The regulation of myogenin gene expression during the embryonic development of the mouse

Siu-Pok Yee and Peter W.J. Rigby

Laboratory of Eukaryotic Molecular Genetics, Medical Research Council, National Institute for Medical Research, The Ridgeway, Mill Hill, London, NW7 1AA UK

The myogenic program can be activated in cultured cells by each of four basic–helix–loop–helix (bHLH) transcription factors, the expression of which is strictly controlled, both temporally and spatially, during embryonic development. To begin to understand the mechanisms by which these regulators are regulated themselves, we have used transgenic animals to define the minimal sequences required for the complete recapitulation of the temporal and spatial expression pattern of the myogenin gene during embryogenesis. We show that this can be achieved with only 133 bp of 5′-flanking DNA and identify two essential motifs, which are consensus binding sites for the bHLH proteins and for the proteins of the RSRF family. We show further that these sequences, when juxtaposed to a heterologous promoter, are capable of imposing the myogenin expression pattern. We conclude that the proper regulation of myogenin requires a bHLH protein, most probably Myf-5, the only myogenic bHLH factor known to be present in the embryo at the time that myogenin is activated, and an RSRF-like binding activity. Furthermore, the expression pattern of a mutant myogenin promoter lacking the RSRF site reveals the existence of at least two populations of cells within the myotomes and of novel rostrocaudal gradients of expression.

[*Key Words:* Myogenin gene; embryogenesis; bHLH transcription factor]

Received February 23, 1993; revised version accepted April 5, 1993.

To understand the mechanisms by which the skeletal muscles of the vertebrate embryo are produced it will be necessary to elucidate both the processes by which the cells in newly born somites first become committed to the myogenic lineage and the ways in which the expression of the genes that encode the proteins of the contractile apparatus is subsequently activated.

The analysis of both of these issues has been greatly advanced by the discovery of four muscle regulatory genes, *myf-5* (Braun et al. 1989b), myogenin (Braun et al. 1989a; Edmondson and Olson 1989; Wright et al. 1989), herculin (also called *myf-6* and *MRF-4*; Rhodes and Konieczny 1989; Braun et al. 1990; Miner and Wold 1990), and *myoD1* (Davis et al. 1987), each of which encodes a transcription factor of the basic–helix–loop–helix (bHLH) family, which has the remarkable property of being able to activate the myogenic program in a variety of cultured cell types. Each of these genes is expressed only in skeletal muscle, and their products bind—most efficiently as heterodimers with the products of the ubiquitously expressed E2 genes—to sequences containing the core motif CANNTG (the E-box), which are found in the promoters and/or enhancers of many muscle-specific genes (Murre et al. 1989; Blackwell and Weintraub 1990; Lassar et al. 1991). Transfection experiments have shown that, in certain types of cultured cells, each of these transcription factors can *trans*-activate the expression of genes encoding muscle structural components (for review, see Olson 1990) and, moreover, that they can autoactivate themselves and, in some cases, cross-activate each other (Braun et al. 1989a,b; Edmondson and Olson 1989; Rhodes and Konieczny 1989; Thayer et al. 1989; Miner and Wold 1990).

The expression patterns of these regulatory genes during embryogenesis are complex. The *myf-5* gene is first activated at 8.0 days postcoitum (dpc) in the dorsomedial lips of the dermamyotomes of the somites (Ott et al. 1991), which contain the precursors of the trunk muscles. Half a day later, transcription of the myogenin gene is turned on in the now discrete myotomes (Sassoon et al. 1989), followed by herculin (9 dpc; Bober et al. 1991; Hinterberger et al. 1991) and *myoD1* (10.5 dpc; Sassoon et al. 1989). The expression of herculin and *myf-5* is transient, being turned off at 11 dpc and 12 dpc, respectively, although herculin is reactivated at 15.5 dpc. Herculin, myogenin, and *myoD1* transcripts are present throughout the remainder of embryogenesis and in the adult. All four of these genes are also expressed in the limb buds, although the absolute and relative timings are quite different. In the fore limb buds, *myf-5* and herculin are activated at 10.5 dpc, followed by both myogenin and *myoD1* at 11 dpc (Hannon et al. 1992). In the hind limb bud expression is delayed by about a day.

The mechanisms by which the regulatory genes are

regulated themselves during embryogenesis remain unclear because the cell culture systems that have been employed as models in which to study this issue do not accurately represent the situation in the embryo. The various myogenic cell lines express different subsets of the factors (for review, see Funk et al. 1991); most strikingly, some proliferating myoblasts express *myoD1* but not myogenin, a situation not known to occur during development. More importantly, the expression of one of the bHLH factors in cultured cells leads rapidly to the *trans*-activation of one or more of the others. Because each of the factors has a discrete pattern of expression in the embryo and these patterns are different in the trunk and the limbs, it is clear that simple *trans*-activation circuits are not allowed during development. Moreover, whereas *myf-5* alone can trigger the entire myogenic program in cultured cells, this does not appear to be the case in the embryo, as the *myf-5* gene is transcribed from 8.0 to 10.5 dpc yet many responsive, muscle-specific genes remain silent. If one assumes that the Myf-5 protein is present during this period, which has not been demonstrated experimentally, it follows that the regulatory mechanisms that allow *myf-5* to act in cultured cells are not operative in the embryo.

To overcome this difficulty, we have analyzed the promoter of the mouse myogenin gene in transgenic mice. Our strategy has been to construct animals in which the expression of the *lacZ* reporter gene is driven by myogenin regulatory elements and mutated derivatives thereof. We have thus identified the minimal sequences required for the complete recapitulation of the temporal and spatial program of myogenin gene expression during development. Mutational analyses show that myogenin transcription depends on two motifs, one a consensus binding site for bHLH proteins, and the other a consensus binding site for RSRF (related to serum response factor) proteins, and that these two motifs are likely to be sufficient. Moreover, mutations in the RSRF site reveal hitherto unsuspected heterogeneities within the myotomes.

Results

Definition of sequences required for myogenin expression

We isolated a mouse myogenin genomic clone, which contains a 30-kb insert, including 27-kb of 5′-flanking sequence, by screening a cosmid library with a mouse myogenin cDNA probe. Hybridization with a fragment from the 5′ end of the cDNA identified a 5.6-kb *Eco*RV fragment in the cosmid that was subcloned for further analysis. The subclone contains 4.25 kb of 5′-flanking DNA, together with the first intron of the myogenin gene, and was used to construct the first fusion gene (GZ4250; see Fig. 1) used for transgenic analysis, in which the *lacZ* reporter was fused in-frame with the translational initiation codon of the myogenin gene.

F_0 conceptuses were harvested at 10.5 dpc, a time at which expression of the endogenous myogenin gene is well established in the somites; the embryos were stained for β-galactosidase activity, whereas DNA was extracted from placentae and analyzed by the polymerase chain reaction (PCR) for the presence of the transgene. Twelve transgenic embryos were obtained, of which eight showed somite-specific staining for β-galactosidase (Table 1) together with (in some cases) low levels of ectopic expression. A representative whole-mount stained embryo is shown in Figure 2A; Figure 2B shows a transverse section of the same embryo demonstrating

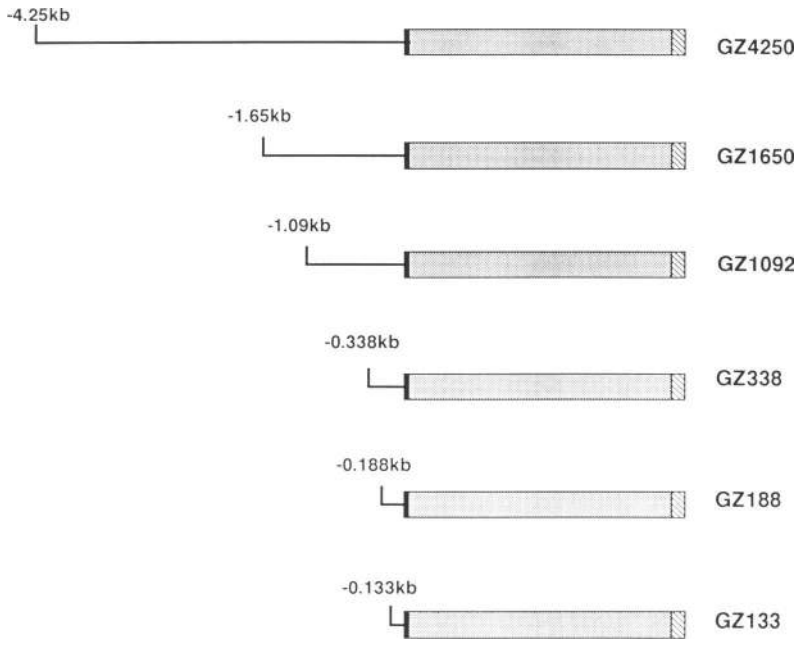

Figure 1. Structures of myogenin–*lacZ* fusion genes. The fusion genes consist of myogenin 5′-flanking sequence fused in-frame at the myogenin-initiating ATG with the *lacZ* reporter gene (see Materials and methods). The fusion genes were excised from the plasmids for microinjection so that no vector sequences were injected. The black lines represent myogenin sequences 5′ of the transcription start, solid boxes represent myogenin 5′-untranslated exon sequences, stippled boxes represent the coding region of the *lacZ* gene, and hatched boxes represent the SV40 polyadenylation signal.

Table 1. *Transgenic analyses of myogenin–*lacZ *fusion constructs*

Construct	Myogenin sequence	Age of embryo (dpc)	Number of blue transgenics[a]	Pattern of expression	Level of expression[b]
GZ4250	4.25 kb	10.5	8/12	myotome	+ + + + +
GZ1650	1.65 kb	10.5	8/8	myotome	+ + + + +
GZ1092	1.092 kb	10.5	6/7	myotome	+ + + + +
GZ338	340 bp	10.5	7/10	myotome	+ + +
GZ188	180 bp	10.5	4/6	myotome	+ +
GZ133	133 bp	10.5	4/6	myotome	+
		12.5	5/10	myotome, limb bud, and musculature	+
		13.5	2/4	myotome, limb bud, and musculature	+ + +
UGHZ[c]	−1092 to −134 bp	10.5	3/7	random	+ + + + +
GMEZ1092[d]	1.092 kb	10.5	11/15	myotome	+ + + + +
		11.5	2/3	myotome, limb bud	+ + + + +
		12.5	4/6	myotome, limb bud, and musculature	+ + + + +
GMEZ133[d]	133 bp	10.5	0/6	none	−
GMEZ188[d]	180 bp	10.5	2/6	myotome	+ +
EGMEZ133[e]	133 bp	10.5	2/4	myotome	+
		12.5	2/2	myotome, limb bud, and musculature	+ +
		13.5	2/5	myotome, limb bud, and musculature	+ + +
GMRZ1092[f]	1.092 kb	10.5	14/24	myotome	+ + +
		11.5	7/10	myotome, limb bud	+ + +
		12.5	6/9	myotome, limb bud, and musculature	+ + +
		13.5	5/7	myotome, limb bud, and musculature	+ + +
RSEGZ[g]	−192 to −55 bp	10.5	7/7	myotome	+ +
		12.5	2/2	myotome, limb bud, and musculature	+ +
BGZ[h]	−	10.5	1/1	random	+

[a]Proportion of *lacZ*-expressing embryos as a fraction of founder embryos containing the transgene as assayed by PCR.
[b]The level of *lacZ* expression for the GZ1092 lines was set as + + + + +. The relative level of expression for the other embryos was determined visually by comparison to equivalent stages of GZ1092.
[c]Myogenin promoter sequence from −1.09 to −0.13 kb was placed 5′ upstream of a 300-bp heat shock promoter–*lacZ* fusion gene.
[d]The myogenin promoter proximal E-box CAGTTG (−17 to −12) was mutated to TGGTCA.
[e]Myogenin promoter sequence GCAGTTG (−18 to −12) was placed 5′ to the 133-bp myogenin promoter with the proximal E-box CAGTTG (−17 to −12) mutated to TGGTCA.
[f]The myogenin promoter AT-rich sequence CTATATTTAT (−69 to −60) was mutated to CTTTATTAAT.
[g]Myogenin promoter from −192 to −55 containing an E-box and the AT-rich sequence was placed 5′ upstream of a 37-bp human β-globin promoter–*lacZ* fusion gene.
[h]*lacZ* reporter gene under the transcriptional control of human β-globin sequences from −37 to +12.

that expression of the reporter gene was confined to the myotomal cells of the somites. The low-level ectopic expression was at different sites in different embryos and is presumed to result from effects of the chromosomal site of integration.

We then constructed a series of 5′ deletions derived from GZ4250 (see Fig. 1) and again analyzed their activities in F_0 embryos. Both GZ1650 and GZ1092, which contain 1.65 and 1.09 kb of 5′-flanking DNA, respectively, gave, at high efficiency (Table 1), the same pattern of myotome-specific staining observed with GZ4250. We therefore chose to undertake detailed analyses of the smaller construct, GZ1092.

Sequence analysis of myogenin 5′-flanking DNA

The nucleotide sequence of the 1.09 kb of DNA upstream of the myogenin gene in our cosmid clone is presented in Figure 3. A very similar sequence for this same region of DNA has been reported recently by Edmondson et al. (1992). The 5′-most 250 bp correspond to a LINE-1 repetitive element that is not present in the correspond-

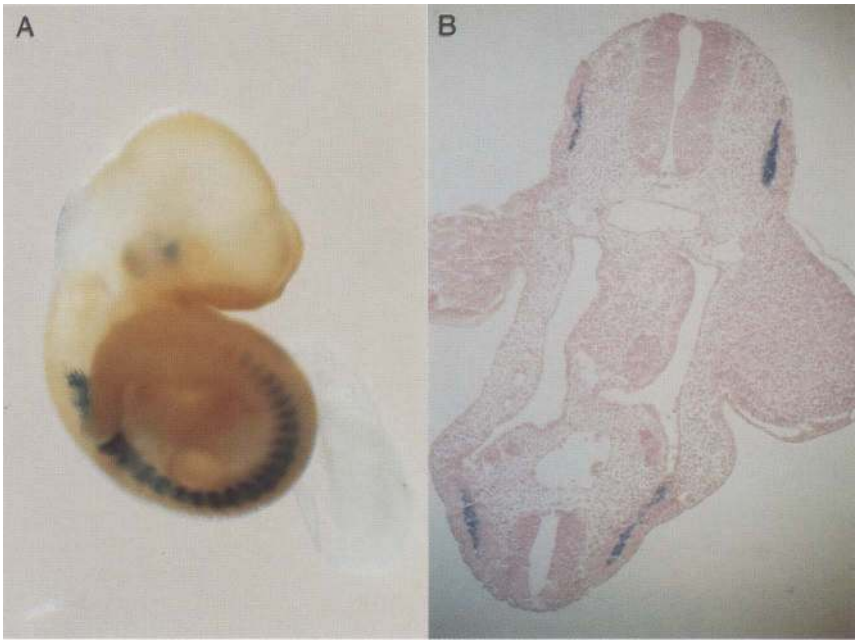

Figure 2. β-Galactosidase expression in a GZ4250 transgenic embryo. (*A*) Lateral view of a whole-mount 10.5-dpc GZ4250 transgenic embryo stained with X-gal. Blue staining is localized in the somites; the more rostral somites are more heavily stained than the more caudal somites. (*B*) Transverse section of the same embryo showing that β-galactosidase activity is restricted to the myotomes of the somites.

ing human sequence (Salminen et al. 1991) and is therefore unlikely to play a role in the regulation of transcription. The murine and human sequences are otherwise highly homologous except for a stretch of 19 CA dinucleotide repeats found in the mouse sequence but not in the human (Fig. 3). The start site for transcription was mapped by both primer extension and RNase protection experiments to a position 48 nucleotides upstream of the translation initiation codon (data not shown). Discounting the LINE element, the 5′-flanking sequence contains six potential E-box (CANNTG) sequences, at positions −770 to −765, −645 to −640, −381 to −376, −317 to −312, −143 to −138 and −17 to −12; the most proximal E-box is located between the TATA box and the start site. The sequence also contains, at position −69 to −60, an AT-rich sequence that represents a consensus binding site for the human transcription factors of the RSRF family of proteins originally isolated by Pollock and Treisman (1991) by virtue of their homology to serum response factor.

Analysis of transgene expression during the development of mice carrying the GZ1092 construct

To analyze in detail the transcriptional activity of the 1.09 kb of myogenin 5′-flanking DNA, we generated three lines of mice carrying the GZ1092 fusion gene. All three lines displayed identical patterns of β-galactosidase activity and similar levels of transgene expression. In Figure 4, we show the developmental profile of transgene expression in one of these lines. β-galactosidase staining was detected first in the newly formed somites of 8.5-dpc embryos; sectioning showed that it was most intense in the myotomal cells of the most rostral, and thus the most mature, somites (Fig. 4A,B). One day later, the levels of transgene expression had risen markedly, and there was a clear intensity gradient along the rostrocaudal axis, with the most intense staining located in the most rostral somites (Fig. 4C). We assume that this gradient simply reflects the relative ages of the somites, the transgene having been expressed for longer in the more mature, rostral somites. It was not apparent in later stage embryos. At 10.5 dpc, the intensity of the staining had increased further in the somites but limb bud expression was not yet apparent (Fig. 4D). In 11.5-dpc embryos, we first detected staining in the limb buds, the intensity being much higher in the forelimb than in the hindlimb buds (Fig. 4E). The section in Figure 4F shows staining in the proximal dorsal muscle mass that will eventually form the triceps and in a more anterior muscle mass that will form the preaxial forearm extensors. One day later, the segmented somitic structures were no longer apparent, and the stained cells had migrated dorsally and ventrally away from the location of the somitic blocks (Fig. 4G). Expression in the hindlimb bud was clearly apparent at this time (12.5 dpc). In 13.5- and 14.5-dpc embryos, β-galactosidase activity was detected in the expected locations in most parts of the fetus including the muscles of the face (Fig. 4H,I). In adult animals, staining was restricted to skeletal muscle; no staining was observed in cardiac muscle nor in any other tissue (data not shown). In all of its aspects, the pattern of β-galactosidase staining was identical to the pattern of myogenin mRNA accumulation determined by in situ hybridization (Sassoon et al. 1989). These data show that the 1.09 kb of 5′-flanking DNA contain all of the regulatory elements necessary to recapitulate faithfully the pattern of myogenin expression during development. Northern analyses of RNA from a line homozygous for the transgene indicated that the amount of *lacZ* mRNA was ~25% of

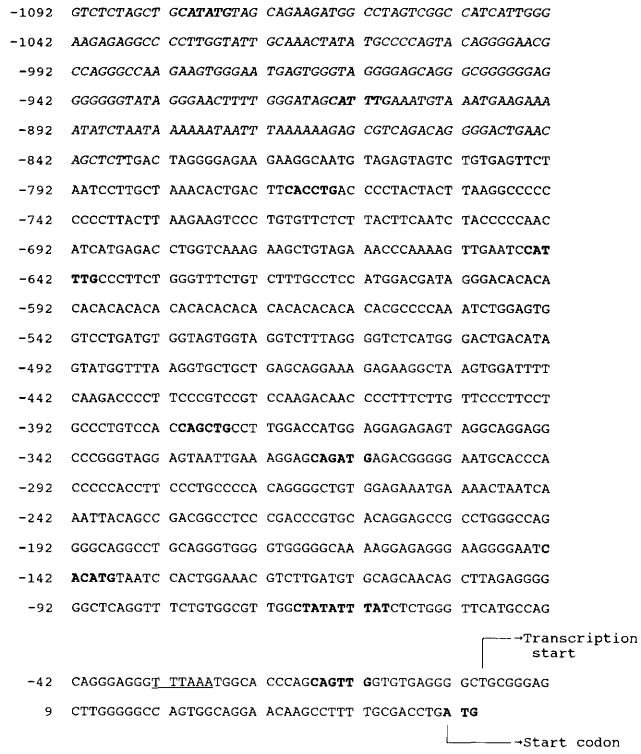

Figure 3. Nucleotide sequence of the mouse myogenin promoter and 5′-untranslated region. The transcriptional start is referred to as +1 and indicated by an arrow. The putative TATA box sequence is underlined, the AT-rich RSRF site and the E-boxes are in boldface type, and the sequence from −1092 to −837, which corresponds to a murine LINE1 element, is italicized.

the amount of mRNA from the endogenous myogenin genes (data not shown). It is therefore possible that there are sequences not contained in the GZ1092 construct that affect the efficiency of myogenin transcription but that are not required for correct spatial and temporal regulation, and/or that structural features of the endogenous myogenin locus are necessary for full efficiency.

Further deletion analysis of myogenin 5′-flanking DNA

To define more precisely the regulatory sequences required for the recapitulation of the myogenin expression pattern, we generated a series of deletions in which the reporter gene was driven by 338 (GZ338), 188 (GZ188), and 133 (GZ133) bp of 5′-flanking DNA. These constructs were analyzed in 10.5-dpc F_0 transgenic embryos (Table 1). In each case, we detected expression of the transgene in the somites, and this was confined to the myotomal cells. The level of expression, however, was markedly reduced at each step of the deletion series, and ectopic expression became more apparent. Nonetheless, it is clear that only 133 bp of 5′-flanking DNA was capable of correctly targeting expression to the myotomes at 10.5 dpc (Fig. 5B), and to the myotomally derived musculature and to the fore- and hindlimb buds at 12.5 and 13.5 dpc (Table 1; Fig. 5C). The fact that the GZ133 construct was expressed at a dramatically lower level than GZ1092 (Fig. 5A) indicated that the sequences between −1092 and −134 exert a marked effect on the efficiency of expression. We asked whether this region has muscle-specific enhancer activity by placing sequences from −1092 to −134 upstream of a heterologous promoter, that of the *hsp68* gene. F_0 transgenic embryos carrying this construct did not show expression in the somites; expression was efficient but random (Table 1). We conclude that these distal sequences contain elements that affect the efficiency of expression but are not sufficient for the proper targeting of myogenin expression.

Mutational analysis of the proximal sequence elements

The 133-bp proximal fragment that is capable of targeting expression to the myotomes contains a single E-box, CAGTTG, located between the TATA box and the transcriptional start site, that is conserved in the human sequence and can bind bHLH factors in C2C12 myotube extracts (data not shown). We therefore introduced mutations into this E-box (CAGTTG → TGGTCA) such that the binding site was destroyed and tested the activity of such mutant constructs in F_0 transgenic embryos. We observed no expression (Table 1), even when embryos were stained for 1 week with several changes of staining solution and therefore conclude that this sequence element is essential. Given its location, it is possible that the mutations simply inactivate the promoter rather than prevent the binding of a regulatory protein required for proper expression. Two lines of evidence show that this is not the case. First, we introduced the E-box mutations into the construct GZ1092 in which expression is driven by 1.09 kb of 5′-flanking sequence. F_0 transgenic embryos carrying the mutated derivative, GMEZ1092, gave an expression pattern very similar to that obtained with the parental construct at 10.5, 11.5, and 12.5 dpc (Table 1; Fig. 5D–F). It did, however, appear that expression in the forelimb bud at 11.5 dpc, although detectable, was lower than with the wild-type construct (Fig. 5E) but that it was normal by 12.5 dpc (Fig. 5F). There was no detectable effect on the activation of expression in the hindlimb bud at 12.5 dpc (Fig. 5F). Second, the GMEZ133 construct was active when cotransfected with myogenin or *myf-5* expression plasmids into 10T½ cells, and it is more active in C2 myotubes than in C2 myoblasts (data not shown). These results demonstrate that mutation of the proximal E-box does not inactivate the promoter and led us to conclude that one E-box is required for proper regulation. We tested this further by first examining the activity of the construct GMEZ188, in which the proximal E-box is mutated but the addition of sequences from −180 to −134 leads to the acquisition of a second E-box, located between −143 and −138. In 10.5-dpc F_0 transgenic embryos, GMEZ188 gave a pattern of expression indistinguishable from that of GZ1092 (Table 1; Fig. 5G). The activity of GMEZ188

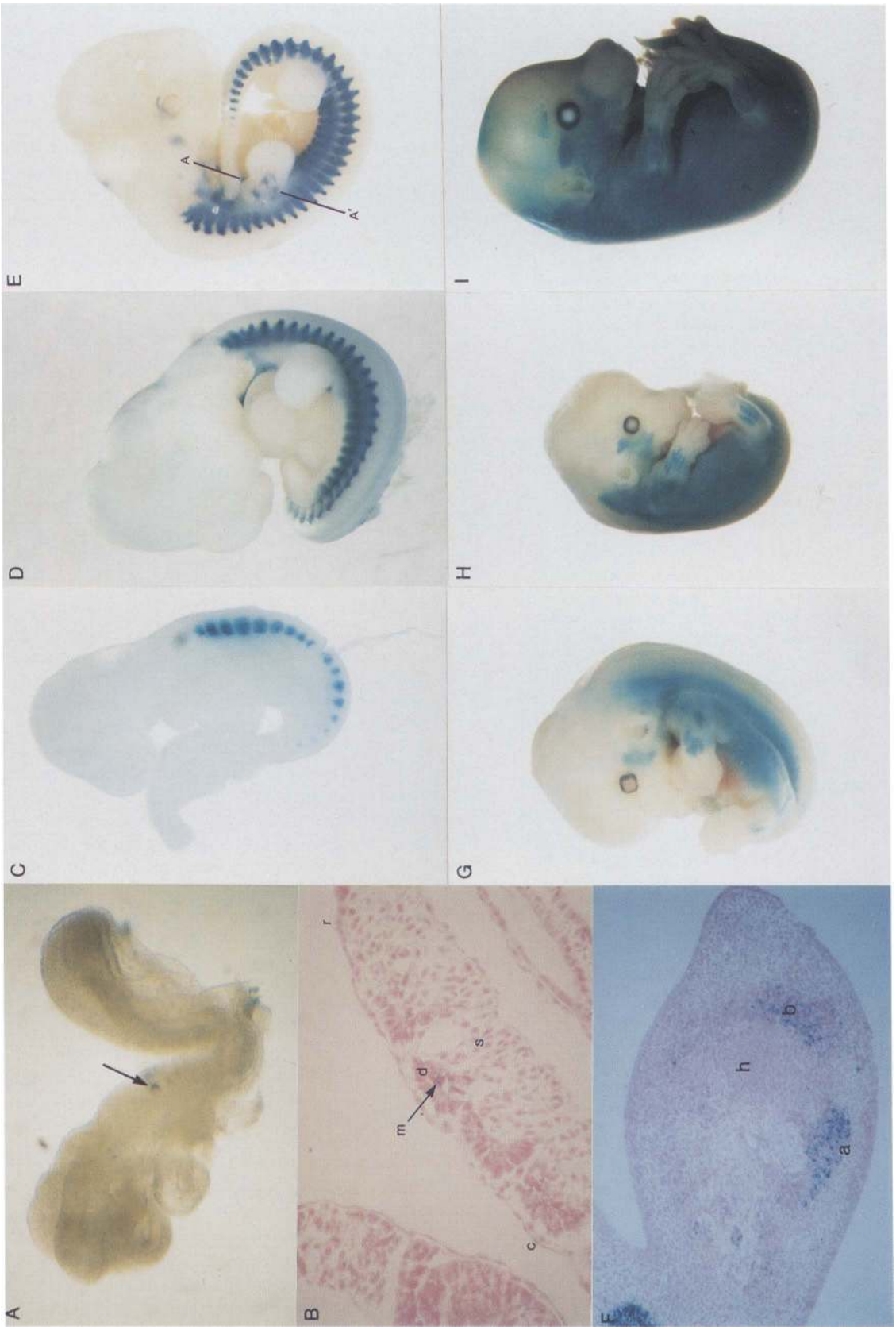

Figure 4. Transgenic embryos from a GZ1092 line stained for β-galactosidase activity at different stages of development. (*A*) An 8.5-dpc whole-mount stained embryo. (*B*) A sagittal section of the same 8.5-dpc embryo. (r) Rostral, (c) caudal, (d) dermatome, (m) myotome, and (s) sclerotome. (*C*) A 9.5-dpc embryo. (*D*) A 10.5-dpc embryo. (*E*) A 11.5-dpc embryo; (A A') The plane of the section shown in *F*. (*F*) A diagonal section of the forelimb bud of a sibling 11.5-dpc embryo. (a) The proximal dorsal muscle mass; (b) the preaxial forearm muscle mass; (h) the humerus. (*G*) A 12.5-dpc embryo. (*H*) A 13.5-dpc embryo. (*I*) A 14.5-dpc embryo.

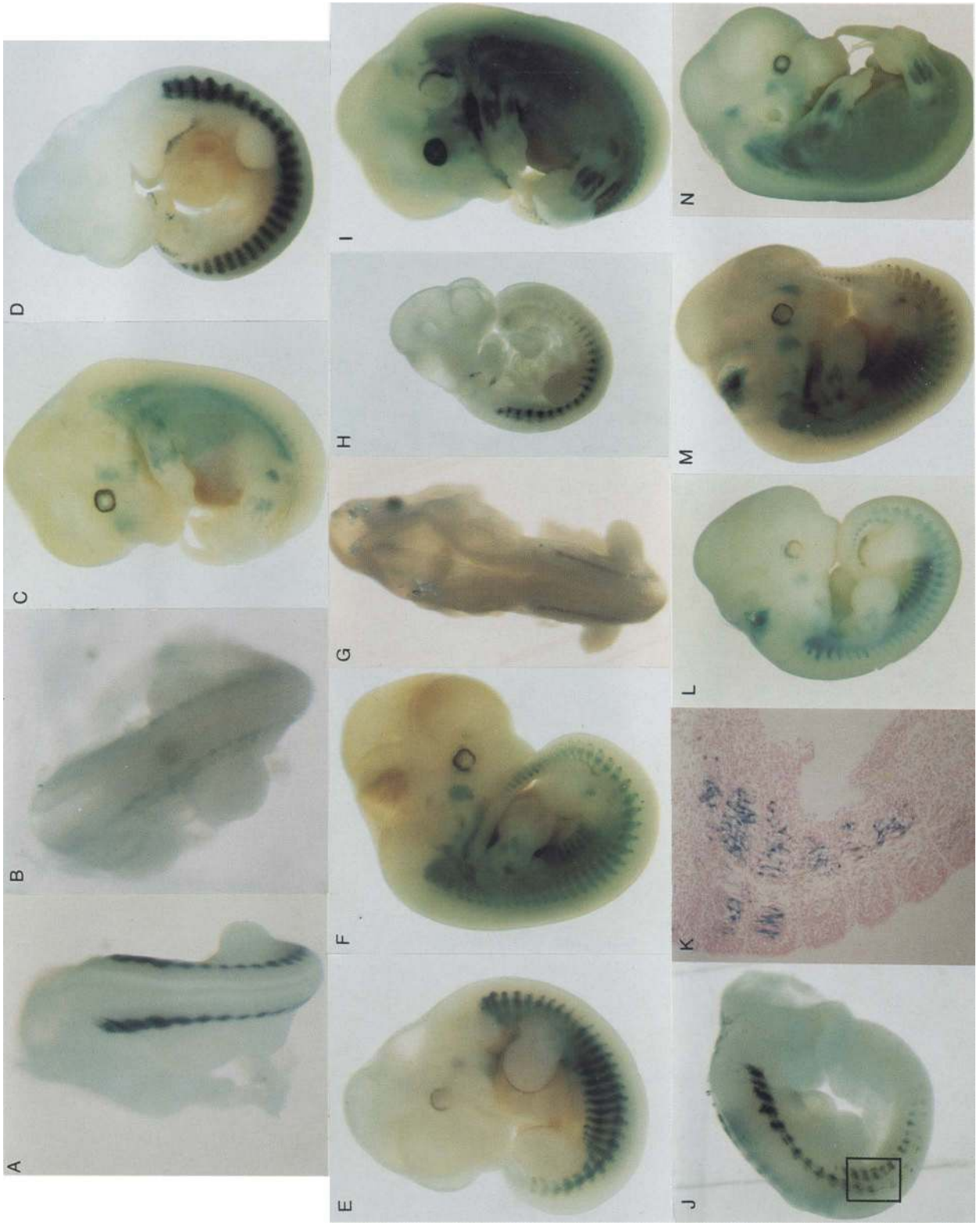

Figure 5. (*See following page for legend.*)

could be attributable to sequences between −180 and −134 other than the E-box. To eliminate this possibility and to unequivocally test the function of the E-box, we constructed EGMEZ133, in which a 7-bp oligonucleotide (5′-GCAGTTG), corresponding to sequences from −18 to −12 of the myogenin promoter, was added to the 5′ end of GMEZ133; EGMEZ133 thus contains a mutated, inactive copy of the proximal E-box in its normal location (−17 to −12) and a wild-type copy of the proximal E-box at −139 to −134. In F_0 embryos harvested at 10.5, 12.5, and 13.5 dpc, EGMEZ133 gave an expression pattern indistinguishable from that of GZ1092 (Table 1; Fig. 5H,I). These data show that a single E-box is required for proper expression in both the trunk and the limbs but that its position is not crucial.

Mutational analysis of the AT-rich motif

About 30 nucleotides upstream of the TATA box, between positions −69 and −60, the mouse myogenin promoter contains an AT-rich sequence, which is conserved in the human sequence, that corresponds to a consensus binding site for the human RSRF family of transcription factors. To investigate the role of this sequence motif in the transcription of the myogenin gene, we introduced into it two mutations, A → T at −67 and T → A at −62, known to prevent RSRF binding (Pollock and Treisman 1991).

The transgenic embryos carrying this mutant construct, GMRZ1092, showed an interesting pattern of *lacZ* expression (Fig. 5J). The most rostral myotomes stained uniformly, as has been observed for all of the other expression constructs. At the level of the forelimb bud, and immediately rostral to it, only the more dorsal cells in each myotome expressed. At the level of the caudal end of the forelimb bud, a dramatic change in pattern occurred. Both the more dorsal and the more ventral cells in each myotome expressed, whereas the cells in the center of each myotome did not. This is shown clearly in the histological section of Figure 5K. In the caudal portion of the embryo, expression was predominantly in the more ventral cells of each myotome (Fig. 5J). There were, moreover, two rostrocaudal gradients of expression, one in the rostral third of the trunk, and the other caudal to the forelimb buds in the ventral halves of the myotomes (Fig. 5J). At 11.5 dpc, this pattern, particularly the predominance of expression in the ventral halves of the more caudal myotomes, was still apparent (Fig. 5L). By 12.5 dpc, however, the expression pattern appeared to be more like that observed with the wild-type construct (Fig. 5M), and in a 13.5-dpc embryo, we could detect no differences from one carrying the wild-type construct (cf. Fig. 5N with Fig. 4H). In both fore- and hindlimb buds, expression of the GMRZ1092 transgene activated at the normal times, 11.5 and 12.5 dpc, respectively (Fig. 5L,M).

An E-box and an RSRF site will impose the myogenin expression pattern

The above experiments demonstrate that both an E-box and a RSRF site are required for the proper expression of the myogenin gene. We then asked whether these sequences are sufficient. To do this we cloned the myogenin sequences from −192 to −55, which contain an E-box and the RSRF site, upstream of the human β-globin TATA box and the *lacZ* reporter gene (Fig. 6A). Although the transgene was not expressed efficiently, the levels were comparable to those obtained with GZ133. More importantly, the pattern of expression was indistinguishable from that of the myogenin–*lacZ* fusion gene, indicating that the cloned sequences were sufficient to impose the correct expression pattern of myogenin (Fig. 6B,C).

Discussion

The data that we have obtained show that only 1.09 kb of 5′-flanking DNA is required to recapitulate efficiently the entire temporal and spatial expression pattern of myogenin during development and to maintain expression in the adult organism. The sequences between −1092 and −134 affect efficiency but they are not capable of imposing muscle-specific expression. The proximal 133 bp are necessary for the proper targeting of expression in the embryo. They contain two sequence motifs known to bind transcription factors expressed in the somites, one for the bHLH proteins and one for the RSRF proteins, and each of these is necessary. Juxtaposition of a fragment containing one of each of these motifs to a heterologous promoter imposes the myogenin expression pattern in both the trunk and the limbs, indicating that they are likely to be sufficient.

The identification of two *cis*-acting sequence motifs

Figure 5. β-Galactosidase activity in myogenin–*lacZ* transgenic embryos. (*A*) Dorsal view of a 10.5-dpc whole-mount stained embryo carrying GZ1092. (*B*) A similar view of a stage-matched embryo carrying GZ133. Expression is less efficient but correctly targeted to the myotomes of the somites. (*C*) Lateral view of a 13.5-dpc whole-mount stained embryo carrying GZ133 showing expression in the trunk musculature and the limbs. (*D–F*) Lateral views of 10.5-, 11.5-, and 12.5-dpc whole-mount stained embryos carrying GMEZ1092. (*G*) Dorsal view of a 10.5-dpc whole-mount stained embryo carrying GMEZ188. (*H,I*) Lateral views of 10.5- and 13.5-dpc whole-mount stained embryos carrying EGMEZ133. The patterns of *lacZ* expression in the transgenic embryos shown in *D–I*, in which the *lacZ* gene is under the control of a myogenin promoter mutated in the proximal E-box, are very similar to those of the stage-matched wild-type controls shown in Fig. 4. (*J*) Lateral view of a 10.5-dpc whole-mount stained embryo carrying GMRZ1092; the box indicates the area of the section shown in *K*, (*K*) Parasaggital section of the embryo in *J* showing expression in the more dorsal and more ventral, but not in the central, cells of the myotomes; (*L–N*) Lateral views of 11.5-, 12.5-, and 13.5-dpc whole-mount stained embryos carrying GMRZ1092.

A

GGGTGGGGTG GGGGCAAAAG GAGAGGGAAG GGGAATCACA TGTAATCCAC TGGAAACGTC
TTGATGTGCA GCAACAGCTT AGAGGGGGGC TCAGGTTTCT GTGGCGTTGG CTATATTTAT

Transcriptional Start

CTCTGGGGCT GGGCATAAAA GTCAGGGCAG AGCCATCTAT TGCTTACATT GCTTCT-LacZ

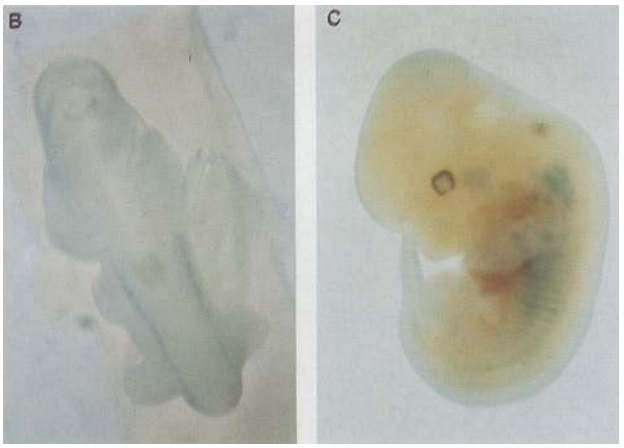

Figure 6. (*A*) Structure of myogenin–human β-globin–*lacZ* fusion gene (RSEGZ). Myogenin 5′ upstream sequences from −192 to −55 were placed upstream of the human β-globin sequence from −37 to +12 containing the TATA box and *lacZ*-coding sequence followed by an SV40 polyadenylation signal. The myogenin sequences are in boldface type, with the E-box underlined and the AT-rich RSRF site double-underlined. The human β-globin TATA box is in italics and the transcriptional start is indicated by an arrow; (*B*) 10.5-dpc and (*C*) 12.5-dpc transgenic embryos stained for β-galactosidase activity showing the *lacZ* expression pattern imposed on a heterologous promoter by the myogenin 5′ upstream sequence from −192 to −55.

required for the proper expression of myogenin leads to clear predictions regarding the nature of the transcription factors that control this muscle regulatory gene. It has to be noted, however, that the necessary and sufficient sequences have not been subjected to saturation mutagenesis, and it therefore remains possible that other, presently unidentified, transcription factors are also required. Moreover, our analyses have focused on temporal and spatial control; other factors may well play a role in achieving correct levels of transcription. The necessity of an E-box implicates a member of the bHLH family, which we presume to be Myf-5, simply on the basis that it is the only regulatory factor of this class known to be expressed in the myogenic precursors before the onset of myogenin transcription. Edmondson et al. (1992) have shown that the proximal E-box is not necessary for muscle-specific expression in cultured cells, a result that we have confirmed with our construct. Such data make it clear that the requirements for correct gene expression during embryogenesis are not recapitulated by cultured cells.

The necessity for the AT-rich sequence implicates the DNA-binding activity MEF-2, which is enriched in myotubes compared with myoblasts and binds to AT-rich sequences in the promoters and/or enhancers of a number of muscle-specific genes (Gossett et al. 1989). Other work has characterized very similar DNA-binding activities, which are ubiquitously distributed, and their relationship to MEF-2 remains controversial (Horlick et al. 1990; Cserjesi and Olson 1991). The DNA-binding specificity of MEF-2 is indistinguishable from that of the RSRF family of proteins, the cDNAs of which were first cloned by Pollock and Treisman (1991) on the basis of sequence homology to serum response factor. There are at least four RSRF genes (Pollock and Treisman 1991, Chambers et al. 1992; Yu et al. 1992), the products of which bind to DNA as dimers. The mobility retardation activity that defines MEF-2 contains one or more members of the RSRF family (Edmondson et al. 1992), and screening of cDNA expression libraries with the MEF-2-binding site (Yu et al. 1992) led to the isolation of cDNAs that correspond to the two RSRF genes defined by Pollock and Treisman (1991). The transcripts of the RSRF genes undergo alternative splicing, and Yu et al. (1992) demonstrate that for one of the genes there is an exon used only in skeletal and cardiac muscle, and brain. MEF-2 is itself induced by the expression of the bHLH factors (Cserjesi and Olson 1991; Lassar et al. 1991, Edmondson et al. 1992). It will be important to determine how many RSRF genes there are in the mouse and to document their expression patterns during development. It is clear, however, that in *Xenopus* embryos, at least two of the genes are expressed specifically in the somites at the stages of development during which the muscle precursors are differentiating (Chambers et al. 1992).

The pattern of β-galactosidase staining activity that we have observed exactly parallels the pattern of myogenin transcript accumulation in the embryo. It is interesting to note that Cusella-De Angelis et al. (1992) have shown that whereas myogenin mRNA first appears at 8.5 dpc, the protein is not detectable until 10.5 dpc, indicating that its appearance is subject to some form of post-transcriptional control. Our fusion genes retain the 5′-untranslated region, ribosome-binding site and ATG of the myogenin gene, and it therefore seems unlikely that straightforward translational control is operating, a conclusion consistent with the observation that the myogenin mRNA is loaded onto polysomes at 9.5 dpc (Cusella-De Angelis et al. 1992). The levels of myogenin protein may therefore be regulated at the level of protein stability. The fact that the myogenin protein is undetectable until 10.5 dpc indicates that autoregulation cannot be a factor in the establishment of the pattern of transgene expression. We conclude that the absolute requirement for an E-box must therefore reflect the role of another bHLH protein that is normally necessary for the activation of myogenin transcription. The fact that this E-box need not be in its normal position, between the TATA box and the start site of transcription, is of interest with regard to the mechanisms by which the regulatory bHLH factors interact with the basal transcription machinery. Since the submission of this paper, Cheng et al. (1992) have reported preliminary transgenic studies confirming that the position of the proximal E-box is not

crucial; their other data are also entirely consistent with ours. It is of interest that mutation of the proximal E-box affects expression in the forelimb bud, causing either a slight delay in the activation of the transgene or a decrease in the initial level of expression. However, forelimb bud expression is normal by 12.5 dpc and we observe no effect on expression in the hindlimb bud. It will be important to investigate the reasons for the apparently specific effect of the mutation on forelimb bud expression.

On the basis of their temporal patterns of expression it has been surmised that the bHLH factors are members of a regulatory hierarchy. An alternative view is that each factor is part of a distinct regulatory pathway characteristic of different cell populations within the myotome. Our data provide definitive evidence of a regulatory hierarchy in which myogenin is positioned below another bHLH protein. Although we cannot eliminate a role for an unknown or ubiquitously expressed factor, two lines of evidence suggest that the regulatory protein is likely to be Myf-5. First, this is the only gene of this class known to be expressed in the relevant cells of the embryo at the appropriate time. Second, Braun et al. (1992a) showed that the inhibition of myogenesis by adenovirus E1a gene products in rat L6 cells affects neither the levels of Myf-5 protein nor its DNA-binding activity but, rather, blocks the ability of the protein to *trans*-activate transcription. In cells expressing E1a, the myogenin gene cannot be activated, suggesting that in this system also, the activity of Myf-5 is necessary for the expression of myogenin. Moreover, the accumulated data on the expression of the bHLH factors in cultured cells are compatible with the view that *myf-5* acts early and that the myogenin gene is activated only at the onset of overt differentiation (for review, see Funk et al. 1991). The myogenin gene, however, is expressed in mice that lack the Myf-5 protein, although it is not clear when it is activated (Braun et al. 1992b). The fact that somitogenesis is delayed by ~2 days in the homozygous knockout mice makes it almost inevitable that myogenin transcription does not begin at the correct time. Whether myogenin is activated by another bHLH protein that does not depend on *myf-5*, or by a quite separate pathway, when *myf-5* is absent requires further investigation. Although *myf-5* is likely to be above myogenin in the hierarchy in the mouse, this order cannot be fundamental to the process of muscle differentiation, as in both *Xenopus* and avian embryos the order in which the bHLH factors is expressed is different from that in the mouse (Jennings 1992; Pownall and Emerson 1992).

It is striking that such an exquisitely specific pattern of gene expression can be achieved with only two closely spaced transcription factor-binding sites. This compact organization of the myogenin promoter is not the general case for the promoters of genes encoding muscle regulatory factors, although they are all expressed in similar patterns, albeit with differences in timing. Goldhamer et al. (1992) have recently shown that expression of the *myoD1* gene in transgenic mice requires an enhancer element located 18–22 kb upstream of the transcription start site, whereas Tapscott et al. (1992), using transfection experiments, have localized an element essential for *myoD1* expression to between −5.5 and −4.5 kb. Moreover, we have demonstrated (S.-P. Yee and P.W.J. Rigby, unpubl.) that 5 kb of 5′-flanking DNA from the *myf-5* gene are incapable of correctly activating the expression of a linked *lacZ* reporter gene. Whether the transcriptional control regions of the other genes are as apparently simple as those of myogenin is not known.

Detailed histological studies of embryos derived from a line of mice in which the *lacZ* reporter gene is driven by the 1.09 kb of 5′-flanking DNA show that β-galactosidase activity is detectable in all of the locations that have been shown previously by in situ hybridization analysis (Sassoon et al. 1989), to express the myogenin gene. These include the axial muscles of the trunk, the muscle masses of the limbs, the branchial arches, and the extraoccular muscles of the face. In birds, it has been shown clearly that the muscles of the limbs are derived from the somites (Chevallier et al. 1977; Christ et al. 1977; Ordahl and Le Douarin 1992). We have looked carefully for blue (myogenin-expressing) cells that might be in the process of migrating from the somites to the limbs and have failed to find them. Cusella-De Angelis et al. (1992) have recently described cells isolated from 8.5-dpc embryos that express myosin heavy chain, but neither myogenin nor *myoD1*, and have suggested that these represent the migratory cells. Although our data are consistent with the notion that the migrating cells do not express myogenin, they cannot address the possibility that these cells might have briefly expressed the gene while still located within the somites.

Mutation of the RSRF site, which has been shown previously to affect the efficiency of expression in cultured muscle cells (Edmondson et al. 1992), reveals hitherto unsuspected heterogeneities within the myotomes. At early times (10.5 and 11.5 dpc) embryos carrying this mutant transgene show expression in all of the most rostral myotomes. Moving caudally through the embryo there is first a dorsalization of expression within the myotomes and then, at the level of the fore limb bud, we observe expression in the more dorsal and more ventral cells but not in the central cells of each myotome. Finally, in the caudal portion of the embryo, expression is predominantly in the ventral halves of the myotomes. Figure 7 shows a schematic representation of this expression pattern derived from histological examination of several embryos. We have examined a considerable number of F_0 transgenic embryos carrying this construct (see Table 1 and Fig. 5). Moreover, we have made six lines of such mice. In four of these, the expression pattern is exactly that documented in Figure 5 and schematized in Figure 7. In the other two the pattern is distorted, presumably as a result of integration site effects, but the predominance of expression in the ventral halves of the caudal myotomes remains apparent. Mutation of the RSRF site has no significant effect on limb bud expression. The evidence thus clearly shows that at early times the RSRF site mutation leads to an expression pattern quite distinct from that seen with any of the other con-

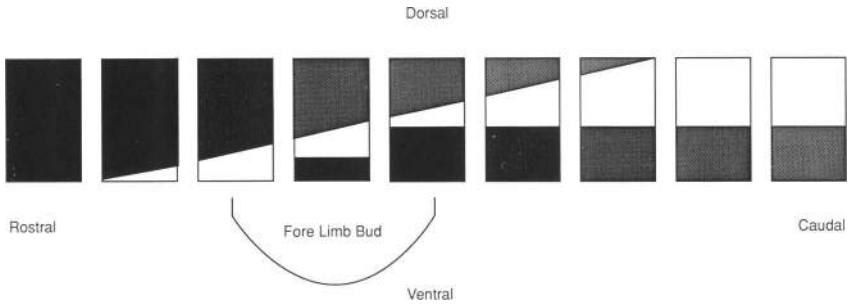

Figure 7. Schematic illustration of the *lacZ* expression pattern in 10.5-dpc GMRZ1092 transgenic embryos.

structs. The effect of the RSRF site can be observed only at early times; by 13.5 dpc, the expression pattern of the mutant construct cannot be distinguished from that of the wild type. Therefore, it appears that in some of the cells of the myotome, the RSRF site is essential for the activation of myogenin transcription at the proper time but that it is not required at later times. Alternatively, the cells that require the RSRF site may not contribute to the final musculature. The expression pattern of the mutant transgene shows that there are at least two populations of myotomal cells and that the RSRF site is necessary in only one of them. It will be of great interest to use animals carrying the mutant transgene to characterize the various cell populations. The rostrocaudal gradients of expression in both the dorsal and ventral sections of the myotomes could reflect the distribution of an unknown factor that interacts with the mutant RSRF site or the distribution of other factors that affect the efficiency of expression and become rate limiting in the absence of RSRF binding.

Our work so far has concentrated on the regulation of myogenin transcription during embryogenesis. In the adult organism, however, there are significant changes in the level of myogenin mRNA following denervation (Duclert et al. 1991) and muscle injury (Grounds et al. 1992). The lines of transgenic mice that we have described may well prove to be useful in the analysis of these regulatory systems.

The identification of binding sites for a bHLH protein, likely to be Myf-5, and for the RSRF proteins as being essential for the proper expression of myogenin focuses attention on the mechanisms by which the expression of these upstream factors is activated. Such information will be helpful in understanding the process by which cells first become committed to the myogenic lineage. Transgenic animals of the sort that we have described will also be extremely useful in studying the cell biology of myogenesis. They reveal subsets of myogenic cells, knowledge of which will contribute to an understanding of the derivation of the various types of mature muscle cells that are eventually produced, and they will also provide sources of autonomously marked cells for grafting experiments.

Materials and methods

Construction of fusion genes

The cosmid M8B was isolated by screening a library of DNA from a T-cell clone of C57BL6 origin in the vector cos202 (kindly provided by Dimitris Kioussis, Laboratory of Molecular Immunology, NIMR, London, UK) with a mouse myogenin cDNA probe. A 5.63-kb *Eco*RV fragment encompassing 5′-flanking sequences and the first intron of the myogenin gene was subcloned in Bluescript pKS+ (Stratagene) to generate pGVB. Another subclone, pGSM, was made by inserting a 486-bp *Sma*I fragment containing 5′-flanking sequence and the first 98 bp of coding sequence into pKS+. To align the *lacZ* reporter gene in-frame with the myogenin initiation codon, PCR was employed to introduce a *Bgl*II site immediately after the initiating ATG with pGSM as template, the universal T7 primer, and oligonucleotide MGB (5′-GAG ATC TTC CAT CAG GTC GCA AAA GGC -3′) as the second primer. The PCR product was purified, digested with *Pst*I and *Bgl*II and then cloned into a derivative of pKS+, pKSB, in which the *Sma*I site had been replaced with a *Bgl*II linker to generate pGPS. A *Bam*HI fragment from pZA (the gift of Dr. R. Kothary, AFRC, IAPGR, Cambridge, UK), which contains the *lacZ* gene followed by an SV40 polyadenylation sequence, was then inserted into the *Bgl*II site of pGPS to generate pGZ188. The orientation of the *lacZ* gene was confirmed by restriction endonuclease digestion, and the sequence of the junction between the myogenin and *lacZ* sequences was determined to confirm that the construction had proceeded correctly. To construct pGZ4250 a 4-kb *Kpn*I–*Pst*I fragment of myogenin, 5′-flanking DNA was cloned into pGZ188. These first two plasmids were digested with *Kpn*I and *Bam*HI to release the myogenin–*lacZ* fusion gene for microinjection. pGZ4250 was digested with *Sma*I and *Bam*HI to release the GZ338 fusion gene and partially digested with *Sac*I and *Bam*HI to release the GZ1092 fusion gene for microinjection. The GZ133 fusion gene was constructed by PCR to introduce an *Xba*I site at −133, using pGPS as template, the universal T3 primer, and oligonucleotide GX5 (5′-CTC TAG ACC ACT GGA AAC GTC TTG-3′) as the second primer. The PCR product was purified, digested with *Xba*I and *Bgl*II, and inserted into pKSB. The *Bam*HI *lacZ*–SV40 fragment from pZA was then inserted into the unique *Bgl*II site. The plasmid was digested with *Xba*I and *Bam*HI to release the GZ133 fusion gene for microinjection.

All mutant myogenin promoters were generated by PCR. The mutant promoter of GMRZ1092, which harbors A→T and T→A mutations at −67 and −62, respectively, was made with pGPS as template, the universal T7 primer, and oligonucleotide MRS5 (5′-GTT GGC TTT ATT AAT CTC TC-3′) as the second primer in one reaction and, in a second reaction, the same template, the universal T3 primer, and oligonucleotide MRS3 (5′-CAG AGA TTA ATA AAG CCA AC-3′) as the second primer. The two PCR products were purified, combined in equimolar ratio, and two further cycles of PCR were performed without primers. Fresh Vent DNA polymerase (New England BioLabs), together with the T3 and T7 primers were then added and PCR was carried out for an additional 25 cycles. The final product

was purified, digested with *Pst*I and *Bgl*II, and cloned into pKSB. The subsequent steps to obtain the final pGMRZ1092 construct were identical to those described above. The E-box mutant promoter, which contains G → A, T → C, A → G, and C → T mutations at −12, −13, −16, and −17, respectively, was also generated by PCR, essentially as described above, using oligonucleotides MGE5 (5′-GGC ACC CAG TGG TCA GTC TGA GG-3′) and MGE3 (5′-CCC TCA CAC TGA CCA CTG GGT GCC-3′). To construct EGMEZ133, PCR was performed using pGMEZ1092 as template and oligonucleotides EMGX5 (5′-CTC TAG AGC AGT TGC CAC TGG AAA CGT CTT G-3′) and MGB as primers. The final product was purified, digested with *Xba*I and *Bgl*II, and cloned into pKSB. The *Bam*HI fragment from pZA was then inserted into the *Bgl*II site to generate pEGMEZ133, and the plasmid was digested with *Xba*I and *Bam*HI to release the fusion gene for microinjection.

To construct RSEGZ, the myogenin–human β-globin–*lacZ* fusion gene, the myogenin 5′-flanking sequences between −192 and −55 were generated by PCR with pGSM as template, and oligonucleotides MGP3 (5′-GGG CAG GCC TGC AGG-3′) and RSR3 (5′-CAG AGA TAA ATA TAG CCA ACG CC- 3′) as primers. The PCR product was purified and cloned into the *Sma*I site of pKS+ to generate pRSEG. Two complementary oligonucleotides, GLOB5 (5′-AGC TTC CCG GGC TGG GCA TAA AAG TCA GGG CAG AGC CAT CTA TTG CTT ACA TTT GCT TCT GCA-3′) and GLOB3 (5′-CAA GCA AAT GTA AGC AAT AGA TGG CTC TGC CCT GAC TTT TAT GCC CAG CCC GGG A-3′), encompassing the human β-globin sequence from −37 to +12, were phosphorylated by use of ATP and T4 polynucleotide kinase, annealed, and cloned between the *Hin*dIII and *Pst*I sites of pZA to generate BGZ40. The human β-globin–*lacZ* fusion gene from this plasmid was excised by *Sma*I digestion and then cloned into the *Eco*RV site of pRSEG to generate pRSEGZ. This plasmid was digested with *Bam*HI and *Kpn*I to release the fusion gene for microinjection. All PCR products and mutant constructions derived from them were sequenced to ensure integrity.

Production and analysis of transgenic mice

The production of transgenic mice was carried out essentially as described by Hogan et al. (1986). All of the mice used were F_1 (CBA × C57BL10) hybrids obtained from the NIMR SPF breeding unit. Embryos to be analyzed were removed from pregnant females, fixed with freshly prepared, ice-cold 4% (wt/vol) paraformaldehyde in PBS for 30 min or longer depending on size. They were then washed twice for 15 min in PBS containing 2 mM $MgCl_2$, 5 mM EGTA, 0.1% (wt/vol) sodium deoxycholate and 0.01% (vol/vol) NP-40 at room temperature, and stained in the same solution containing 1 mg/ml of X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside), 5 mM $K_3Fe(CN)_6$, and 5 mM $K_4Fe(CN)_6$, in the dark, at 37°C. Embryos to be sectioned were washed 3 times in PBS.

Analyses for transgenesis were performed on DNA isolated from placenta or tail biopsies. Samples were incubated with 0.1–0.2 ml of 100 mM NaCl, 1 mM EDTA, 1% (wt/vol), SDS, and 1 mg/ml of proteinase K at 55°C overnight and then extracted twice with an equal volume of phenol/chloroform (1 : 1). Each sample (1 μl) was then analyzed by PCR with two pairs of oligonucleotides as primers in the same reaction. The first pair, M5′ (5′-CAG TGG CAG GAA CAA GCC-3′) and LZ3 (5′-GAT GGG CGC ATC GTA ACC GTG C-3′), detect a 309-bp fragment specific to the transgene. The second pair, MGP1 (5′-CCA AGT TGG TGT CAA AAG CC-3′) and MGP2 (5′-CTG ACT CCT TAA AGC AGA GAG-3′) detect a 171-bp fragment specific to the endogenous myogenin gene as an internal control.

Acknowledgments

We are especially grateful to Tim Mohun for much helpful advice and encouragement and for determining the sequence of the myogenin 5′-flanking DNA. We thank Richard Treisman for advice on RSRFs, Judith Boer for performing the transfection experiments, David Stott for help with the initial transgenic experiments, Ruedi Allemann, Jo Chan, and Dennis Summerbell for discussions, and Dimitris Kioussis for the cosmid library. We acknowledge Eric Olson for communicating his unpublished results and for helpful discussions. S.-P.Y. was supported by a Centennial Fellowship from the Medical Research Council of Canada. This work was paid for by the United Kingdom Medical Research Council.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

Note added in proof

The DNA sequence presented in Figure 3 has been submitted to the EMBL data library under accession number X71910.

References

Blackwell, K.T. and H. Weintraub. 1990. Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **250:** 1104–1110.

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

Braun, T., E. Bober, G. Buschhausen-Denker, S. Kotz, K. Grzeschik, and H.H. Arnold. 1989a. Differential expression of myogenic determination genes in muscle cells: Possible autoactivation by the Myf gene products. *EMBO J.* **8:** 3617–3625.

Braun, T., G. Buschhausen-Denker, E. Bober, E. Tennich, and H.H. Arnold. 1989b. 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., E. Bober, B. Winter, N. Rosenthal, and H.H. Arnold. 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., E. Bober, and H.H. Arnold. 1992a. Inhibition of muscle differentiation by the adenovirus E1a proteins: Repression of the transcriptional activating function of the HLH protein Myf-5. *Genes & Dev.* **6:** 888–902.

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

Chambers, A., S. Kotecha, N. Towers, and T. Mohun. 1992. Muscle-specific expression of SRF-related genes in the early embryo of *Xenopus laevis*. *EMBO J.* **11:** 4981–4991.

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

Chevallier, A., M. Kieny, and A. Mauger. 1977. Limb-somite relationship: Origin of the limb musculature. *J. Embryol. Exp. Morph.* **41:** 245–248.

Christ, B., H. Jacob, and M. Jacob. 1977. Experimental analysis of the origin of the wing musculature in avian embryos. *Anat. Embryol.* **150:** 171–186.

Cserjesi, P. and E.N. Olson. 1991. Myogenin induces the myocyte-specific enhancer-binding factor MEF-2 independently of other muscle-specific gene products. *Mol. Cell. Biol.* **11:** 4854–4862.

Cusella-De Angelis, M.G., G. Lyons, C. Sonnino, L. De Angelis, E. Vivarelli, K. Farmer, W.E. Wright, M. Molinaro, M. Bouche, M. Buckingham, and G. Cossu. 1992. MyoD, myogenin independent differentiation of primordial myoblasts in mouse somites. *J. Cell Biol.* **116:** 1243–1255.

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

Duclert, A., J. Piette, and J.-P. Changeux. 1991. Influence of innervation on myogenic factors and acetylcholine receptor α-subunit mRNAs. *Neuroreport* **2:** 25–28.

Edmondson, D.G. and E.N. Olson. 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., T-C. Cheng, P. Cserjesi, T. Chakraborty, and E.N. Olson. 1992. Analysis of the myogenin promoter reveals an indirect pathway for positive autoregulation mediated by the muscle-specific enhancer factor MEF-2. *Mol. Cell. Biol.* **12:** 3665–3677.

Funk, W.D., M. Ouellette, and W.E. Wright. 1991. Molecular biology of myogenic regulatory factors. *Mol. Biol. Med.* **8:** 185–195.

Goldhamer, D.J., A. Faerman, M. Shani, and C.P. Emerson, Jr. 1992. Regulatory elements that control the lineage-specific expression of *myoD*. *Science* **256:** 538–542.

Gossett, L.A., D.J. Kelvin, E.A. Sternberg, and E.N. Olson. 1989. A new myocyte-specific enhancer-binding factor that recognises a conserved element associated with multiple muscle-specific genes. *Mol. Cell. Biol.* **9:** 5022–5033.

Grounds, M.D., K.L. Garrett, M.C. Lai, W.E. Wright, and M.W. Beilharz. 1992. Identification of muscle precursor cells *in vivo* by use of MyoD1 and myogenin probes. *Cell Tiss. Res.* **267:** 99–104.

Hannon, K., C.K. Smith, II, K.R. Bales, and R.F. Santerre. 1992. Temporal and quantitative analysis of myogenic regulatory and growth factor gene expression in the developing mouse embryo. *Dev. Biol.* **151:** 137–144.

Hinterberger, T.J., D.A. Sassoon, S.J. Rhodes, and S.F. Konieczny. 1991. Expression of the muscle regulatory factor MRF4 during somite and skeletal myofiber development. *Dev. Biol.* **147:** 144–156.

Hogan, B., F. Constantini, and E. Lacy. 1986. *Manipulating the mouse embryo*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Horlick, R.A., G.M. Hobson, J.H. Patterson, M.Y. Mitchell, and P.A. Benfield. 1990. Brain and muscle creatine kinase genes contain common TA-rich recognition protein-binding regulatory elements. *Mol. Cell. Biol.* **10:** 4826–4836.

Jennings, C.G.B. 1992. Expression of the myogenic gene MRF4 during *Xenopus* development. *Dev. Biol.* **150:** 121–132.

Lassar, A.B., R.L. Davis, W.E. Wright, T. Kadesch, C. Murre, A. Vorenova, D. Baltimore, and H. Weintraub. 1991. Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins *in vivo*. *Cell* **66:** 305–315.

Miner, J.H. and B. Wold. 1990. Herculin, a fourth member of the *MyoD* family of myogenic regulatory genes. *Proc. Natl. Acad. Sci.* **87:** 1089–1093.

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

Olson, E.N. 1990. MyoD family: A paradigm for development? *Genes & Dev.* **4:** 1454-1461.

Ordahl, C.P. and N.M. Le Douarin. 1992. Two myogenic lineages within the developing somite. *Development* **114:** 339–353.

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

Pollock, R. and R. Treisman. 1991. Human SRF-related proteins: DNA-binding properties and potential regulatory targets. *Genes & Dev.* **5:** 2327–2341.

Pownall, M.E. and C.P. Emerson, Jr. 1992. Sequential activation of three myogenic regulatory genes during somite morphogenesis in quail embryos. *Dev. Biol.* **151:** 67–79.

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

Salminen, A., T. Braun, A. Buchberger, S. Jurs, B. Winters, and H.H. Arnold. 1991. Transcription of the muscle regulatory gene *MYF4* is regulated by serum components, peptide growth factors and signaling pathways involving G proteins. *J. Cell Biol.* **115:** 905–917.

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

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

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

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

Yu, Y.-T., R.E., Breitbart, L.B. Smoot, Y. Lee, V. Mahdavi, and B. Nadal-Ginard. 1992. Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted MADS box transcription factors. *Genes & Dev.* **6:** 1783–1798.

The regulation of myogenin gene expression during the embryonic development of the mouse.

S P Yee and P W Rigby

Genes Dev. 1993, **7:**
Access the most recent version at doi:10.1101/gad.7.7a.1277