG00000026848	Tor1b	0.724	2.42E-03	-1.6	0.0565
ENSMUSG00000006369	Fbln1	0.727	2.08E-03	-1.67	0.0409
ENSMUSG00000030417	Pdcd5	0.731	2.65E-03	-1.56	0.0562
ENSMUSG00000032555	Topbp1	0.732	2.33E-03	-1.61	0.0443
ENSMUSG00000020069	Hnrnph3	0.737	2.95E-03	-1.5	0.0584
ENSMUSG00000020290	Xpo1	0.739	2.35E-03	-1.61	0.0359
ENSMUSG00000019873	Reep3	0.741	3.04E-03	-1.49	0.0548
ENSMUSG00000006403	Adamts4	0.753	3.20E-03	-1.48	0.0422
ENSMUSG00000027955	1110032E23Rik	0.769	2.84E-03	-1.53	0.014
ENSMUSG00000044080	S100a1	1.357	2.10E-03	1.67	0.0288
ENSMUSG00000039542	Ncam1	1.363	2.31E-03	1.62	0.0416
ENSMUSG00000017466	Timp2	1.369	1.70E-03	1.77	0.0213
NA	Tmem170b	1.371	2.01E-03	1.68	0.0356
ENSMUSG00000008226	Scrn3	1.372	1.77E-03	1.75	0.026
ENSMUSG00000054364	Rhob	1.374	1.98E-03	1.69	0.036
ENSMUSG00000038271	Iffo1	1.374	2.28E-03	1.63	0.0467
ENSMUSG00000023927	Satb1	1.376	2.21E-03	1.64	0.0455
ENSMUSG00000028517	Ppap2b	1.378	1.90E-03	1.71	0.0349
ENSMUSG00000057439	Kir3dl2	1.39	2.08E-03	1.67	0.0493
ENSMUSG00000024937	Ehbp1l1	1.39	2.18E-03	1.65	0.0526
ENSMUSG00000042686	Jph1	1.395	1.65E-03	1.78	0.0345
ENSMUSG00000033788	Dysf	1.398	1.72E-03	1.76	0.0397
ENSMUSG00000019960	Dusp6	1.407	1.93E-03	1.71	0.0537
ENSMUSG00000032009	Sesn3	1.408	2.27E-03	1.63	0.0684
ENSMUSG00000041261	Car8	1.411	2.19E-03	1.65	0.066
ENSMUSG00000027984	Hadh	1.412	1.45E-03	1.84	0.0352
ENSMUSG00000031762	Mt2	1.419	1.73E-03	1.76	0.0523
ENSMUSG00000026383	Epb4.1l5	1.421	2.21E-03	1.64	0.074
ENSMUSG00000062908	Acadm	1.422	1.53E-03	1.82	0.0445
ENSMUSG00000032479	Mtap4	1.422	2.01E-03	1.68	0.0668
ENSMUSG00000025488	Cox8b	1.423	2.24E-03	1.63	0.0762
ENSMUSG00000028108	Ecm1	1.424	1.40E-03	1.86	0.0391
ENSMUSG00000020182	Ddc	1.424	2.33E-03	1.62	0.081
ENSMUSG00000006435	Neurl1a	1.426	1.77E-03	1.75	0.058

ENSMUSG00000029718	Pcolce	1.436	1.20E-03	1.96	0.0311
ENSMUSG00000001131	Timp1	1.438	1.05E-03	2.07	0.019
ENSMUSG00000020393	Kremen1	1.438	1.50E-03	1.83	0.0513
ENSMUSG00000058070	Eml1	1.438	1.79E-03	1.74	0.0661
ENSMUSG00000026208	Des	1.439	1.30E-03	1.93	0.038
ENSMUSG00000022905	Kpna1	1.439	1.81E-03	1.74	0.067
ENSMUSG00000018845	Unc45b	1.439	2.04E-03	1.68	0.0783
ENSMUSG00000025348	Itga7	1.441	1.99E-03	1.69	0.0771
ENSMUSG00000019806	Aig1	1.442	1.82E-03	1.74	0.0691
ENSMUSG00000015944	Gatsl2	1.451	1.64E-03	1.78	0.0662
ENSMUSG00000071757	Zhx2	1.453	2.12E-03	1.67	0.0886
ENSMUSG00000074063	Osgin1	1.455	1.34E-03	1.9	0.0509
ENSMUSG00000024059	Clip4	1.455	2.29E-03	1.62	0.0982
ENSMUSG00000031778	Cx3cl1	1.457	1.96E-03	1.7	0.0849
ENSMUSG00000006638	Abhd1	1.458	1.33E-03	1.91	0.0497
ENSMUSG00000030592	Ryr1	1.459	7.95E-04	2.17	0.0166
NA	NA	1.46	2.23E-03	1.64	0.0989
NA	NA	1.461	1.64E-03	1.78	0.0719
ENSMUSG00000024130	Abca3	1.461	2.05E-03	1.68	0.0914
ENSMUSG00000050578	Mmp13	1.461	2.14E-03	1.66	0.094
ENSMUSG00000031461	Myom2	1.465	1.57E-03	1.81	0.0702
ENSMUSG00000060284	Sp7	1.466	2.20E-03	1.64	0.1009
ENSMUSG00000021770	Samd8	1.468	1.04E-03	2.07	0.0332
ENSMUSG00000032060	Cryab	1.468	1.55E-03	1.81	0.0715
ENSMUSG00000039782	Cpeb2	1.468	1.82E-03	1.74	0.0841
ENSMUSG00000029070	Mxra8	1.468	1.91E-03	1.71	0.0892
ENSMUSG00000019139	Isyna1	1.47	2.06E-03	1.68	0.097
ENSMUSG00000029071	Dvl1	1.471	9.93E-04	2.09	0.0324
ENSMUSG00000031972	Acta1	1.474	8.88E-04	2.13	0.0279
ENSMUSG00000024486	Hbegf	1.474	1.33E-03	1.9	0.0593
ENSMUSG00000029484	Anxa3	1.474	1.35E-03	1.88	0.0634
ENSMUSG00000018893	Mb	1.474	1.39E-03	1.87	0.0649
ENSMUSG00000027797	Dclk1	1.474	1.51E-03	1.83	0.072
ENSMUSG00000030257	Srgap3	1.476	7.83E-04	2.17	0.0242
ENSMUSG00000058056	Palld	1.477	1.56E-03	1.81	0.0765
ENSMUSG00000024885	Aldh3b1	1.478	1.36E-03	1.87	0.0658
NA	NA	1.483	1.52E-03	1.83	0.0765
ENSMUSG00000019843	Fyn	1.483	2.26E-03	1.63	0.1141
ENSMUSG00000042613	Pbxip1	1.484	2.31E-03	1.62	0.1172
ENSMUSG00000022297	Fzd6	1.486	1.81E-03	1.74	0.0941
ENSMUSG00000029787	Avl9	1.49	1.44E-03	1.85	0.0762
NA	NA	1.49	2.02E-03	1.68	0.107
ENSMUSG00000033065	Pfkm	1.491	1.17E-03	1.99	0.0554
ENSMUSG00000022508	Bcl6	1.492	2.10E-03	1.67	0.1108
ENSMUSG00000031753	Cog4	1.494	1.19E-03	1.97	0.059
ENSMUSG00000026564	Dusp27	1.495	1.03E-03	2.07	0.045
ENSMUSG00000032656	March3	1.497	1.35E-03	1.88	0.0749
ENSMUSG00000028803	Nipal3	1.498	1.52E-03	1.82	0.0846
ENSMUSG00000074743	Thbd	1.5	1.66E-03	1.78	0.0932
ENSMUSG00000025190	Got1	1.501	1.86E-03	1.73	0.1045
ENSMUSG00000026663	Atf6	1.504	1.59E-03	1.8	0.0923

ENSMUSG00000030352	Tspan9	1.505	1.54E-03	1.82	0.09
ENSMUSG00000026854	Usp20	1.506	2.03E-03	1.68	0.1162
ENSMUSG00000042115	Klhdc8a	1.51	1.15E-03	2	0.0628
ENSMUSG00000067071	Hes6	1.514	1.27E-03	1.93	0.075
ENSMUSG00000061816	Myl1	1.514	1.70E-03	1.77	0.1035
ENSMUSG00000070780	Rbm47	1.515	1.84E-03	1.74	0.1103
ENSMUSG00000032420	Nt5e	1.516	2.07E-03	1.67	0.1237
ENSMUSG00000079243	Xirp1	1.517	9.86E-04	2.09	0.0533
ENSMUSG00000020867	Spata20	1.518	1.44E-03	1.85	0.0903
ENSMUSG00000050737	Ptges	1.518	1.84E-03	1.73	0.1125
ENSMUSG00000054675	Tmem119	1.521	1.42E-03	1.86	0.091
ENSMUSG00000026418	Tnni1	1.526	9.97E-04	2.08	0.0579
ENSMUSG00000006457	Actn3	1.528	9.74E-04	2.09	0.0576
NA	NA	1.528	2.00E-03	1.69	0.1277
ENSMUSG00000021759	Ppap2a	1.529	1.13E-03	2.01	0.0694
ENSMUSG00000071847	Apcdd1	1.531	1.10E-03	2.04	0.0661
ENSMUSG00000005583	Mef2c	1.535	8.73E-04	2.14	0.0543
ENSMUSG00000021750	Fam107a	1.535	2.19E-03	1.65	0.1398
ENSMUSG00000055775	Myh8	1.537	2.23E-03	1.64	0.1436
ENSMUSG00000030748	Il4ra	1.538	8.26E-04	2.15	0.054
ENSMUSG00000040694	Apobec2	1.543	1.67E-03	1.78	0.1165
ENSMUSG00000022157	Mcpt8	1.549	1.50E-03	1.83	0.1099
ENSMUSG00000026489	Cabc1	1.553	8.65E-04	2.14	0.0616
ENSMUSG00000046027	Stard5	1.554	6.97E-04	2.23	0.05
ENSMUSG00000044086	Lmod3	1.56	1.24E-03	1.95	0.0941
ENSMUSG00000001027	Scn4a	1.561	7.71E-04	2.18	0.0603
ENSMUSG00000034295	Fhod3	1.571	8.69E-04	2.14	0.0695
ENSMUSG00000034898	Filip1	1.571	1.17E-03	1.98	0.0936
ENSMUSG00000032377	Plscr4	1.576	1.03E-03	2.07	0.0818
ENSMUSG00000033060	Lmo7	1.576	2.14E-03	1.66	0.1598
ENSMUSG00000057280	Musk	1.579	5.92E-04	2.32	0.0497
ENSMUSG00000038239	Hrc	1.587	1.74E-03	1.76	0.1443
ENSMUSG00000020758	Itgb4	1.588	8.92E-04	2.13	0.0783
ENSMUSG00000026459	Myog	1.596	8.42E-04	2.15	0.0792
ENSMUSG00000031700	Gpt2	1.601	6.19E-04	2.3	0.0602
ENSMUSG00000038403	Hfe2	1.608	1.17E-03	1.99	0.1105
ENSMUSG00000002020	Ltbp2	1.612	1.38E-03	1.87	0.134
ENSMUSG00000020629	Adi1	1.615	5.18E-04	2.38	0.0562
ENSMUSG00000031304	Il2rg	1.615	1.45E-03	1.85	0.1396
ENSMUSG00000042918	Mamstr	1.624	8.03E-04	2.16	0.0885
ENSMUSG00000032238	Rora	1.624	1.02E-03	2.08	0.1024
ENSMUSG00000079588	Tmem182	1.626	1.07E-03	2.05	0.1064
ENSMUSG00000030249	Abcc9	1.626	1.18E-03	1.98	0.1196
ENSMUSG00000030218	Mgp	1.628	4.44E-04	2.45	0.052
ENSMUSG00000047547	Cltb	1.628	1.41E-03	1.86	0.1438
ENSMUSG00000029804	Herc3	1.628	1.57E-03	1.81	0.1539
ENSMUSG00000004371	Il11	1.63	5.61E-04	2.35	0.0646
ENSMUSG00000040511	Pvr	1.631	4.52E-04	2.44	0.0544
ENSMUSG00000027332	Ivd	1.632	1.29E-03	1.93	0.1319
ENSMUSG00000004885	Crabp2	1.638	8.06E-04	2.16	0.0943
ENSMUSG00000020067	Mypn	1.64	1.71E-03	1.77	0.1688

ENSMUSG00000027800	Tm4sf1	1.643	5.53E-04	2.36	0.0685
ENSMUSG00000021253	Tgfb3	1.643	9.04E-04	2.12	0.1026
ENSMUSG00000022309	Angpt1	1.643	1.21E-03	1.96	0.1302
ENSMUSG00000027861	Casq2	1.644	1.07E-03	2.06	0.1142
ENSMUSG00000040254	Sema3d	1.65	1.80E-03	1.74	0.1802
ENSMUSG00000025172	Ankrd2	1.651	5.30E-04	2.37	0.0704
NA	NA	1.654	1.06E-03	2.06	0.1183
ENSMUSG00000026185	Igfbp5	1.658	3.39E-04	2.59	0.0467
ENSMUSG00000025429	Pstpip2	1.668	5.22E-04	2.38	0.0759
ENSMUSG00000023367	Tmem176a	1.668	8.61E-04	2.14	0.1098
ENSMUSG00000033207	Mamdc2	1.674	7.40E-04	2.2	0.1035
ENSMUSG00000002007	Srpk3	1.677	5.14E-04	2.38	0.0786
ENSMUSG00000039891	Txlnb	1.678	9.47E-04	2.11	0.1197
ENSMUSG00000034161	Scx	1.682	3.55E-04	2.58	0.0564
ENSMUSG00000001020	S100a4	1.687	7.32E-04	2.2	0.1083
ENSMUSG00000021451	Sema4d	1.688	5.80E-04	2.33	0.0895
ENSMUSG00000021596	Mctp1	1.699	1.08E-03	2.05	0.1378
ENSMUSG00000030257	Srgap3	1.7	1.05E-03	2.07	0.1359
ENSMUSG00000019787	Trdn	1.706	2.30E-03	1.62	0.2401
ENSMUSG00000031791	Tmem38a	1.709	3.43E-04	2.58	0.0642
ENSMUSG00000074971	Fibin	1.711	2.03E-03	1.68	0.2261
ENSMUSG00000026308	Klhl30	1.713	9.78E-04	2.09	0.1372
ENSMUSG00000075307	Kbtbd10	1.72	8.38E-04	2.15	0.1291
ENSMUSG00000020034	Tcp11l2	1.722	1.29E-03	1.93	0.1714
ENSMUSG00000059256	Gzmd	1.728	4.40E-04	2.45	0.0868
NA	NA	1.747	1.06E-03	2.06	0.1553
ENSMUSG00000023960	Enpp5	1.75	5.88E-04	2.32	0.1127
ENSMUSG00000039323	Igfbp2	1.753	6.66E-04	2.27	0.1224
NA	NA	1.753	1.36E-03	1.87	0.1981
ENSMUSG00000019787	Trdn	1.756	1.48E-03	1.83	0.2083
ENSMUSG00000009214	Tmem8c	1.76	8.57E-04	2.14	0.146
ENSMUSG00000022156	Gzme	1.761	5.73E-04	2.34	0.1145
ENSMUSG00000019787	Trdn	1.775	3.82E-04	2.54	0.0911
ENSMUSG00000052374	Actn2	1.776	4.17E-04	2.48	0.0993
ENSMUSG00000032601	Prkar2a	1.783	3.12E-04	2.64	0.0808
ENSMUSG00000027253	Lrp4	1.815	2.22E-04	2.8	0.0725
ENSMUSG00000027199	Gatm	1.815	5.65E-04	2.35	0.1319
ENSMUSG00000022519	Srl	1.83	5.45E-04	2.37	0.1337
ENSMUSG00000042451	Mybph	1.838	8.57E-05	3.34	0.0278
ENSMUSG00000061603	Akap6	1.839	4.91E-04	2.4	0.1312
ENSMUSG00000055027	Smyd1	1.842	3.00E-04	2.66	0.0962
ENSMUSG00000025085	Ablim1	1.854	1.99E-04	2.9	0.0718
NA	NA	1.88	4.68E-04	2.42	0.1408
ENSMUSG00000056091	St3gal5	1.892	2.30E-04	2.78	0.0957
ENSMUSG00000029298	BC057170	1.893	1.56E-04	3.03	0.0692
ENSMUSG00000021640	Naip1	1.895	9.70E-04	2.09	0.2059
ENSMUSG00000037139	Myom3	1.896	1.64E-04	3.02	0.0704
ENSMUSG00000020475	Pgam2	1.9	1.68E-04	2.99	0.0744
ENSMUSG00000038763	Alpk3	1.907	4.56E-04	2.43	0.1479
ENSMUSG00000024190	Dusp1	1.937	1.75E-04	2.98	0.0857
ENSMUSG00000042828	Trim72	1.942	2.14E-04	2.81	0.1064

ENSMUSG00000017300	Tnnc2	1.97	1.05E-04	3.23	0.0676
ENSMUSG00000020908	Myh3	1.975	8.96E-05	3.33	0.0603
ENSMUSG00000078949	R3hdml	1.983	2.77E-04	2.72	0.1289
ENSMUSG00000011148	Adssl1	2.011	1.44E-04	3.06	0.095
ENSMUSG00000026576	Atp1b1	2.021	2.57E-04	2.74	0.1362
ENSMUSG00000031636	Pdlim3	2.023	8.18E-05	3.34	0.0692
ENSMUSG00000021909	Tnnc1	2.049	5.45E-05	3.45	0.0652
ENSMUSG00000001763	Tspan33	2.05	3.90E-05	3.6	0.0531
ENSMUSG00000079363	Gbp4	2.095	2.10E-04	2.87	0.1372
ENSMUSG00000042793	Lgr6	2.152	2.34E-05	3.76	0.0594
ENSMUSG00000059336	Slc14a1	2.174	6.62E-05	3.41	0.0934
ENSMUSG00000044033	2610301F02Rik	2.18	4.29E-05	3.57	0.0802
ENSMUSG00000019787	Trdn	2.203	2.01E-03	1.69	0.4413
ENSMUSG00000028348	Murc	2.291	5.84E-05	3.44	0.1126
ENSMUSG00000023886	Smoc2	2.294	2.73E-05	3.72	0.0876
ENSMUSG00000025582	Nptx1	2.32	3.51E-05	3.69	0.0941
ENSMUSG00000024049	Myom1	2.322	1.95E-05	3.8	0.0854
ENSMUSG00000030672	Mylpf	2.355	1.56E-05	4.24	0.0565
ENSMUSG00000027208	Fgf7	2.411	1.17E-05	4.55	0.0442
ENSMUSG00000038623	Tm6sf1	2.543	9.35E-05	3.31	0.1722
ENSMUSG00000020598	Nrcam	2.587	4.68E-05	3.51	0.1562
ENSMUSG00000007122	Casq1	3.177	7.79E-06	4.61	0.1272

Delta = 0.595