4.2. Pax3 and Pax7, and limb muscle development

Pax genes are a family of developmental control genes characterized by a paired-box, a highly conserved motif of 128 amino acids. Initially, pairedbox genes were detected in the segmentation genes of Drosophila melanogaster (Bopp et al., 1986). On the basis of homology to Drosophila paired-box sequences, nine members of the Pax family have been isolated in higher vertebrates and are classified into four paralog groups (Burri et al., 1989; Deutsch et al., 1988; Gruss and Walther, 1992; Mansouri et al., 1996a; Stuart et al., 1993). Crystal structure data indicate that the pairedbox consists of two linked subdomains, each of which structurally resembles a helix–turn–helix (HTH) motif, a structure found in all homeodomains (Xu et al., 1999). The homeodomians found in paired-box proteins cooperatively dimerize on palindromic sites of the DNA sequence 5^0-TAAT(N)$_{2-3}$ATTA-3^0 (Chi and Epstein, 2002; Wilson et al., 1993). In addition to the paired box, several Pax proteins, including Pax3, Pax4, Pax6, and Pax7, possess two other conserved motifs, an octapeptide and a homeodomain (Epstein et al., 1994; Walther et al., 1991). Expression of Pax genes is temporally and spatially restricted during development, indicating that they are regulatory genes controlling the early steps of muscle development (Mansouri et al., 1999). In humans, several mutations in Pax genes have been associated with congenital diseases (Stuart et al., 1993). Chromosomal translocations involving Pax3 or Pax7 are found in rhabdomyosarcoma, a pediatric tumor believed to arise from skeletal muscle progenitor cells (Galili et al., 1993; Shapiro et al., 1993).

Pax3 and its paralogue Pax7 have been implicated in the specification of cells that will enter into the skeletal muscle development program. Transcripts of Pax3 were first detected in E8.5 mouse embryos in the dorsal part of the neuroepithelium and in the adjacent segmented dermomyotome, and were later found in somites, restricted to muscle progenitor cells (Goulding et al., 1991). From E10 to E11, expression of Pax3 was observed in the undifferentiated mesenchyme of both the forelimb and hindlimb (Goulding et al., 1991). Pax7 was expressed later in the central dermomyotome (Jostes et al., 1990). Expression of Pax3 was not seen in migrating myoblasts, which are known to express Pax-7.

The spontaneous mouse mutation splotch (Sp), having a deletion within Pax3 (Epstein et al., 1991), provided a means to study the function of Pax3 before loss of function artificial mutants became available. In homozygous Splotch mice, Pax3-expressing cells were absent and the diaphragm and limb muscles were lost (Bober et al., 1994; Franz et al., 1993), while no abnormalities were detected in the trunk muscle (Goulding et al., 1994). Forced expression of the Pax3 gene in chick somites in vitro was shown to induce MyoD and Myf5 in the paraxial mesoderm (Maroto et al., 1997),

and therefore indicated that Pax3 functions upstream of these MRFs. Pax3 was shown to bind the promoter of MyoD directly and transactivated enhancers of Myf5 and MyoD (Bajard et al., 2006; Hu et al., 2008). Dominant-negative forms of both Pax3 and Pax7 repressed MyoD, but did not interfere with the expression of Myf5 (Relaix et al., 2006).

Mice with a null allele for Pax7 were created by homologous recombination in embryonic stem cells (Mansouri et al., 1996b). Homozygous Pax7 null offspring developed to term and showed no defects in embryonic or fetal myogenesis, but died within 3 weeks after birth. In all Pax7 null mutants, facial skeletal structures were affected, but no obvious phenotypes were detected in skeletal muscle (Mansouri et al., 1996b; Seale et al., 2000). Pax7 null mice have significant numbers of satellite cells at birth, but the population is progressively depleted as a result of cell cycle defects and increased apoptosis, indicating the requirement of Pax7 for the maintenance of adult satellite cells (Kuang et al., 2006; Oustanina et al., 2004; Relaix et al., 2006). Compound Pax3 and Pax7 null mice have no limb muscles and have no embryonic or fetal trunk muscle, although a primary myotome was found (Relaix et al., 2005).

A battery of mice with the genetic loss-of function of Pax3, Pax7, and MRF transcription factors and their combinations have been created. In an elaborate review article, Murphy and Kardon (2011) compiled and tabulated the phenotypes of these mutant mice, which showed that the development of trunk and limb muscles of embryonic, fetal, and adult mice has different genetic requirements. More recently, Cre-mediated lineage analysis in mice was performed with the Cre-recombinase gene placed under the control of the promoter/enhancer sequences of Pax3, Pax7, or MRFs. For temporal control of labeling and manipulation, tamoxifen-inducible Cre alleles have also been created with the induced expression of reporter genes, or the induction of diphtheria toxin to ablate-specific cell lineages (Hutcheson and Kardon, 2009). Pax3 and Pax7 exhibit divergent functions. Pax3 is required for embryonic and fetal limb muscle development, and Pax7 can substitute for Pax3 in the somite and the trunk, but not in the limbs (Relaix et al., 2004). Pax7 progenitors do not give rise to embryonic muscle, but give rise to all fetal and adult myoblasts and myofibers in the limb (Hutcheson and Kardon, 2009; Lepper and Fan, 2010).

Pax3 expression is downregulated postnatally, except for those found in the diaphragm and a few other muscles (Kuang and Rudnicki, 2008). In the adult, muscle growth and repair rely on the proliferation and differentiation of satellite cells. Pax7 is expressed by the majority of quiescent satellite cells (Seale et al., 2000), and appears to be essential for postnatal maintenance and self-renewal. Pax3 is only present in the

satellite cells of particular muscles (Relaix et al., 2006), and cannot substitute for Pax7 (Lagha et al., 2008; Le Grand and Rudnicki, 2007). Using chromatin immunoprecipitation (ChIP) studies, Pax7 was shown to directly bind a Myf5 enhancer in myoblasts, together with the components of the histone methyltransferase complex directing methylation of histone H3 lysine 4 (H3K4) (McKinnell et al., 2008). In adult muscle, most Pax7þ satellite cells express Myf5 (Beauchamp et al., 2000), but a small subpopulation (about 10%) are Myf5 negative (Kuang et al., 2007), and these Myf5 satellite cells are derived from the embryonic somite Pax3þ, Pax7þ, or MRFs progenitors (Gros et al., 2005). During muscle regeneration, activated Myf5 satellite cells can asymmetrically generate Myf5 cells for self-renewal and Myf5þ committed cells (Kuang et al., 2007). Recently, gene inactivation by Cremediated recombination after delivery showed that Pax7 is required between P0 and P21 for the generation of neonatal satellite cells, but is not required for effective muscle regeneration after P21 (Lepper et al., 2009). Further studies are required regarding the significance of this change in genetic requirement for muscle stem cells during different developmental stages.

4.3. Six and Eya factors

The sine oculis homeobox (SIX) protein family is a group of evolutionarily conserved transcription factors homologous to the Drosophila melanogaster sine oculis (so) gene, which is associated with formation of the compound eye and the entire visual system (Fischbach and Technau, 1984). Two other SIX family members were also identified in flies, with optix functioning in the developing eye, and DSix4 in several mesoderm derivatives including a subset of somatic muscles, the somatic cells of the gonad and the fat body (Clark et al., 2006; Kirby et al., 2001; Seo et al., 1999). Homologs of Drosophila so, optix, and DSix4 have been identified throughout the animal kingdom (Seo et al., 1999). Mice and humans have six Six genes, which can be divided into three subclasses designated as the Six1/2, Six3/6, and Six4/5 subfamilies. The Six4 protein was first identified as the cDNA clones encoding the ARE- (Na, K-ATPase a1 subunit gene regulatory element) binding protein (Kawakami et al., 1996). It was subsequently demonstrated that Six1, Six2, Six4, and Six5 show similar binding specificity to the ARE/ MEF3 site (consensus sequence TCAGGTT) (Spitz et al., 1998).

The SIX family of transcription factors are characterized by the presence of two evolutionarily conserved domains, the SIX domain (SD) with 146 amino acid followed by the homeobox nucleic acid recognition domain (HD) on the C-terminal side (Kumar, 2009). HDs are 60 amino

acids long, contain a HTH motif, and are classified based on their overall sequence similarity and the identity of several critical residues within the recognition helix (Galliot et al., 1999). SIX HDs most closely resemble the Paired HD found in the Pax family (Galliot et al., 1999; Kumar, 2009). DNA-binding properties appear to be confined to the HD, and the SD appears to be critical for mediating protein-protein interactions (Hu et al., 2008).

SIX homeogenes are expressed in several cell types and are involved in different types of organogenesis including the kidney, thymus, and auditory system, as well as displaying a strong disorganization of craniofacial structures (Laclef et al., 2003a,b; Zheng et al., 2003). In mice, Six1, Six4, and Six5 genes are expressed from E8 in overlapping expression patterns in somites, limb buds, dorsal root ganglia, and branchial arches (Oliver et al., 1995). Physical interactions between Six and Eya (eyes absent) proteins, were first described in Drosophila (Pignoni et al., 1997), and were also found conserved in vertebrates (Grifone et al., 2004, 2007; Heanue et al., 1999 ; Ohto et al., 1999). SIX translocates Eya1–3 to the nucleus, and G-proteins can stop this interaction (Fan et al., 2000; Ohto et al., 1999). Four Eya genes have been cloned in mice, humans, and chicks (Borsani et al., 1999; Xu et al., 1997a,b).

To determine the functions of the Six1 gene during muscle development, Six1 null mutants were created which showed a lack of kidneys and the thymus, as well as displaying a strong disorganization of craniofacial structures, including the inner ear, nasal cavity, and the craniofacial skeleton. Six1 null mice died at birth from respiratory failure, and displayed impaired primary myogenesis, as shown by a severe reduction and disorganization of primary myofibers in most body muscles (Laclef et al., 2003a,b). While Six4 null mice appeared to be normal (Ozaki et al., 2001), compound Six1 and Six4 null mutant mice displayed a more pronounced impairment in myogenesis than Six1 null animals (Grifone et al., 2004). Notably, these double-mutant embryos no longer have myogenic progenitor cells in their limb buds, resulting in muscle-less legs (Grifone et al., 2005). Both in Six1 null, Six4 null and Eya1 null, Eya2 null double mutants, Pax3 expression in the hypaxial dermomyotome was lost, leading to cell misrouting and cell death, preventing muscle progenitor cell migration into the limbs (Grifone et al., 2005, 2007). In the genetic hierarchy of myogenesis, the Six and Eya genes lie upstream of Pax3. In the trunk, Six1 and Six4 genes have been shown to control the expression of Mrf4, and Six1 null, Six4 null doublemutant embryos also have a reduced and delayed expression of MyoD, Myogenin, and myotomal markers, whereas early activation of Myf5 in the epaxial somite still takes place (Grifone et al., 2005).

Six1 and Six4 act directly on Myf5 activation through a MEF3 site present in the 145-bp regulatory element that directs Myf5 expression in the limb buds (Giordani et al., 2007). It is known that in addition to its role in the early phase of muscle development, Six1 has influence ranging from establishing muscle fiber type to increasing the number of fast-twitch (glycolytic) muscle fibers, and to increasing expression of stereotypical genes (Grifone et al., 2004).

4.4. Factors interacting with Pax3/Pax7

Pax3 potentiates migration of limb muscle precursors by directly activating the expression of the c-Met tyrosine kinase receptor essential for the migration of muscle progenitor cells from the somite toward the limb buds. The c-Met tyrosine kinase receptor binds hepatocyte growth factor (Bottaro et al., 1991; Trusolino et al., 2010), and c-met null embryos showed that the limb bud, diaphragm, and tip of the tongue were not colonized by myogenic precursor cells (Bladt et al., 1995). The c-met promoter contains a Pax3-binding site, and Pax3 controls the release of migrating muscle precursors in vivo by activating c-met (Christ and BrandSaberi, 2002; Epstein et al., 1996). Another gene expressed at sites of muscle precursor cell detachment and in migratory muscle precursors, is the homeobox gene Lbx1, which requires Pax3 for the expression as indicated by the absence in Splotch mice. Lbx1-deficient mice lack muscles in their limbs due to a defect in migration of muscle precursor cells (Brohmann et al., 2000; Gross et al., 2000).

In vitro cultures of Chick paraxial mesoderm showed that Shh and Wnt1 together induced high expression of Pax3, Pax7, Myf5, MyoD, and the MyHC. These signaling molecules are secreted by the floor plate-notochord for Shh and the neural tube for Wnt1 (Maroto et al., 1997), both neighboring the somites (Tajbakhsh et al., 1997).

4.5. Head muscles

The striated muscles in the head are classified as follows: branchiomeric muscles which control jaw movement, facial expression, and pharyngeal and laryngeal function; the six extraocular muscles, which move and rotate the eye; and neck and tongue muscles, which are derived from myoblasts originating in the most anterior set of somites (Grifone and Kelly, 2007 ; Noden and Francis-West, 2006). While trunk and limb muscles originate from the somites in the mesoderm flanking the neural tube, branchiomeric muscles originate from the branchial arches in the head mesoderm, and extraocular muscles originate from the prechordal mesoderm (Noden et al., 1999; Noden and Francis-West, 2006; Sambasivan et al., 2011; Shih et al.,

2008). In Pax3 null Myf5 (and closely located Mrf4) null mutants, head muscles developed while trunk and limb muscles were absent (Tajbakhsh et al., 1997). This demonstrated that the developmental pathway of head muscles was different to that of trunk and limb muscles (Shore and Sharrocks, 1995).

Unlike developing somites, head muscles did not require the Pax3 gene for development. All head muscles express the myogenic determination transcription factors, myf5, then myoD (Nodan et al., 2006); however, it is believed that either Myf-5 or MyoD is essential for skeletal muscle formation (Rudnicki et al., 1993) and, in Myf-5 null mice, MyoD activation could rescue the myogenic program. The possibility that the lack of Pax3 in Pax3: Myf5 (Mrf4) mutants may have been rescued by Pax7 was investigated. Indeed, Pax7 was expressed in adult head progenitors, but this expression followed the expression of MyoD and myogenin, and Pax7 null mice showed head muscles, implying that Pax7 was not required for the expression of myogenic MRF (Horst et al., 2006). Other transcription factors were investigated for their involvement in head muscle development. Six1: Six4 and Eya1:Eya2 double null mutants showed a deficiency in limb muscles, but developed trunk and head muscles. This indicated that Six1, Six4, Eya1, and Eya2 were not involved in development (Grifone et al., 2007). Musculin (Msc or MyoR) and Capsulin (epicardin, Tcf21, Pod-1), two related bHLH transcription factors, were suggested to play important roles in head muscle development, because double null embryos for these factors revealed a complete absence of the major muscles for mastication (Lu et al., 2002).

In embryos, Tbx1 expression was observed in the pharyngeal region at E9.5, both in endoderm and mesoderm, and at E12.5, several additional sites of Tbxl expression became evident including the myogenic core of the tongue, the incisor tooth buds, and the branching lung epithelium (Chapman et al., 1996). By targeting Tbx1, null mice were created which displayed a wide range of developmental anomalies including hypoplasia of the thymus and parathyroid glands, cardiac outflow tract abnormalities, abnormal facial structures, abnormal vertebrae, and cleft palate (Jerome and Papaioannou, 2001). These abnormalities associated with a homozygous deficiency of Tbx1 were traced to the abnormal development of pharyngeal arches and pouches, the head mesenchyme, and the otic vesicle (Jerome and Papaioannou, 2001). Pitx2 is a paired-related homeobox gene found mutated in Rieger syndrome type I, causing a haploinsufficient disorder that includes tooth anomalies, anterior segment eye defects, and facial dysmorphologies (Diehl et al., 2006; Lu et al., 1999). In Pitx2 null embryos, branchiomeric muscle precursors were initially present, but failed to expand and activate the myogenic program (Dong et al., 2006). Conditional Pitx2 inactivation in mice and knockdown in chick

primary cultures supported a direct role for Pitx2 in branchiomeric muscle development (Dong et al., 2006).

Extraocular muscles are mainly derived from the prechordal mesoderm, while pharyngeal muscle progenitors are present in close proximity to cells of the SHF (Noden and Francis-West, 2006). Transcription factors required for branchiomeric myogenesis, such as Tbx1, MyoR, or capsulin, do not play a role in extraocular muscle development (Kelly et al., 2004; Lu et al., 2002). Furthermore, the MRFs differ between extraocular and branchiomeric muscles (Sambasivan et al., 2009). Myf5 or Mrf4 is necessary to initiate extraocular myogenesis, whereas Mrf4 is dispensable for pharyngeal muscle progenitor fate, and Tbx1 and Myf5 are required for myogenesis in this muscle type (Sambasivan et al., 2009). Although Pitx2 function is required for both extraocular and branchiometric muscles (Dong et al., 2006), extraocular muscle development is clearly subject to different genetic regulation from branchiomeric muscles (Table 4.5).

Head muscle progenitors for branchiomeric muscles are present in close proximity to cells of the SHF, and show overlapping patterns of gene expression, and Tbx1 involvement for the development are common to both. The segregation of two branchiomeric head muscle lineages have been reported, both of which contribute to the myocardium. The first, which is derived from the first branchial arch, gives rise to the masseter muscles, and also contributes to the myocardium of the right ventricle

Table 4.5 Differences in skeletal muscle lineages

Paraxial mesoderm		PITX2		Myf5, Mrf4, MyoD		Myogenin)JawExtraocularmuscle muscle
Branchial arches	)	PITX2, Tbx1	!	Myf5, MyoD	!	
Most anterior	)	PITX2	!	Myf5, MyoD	!	Myogenin)Neck & tongue muscle
somites	)	Pax3, MyoD	!	Myf5, Mrf4,	!	Myogenin)Trunk muscle
Somites	)	Six1/Six2, Pax3;	!	Myf5, Mrf4	!	Myogenin)Limb muscle
Somites	)	Eya1/Eya2	!		!	Myogenin)

References: Bismuth and Relaix (2010), Braun and Gautel (2011), and Sambasivan et al. (2011).

(Lescroart et al., 2010). The second lineage, which is derived from the mesoderm of the second branchial arch, gives rise to the muscles responsible for facial expression, and also contributes to the outflow tract myocardium at the base of the arteries (Lescroart et al., 2010). This reveals a common lineage relationship between head muscles and the SHF for cardiogenesis (Braun and Gautel, 2011; Grifone and Kelly, 2007; Sambasivan et al., 2011)

5. Signal Transduction in the Development of SMs

5.1. SRF and the CArG box

SM cells do not express the MRF family of myogenic transcription factors, and normal development of SM was observed in MyoD, Myf5, or myogenin null mice indicating a differentiation process distinct from striated muscles (Hasty et al., 1993; Nabeshima et al., 1993; Rudnicki et al., 1993). In attempt to define specific markers, genes with restricted expression for SM indicated the role of the promoter region containing CArG boxes, such as three CArG boxes in the SM a-actin gene promoter and two in the SM22a gene promoter (Blank et al., 1992; Kim et al., 1997; Owens, 1995).

The CArG box was originally defined as a part of the proximal enhancer segment of the human c-fos promoter, involved in the immediate and early response to serum stimulation (Treisman, 1986, 1987). In parallel, evolutionarily conserved CArG boxes were also identified as the regulatory element in human a-cardiac actin promoter (Minty and Kedes, 1986 ; Miwa and Kedes, 1987). The CArG box with a 10-bp sequence CC(A/T)$_6$GG, was found to be a specific binding site for a nuclear factor termed SRF, and successive experiments established the role of SRF in the induction of the c-fos immediate early response as well as in the expression of several muscle-specific genes (Johansen and Prywes, 1995). SRF is a 67- kD DNA-binding protein to the c-fos promoter isolated by Norman et al. (1988) , and is a member of the MADS-box, a highly conserved DNA-binding / dimerization domain. MADS is an acronym referring to the original members of the family, the MCM1, Agamous, Deficiens, and SRF (SchwarzSommer et al., 1990). There are numerous MADS-box proteins in plants, but in metazoans, SRF and the four members of the MEF2 family (MEFA, MEFB, MEFC, and MEFD) are the only known MADS-box proteins (Black and Olson, 1998).

Among the genes identified as the markers of SM cells (Owens, 1995), numerous genes contained at least one evolutionarily conserved CArG box within 10 kb of their transcription start sites, and Smoothelin A, SM22a, SM a-actin, SM g-actin, SM-calponin, and SM-MHC genes contained multiple CArG box (Miano, 2003). A study with the SM22a gene showed that a nucleotide sequence on both sides of the CArG box is required for SM specificity with the binding of other factors (Strobeck et al., 2001).

Several MADS-box proteins were shown to specifically recruit other transcription factors into multicomponent regulatory complexes, and such interactions with other proteins appeared to play a pivotal role in the regulation of target genes (Shore and Sharrocks, 1995). The requirement for the association of SRF with other cell-restricted factors was first described for the serum response element of the c-fos gene, resulting in the ternary complex involving the CArG box (Shaw et al., 1989). Growth factor responsiveness of the c-fos promoter was a result of the association of SRF with an Ets family of transcription factors recognizing Ets-specific GGAA/T sequences that immediately flank one of the c-fos CArG boxes (Dalton and Treisman, 1992). SRF is expressed widely, including in nonmuscle cells. Upon induction of SM-specific factors, a switch of the SRF association partner replacing growth-specific Ets factors toggled, the program from cell growth to SM differentiation (Miano, 2003).

Genome-wide screening forCArGboxeshasbeenconducted manytimes. Zhang et al. (2005) used Protein A-tagged SFR transfectant P19 cells, induced with Me$_2$SO, to collect direct SRF-bound gene targets, and uncovered scores of potential SRF-binding CArG and CArG-like boxes containing genes. Among these, 43 genes were functionally validated, and many of these target genes showed the transcription factor binding motifs for NKE (Nkx2–5 and other Nkx2 homeodomains proteins), GATA (GATA1–6, dual C4 zincfinger protein), mTATA, E box, HNF1/4, STATs, Smad, Comp, mTEF, Ets, NF-kB, and YY1 (Niu et al., 2007). YY1 is a negative regulator for SRF binding, antagonizing SRF action (Gualberto et al., 1992). For genome-wide screening, Sun et al. (2006) used a computational approach involving comparable genomic analysis of human and mouse orthologous genes, and uncovered more than 100 hypothetical SRF-dependent genes. Cooper et al. (2007) used a ChIP assay in combination with human promoter microarrays to identify 216 putative SRF-binding sites in the human genome, with 84 validated SRF-binding sites that changed with three different cell types including SM cells, neurons, and lymphocytes. Benson et al. (2011) used in silico CArG sequence conservation screening to interrogate an 8-kb window of the genomic sequence centered at the transcription start sites in each of the 18,925 protein-coding human genes. Of the 142,597 CArG boxes identified, 8252 meet the criteria for conservation within vertebrate

species.   These   data suggestthegenome-
wideusageofCArGboxesforgeneregulationinmultiple cellular processes.
Ubiquitously expressed SRF may serve as an anchor which binds with cell
lineage specific partners to activate cell type genes. Posttranslational
modification of SRF regulates functions (Johansen and Prywes, 1995).
SRF differs at its 3^0 untranslated region by alternative usage of two
polyadenylation signals, and this may add versatility (Belaguli et al., 1997;
Norman et al., 1988; Zhang et al., 2007). SRF null mice stopped
developing at the onset of gastrulation, lacking detectable mesodermal
cells (Arsenian et al., 1998; Weinhold et al., 2000), indicating that SRF is
already involved prior to muscle development.

5.2. Myocardin

A search for SRF partner proteins was attempted using bioinformatics-
based screening. Myocardin was one of the cDNAs identified in the
expressed sequence tag (EST) database for cardiac cDNA libraries.
Myocardin showed the deduced 35 amino acid sequence of the SAP (SAF-
A/B, Acinus, PIAS) domain (Wang et al., 2001), which is a helix–linker–
helix   motif   recognizing
chromosomalregionsknownasscaffoldormatrixattachmentregions(Aravin
d and Koonin, 2000). Subsequent in vitro experiments revealed that
myocardin interacts with SRF physically, and functions as a very powerful
coactivator, inducing CArG box-dependent promoters for the SM22, SM-
calponin, caldesmon, SM myosin light chain kinase, SM a-actin, and SM-
MHC proteins, which are all markers of SM (Chen et al., 2002; Wang et
al., 2002).

In vivo, the expression of myocardin was first detected in the cardiac
crescent at E7.75 concomitant with expression of the homeobox gene
Nkk2–5, the earliest marker for cardiogenic specification (Lints et al.,
1993), and continued throughout the developing heart until birth (Wang et
al., 2001). Myocardin was also expressed in a subset of embryonic vascular
and visceral SM cells starting from E13.5, but was not expressed in
skeletal muscle (Wang et al., 2001). When myocardin was expressed in
cells already committed to myogenic lineage, the expression of SM
markers like SM a-actin and calponin was elevated (Chen et al., 2002).
When myocardin was expressed in nonmyogenic 10T1/2 cells, markers for
SM were elevated, but not those of skeletal or cardiac muscle, whereas
MyoD expression elevated the expression of skeletal muscle markers
(Wang et al., 2003). Consequently, myocardin was characterized as a
master regulator of SM gene expression (Wang et al., 2003; Yoshida et al.,
2003). Myocardin null mice lacking a functional myocardin gene revealed

no abnormalities before E8.0, but died by E10.5 with the absence of vascular SMC differentiation (Li et al., 2003). Cardiac structure and gene expression were apparently normal in myocardin null mice, indicating that myocardin is dispensable for heart development. Surprisingly, the expression of SM22, SM a-actin, and atrial natriuretic factor, genes known to be direct targets of myocardin, was normal in myocardin null mice, suggesting the possibility of compensatory factors (Li et al., 2003).

5.3. Myocardin-related transcription factors (MRTFs)

Two myocardin-related transcription factors, MRTF-A (MKL1, MAL, BSAC) and MRTF-B (MKL2), were obtained from mouse embryo cDNA libraries having nucleotide homology (Miralles et al., 2003 ; Sasazuki et al., 2002; Wang et al., 2002). Myocardin, MRTF-A, and MRTF-B share more than 60% amino acid identity within the N-terminal 400 amino acid stretch, all having a conserved N-terminal domain composed of a RPEL domain, a SAP domain for DNA binding, and a leucine zipper-like domain for homo- or heterodimerization (Wang and Olson, 2004). A RPEL domain consists of three actin binding RPEL repeats that have been implicated in nucleocytoplasmic shuttling of MRTF-A and - B, and also constitutive nuclear localization of myocardin (Guettler et al., 2008; Miralles et al., 2003). The constitutive nuclear accumulation of myocardin was due to its strong affinity for the nuclear importing importin a1/b1 heterodimer, and the low nuclear import abilities of MRTF-A and -B was attributed to their weak binding affinities (Nakamura et al., 2010).

While myocardin is expressed specifically in cardiac and SM cells, MRTFA and -B are expressed in numerous embryonic and adult tissues (Wang et al., 2002). Targeted mutation in the MRTF-B gene resulted in complete lethality between E13.5 and 14.5 (Oh et al., 2005). The phenotypes observed in MRTF-B null mice were malformations of the branchial arch arteries and cardiac outflow tract, specifically a reduction in SM within the walls of the third, fourth, and sixth branchial arch arteries at E11.5. Mice homozygous for a null mutation in the MRTF-A gene were viable, but the female mice were unable to effectively nurse their offspring due to defects in their mammary myoepithelial cells (Li et al., 2006). Mammary myoepithelial cells surround the epithelial layer of milk-producing cells, providing structural support and contractility required for lactation (Gudjonsson et al., 2005). In contrast to SM cells, which are derived from mesodermal precursors and neural crest cells, myoepithelial cells of the mammary gland are derived from the ectoderm. Myoepithelial cells, however, possess several SM structural proteins and contractile ability.

Apart from their role in development, myocardin and MRTFs are involved in promoting the rearrangement of the actin cytoskeleton in response to numerous physiological and pathological stimuli. For their

6. Regulatory   ncRNAs   in   Muscle   Development

role in an interaction with the environment, readers are asked to refer to Olson and Nordheim (2010) and Parmacek (2010).

Recent transcriptome analyses have revealed that eukaryotic genomes transcribe up to 90 % of genomic DNA, but only 1–2% of these transcripts encode for proteins with the vast majority being transcribed as protein ncRNAs (Kaikkonen et al., 2011). ncRNAs can be divided into structural ncRNAs and regulatory ncRNAs. Constitutively expressed structural ncRNAs include rRNA, tRNA, snRNA, and snoRNA. Regulatory ncRNAs can be classified into miRNAs, Piwi-interacting RNAs, small interfering RNAs, and long ncRNAs. There is increasing evidence for the regulatory roles of some ncRNAs during development (Amaral and Mattick, 2008; Dinger et al., 2008). A novel class of promoter-associated RNAs and enhancer RNAs has been described recently (De Santa et al., 2010; Kaikkonen et al., 2011; rom and Shiekhattar, 2011; rom et al., 2010; Ponting et al., 2009), and endogenous small interfering RNAs and a Piwi-interacting RNA were also recently uncovered in mice (Babiarz et al., 2008; Tam et al., 2008; Watanabe et al., 2008a,b; Wu et al., 2010). Examples of muscle-specific long ncRNA are limited at present with antisense b-MyHC transcript (Haddad et al., 2003, 2008) and linc-MD1 (Cesana et al., 2011). Due to the emerging nature of regulatory ncRNAs, we limit our scope here only to relatively well-described miRNAs.

6.1. Developmental roles of miRNAs

miRNA was first discovered in the nematode Caenorhabditis elegans as a gene lin-4, that controlled the timing of C. elegans larval development (Lee et al., 1993), subsequently in flies for the control of cell proliferation, death, and fat metabolism (Brennecke et al., 2003), and in plants for the control of leaf and flower development (Aukerman and Sakai, 2003). microRNAs are small single-stranded ncRNA that are processed from long primiRNAs transcribed by RNA polymerase II or from introns of spliced mRNAs. These precursors are cleaved inside the nucleus into shorter pre-miRNAs by Drosha protein complexes, and then bound by a specific nuclear

transporter, Exportin-5, to diffuse through nuclear pores into the cytoplasm. In the cytoplasm, they are cleaved by Dicer, a double-stranded RNA-specific RNase III endonuclease, to generate 21–22 base doublestranded miRNAs. One of the two strands complexes with proteins called Argonautes and forms a RNA-induced silencing complex (RISC). Mature miRNA–RISC complexes base pair complimentary with the target mRNAs, and serve as the posttranscriptional level of spatial and temporal regulations by altering the amount of regulator proteins for development and homeostasis (Guo et al., 2010).

The first miRNAs from Drosophila melanogaster and from vertebrate species were reported by Lagos-Quintana et al. (2001), and subsequent cloning out of total RNA from 18.5-week-old C57BL.5 mice gave a number of miRNAs, which were conserved in humans and other vertebrates (Lagos-Quintana et al., 2002). Some clues about the function of miRNAs in embryo development come from experiments that disrupt the formation of cell miRNAs by inactivating Dicer, an enzyme essential for miRNA production. The loss of Dicer leads to lethality early in development, with Dicer-null embryos dying at E7.5 (Bernstein et al., 2003). In muscle development, the specific requirement for miRNAs was demonstrated by tissue-specific knockout mice of the Dicer gene in skeletal muscle (O'Rourke et al., 2007), and in myocardial (Chen et al., 2008) and vascular lineages (Albinsson et al., 2010). The multiplicity of miRNAs and the target for each miRNA appear to promote complex combinatorial regulation with redundancy (Ambros, 2010; Lewis et al., 2005; Matkovich et al., 2011). A number of miRNA-based screenings have been completed, and are compiled in Serva et al. (2011). It has become clear that miRNAs play diverse roles in fundamental biological processes, such as cell proliferation, differentiation, apoptosis, stress response, and tumorigenesis. Numerous miRNAs were shown to inhibit, while some stimulated cell proliferation.

6.2. miRNA in muscle development

To define the contribution made by miRNAs for maintaining cardiomyocytes, the adult heart RNA library was used to obtain a profile of miRNAs, which revealed that miR1 was the predominantly expressed miRNA comprising 45% of all mouse miRNAs found in heart (Lagos-Quintana et al., 2002). Also, more than 90% of all cardiac miRNAs in the adult heart consisted of the 18 most abundant miRNAs in the heart (Rao et al., 2009), but few of these miRNAs appeared to be tissue specific. miR-1 and miR-133a were found to be expressed specifically in cardiac and skeletal muscles during development and in the adult human. They arise from two precursors, coded in two highly conserved chromosomal loci.

miR-1–1 and miR-133a-2 are closely linked on mouse chromosome 2 or human chromosome 20. miR-1–2 and miR-133a-1 are closely linked on mouse chromosome 18 or human chromosome 18, and are located in the intron between exons 12 and 13 of the Dip1 gene, or DAPK (death associated protein kinase)-interacting protein (Rao et al., 2006). The miR-133 family contains a third member, miR-133b, which is also closely linked to miR-206 on mouse chromosome 1 or human chromosome 6, and transcribed together as bicistronic transcripts (Liu et al., 2007). MiR-206 is unique in that it is only expressed in skeletal muscle (reviewed in McCarthy, 2008).

During early embryonic stages, miR-1 and miR-133 function in concert to promote mesoderm induction, while suppressing differentiation into the ectodermal or endodermal lineages (Ivey et al., 2008). However, miR-1 and miR-133 have antagonistic effects for subsequent differentiation, with miR-1 promoting differentiation toward a cardiac fate, whereas miR-133 inhibits differentiation into cardiac muscle (Ivey et al., 2008). The cardiac transcription of miR-1/miR-133 bicistronic miRNAs is directly regulated by Mef2 and SRF (Liu et al., 2007; Zhao et al., 2005). Both miR-1 and miR-133 are coexpressed in cardiac and skeletal muscle throughout mouse development, and Hand2 was found as the target of miR-1 (Zhao et al., 2005).

In skeletal muscle, Dicer conditional alleles were generated to inactivate Dicer in specific tissues (O'Rourke, et al., 2007), which showed that Dicer activity is essential for skeletal muscle development during both embryogenesis and postnatal life. Dicer inactivation in skeletal muscle results in lower levels of muscle-specific miRNAs and Dicer muscle mutants die perinatally. Dicer skeletal muscle mutants are observed to have reduced skeletal muscle in the presence of increased levels of MRFs. MyoD, myogenin, and Mef2, promote miR-1, miR-133, and miR-206 expression by binding to upstream cis elements (Rao et al., 2006). It is possible that MRFs are upregulated to compensate for reduced miRNA levels in Dicer mutants, but this increased level of MRFs fails to promote myogenesis in Dicer mutant skeletal muscle. Another possibility is that myogenic inhibitory factors such as MyoR and Id family members are upregulated in Dicer mutants and consequently antagonize MRFs (Benezra et al., 1990; Lu et al., 1999). These proteins function by inhibiting binding of MRFs to E boxes in the promoters of muscle-specific genes. Indeed, overexpression of miR-206 in cultured myoblasts resulted in decreased expression of MyoR, Id1, Id2, and Id3 (Kim et al., 2006). In the wild type, muscle miRNAs may promote myogenesis by downregulating expression of inhibitory factors (O'Rourke et al., 2007). In skeletal lineage C2C12 cells, overexpression of miR-1 led to the downregulation of endogenous HDAC4 protein, whereas overexpression of miR-133 repressed expression of endogenous SRF proteins. By contrast,

mRNA levels of SRF and HDAC4 were not altered by these miRNAs, supporting the notion that miRNAs repress the function of their target genes mainly by inhibiting translation (Chen et al., 2006).

The sequences of miR-1 and miR-133 are quite divergent (sharing only four nucleotides). Williams et al. (2009) described a bioinformatic program for miRNA target prediction, TargetScan (Lewis et al., 2005), which gave 480 total predicted target genes for miR-1 and 351 total predicted target genes for miR-133, with only 38 target genes predicted to be shared between miR-1 and miR-133, indicating that these two miRNAs have divergent functions with few overlaps. Homozygous deletion of one miR-1 gene, miR-1-2, in mice resulted in a mortality of up to 50% of offspring in late embryonic-postnatal stages with ventricular septal defects (Zhao et al., 2007). Mice lacking either miR-133a-1 or miR-133a-2 were normal, whereas deletion of both miRNAs caused lethal defects in approximately half of the double-mutant embryos or neonates (Liu et al., 2008a,b). miR133a double-mutant mice that survived to adulthood ultimately died of heart failure or sudden death. Deletion of both miR-133a genes resulted in the aberrant expression of SM genes in the heart, and these abnormalities can be attributed, at least in part, to elevated expression of SRF and cyclin D2, which are targets for repression by miR-133a (Liu et al., 2008a,b).

Three miRNAs are known to be encoded within MyHC genes and they are: miR-208a encoded in an intron 27 of the mouse a-MyHC (Myh6) gene, miR-208b encoded in intron 31 of the mouse b-MyHC (Myh7) gene, and miR-499 encoded in intron 19 of the mouse b-MyHC like Myh14 (or

Table 4.6 Muscle-specific microRNAs with linkage to another microRNA or myosin heavy chain gene

MicroRNA	Expressed muscle type	Mouse chromosome	Human chromosome
miR-1-1– miR133a-2	Cardiac, skeletal	Chr2 (9.3 kbp)	Chr20q13.33 (10.5 kbp)
miR-1-2– miR133a-1	Cardiac, skeletal	Chr18 (2.5 kbp)	Ch18q11.1 (3.2 kbp)
miR-206– miR133 b	Skeletal	Chr 1 (3.6 kbp)	Ch 6p12.2 (4.4 kbp)
miR-208a: Myh6 intron	Cardiac	Chr14	Ch14q11.2
miR-208b: Myh7 intron	Cardiac, slow skeletal	Chr14	Ch14q11.2
miR-499: Myh14 intron	Cardiac, slow skeletal	Chr 2	Ch20q11.22

| miR-143–miR145 | Smooth | Chr18 (1.3 kbp) | Ch 5q32 (1.6 kbp) |

Linked microRNAs are shown with dash with the distance between two microRNA indicated in parenthesis. MicroRNAs in Myosin heavy chain introns are shown with: References: Rao et al. (2006), van Rooij et al. (2008), Cordes et al. (2009), and Rangrez et al. (2011).

Myh7b) gene (van Rooij et al., 2008). a-MyHC is the major contractile myosin in the adult murine heart, while it is little represented in the human heart. b-MyHC is highly abundant in the human heart, but is expressed in the embryonic and failing heart in mice, and is the major contractile protein in type I slow skeletal muscle. b-MyHC like Myh14 (or Myh7b) was cloned in humans (Desjardins et al., 2002) and its sequence was analyzed phylogenically (McGuigan et al., 2004), but is still a little studied protein. miR-208 is important in regulating cardiac gene expression in response to stress (Callis et al., 2009; van Rooij et al., 2007). miR-208b is coexpressed with b- MyHC, showing highest expression in slow type I myofibers (Table 4.6).

miR-143 and miR-145 are the most highly enriched miRNAs in vascular SM. These two miRNAs are cotranscribed from a single locus under the transcriptional control of SRF, myocardin, and Nkx2–5 (Cordes, et al., 2009). miR-143 and miR-145 target a network of transcription factors, including Klf4, Klf5, myocardin, MRTFB, and Elk-1, which is consistent with a role for these miRNAs in regulating the quiescent versus proliferative phenotypes of SM (Cordes, et al., 2009; Rangrez et al., 2011).

Numerous miRNAs are implicated in muscle disease, and readers are asked to refer to excellent reviews (for cardiac disease: Thum et al., 2008 ; Small et al., 2010; Small and Olson, 2011. For skeletal disease: Eisenberg et al., 2009; Gu¨ller and Russell, 2010. For SM disease; Alexander and Owens, 2012) (Table 4.6).

7. Concluding   Remarks

GATA4, 5, and 6 have emerged as key factors in heart development across species. Although much has recently been clarified on the general molecular mechanisms of their roles, a full perspective of the regulatory mechanisms of GATA factors remains largely unknown because of their redundancy and a huge variety of interactions with other factors. Together, an increased knowledge of the downstream targets and interacting cofactors of cardiac-expressed GATA factors is likely to reveal more on their requirements and the mechanisms of their critical roles for cardiac development.

Hand genes may have a common function and, at least a partial, genetic redundancy during development. They are required for ventricular

development especially in the ballooning step. The complementary expression pattern of Hand genes during ventricular development may be necessary for the proper regionalization and formation of the right and left ventricles. The function of Hand2 in the SHF is crucial for the development of the right ventricle and the outflow tract. Hand function may be important for developing ventricles in humans and further investigation is necessary to elucidate the underlying epigenetic cause affecting Hand proteins that might be implicated in congenital heart diseases.

The developmental biology of myogenesis has progressed greatly over the last few decades. We can now include the muscle as one of the wellunderstood tissues at the molecular, physiologic, and anatomic levels. Recombinant technology in the past three decades has been instrumental in revealing the factors governing each muscular type development. Complex genetic networks for transcriptional regulators and signaling pathways have been documented for various types of myogenesis, namely the development of cardiac, skeletal, and SM as reviewed above.

Genome-wide studies are beginning to reveal many potentially functional elements such as CarG boxes, and uncover (protein) ncRNAs that are often derived from genome regions other than protein coding sequences and their regulatory regions. Among these RNAs, miRNAs are gaining attention, with others waiting to be uncovered. We now know that the genome regions referred to as "Junk" in the last century could house yet undiscovered biological functions. Consequently, it is appropriate to carefully recheck knockout constructs used in the past for inadvertent alterations of functional elements that may have affected the interpretation of experimental results. The developmental and regulatory history of myogenesis raises many questions in terms of lineage, cell behavior, and gene networks. For apparently confusing observations on the interactions between signaling pathways and effects on proliferation or differentiation, quantitative data are very important in determining the impact of signaling pathways or transcriptional regulators. For example, the level of Wnt signaling in skeletal myogenesis is probably not the same as the level that promotes myocardial differentiation. Moving from a qualitative to a quantitative level of description and extending this to the cellular level is a major challenge for developmental biology in general.

Toward achieving this goal, gene expression analyses with transcription profiling (transcriptome) and microarrays have been conducted. These mRNA analyses, however, have limitations because several investigations revealed that protein expression levels correlated poorly with mRNA levels due to multiple regulatory factors including mRNA stability, rate of translation, and protein degradation (Gygi et al., 1999; Unwin and Whetton, 2006). Furthermore, microarrays cannot cover

the entire transcriptome, especially of unknown products. Protein profiling (proteomics), on the other hand, could cover a wide range of expressed proteins and also has the potential to provide information on posttranscriptional modifications and subcellular localizations. Over the past few years, muscle proteomics has successfully cataloged abundant and soluble proteins (Chen et al., 2010 ; McGregot and Dunn, 2006; Ohlendieck, 2010, 2011). Although present proteomic technology requires further technical refinements to answer mounting questions, its ability to catalog entire protein complements comprehensively is opening an exciting new horizon of research (Doran et al., 2007; Guevel et al., 2011). We recently conducted a comparative proteomic analysis of a cardiosphere cell clone in two different reversible morphological statuses and observed a drastic protein profile shift with many unexpected findings (Machida et al., 2011).

Advances in human genetic tools have also increased the understanding of the importance of developmental pathways in myogenesis implicated in human diseases. A deeper understanding of the interactions between various signaling and transcriptional networks, and their ultimate downstream targets, will be necessary to identify potential approaches in parents at risk. Regarding disease-related developmental biology, exciting future technologies may generate disease-specific embryonic stem cell lines or induced pluripotent stem cells for mechanistic studies of disease etiology and the development of patient-specific stem cells as therapeutics. Although issues such as stem cell expansion, delivery, incorporation, electrical coupling, and safety remain to be addressed, it may now become possible to guide stem or progenitor cells into a muscle lineage, based on our knowledge of early developmental pathways, in preparation for future regenerative medicine.

ACKNOWLEDGMENTS

REFERENCES

Abu-Issa, R., Kirby, M.L., 2007. Heart field: from mesoderm to heart tube. Annu. Rev. Cell. Dev. Biol. 23, 45–68.

Abu-Issa, R., Smyth, G., Smoak, I., Yamamura, K., Meyers, E.N., 2002. Fgf8 is required for pharyngeal arch and cardiovascular development in the mouse. Development 129 , 4613–4625.

Ai, D., Fu, X., Wang, J., Lu, M.F., Chen, L., Baldini, A., et al., 2007. Canonical Wnt signaling functions in second heart field to promote right ventricular growth. Proc. Natl. Acad. Sci. USA 104, 9319–9324.

Albinsson, S., Suarez, Y., Skoura, A., Offermanns, S., Miano, J.M., Sessa, W.C., 2010. MicroRNAs are necessary for vascular smooth muscle growth, differentiation, and function. Arterioscler. Thromb. Vasc. Biol. 30, 1118–1126.

Alexander, M.R., Owens, G.K., 2012. Epigenetic control of smooth muscle cell differentiation and phenotypic switching in vascular development and disease. Annu. Rev. Physiol. 74, 8.1–8.28.

Alsan, B.H., Schultheiss, T.M., 2002. Regulation of avian cardiogenesis by Fgf8 signaling. Development 129, 1935–1943.

Amaral, P.P., Mattick, J.S., 2008. Noncoding RNA in development. Mamm. Genome 19 , 454–492.

Ambros, V., 2010. MicroRNA: genetically sensitized worms reveal new secrets. Curr. Biol. 20, R598–R600.

Araki, T., Chan, G., Newbigging, S., Morikawa, L., Bronson, R., Neel, B., 2009. Noonan syndrome cardiac defects are caused by PTPN11 acting in endocardium to enhance endocardial-mesenchymal transformation. Proc. Natl. Acad. Sci. USA 106, 4736–4741.

Aravind, L., Koonin, E.V., 2000. SAP-a putative DNA-binding motif involved in chromosomal organization. Trends Biochem. Sci. 25, 112–114.

Arsenian, S., Weinhold, B., Oelgeschla¨ger, M., Ru¨ther, U., Nordheim, A., 1998. Serum response factor is essential for mesoderm formation during mouse embryogenesis. EMBO J. 17, 6289–6299.

Aukerman, M.J., Sakai, H., 2003. Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. Plant Cell 15, 2730–2741.

Aziz, A., Liu, Q.-C., Dilworth, F.J., 2010. Regulating a master regulator: establishing tissuespecific gene expression in skeletal muscle. Epigenetics 5, 691–695.

Babiarz, J.E., Ruby, J.G., Wang, Y., Bartel, D.P., Blelloch, R., 2008. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. Genes Dev. 22, 2773–2785.

Bajard, L., Relaix, F., Lagha, M., Rocancourt, D., Daubas, P., 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.

Bamshad, M., Lin, R.C., Law, D.J., Watkins, W.C., Krakowiak, P.A., Moore, M.E., et al., 1997. Mutations in human TBX3 alter limb, apocrine and genital development in ulnarmammary syndrome. Nat. Genet. 16, 311–315.

Beauchamp, J.R., Heslop, L., Yu, D.S.W., Tajbakhsh, T., Kelly, R.G., Wernig, A., et al., 2000. Expression of CD34 and Myf5 defines the majority if quiescent adult skeletal muscle satellite cells. J. Cell Biol. 151, 1221–1233.

Belaguli, N.S., Schildmeyer, L.A., Schwartz, R.J., 1997. Organization and myogenic restricted expression of the murine serum response factor gene: a role for autoregulation. J. Biol. Chem. 272, 18222–18231.