Cell, Vol. 51, 987–1000, December 24, 1987, Copyright © 1987 by Cell Press

Expression of a Single Transfected cDNA Converts Fibroblasts to Myoblasts

Robert L. Davis,*† Harold Weintraub,* and Andrew B. Lassar*
* Department of Genetics
Hutchinson Cancer Research Center
1124 Columbia Street
Seattle, Washington 98104
† Department of Pathology
School of Medicine
University of Washington
Seattle, Washington 98195

Summary

5-azacytidine treatment of mouse C3H10T1/2 embryonic fibroblasts converts them to myoblasts at a frequency suggesting alteration of one or only a few closely linked regulatory loci. Assuming such loci to be differentially expressed as poly(A)$^+$ RNA in proliferating myoblasts, we prepared proliferating myoblast-specific, subtracted cDNA probes to screen a myocyte cDNA library. Based on a number of criteria, three cDNAs were selected and characterized. We show that expression of one of these cDNAs transfected into C3H10T1/2 fibroblasts, where it is not normally expressed, is sufficient to convert them to stable myoblasts. Myogenesis also occurs, but to a lesser extent, when this cDNA is expressed in a number of other cell lines. The major open reading frame encoded by this cDNA contains a short protein segment similar to a sequence present in the *myc* protein family.

Introduction

Jones and colleagues originally showed that myogenic, adipogenic, and chondrogenic clones can be derived from the embryonic mouse fibroblast line C3H10T1/2 (10T1/2 cells) following a brief treatment with 5-azacytidine (Taylor and Jones, 1979). Experiments using other cytidine analogs, and the correlation of this phenomenon with incorporation of 5-azacytidine into DNA, point to DNA demethylation as having a significant role (Jones and Taylor, 1980). Moreover, based on the high frequency of myogenic conversion (25%–50%), it appears that demethylation of only one or a few closely linked loci are necessary to convert 10T1/2 cells into myoblasts (Konieczny and Emerson, 1984). This possibility has been substantiated by the observation that transfected DNA from 5-azacytidine-derived, myogenic 10T1/2 clones can convert normal 10T1/2 cells to myoblasts at a frequency (10^{-4}) consistent with the transfer of one or only a few closely linked demethylated loci, while transfected DNA from normal 10T1/2 cells has no effect (Lassar et al., 1986; see also Konieczny et al., 1981). Myogenic lines, derived after 5-azacytidine treatment or genomic myoblast DNA transfection of 10T1/2 cells, express a repertoire of muscle-specific genes when these cells are induced to differentiate.

We wish to identify the loci involved in converting 10T1/2 cells to determined myoblasts. Both immunological and biochemical studies have indicated that proliferating myoblasts express lineage-specific markers prior to terminal differentiation (Walsh and Ritter, 1981; Konieczny and Emerson, 1984; Sasse et al., 1984; Kaufman et al., 1985). We hypothesized that, unlike uncommitted 10T1/2 cells, proliferating myoblasts derived from them continually express, as poly(A)$^+$ RNA, regulatory loci that both activate the expression of lineage-specific markers in proliferating myoblasts and make these cells competent to express other muscle-specific genes when induced to differentiate. We have employed subtracted cDNA hybridization between proliferating myoblast cDNA and proliferating 10T1/2 poly(A)$^+$ RNA to prepare probes for screening a myocyte cDNA library. A number of cDNA clones have thus been isolated and tested for expression specific to myogenic cells and muscle tissue. Use of several assumptions regarding the expression of myogenic determination loci restricted further analysis to three of these clones. We show that expression of one of these cDNAs in transfected 10T1/2 cells is sufficient to convert them at high frequency to stable myoblasts. Myogenesis is also induced, but to a lesser extent, in mouse 3T3 embryonic fibroblast lines and in a variety of other cell lines transfected with this cDNA. This gene, which we call MyoD, is normally expressed only in mouse skeletal muscle in vivo and myogenic cell lines in vitro; it is not detectable in other tested adult or newborn mouse tissue, nor in several nondifferentiating variant myogenic cell lines. The single major open reading frame encoded by this cDNA contains a region with sequence similarity to a region of the *myc* protein family. This region has been shown to be essential for c-*myc* transforming activity in vitro (Stone et al., 1987). A similar sequence is shared by two transcription units of the Drosophila *achaete-scute* complex (Villares and Cabrera, 1987), which is involved in neuroblast commitment.

Results

Isolation of Myoblast-Specific cDNAs

To prepare subtracted probes enriched for putative myogenic regulatory sequences, ^{32}P-labeled cDNA was prepared using poly(A)$^+$ RNA from proliferating myoblasts of two types of myogenic cell lines. One line was derived from 5-azacytidine treated 10T1/2 cells (F3 aza-myoblasts); the other was the normal diploid mouse myogenic line C2C12 (Blau et al., 1983). In preparing these probes, we wanted to select against 5-azacytidine-induced sequences irrelevant to myogenic determination, since 5-azacytidine induces the expression of a variety of genes in various cell types (Groudine et al., 1981; Hsiao et al., 1986). We suspected that mRNAs which were not common to both aza-myoblasts and C2C12 cells might represent such 5-aza-

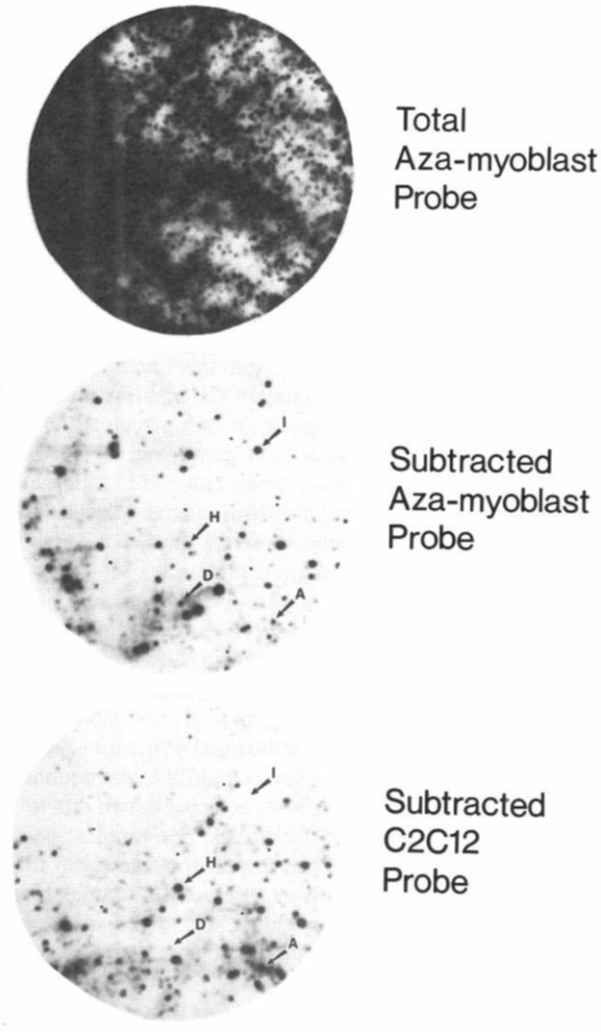

Figure 1. Screening a Total Aza-Myoblast/Myotube cDNA Library with Myoblast-Specific Probes

Autoradiographs of approximately 10^4 clones of an aza-myoblast/myotube cDNA library screened with ^{32}P-labeled total myoblast (overnight exposure) or subtracted, myoblast-specific probes (8 day exposure). The bottom two autoradiographs are duplicate filters. Representative phages containing MyoD (D), MyoA (A), MyoH (H), and intracisternal type A particle (I) cDNAs are indicated.

cytidine-induced messages as well as other sequences irrelevant to myogenesis. We prepared subtracted probes such that isolated cDNA clones would represent sequences that are absent in proliferating 10T1/2 cells and present in both myogenic cell lines. Aza-myoblast and C2C12 ^{32}P-cDNAs were each exhaustively hybridized ($C_0t \geqslant 3000$) twice to 10T1/2 poly(A)$^+$ RNA, with the myoblast-specific single-stranded material isolated each time after passage over hydroxyapatite. The myoblast-specific cDNA probes were then positively selected by hybridizing the remaining aza-myoblast cDNA to C2C12 poly(A)$^+$ RNA, and the C2C12 cDNA to aza-myoblast poly(A)$^+$ RNA, followed by recovery of double-stranded material (see Experimental Procedures for details).

The probes in final form were used to screen a λgt10 myoblast/myotube cDNA library prepared using cDNA synthesized from a 1:1 mixture of proliferating myoblast and differentiated myotube poly(A)$^+$ RNA of the F3 aza-myoblast line. Approximately 1% of the recombinant phage hybridized with the subtracted probes. Figure 1 shows a comparison of autoradiographs from plaque hybridizations of the two subtracted probes versus an unsubtracted aza-myoblast cDNA probe. The positive plaques from this initial screening were divided into three groups: those showing stronger hybridization to the aza-myoblast probe; those showing stronger hybridization to the C2C12 probe; or those showing equal hybridization to both probes. Plaques from each group were picked from the primary filters and rescreened (after clonal gridding) on secondary filters. From 10^4 cDNA clones screened, 92 cDNA clones were isolated, which represent 26 different sequences.

We used the following criteria to select cDNA clones for further analysis: complete lack of expression in proliferating 10T1/2 cells; little or no increase in expression in differentiated myotubes versus proliferating myoblasts to exclude markers characteristic of terminally differentiated myotubes; and reduced to absent expression in non-differentiating variants derived from an aza-myoblast line. The validity of employing this last criterion is based upon several characteristics of the variant lines, which are outlined below. Classes of isolated clones failing our criteria included intracisternal type A particle RNAs, whose expression is induced by 5-azacytidine but is not diminished in nondifferentiating variants (data not shown). In addition, since a small percentage of cells in the proliferating myoblast cultures used to prepare poly(A)$^+$ RNA do differentiate, we expected to, and did, isolate some clones with expression characteristic of terminal differentiation markers (three of which were desmin, myosin heavy chain, and myosin light chain 2). Eventually, the focus narrowed to the three cDNAs described below (see Experimental Procedures for a detailed account of how the analysis was simplified).

Characterization of Selected cDNA Clones

The three cDNA clones are referred to as MyoA, MyoD, and MyoH, representing RNAs of approximately 5.0, 2.0, and 2.4 kb, respectively. Their representations in the F3 aza-myoblast/myotube library are 29 for MyoA, 20 for MyoD, and 488 for MyoH, per 10^5 cDNAs. For comparison, we screened the F3 aza-myoblast/myotube library with probes to genes thought to represent the low abundance class of mRNAs. These included c-*myc*, at 4 per 10^5 cDNAs and hypoxanthine-guanine phosphoribosyl transferase, at 3 per 10^5 cDNAs. The lowest abundance cDNA we isolated with the subtracted probe has a frequency of 4 per 10^5 cDNAs. We feel that, while we did not perform a saturating screen (since we screened only 10^4 cDNAs and a typical somatic cell is thought to contain about 10^5 mRNAs, half or more of which are of the low abundance class; Hedrick et al., 1984), we were able to detect and isolate some cDNAs in the 10^4 clones screened that represent low abundance poly(A)$^+$ RNAs. Representative phage containing the Myo cDNAs in a

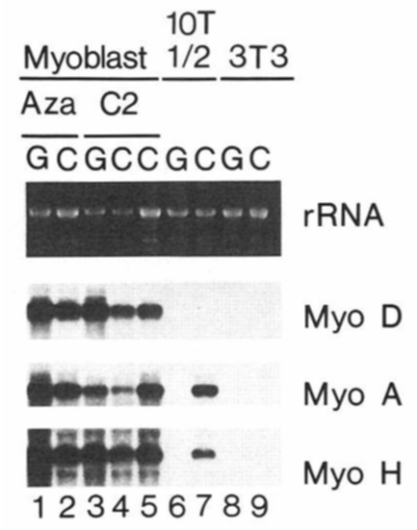

Figure 2. Expression of Myo mRNAs in Myoblasts, 10T1/2 Cells, and Swiss 3T3 Cells

Northern analysis (total RNA) of MyoD, MyoA, and MyoH expression in growing (G) and confluent (C) P2 aza-myoblasts (lanes 1 and 2), C2C12 myoblasts (lanes 3, 4, and 5), 10T1/2 cells (lanes 6 and 7), and Swiss 3T3 cells (lanes 8 and 9). Confluent myoblast cultures were composed of a mixture of differentiated myotubes and undifferentiated cells. In lane 5, C2C12 myotubes were purified by incubating the culture in differentiation medium (2% horse serum) plus 10^{-5} M cytosine arabinoside for 4 days. The top panel shows an ethidium bromide stain (of rRNA) of the RNA sample used for Northern analysis.

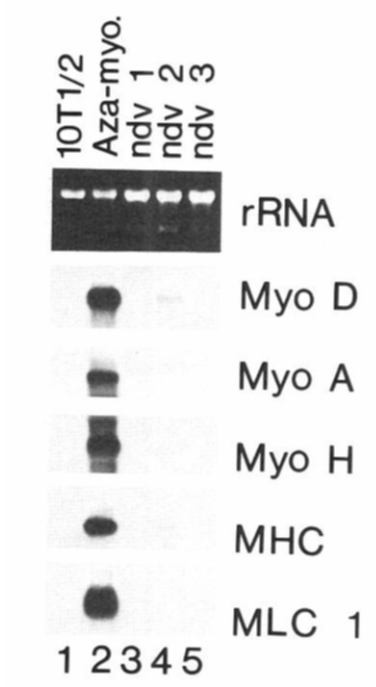

Figure 3. Expression of Myo mRNAs and Muscle Differentiation Markers in Nondifferentiating Aza-Myoblast Variants

Northern analysis (total RNA) of MyoD, MyoA, MyoH, myosin heavy chain (MHC), and myosin light chain 1 (MLC1) in 10T1/2 cells (lane 1), P2 aza-myoblasts (lane 2), and in three lines of nondifferentiating variants (ndv) derived from P2 aza-myoblasts (lanes 3–5). Levels of the Myo mRNAs were analyzed in proliferating cells whereas MHC and MLC1 levels were analyzed in cells placed in differentiation medium (2% horse serum) for 4 days. The top panel shows an ethidium bromide stain (of rRNA) of the RNA samples (from proliferating cells) used for Northern analysis of the Myo mRNAs.

screen of the F3 azamyoblast/myotube library are indicated in Figure 1.

Figure 2 shows the expression of the three corresponding RNAs by Northern analysis. All three messages are expressed in proliferating myoblasts and differentiated myotubes (lanes 1–5). As expected, the RNAs are not expressed in proliferating 10T1/2 cells (lane 6). Whereas MyoD is not expressed in confluent, growth-arrested 10T1/2 cells, MyoA and MyoH are expressed in these cells (lane 7). In separate experiments, we have also shown that induction of MyoA and MyoH occurs in nonconfluent 10T1/2 cells when serum is removed. MyoA and MyoH are not genes generally responsive to growth arrest in other cell types, as shown by the lack of induction in growth-arrested Swiss 3T3 fibroblasts (lanes 8 and 9), as well as a variety of other fibroblast and adipoblast cell lines tested (data not shown).

Figure 3 shows expression of the Myo RNAs in nondifferentiating variants derived from the P2 aza-myoblast line shown in Figure 2. These nondifferentiating lines, which fail to express a number of muscle differentiation markers (see Figure 3), were derived by continual passage of proliferating cells after most of the cells on a plate had terminally withdrawn from the cell cycle and had differentiated. Cells able to proliferate were passaged several times and then cloned (see Experimental Procedures for a more detailed description of their derivation). We reasoned that these lines had either lost their ability to differentiate or in fact were no longer determined myoblasts. That expression of the three Myo RNAs is either absent (variants 1 and 3) or highly attenuated (variant 2) in these cells (when proliferating) suggests that these lines lack a factor(s) required for expression of muscle-specific markers prior to differentiation. The absence of MyoA and MyoH expression in confluent variants (data not shown) suggests that these cells are not "simple" revertants to the 10T1/2 phenotype. Perhaps the variant lines receive a continuous growth signal even in the absence of serum factors. In contrast to the Myo RNAs, intracisternal type A particle (IAP) RNA, whose expression is induced by 5-azacytidine, remains at high levels in the nondifferentiating variants (data not shown). In addition to IAP RNA, we isolated five other cDNAs whose expression was maintained in the nondifferentiating variants. Although it is possible that these sequences play a role in myogenesis, we decided to initially concentrate on those muscle-specific cDNAs (MyoA, D, and H) whose expression is absent or highly attenuated in these variant lines.

Transfection of 10T1/2 Cells with a MyoD Expression Vector Converts Them to Stable Myoblasts

Having obtained only three cDNAs that satisfied our screening criteria, we focused our efforts on identifying phenotypes associated with their expression in proliferating 10T1/2 cells, where they are not normally expressed.

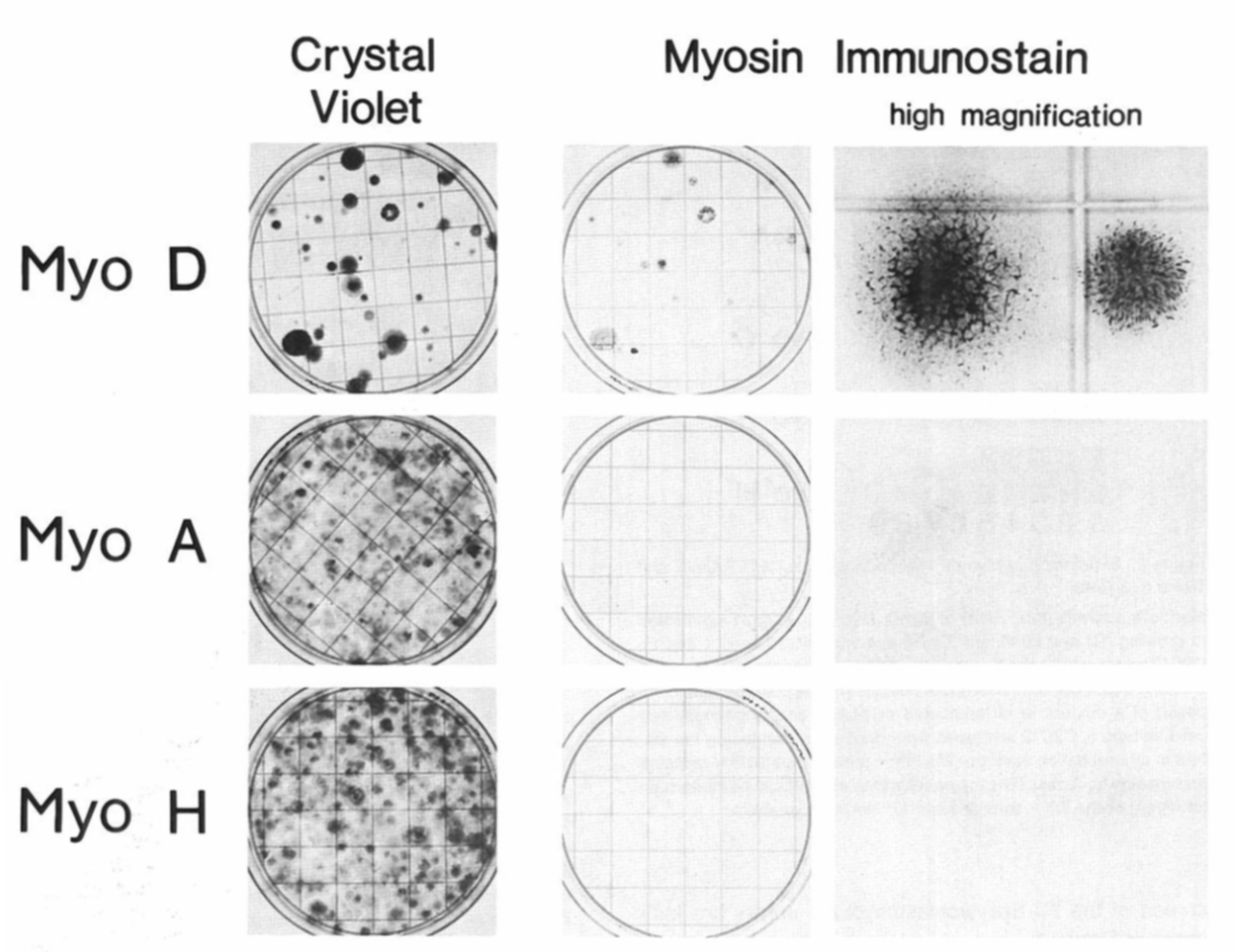

Figure 4. Transfection of 10T1/2 Cells with Myo cDNA Expression Vehicles

10T1/2 cells were transfected with 10 μg of the expression vehicles (derived from pEMSVscribe), containing the Myo cDNAs in the sense orientation, 0.1 μg of pLNSAL (a plasmid encoding G418 resistance), and 10 μg of genomic 10T1/2 DNA. The left-hand panels show a crystal violet stain of the resultant G418 resistant colonies, the middle panels show the same dishes as on the left stained for myosin heavy chain, and the right-hand panels show a high magnification (12×) micrograph of myosin-stained colonies on these dishes.

The cDNAs were cloned into an expression vector under control of the Moloney sarcoma virus LTR with an SV40 poly(A) addition signal (see Experimental Procedures). Whereas both MyoD and MyoH cDNAs were nearly full length, the MyoA cDNA represented only 2 kb out of a 5 kb message.

Using the "sense" expression orientation (determined by probing Northerns with strand-specific probes), these vectors were cotransfected into 10T1/2 cells at approximately a 100:1 molar ratio with a selection vector encoding the bacterial neomycin phosphotransferase gene (*neo*). The cells were selected for G418 resistance 24 hr after transfection. Cotransfection with the MyoD expression vector gave a noticeable reduction in the number of G418 resistant colonies (stained with crystal violet) compared to the number of colonies resulting after cotransfection with MyoA or MyoH (Figure 4, left). The transfected cultures were fixed and scored for myosin heavy chain expression using alkaline phosphatase-linked immunostaining (Figure 4, middle and right). We observed that about 50% (an average of several transfections) of those colonies obtained from transfection with MyoD demonstrated myosin-positive, multinucleate syncitia, morphologically indistinguishable from myogenic colonies derived by 5-azacytidine treatment of 10T1/2 cells. No notable phenotypic changes were observed using MyoA or MyoH vectors. Figure 4 (top right) shows representative myogenic G418 resistant colonies derived by cotransfection with the MyoD expression vector and immunostained for myosin heavy chain. Fusion of cells to form syncitia was coincident with the presence of the myosin heavy chain gene product, showing coordination of the muscle differentiation program. In addition, we have determined that these cells also express myosin alkali light chain 1,3 by immunostaining (data not shown) and myosin light chain 2 by Northern analysis (see below). Control transfections of 10T1/2 cells with MyoD in a promoterless plasmid or in the antisense orientation in the same expression vector showed no effect, either on colony formation or myogenesis (data not shown).

The reduction in number of G418 resistant colonies on MyoD transfection of 10T1/2 cells is a reproducible pheno-

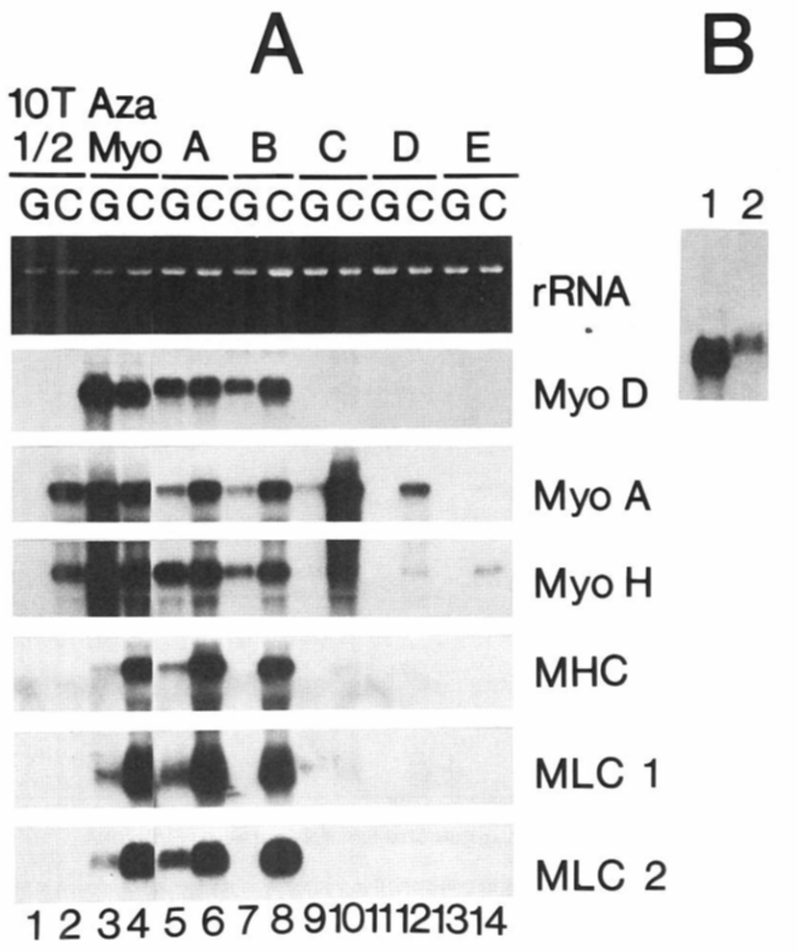

Figure 5. Expression of the Myo RNAs and Muscle Differentiation Markers in 10T1/2 Clones Transfected with the MyoD Expression Vehicle

(A) Northern analysis (total RNA) of MyoD, MyoA, MyoH, myosin heavy chain (MHC), myosin light chain 1 (MLC1), and myosin light chain 2 (MLC2) in growing (G) and confluent (C) 10T1/2 cells (lanes 1 and 2), P2 aza-myoblasts (lanes 3 and 4), and five clones (A, B, C, D, and E) of 10T1/2 cells transfected with the MyoD expression vehicle, picked from a parallel plate as that shown in Figure 4 (lanes 5–14). Clones A and B show extensive fusion when cultured in differentiation medium, clone C initially fused but lost this capacity during subsequent passage, and clones D and E never displayed any fusion. The top panel shows an ethidium bromide stain (of rRNA) of the RNA sample used for Northern analysis.

(B) Northern analysis of MyoD expression in proliferating P2 aza-myoblasts (lane 1) and myogenic transfectant clone A cells (lane 2).

type that becomes noticeable when the input ratio of MyoD expression vector to selection vector exceeds 10:1. One attractive explanation is that beyond a threshold level of MyoD expression some transfected cells differentiate early and fail to proliferate, thereby not contributing to the final colony scoring. As the colonies were growing, we often observed isolated multinucleate cells, or cells in small clusters, that had the appearance of myotubes, although it is generally difficult to discriminate between them and the morphological effects of G418 selection. However, after immunostaining plates with fully grown colonies, we again saw isolated multinucleate cells strongly positive for myosin heavy chain, which suggested that these isolated clusters of cells were, in fact, prematurely differentiating transfectants. This point is discussed further below.

About half of the G418 resistant clones cotransfected with MyoD differentiate into muscle. Several differentiating and nondifferentiating clones were isolated and the expression of MyoD in each was examined by Northern analysis. Figure 5A shows that clones A and B, which express muscle structural genes, express MyoD at levels comparable to aza-myoblasts (lanes 3–8). In contrast clones D and E, which do not differentiate, lack MyoD expression (lanes 11–14). Clone C was derived from a fused, apparently myogenic colony, which after subsequent passage lost its capacity to fuse. Muscle differentiation markers and MyoD are both expressed at only trace levels in this clone (lanes 9 and 10). For each clone, the expression of MyoD and the ability to differentiate are correlated with expression of myosin heavy chain, myosin light chain 1, and myosin light chain 2 by Northern analysis. Using low percentage gels, the exogenous and endogenous MyoD transcripts can be distinguished, due to a slight difference in electrophoretic mobility. We can detect only the exogenous transcript of MyoD in transfected myogenic colonies (Figure 5B), which suggests that the endogenous gene cannot be transactivated by its own gene product. The expression of MyoA and MyoH in proliferating cells, which is characteristic of myoblasts, is evident in transfected clones that express MyoD (clones A and B, and to a lesser extent clone C; Figure 5A). This suggests that MyoD expression alters the growth-arrest restricted expression of these genes characteristic of parental 10T1/2 cells.

Expression of MyoD in Other Cell Lines Induces Myogenesis

Myogenesis has been reported to occur in 5-azacytidine-treated Swiss 3T3 cells and CVP3SC6, a mouse prostate cell line (Taylor and Jones, 1982). However, the frequency of myogenesis in 5-azacytidine-treated 10T1/2 cells is at least two orders of magnitude greater than in Swiss 3T3 cells (Taylor and Jones, 1979). This observation led us to

Table 1. Frequency of Myogenic Colonies in Different Cell Lines Transfected with MyoD

	Transfected Cell Line	# of Myogenic Colonies[a] per Total G418^R Colonies Transfected with Expression Vehicle (without insert)	# of Myogenic Colonies[b] per Total G418^R Colonies Transfected with MyoD Expression Vehicle	Range of % of Myosin-Positive Cells per Myogenic Colony	Extent of Myogenesis per Transfection with MyoD Expression Vehicle[c]
(A)	C3H10T1/2 (mouse fibroblast)	0/456	81/154 (53%)	1%–50%	high
(B)	NIH 3T3 (mouse fibroblast)	0/720	23/250 (9%)	0.1%–1%	low
(C)	Swiss 3T3 (mouse fibroblast)	0/48	17/38 (45%)	1%–10%	medium
(D)	Swiss 3T3 Clone 2 (C2) (mouse fibroblast)	0/61	20/50 (40%)	1%–10%	medium
(E)	L Cell (mouse fibroblast)	0/900	23/800 (3%)	1%–10%	low
(F)	3T3-L1 (Swiss 3T3–derived adipoblast)	0/118	37/120 (31%)	1%–50%	high
(G)	3T3-F442A (Swiss 3T3–derived adipoblast)	0/61	30/63 (48%)	1%–10%	medium
(H)	TA1 (C3H10T1/2-derived adipoblast)	15/118[d] (13%)	104/160 (65%)	1%–50%	high
(I)	CV1 (monkey kidney cell)	0/190	0/186 (0%)	—	none

[a] Cells were transfected with 0.2 μg pCMV neo, 2.0 μg pEMSVscribe, and 20.0 μg C3H10T1/2 high molecular weight carrier DNA.
[b] Cells were transfected with 0.2 μg pCMVneo, 2.0 μg pEMSVscribe-MyoD, and 20.0 μg C3H10T1/2 high molecular weight carrier DNA.
[c] Determined by noting the percentage of myosin-positive cells per polyclonal culture.
[d] TA1 cells show the same level of spontaneous myogenesis in the absence of DNA transfection. This frequency of myogenic conversion is increased by transfection with the MyoD expression vehicle.

think that MyoD may operate in a permissive environment of 10T1/2 cells such that it activates a predetermined state (as if 10T1/2 cells are only one step away from myogenic determination). If so, one would not expect MyoD expression alone in other lines to result in myogenic determination. Alternatively, the low frequency of myogenesis in 5-azacytidine-treated Swiss 3T3 cells, for example, may be due to a deficiency of factors necessary for MyoD expression itself, rendering the locus relatively insensitive (directly or indirectly) to genomic demethylation. In this case, MyoD expression in Swiss 3T3 cells, driven by a viral LTR, might efficiently convert these cells into myoblasts. We cotransfected the MyoD expression vehicle with a vector encoding *neo* into the cell types listed in Table 1 and selected for G418 resistance. Colonies were immunostained for myosin heavy chain (Figure 6). With the exception of CV1 cells, myosin heavy chain synthesis and cell fusion were detected in all cell lines transfected with the MyoD expression vehicle. However, the frequency of myogenic colonies and the percentage of myosin-positive cells within these colonies varied with each line. Whereas 53% of transfected 10T1/2 colonies were myogenic, 3%, 9%, and 45% of the G418 resistant L cell, NIH3T3, and Swiss 3T3 colonies (respectively) were myogenic (see Table 1). In the latter three cell lines, the percentage of myosin-positive cells per myogenic colony is considerably less than in myogenic 10T1/2 colonies. The low percentage of differentiated cells in these colonies probably accounts for the low degree of cell fusion (due to the lack of adjacent fusion partners). Myogenesis also occurs in BALB/C 3T3 cells transfected with the MyoD expression vehicle (data not shown). No myosin heavy chain was detectable in the above cell lines transfected with the parental expression vehicle lacking the MyoD cDNA insert.

In addition to transfected "fibroblast" lines, MyoD expression also induces myogenesis in adipoblast cell lines derived from either Swiss 3T3 (3T3–L1 and 3T3–F442A) or 5-azacytidine-treated 10T1/2 cells (TA1). Some of the 3T3–L1 and TA1 myogenic colonies show as high a degree of differentiation as transfected 10T1/2 colonies (Figures 6F and 6H, respectively). We are currently examining whether the adipocyte differentiation program is concomitantly expressed in these apparently myogenic colonies. In addition to myosin heavy chain, myosin light chain 1,3 has also been detected by immunostaining in BALB/C 3T3, 3T3–L1, and 3T3–F442A cells transfected with the MyoD expression vehicle. CV1, a monkey kidney cell line, is the only line that did not display detectable myogenesis after MyoD transfection. It is unclear if the lack of myogenesis in this cell line is due to the epithelial origin of CV1 cells or to a species incompatibility of mouse MyoD in a monkey cell. Northern analysis revealed that G418 resis-

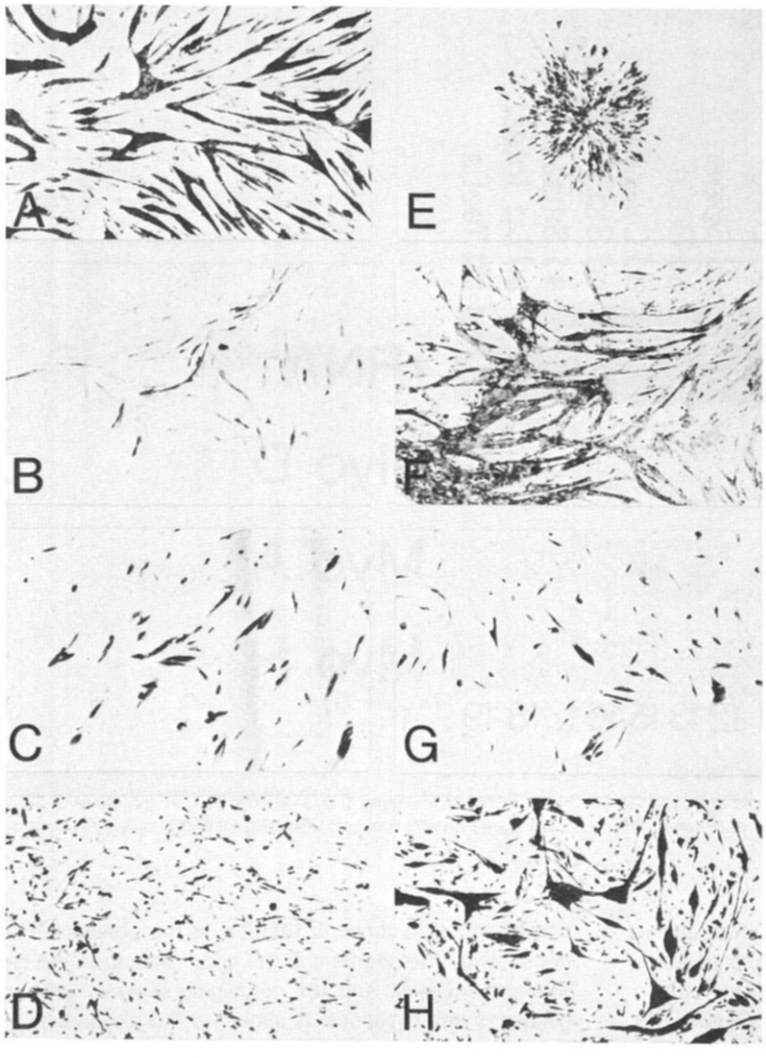

Figure 6. Myosin Staining of Different Cell Lines Transfected with the MyoD Expression Vehicle

Immunostained myosin heavy chain–positive colonies (shown at 50× magnification) resulting after transfection of 2.0 μg of the MyoD expression vehicle (pEMSVscribe MyoD) 0.2 μg of pCMVneo (a plasmid encoding G418 resistance), and 20 μg of genomic 10T1/2 DNA into the following cell lines: C3H10T1/2 (A), NIH 3T3 (B), Swiss 3T3 (C), Swiss 3T3-clone 2 (D), L cells (E), 3T3-L1 adipocytes (F), 3T3-F442A adipocytes (G), and TA1 adipocytes (H).

tant polyclones of transfected CV1 cells were expressing MyoD RNA at levels comparable to transfected 10T1/2 polyclones (data not shown). The effects of expressing MyoD in more defined differentiated cells (e.g., hepatoma cells, neuroblastomas, and myelomas), as well as in non-differentiating muscle variants, is currently under study.

Expression of MyoD, MyoA, and MyoH in Mouse Tissues

We examined the expression of MyoD, MyoA, and MyoH RNAs in various mouse adult tissues and some neonatal tissues by Northern analysis. Figure 7 shows that MyoD RNA is expressed in neonatal and adult skeletal muscle (lanes 2, 3, 6, and 7), although not in neonatal or adult cardiac muscle (lanes 4 and 8), nor in any other tissue tested. In addition, MyoD RNA is not detectable in the mouse smooth muscle cell line, BC3H1 (data not shown). The level of expression of MyoD RNA in skeletal muscle tissue is about 10% of that in skeletal muscle cell lines. It may be significant that in chicken primary cultures MyoD expression is highly reduced in myotubes as compared with proliferating myoblasts (data not shown).

MyoA RNA is expressed in neonatal and adult skeletal and cardiac muscle, as well as adult kidney, lung, brain, ovary, and stomach (trace). MyoH RNA is expressed in neonatal and adult skeletal and cardiac muscle, as well as neonatal liver, and adult lung, ovary, and stomach (trace). The expression pattern for MyoH RNA is reminiscent of a previously characterized gene, H19, described by Tilghman and colleagues in their studies of loci under control of the *raf* and *Rif* loci in developing mouse liver (Pachnis et al., 1984). In fact, MyoH cross-hybridizes to an H19 cDNA generously provided by S. Tilghman (data not shown). The function of H19 in skeletal and cardiac muscle, where it is expressed constitutively (Pachnis et al., 1984; this study), is unknown. We can also detect transcripts homologous to MyoD, MyoA, and MyoH in human and chicken myogenic cell cultures (data not shown).

Sequence of MyoD

We have sequenced MyoD by generating nested deletions (Henikoff, 1987) and using dideoxy sequencing procedures (Sanger et al., 1977). This sequence is presented in Figure 8. The sense orientation encodes only

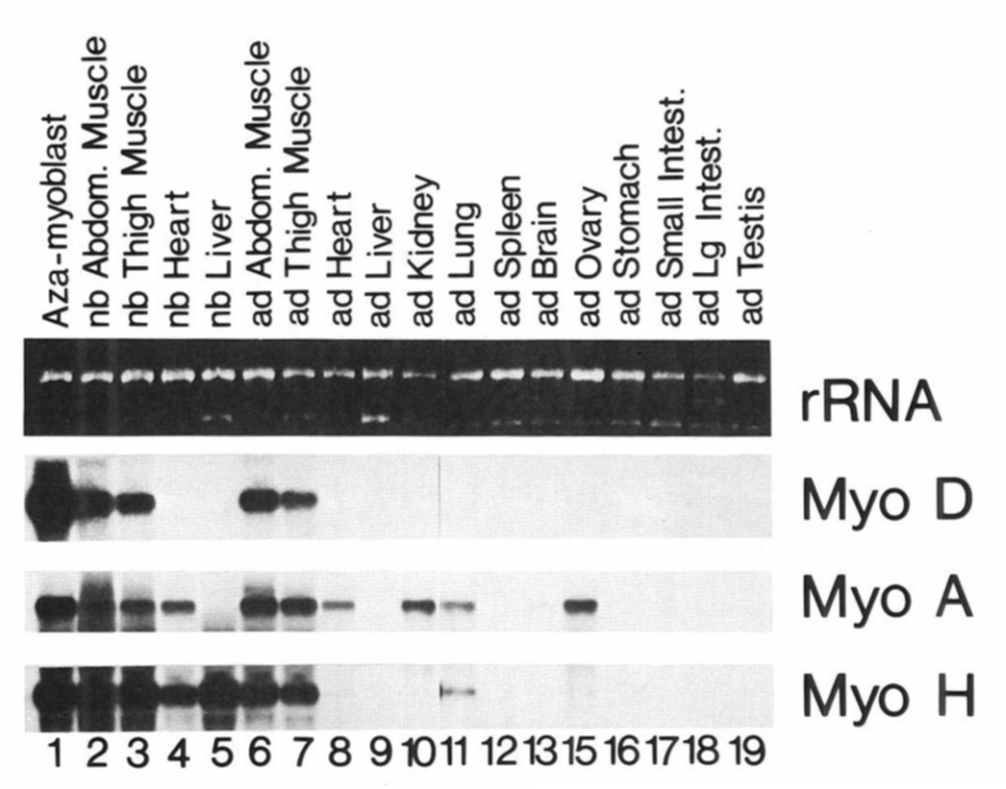

Figure 7. Expression of Myo mRNAs in Various Mouse Tissues

Northern analysis (total RNA) of MyoD, MyoA, and MyoH expression in the P2 aza-myoblast line (lane 1) and various newborn (nb; lanes 2–5) and adult (ad; lanes 6–19) mouse tissues (indicated above each lane). The top panel shows an ethidium bromide stain (of rRNA) of the RNA sample used for Northern analysis.

one open reading frame that is greater than 100 residues and has a good initiation sequence consensus (*GATATGG*). This 318 amino acid reading frame is also consistent with the observation that a shorter MyoD cDNA clone (a 300 bp 3′ deletion) which contains this reading frame intact also converts 10T1/2 cells to myoblasts (data not shown). The segment of the open reading frame between residues 62 and 101, containing several clustered histidines and cysteines, is similar to the putative metal binding domains of the "zinc-finger" class (Berg, 1986), al-

```
ACACCTCTGACAGGACAGGACAGGGAGGAGGGGTAGAGGACAGCCGGTGTGCATTCCAACCCACAGAACCTTTGTCATTGTACTGTTGGGGTTCCGGAGTGGCAGAAAGTTAAGACGACT     120

CTCACGGCTTGGGTTGAGGCTGGACCCAGGAACTGGGATATGGAGCTTCTATCGCCGCCACTCCGGGACATAGACTTGACAGGCCCCGACGGCTCTCTCTGCTCCTTTGAGACAGCAGAC     240
                                                    MetGluLeuLeuSerProProLeuArgAspIleAspLeuThrGlyProAspGlySerLeuCysSerPheGluThrAlaAsp      27

GACTTCTATGATGATCCGTGTTTCGACTCACCAGACCTGCGCTTTTTTGAGGACCTGGACCCCGCCTGGTGCACGTGGGAGCCCTCCTGAAACCCGCAGCAGCACGCACACTTCTCTACT     360
AspPheTyrAspAspProCysPheAspSerProAspLeuArgPhePheGluAspLeuAspProArgLeuValHisValGlyAlaLeuLeuLysProGluGluHisAlaHisPheSerThr      67

GCGGTGCACCCAGGCCCAGGCGCTCGTGAGGATGAGCATGTGCGCGCGCCCAGCGGGCACCACCAGGCGGGTCGCTGCTTGCTGTGGGCCTGCAAGGCGTGCAAGCGCAAGACCACCAAC     480
AlaValHisProGlyProGlyAlaArgGluAspGluHisValArgAlaProSerGlyHisHisGlnAlaGlyArgCysLeuLeuTrpAlaCysLysAlaCysLysArgLysThrThrAsn     107

GCTGATCGCCGCAAGGCCGCCCACCATGCGCGAGCGCCGCCGCCTGAGCAAAGTGAATGAGGCCTTCGAGACGCTCAAGCGCTGCACGTCCAGCAACCCGAACCAGCGGCTACCCAAGGTG     600
AlaAspArgArgLysAlaAlaThrMetArgGluArgArgArgLeuSerLysValAsnGluAlaPheGluThrLeuLysArgCysThrSerSerAsnProAsnGlnArgLeuProLysVal     147

GAGATCCTGCGCAACGCCATCCGCTACATCGAAGGTCTGCAGGCTCTGCTGCGCGACCAGGACGCCGCGCCCCCTGGCGCCGCTGCCTTCTACGCACCTGGACCGCTGCCCCCAGGCCGT     720
GluIleLeuArgAsnAlaIleArgTyrIleGluGlyLeuGlnAlaLeuLeuArgAspGlnAspAlaAlaProProGlyAlaAlaAlaPheTyrAlaProGlyProLeuProProGlyArg     187

GGCAGCGAGCACTACAGTGGCGACTCAGACGCGTCCAGCCCGCGCTCCAACTGCTCTGATGGCATGATGGATTACAGCGGCCCCCCAAGCGGCCCCCGGCGGCAGAATGGCTACGACACC     840
GlySerGluHisTyrSerGlyAspSerAspAlaSerSerProArgSerAsnCysSerAspGlyMetMetAspTyrSerGlyProProSerGlyProArgArgGlnAsnGlyTyrAspThr     227

GCCTACTACAGTGAGGCGGTGCGCGAGTCCAGGCCAGGGAAGAGTGCGGCTGTGTCGAGCCTCGACTGCCTGTCCAGCATAGTGGAGCGCATCTCCACAGACAGCCCCGCTGCGCCTGCG     960
AlaTyrTyrSerGluAlaValArgGluSerArgProGlyLysSerAlaAlaValSerSerLeuAspCysLeuSerSerIleValGluArgIleSerThrAspSerProAlaAlaProAla     267

CTGCTTTTGGCAGATGCACCACCAGAGTCGCCTCCGGGTCCGCCAGAGGGGGCATCCCTAAGCGACACAGAACAGGGAACCCAGACCCCGTCTCCCGACGCCGCCCCTCAGTGTCCTGCA    1080
LeuLeuLeuAlaAspAlaProProGluSerProProGlyProProGluGlyAlaSerLeuSerAspThrGluGlnGlyThrGlnThrProSerProAspAlaAlaProGlnCysProAla     307

GGCTCAAACCCCAATGCGATTTATCAGGTGCTTTGAGAGATCGACTGCAGCAGCAGAGGGCGCACCACCGTAGGCACTCCTGGGGATGGTGCCCCTGGTTCTTCACGCCCAAAAGATGAA    1200
GlySerAsnProAsnAlaIleTyrGlnValLeu                                                                                              318

GCTTAAATGACACTCTTCCCAACTGTCCTTTCGAAGCCGTTCTTCCAGAGGGAAGGGAAGAGCAGAAGTCTGTCCTAGATCCAGCCCCAAAGAAAGGACATAGTCCTTTTTGTTGTTGTT    1320

GTTGTAGTCCTTCAGTTGTTTGTTTGTTTTTCATGCGGCTCACAGCGAAGGCCACTTGCACTCTGGCTGCACCTCACTGGGCCAGAGCTGATCCTTGAGTGGCCAGGCGCTCTTCCTTT    1440

CCTCATAGCACAGGGGTGAGCCTTGCACACCTAAGCCCTGCCCTCCACATCCTTTTGTTTGTCACTTTCTGGAGCCCTCCTGGCACCCACTTTTCCCCACAGCTTGCGGAGGCCACTCAG    1560

GTCTCAGGTGTAACAGGTGTAACCATACCCCACTCTCCCCCTTCCCGCGGTTCAGGACCACTTATTTTTTTATATAAGACTTTTGTAATCTATTCGTGTAAATAAGAGTTGCTTGGCCAG    1680

AGCGGGAGCCCCTTGGGCTATATTTATCTCCCAGGCATGCTGTGTAGTGCAACAAAAACTTTGTATGTTTATTCCTCAAGCGGGCGAGTCAGGTGTTGGAAATCC                    1785
```

Figure 8. Nucleotide Sequence of the MyoD1 cDNA and Its Predicted Amino Acid Sequence

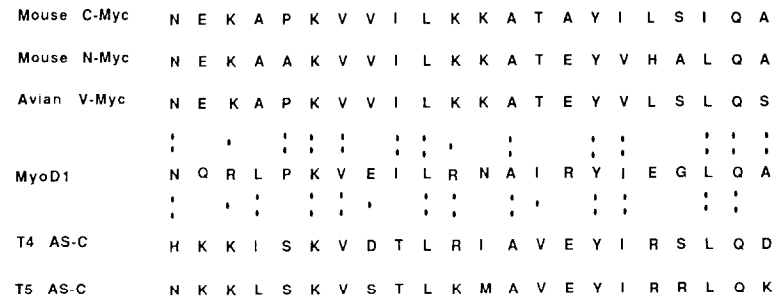

Figure 9. Amino Acid Similarities in MyoD1, Myc, and *achaete-scute* Genes

Amino acids shared by MyoD1 (aa. 141–162) mouse c-*myc* (aa. 387–408, Bernard et al., 1983), mouse N-*myc* (aa. 412–433; DePinho et al., 1986), avian V-*myc* (aa. 373–394; Alitalo et al., 1983), and genes in the *achaete-scute* locus (aa. 143–164 in T4, aa. 71–92 in T5; Villares and Cabrera, 1987) are indicated. Identical amino acids shared by MyoD1 and at least one *myc* or *achaete-scute* gene are indicated by two dashes; conservative amino acid substitutions in MyoD1 are indicated by a single dash.

though it lacks certain conserved residues. This region is followed by a highly basic region (residues 102–124). In a homology search of the Protein Identification Resource (National Biomedical Research Foundation at Georgetown University Medical Center, Washington, D.C.), it was found that the basic region is just amino-terminal to a short segment (22 residues) that shows strong similarity to a conserved segment of mouse, chicken, and human c-*myc* proteins, as well as avian v-*myc* and mouse N-*myc*. Figure 9 outlines this segment. This region of the human c-*myc* protein has been shown, by insertion and deletion mutagenesis, to be critical to its transforming potential in vitro (Stone et al., 1987). We also noticed that this same segment is shared by two recently sequenced homologous transcription units of the Drosophila *achaete-scute* complex (Villares and Cabrera, 1987), whose expression is correlated with neurogenic lineage development.

Discussion

MyoD1, a Gene that Converts Fibroblasts to Myoblasts

Using a subtracted cDNA screening procedure, we isolated a single myoblast-specific cDNA whose expression in transfected fibroblast-like 10T1/2 cells converts them to stable myoblasts. We would like to name the gene encoding this sequence MyoD1, for myoblast determination gene number 1. Myogenesis is also evident, although to a lesser extent, in several other fibroblast and adipoblast cell lines when transfected with MyoD1. MyoD1 RNA is expressed in normal skeletal muscle and has not been detected in other tissues of the mouse. Hence, we believe (but have not proven) that MyoD1 also functions to activate normal myogenesis, where it is also known that fibroblasts and myoblasts are derived from a common precursor (Dienstman and Holtzer, 1975).

Several factors aided us in identifying MyoD1. A key assumption in our analysis was that a regulator of myogenic determination would be present in undifferentiated myoblasts. This assumption allowed us to screen out differentiation-specific muscle structural markers. In addition, the nondifferentiating myoblast variants were extremely useful for screening out intracisternal type A particle and other "nonspecific" genes activated by 5-azacytidine. Finally, we were fortunate that the abundance of MyoD1 mRNA in the cell allowed its detection (four times) in the 10^4 cDNA clones screened. How general can one expect the subtracted hybridization procedure to be for identifying sequences that regulate determination or differentiation? We believe that the real limitations are biological—to find two appropriate cell types, or preferably, noninducing variants of a single cultured cell type. Clearly, in our case, the myogenic conversion of 10T1/2 cells induced by genomic myoblast DNA transfection and the high frequency of myogenesis induced by 5-azacytidine (Taylor and Jones, 1982; Konieczny and Emerson, 1984) were important clues for subsequent functional analysis, since they suggested that only one event was required. In this respect, it is of interest to point out that the comparison between 10T1/2 cells and myoblasts is very different from, for example, more differentiated cultured cell lines (e.g., MEL cells, HL60 cells, or pre-B cells) in induced versus uninduced conditions. In these instances, regulatory changes may depend more on modification of a repertoire of regulatory proteins already present (Sen and Baltimore, 1986).

MyoD1 Seems to Antagonize Cell Growth

The "phenotype" of 10T1/2 cultures cotransfected with MyoD1 presents two complications: the G418 resistance transformation frequency is about 1/10 that of controls, and only 50% of all viable clones are positive for myogenesis. The second point is best explained by variability in expression of the cotransfected MyoD1 expression vehicle since nonmyogenic transfected colonies do not express significant levels of MyoD1 mRNA while myogenic colonies do. This variability in expression is probably a consequence of both a failure of successful cotransfection and a failure to express the cotransfected vector due, for example, to chromosomal position effects. One clone that initially fused eventually lost the capacity to fuse, and, when assayed, failed to express significant levels of MyoD1. Nondifferentiating aza-myoblast variants also failed to express significant levels of MyoD1. It would seem then that MyoD1 expression is required to maintain myogenic determination potential, at least for 10T1/2 derived myoblasts.

Why is the transformation frequency reduced with MyoD1 cotransfection? We would speculate that one possibility is that MyoD1 expression rapidly induces most recipient 10T1/2 cells to commit to myogenesis. As a consequence they withdraw from the division cycle and fail to form colonies. Implicit in this explanation is that there is a competition between growth and myogenesis and that

when MyoD1 is expressed at high levels, those initial transfectants differentiate, even in the presence of high serum. When MyoD1 is expressed at moderate levels, differentiation occurs only in low serum, and when expressed at low levels, myogenesis does not occur at all. In support of this notion, we observed that the level of exogenous MyoD1 mRNA in two transfected myogenic clones is nearly equivalent to the endogenous mRNA levels expressed in both aza and C2C12 myoblasts.

In preliminary experiments, we tested to see if MyoD1 expression inhibited colony formation because of early differentiation by using factors that promote continued proliferation by an independent mechanism and block myogenic differentiation (Linkhart et al., 1981; Lathrop et al., 1985; Clegg et al., 1987). Using one such reagent, fibroblast growth factor (FGF), we were able to increase the transformation frequency to control levels. Also, in parallel experiments, the number of myogenic colonies increased. That some of the MyoD1 transfected 10T1/2 cells differentiated early implies that in the absence of FGF we should be able to detect them as small clusters of cells expressing muscle markers. Indeed, isolated clusters of differentiated cells were observed, and their frequency was roughly what was expected based on the transformation frequency using control DNA. Finally, we performed transient transfections of 10T1/2 cells with MyoD1 and used immunofluorescence to detect a small percentage of cells expressing desmin and myosin heavy chain.

Interestingly, MyoD1 transfection activates the continued expression of two mRNAs (MyoA and MyoH) in proliferating myoblasts that are absent in proliferating 10T1/2 cells and are otherwise expressed only in growth-arrested 10T1/2 cells. The expression in proliferating myoblasts of these growth regulated mRNAs further suggests that MyoD1 may antagonize or dampen mitogenic signals (see below).

MyoD1 Shares a Region of Amino Acid Similarity with *myc*

The similarity of a MyoD1 coding region segment with *myc* proteins may be important in ultimately explaining the mutual antagonism between growth and differentiation illustrated in this system. It has been observed that c-*myc* is induced in a variety of cell types when they enter the cell cycle from a quiescent state (Kelly et al., 1983; Thompson et al., 1985). In addition, c-*myc* is often down-regulated as cells terminally differentiate in vitro (Westin et al., 1982; Lachman and Skoultchi, 1984; Dony et al., 1985; Sejersen et al., 1985; Endo and Nadal-Ginard, 1986) and in vivo (Conklin and Groudine, 1986; Schneider et al., 1986). Recent experiments employing a transfected, constitutively expressed *myc* gene suggest that this down-regulation may be a necessary factor in differentiation (Coppola and Cole, 1986; Dmitrovsky et al., 1986; Prochownik and Kukowska, 1986). Additionally, v-*myc* expression in MC29 transformed quail myoblasts inhibits myogenic differentiation (Falcone et al., 1985). Insertion and deletion mutation analysis of the region of human c-*myc* protein containing the shared sequence with MyoD1 has revealed that the loss of this segment or small nearby insertions abolishes *myc* transforming activity in vitro (Stone et al., 1987). While direct or indirect (i.e., via binding to a common protein or ligand) interaction of MyoD1 with c-*myc* based on a single, small conserved segment is highly speculative, it does provide one focus for further analysis of MyoD1 function by site-directed mutagenesis. It is striking that the gene products of the *achaete-scute* locus, which play a role in Drosophila neurogenic cell determination, also share the same common region of *myc* homology with MyoD1. A common amino acid sequence in determination factors of at least two cell lineages (myogenic and neurogenic) raises the possibility that other cell lineages may use analogous determination factors that share this sequence.

The Genetic Basis for Myogenic Conversion

Our present data neither support nor refute the possibility that MyoD1 is the gene directly activated by 5-azacytidine in 10T1/2 cells, and additional experiments are in progress to address this question. Some evidence comes from preliminary observations that a specific HpaII fragment detected by MyoD1 (in a Southern blot) is observed in aza-myoblasts but not in 10T1/2 cells, suggesting preferential methylation of this site(s) in 10T1/2 cells.

Is myogenesis controlled by a single master gene? At present, it is unclear if MyoD1 expression alone is sufficient to activate the entire myogenic program. Whereas MyoD1 expression induces myogenesis in several "fibroblast" and adipoblast cell lines, the extent of differentiation within a myogenic colony differs with each line and is in most cases much less than that observed in transfected 10T1/2 cells. Furthermore, a monkey kidney cell line (CV1), expressing transfected MyoD1 mRNA, shows no evidence at all of myogenic transformation. We do not understand this variability in "penetrance" between different recipient cell lines. It is possible that 10T1/2 cells already contain some additional elements that cooperate with MyoD1 for determining the myogenic phenotype. Alternatively, it is possible that other cell lines contain an inhibitor of myogenesis which results in a decreased probability of differentiation (depending on the "balance" between positive and negative factors in individual cells of a colony) and that the argument for several positive myogenic conversion factors is incorrect. Heterokaryon experiments in which myotubes were fused with a variety of cell types have provided evidence for both positive factors in muscle cells (Blau et al., 1983; Wright, 1984a; Blau et al., 1985) as well as inhibitory factors in nonmyogenic cells (Wright and Aronoff, 1983; Lawrence and Coleman, 1984; for a discussion of positive and negative differentiation factors see Wright, 1984b, 1985).

The significant reduction of MyoD1 expression in *all* nondifferentiating aza-myoblast variants suggests that MyoD1 expression (or factors that regulate it) may alone be necessary for myogenic determination of these cells. Alternatively, this locus and/or one controlling it may be unusually sensitive to inactivation, either because it is the most distal (and therefore dependent) element in a series of operations or because it is simply more sensitive due to its molecular properties. We are currently researching

if reexpression of this gene in the variant lines will reactivate the muscle program. In so far as MyoD1 does not promote increased adipogenesis in transfected 10T1/2 cells (data not shown), it is unlikely that MyoD1 is a general differentiation agent. That MyoD1 expression is limited (as far as we can detect) to skeletal muscle and myoblasts supports this notion. It is possible, however, that MyoD1 is expressed transiently in many lineages during development, yet its continued expression is detectable only in skeletal muscle, which is known to retain muscle stem cells or satellite cells.

Previously, we suggested the possibility that 10T1/2 cells may be determined myoblasts that have lost their ability to differentiate as a result of the activation of an oncogene (Lassar et al., 1986). In this particular instance, myogenesis in 10T1/2 cells could result from expression of a suppressor of such an activated oncogene. Although we cannot rule out this scenario, both the restriction of MyoD1 expression to skeletal muscle and its ability to induce myogenesis in a number of different cell lines argue against this gene being simply an "anti-oncogene."

There is a great deal of evidence in C. elegans and Drosophila that regulatory genes are expressed in several cell lineages, supporting current and classical notions that cell type would be encoded by a subset of combinations from a pool of genomic regulatory genes. There is less evidence for cell type-specific "determination genes" being expressed in only *one* lineage or cell type and also being cell autonomous in that lineage. On the other hand, current DNA binding studies do detect DNA binding activities that are present only in specific differentiated cells. There are also a few "determination" genes that do seem to affect only a single type of cell (e.g., mec3 in C. elegans, *achaete-scute* and sevenless in Drosophila, and genes involved in germ-line sex determination). Possibly, combinatorial schemes (as exemplified in Drosophila segmentation) and cell lineage (as illustrated in C. elegans) are both used to generate positional identity (for example, see Wolpert and Stein, 1984) in the embryo for each individual cell. Eventually, this information, together with input from local cues, could focus down to the activation of only one or a few specific genes that are ultimately responsible for determining a specific cell type.

Experimental Procedures

Cells

C3H10T1/2, clone 8, Swiss 3T3, and Swiss 3T3-L1 adipocytes (Green and Kehinde, 1974) were from the American Type Culture Collection. 5-azacytidine-derived myoblast lines were generated by exposing C3H10T1/2 cells to 3 μM 5-azacytidine for 24 hr (Lassar et al., 1986). Resulting myogenic colonies were subcloned five times, and two independently isolated aza-myoblast lines (F3 and P2) were established. C2C12 myoblasts (Blau et al., 1983) were provided by Helen Blau. 3T3 Swiss-C2 and 3T3-F442A cells (Green and Kehinde, 1976) were provided by Howard Green. L cells (TK^-, $APRT^-$) were provided by Richard Axel. TA1 cells (Chapman et al., 1984) were obtained from Gordon Ringold. Other cell lines are described in the figure legends. All cell lines were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 15% fetal calf serum (Hyclone Laboratories). Myogenic differentiation was induced by incubation in DMEM with 2% heat-inactivated horse serum (differentiation medium).

Subtracted cDNA Probe Preparation

Myoblast cDNA was prepared using poly(A)$^+$ RNA isolated from cloned myoblast lines. Synthesis included ^{32}P-dCTP (800Ci/mmol, NEN) and AMV reverse transcriptase (Siekagaku America, Inc.) as described (Davis et al., 1984). Initial specific activities of these probes were approximately 3 × 10^8 dpm/μg. Subtracted hybridizations were performed (following the protocol of Davis et al., 1984) in 8 μl of 0.5 M NaH_2PO_4/Na_2HPO_4 (pH 7.0), 0.1% SDS, 5 mM EDTA at 68°C under paraffin oil. Reactions achieved C_0t values of 3000 or greater. Selection for single-stranded material was performed over hydroxyapatite equilibrated in 0.12 M NaH_2PO_4/Na_2HPO_4 (pH 7.0), 0.1% SDS at 60°C in a water-jacketed column (Bio-Rad). Selection for double-stranded material in the positive selection was performed after binding to hydroxyapatite by elution with 0.5 M NaH_2PO_4/Na_2HPO_4 (pH 7.0), 0.1% SDS at 60°C. The percentage of remaining material from the initial screening (twice subtracted with C3H10T1/2 poly(A)$^+$ RNA and once positively selected with myoblast poly(A)$^+$ RNA) was 1.15% for the aza-myoblast probe and 2.40% for the C2C12 myoblast probe.

Library Screening

The F3 aza-myoblast cDNA library was prepared by standard methods in λgt10 (Gubler and Hoffman, 1983; Huynh et al., 1985). Plaque hybridizations were performed using replica nitrocellulose filters from plates containing 1–1.5 × 10^4 cDNA clones. Approximately 1 × 10^5–3 × 10^5 dpm of subtracted or unsubtracted probes were utilized. The hybridizations were carried out for three days at 42°C in 25 mM NaH_2PO_4/Na_2HPO_4 (pH 7.0), 5× SSC (1× SSC is 150 mM NaCl, 15 mM Na-citrate), 1× Denhardt's solution (0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% BSA), 250 μg/ml total Torula RNA, 50% Formamide, with 10% dextran sulfate and 2% SDS. Washings were performed in 2× SSC, 0.1% SDS for 45 min at 60°C and in 0.4× NPE (1× NPE is 150 mM NaCl, 50 mM NaH_2PO_4/Na_2HPO_4 [pH 7.0], 1 mM EDTA), 0.1% SDS for 45 min at 60°C, followed by autoradiography.

cDNA Clone Selection

Ninety-two cDNA clones, representing about 10^4 cDNA clones plated, were initially picked. The plaque density of the initial filters required a secondary screen to isolate individual clones. The secondary screen involved a subtracted "plus/minus" type of procedure. The 92 initial phage isolates were individually gridded (75 times each) and screened on triplicate nitrocellulose filters with either ^{32}P-labeled aza-myoblast cDNA subtracted once with 10T1/2 poly(A)$^+$ RNA, ^{32}P-labeled C2C12 cDNA subtracted similarly (plus probes), or ^{32}P-labeled 10T1/2 cDNA (minus probe). These probes were prepared as described above. Phage DNA was purified from clones showing myoblast-specific hybridization, and the cDNA inserts isolated by EcoRI digestion and agarose gel electrophoresis, followed by binding to and elution from DE81 paper (Whatman). We used the following initial criteria in selecting cDNAs for further analysis: lack of expression in proliferating 10T1/2 cells; little to no increase in expression in differentiated myotubes versus proliferating myoblasts (which is characteristic of terminal differentiation markers); and reduced or absent expression in nondifferentiating variants of myogenic cells. The 92 cDNAs are listed by frequency of actual isolation as follows: MyoH (37); intracisternal type A particle RNA (21); MyoA (3); MyoD (4); six different clones showing expression characteristic of terminal differentiation markers (8); five different clones expressed at low levels in 10T1/2 cells, higher levels in myogenic cells, and diminishing only slightly or not at all in differentiation-defective variants (5); one clone induced with growth arrest in 10T1/2 cells and several other 3T3 lines (1); eight different clones that showed little to no differential expression on Northern analysis (9); and two different clones containing apparently repetitive sequences by Northern analysis (4).

Derivation of Nondifferentiating Variants of Aza-Myoblasts

A 5-azacytidine-derived myogenic line (P2) was allowed to reach confluence in DMEM plus 15% fetal calf serum. The media was then switched to DMEM plus 2% heat-inactivated horse serum. After 3 days, when most of the cells had terminally withdrawn from the cell cycle, fused, and fully differentiated, the cells were trypsinized and replated at low density. Cells able to proliferate were passaged several more times following the above regimen and then cloned. When main-

tained at confluence in mitogen-depleted media, most of these clones failed to fuse or immunostain for myosin heavy chain.

RNA Preparation

Total cellular RNA from adherent tissue culture cells was prepared by washing plates with Tris-buffered saline followed by overlaying guanidine lysis buffer of 5 M guanidine isothiocyanate, 25 mM Na-Citrate (pH 7.0), 0.1 M 2-mercaptoethanol, 0.5% Sarkosyl. Lysates were collected and centrifuged over cushions of 5.7 M CsCl (pH 7.0), 0.1 M EDTA at 25 Krpm in a Beckman SW28.1 rotor for 18 hr at 20°C. Pellets were resuspended in 10 mM Tris (pH 7.0), 0.5% SDS, 5 mM EDTA, extracted once with phenol:chloroform:isoamyl alcohol (25:24:1) and once with chloroform:isoamyl alcohol (24:1), ethanol-precipitated, and stored in diethylpyrocarbonate treated water at −70°C. Whole mouse tissue RNA was prepared similarly. Six-day or 9 week old BALB/C mice were sacrificed, and dissected tissues were washed in Tris-buffered saline, frozen in liquid nitrogen, and homogenized in guanidine lysis buffer. Poly(A)$^+$ RNA selection of total cellular RNA was performed by standard methods (Maniatis et al., 1982).

RNA Analysis

Five micrograms of total cell RNA were used for Northern analysis on 1.5% agarose gels containing 6.7% formaldehyde. Integrity and relative amounts of RNA were checked by ethidium-bromide staining of ribosomal RNA on parallel gels. Gels were treated with 50 mM NaOH, 10 mM NaCl for 40 min, neutralized in 100 mM Tris (pH 7.4), 20× SSC for 40 min, and transferred for 16 hr to Gene Screen (NEN Research Products) in 20× SSC. RNA was cross-linked by exposure to UV. Hybridizations and washings were performed as described for the library screening, using 5–10 × 10^6 dpm of probes prepared by random primer synthesis (Boehringer Mannheim). The mouse myosin heavy chain, myosin light chain 2, and myosin light chain 1 probes had each been isolated from the F3 aza-myoblast/myotube cDNA library. The identity of these cDNAs was determined by hybridization to rat probes kindly provided by Bernardo Nadal-Ginard. Intracisternal type A (IAP) cDNAs were identified by cross-hybridization to an IAP probe kindly provided by Kira Lueders.

Expression of cDNAs

MyoD, MyoA, and MyoH cDNAs were isolated from λgt10 clones. EcoRI cut λDNA was subjected to low melting point agarose gel electrophoresis, followed by extraction with 1% CTAB (hexadecyltrimethylammonium bromide) in water-saturated n-butanol. Fragments were cloned into EcoRI cut pEMSVscribe (for expression) or EcoRI cut pVZ-1 (for sequencing). Orientations were determined by preparing strand-specific ribo-probes for MyoA, MyoD, and MyoH. pEMSVscribe was generated by first digesting Bluescribe M13(+) (Stratagene) with EcoRI, filling in the ends with the Klenow fragment of E. coli pol I (Boehringer Mannheim), and religating to create Bluescribe M13(+)EcoRI−. This vector was then cut with HindIII, and the HindIII fragment of EMSV-33 (Harland and Weintraub, 1985) containing the Moloney sarcoma virus LTR, and SV40 poly(A) addition signal was inserted by religation. This construct allows unique EcoRI insertion between the LTR and poly(A) signal sequence, which are themselves flanked by T7 and T3 promoters. pVZ-1 (Henikoff and Eghtedarzadeh, 1987) is a modified version of Bluescribe M13(+) employed to generate unidirectional plasmid deletions.

Transfections

pCMV-*neo*, 0.1–0.2 μg (Linial, 1987) or pLNSAL (Palmer et al., 1987) were used for cotransfection to G418 resistance. One to ten micrograms of the expression vehicles or control vehicles and 10–20 μg 10T1/2 high molecular weight DNA as carrier was used. Actual ratios are described in the text. Cells, 1.5–3.0 × 10^5, were plated 2 days prior to transfection on 10 cm diameter dishes. Cells were refed with DMEM containing 10% fetal calf serum 3 hr prior to transfection. The DNAs were combined and added to cells as a Ca–PO_4 precipitate (Wigler et al., 1979) with 10 μM chloroquine. The DNA was removed 24 hr later with addition of fresh medium as above. Twenty-four hours after this, the cells were split 1:10 to 10 cm diameter or 15 cm diameter gelatin coated dishes with DMEM containing 15% fetal calf serum and 0.4 mg/ml G418 (Geneticin, Gibco). To score for myogenesis, after colonies of about 10^4 cells had grown, the media was switched to DMEM containing 2% heat-inactivated horse serum with 0.2 mg/ml G418. To score for adipogenesis, the media was switched to DMEM containing 15% fetal calf serum, 1 μM Dexamethasone, 5 μg/ml insulin with 0.2 mg/ml G418.

Fixing and Immunostaining of Colonies

To score for myogenesis, plates were fixed 3–5 days after switching to horse serum containing media. Plates were washed once with Tris-buffered saline and incubated for 1 min in AFA (70% ethanol, 3.7% formaldehyde, 5% glacial acetic acid), followed by extensive washing with phosphate-buffered saline (PBS). Plates were stored at 4°C with PBS and 0.02% sodium azide. Fixed colonies were incubated with the anti-myosin heavy chain monoclonal antibody MF-20 (Bader et al., 1982). Immunostaining was performed with the alkaline phosphatase Vectastain ABC kit (Vector Laboratories, Inc.). Plates were examined at 50× and 20× magnification. For adipogenesis scoring, plates were fixed 7 days after switching to dexamethasone insulin-containing media as above, except that formalin was used instead of AFA. Adipogenic colonies were visualized using Oil Red-O.

Sequencing

Overlapping deletions of MyoD cloned into pVZ-1 in both orientations were obtained by unidirectional digestion methods (Henikoff, 1987). Deletion clones were sized on agarose gels and selected clones were used to prepare single-stranded DNA using an M13 helper phage and the E. coli strain TG1/λ (kindly provided by T. St. John). Single-stranded DNA was used for dideoxy sequencing using the T7 Sequenase kit (United States Biochemical Corp.) and α-thio[^{35}S]dATP (1000 Ci/mmol, NEN). Reactions were run on 6% acrylamide gradient gels, followed by autoradiography. Sequence recording and analysis was performed using a digitizer and Genepro software (Joe Brown, Oncogen, Seattle, WA, and Jim Wallace, Fred Hutchinson Cancer Research Center, Seattle, WA; version 4.1, 1987).

Acknowledgments

We thank Mark Groudine and Richard Harland for help with constructing the cDNA library; John McKay for assistance with cloning the aza-myoblast lines; Helen Blau, Howard Green, and Gordon Ringold for cell lines; Steve Henikoff, Maxine Linial, Kira Lueders, Dusty Miller, Bernardo Nadal-Ginard, and Shirley Tilghman for plasmids; Don Fishman and Frank Stockdale for antibodies; Pei Feng Cheng for outstanding technical assistance; Helen Devitt for typing this manuscript; and our colleagues in the Department of Genetics for their thoughtful comments on the manuscript. A. B. L. expresses his appreciation to Hazel Sive and Rufus Mandelbaum for advice and encouragement. R. L. D. is supported by the National Institutes of Health Medical Scientist Training Program. During the course of this study A. B. L. was supported by postdoctoral fellowships from the Jane Coffin Childs Memorial Fund for Medical Research and the Muscular Dystrophy Association; he is currently a Lucille P. Markey Scholar and this work was supported in part by a grant from the Lucille P. Markey Charitable Trust. H. W. would like to dedicate this paper to B. W.; R. L. D. would like to dedicate to H. D. This work was supported by a grant to H. W. from the National Institutes of Health.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 20, 1987.

References

Alitalo, K., Bishop, J. M., Smith, D. H., Chen, E. Y., Colby, W. W., and Levinson, A. D. (1983). Nucleotide sequence of the v-*myc* oncogene of avian retrovirus MC29. Proc. Natl. Acad. Sci. USA *80*, 100–104.

Bader, D., Masaki, T., and Fischman, D. A. (1982). Immunochemical analysis of myosin heavy chain during avian myogenesis in vivo and in vitro. J. Cell Biol. *95*, 763–770.

Berg, J. M. (1986). Potential metal-binding domains in nucleic acid binding proteins. Science *232*, 485–487.

Bernard, O., Cory, S., Gerondakis, S., Webb, E., and Adams, J. M. (1983). Sequence of the murine and human cellular *myc* oncogenes and two modes of *myc* transcription resulting from chromosome translocation in B lymphoid tumours. EMBO J. *2*, 2375–2383.

Blau, H. M., Chiu, C.-P., and Webster, C. (1983). Cytoplasmic activation of human nuclear genes in stable heterocaryons. Cell *32*, 1171–1180.

Blau, H. M., Pavlath, G. K., Hardeman, E. C., Chiu, C.-P., Silberstein, L., Webster, S. G., Miller, S. C., and Webster, C. (1985). Plasticity of the differentiated state. Science *230*, 758–766.

Chapman, A. B., Knight, D. M., Dieckmann, B. S., and Ringold, G. M. (1984). Analysis of gene expression during differentiation of adipogenic cells in culture and hormonal control of the developmental program. J. Biol. Chem. *259*, 15548–15555.

Clegg, C. H., Linkhart, T. A., Olwin, B. B., and Hauschka, S. D. (1987). Growth factor control of skeletal muscle differentiation: commitment to terminal differentiation occurs in G1 phase and is repressed by fibroblast growth factor. J. Cell Biol. *105*, 949–956.

Conklin, K. F., and Groudine, M. (1986). Varied interactions between proviruses and adjacent host chromatin. Mol. Cell. Biol. *6*, 3999–4007.

Coppola, J. A., and Cole, M. D. (1986). Constitutive c-*myc* oncogene expression blocks mouse erythroleukaemia cell differentiation but not commitment. Nature *320*, 760–763.

Davis, M. M., Cohen, D. I., Nielsen, E. A., Steinmetz, M., Paul, W. E., and Hood, L. (1984). Cell-type-specific cDNA probes and the murine I region: the localization and orientation of A_α^d. Proc. Natl. Acad. Sci. USA *81*, 2194–2198.

DePinho, R. A., Legouy, E., Feldman, L. B., Kohl, N. E., Yancopoulos, G. D., and Alt, F. W. (1986). Structure and expression of the murine N-*myc* gene. Proc. Natl. Acad. Sci. USA *83*, 1827–1831.

Dienstman, S. R., and Holtzer, H. (1975). Myogenesis: a cell lineage interpretation. In Results and Problems in Cell Differentiation, Volume 7, J. Reinert and H. Holtzer, eds. (New York: Springer-Verlag), pp. 1–25.

Dmitrovsky, E., Kuehl, W. M., Hollis, G. F., Kirsch, H. R., Bender, T. P., and Segal, S. (1986). Expression of a transfected human c-*myc* oncogene inhibits differentiation of a mouse erythroleukaemia cell line. Nature *322*, 748–750.

Dony, C., Kessel, M., and Gruss, P. (1985). Post-transcriptional control of *myc* and p53 expression during differentiation of the embryonal carcinoma cell line F9. Nature *317*, 636–639.

Endo, T., and Nadal-Ginard, B. (1986). Transcriptional and posttranscriptional control of c-*myc* during myogenesis: its mRNA remains inducible in differentiated cells and does not suppress the differentiated phenotype. Mol. Cell. Biol. *6*, 1412–1421.

Falcone, G., Tato, F., and Alema, S. (1985). Distinctive effects of the viral oncogenes *myc*, *erb*, *fps*, and *src* on the differentiation program of quail myogenic cells. Proc. Natl. Acad. Sci. USA *82*, 426–430.

Green, H., and Kehinde, O. (1974). Sublines of mouse 3T3 cells that accumulate lipid. Cell *1*, 113–116.

Green, H., and Kehinde, O. (1976). Spontaneous heritable changes leading to increased adipose conversion in 3T3 cells. Cell *7*, 105–113.

Groudine, M., Eisenman, R., and Weintraub, H. (1981). Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation. Nature *292*, 311–317.

Gubler, U., and Hoffman, B. J. (1983). A simple and very efficient method for generating cDNA libraries. Gene *25*, 262–269.

Harland, R., and Weintraub, H. (1985). Translation of mRNA injected into Xenopus oocytes is specifically inhibited by antisense RNA. J. Cell Biol. *101*, 1094–1099.

Hedrick, S. M., Cohen, D. I., Nielsen, E. A., and Davis, M. M. (1984). Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature *308*, 149–153.

Henikoff, S. (1987). Unidirectional digestion with exonuclease III in DNA sequence analysis. Meth. Enzymol. *155*, 156–165.

Henikoff, S., and Eghtedarzadeh, M. K. (1987). Conserved arrangement of nested genes at the Drosophila *Gart* locus. Genetics, in press.

Hsiao, W.-L.W., Gattoni-Celli, S., and Weinstein, I. B. (1986). Effects of 5-azacytidine on expression of endogenous retrovirus-related sequences in C3H10T1/2 cells. J. Virol. *57*, 1119–1126.

Huynh, T. V., Young, R. A., and Davis, R. W. (1985). DNA cloning: a practical approach, Volume 1, D. M. Glover, ed. (Oxford: IRL Press), pp. 49–78.

Jones, P. A., and Taylor, S. M. (1980). Cellular differentiation, cytidine analogs and DNA methylation. Cell *20*, 85–93.

Kaufman, S. J., Foster, R. F., Haye, K. R., and Faiman, L. E. (1985). Expression of a developmentally regulated antigen on the surface of skeletal and cardiac muscle cells. J. Cell Biol. *100*, 1977–1987.

Kelly, K., Cochran, B. H., Stiles, C. D., and Leder, P. (1983). Cell-specific regulation of the c-*myc* gene by lymphocyte mitogens and platelet-derived growth factor. Cell *35*, 603–610.

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

Konieczny, S. F., Baldwin, A. S., and Emerson, C. P., Jr. (1985). Myogenic determination and differentiation of 10T1/2 cell lineages: evidence for a simple genetic regulatory system. In Molecular Biology of Muscle Development, UCLA Symposium on Molecular and Cellular Biology, New Series Vol. 29, C. Emerson, D. A. Fischman, B. Nadal-Ginard, and M. A. Q. Siddiqui, eds. (New York: Alan R. Liss, Inc.), pp. 21–34.

Lachman, H. M., and Skoultchi, A. I. (1984). Expression of c-*myc* changes during differentiation of mouse erythroleukemia cells. Nature *310*, 592–594.

Lassar, A. B., Paterson, B. M., and Weintraub, H. (1986). Transfection of a DNA locus that mediates the conversion of 10T1/2 fibroblasts to myoblasts. Cell *47*, 649–656.

Lathrop, B., Olson, E., and Glaser, L. (1985). Control by fibroblast growth factor of differentiation in the BC_3H1 muscle cell line. J. Cell Biol. *100*, 1540–1547.

Lawrence, J. B., and Coleman, J. R. (1984). Extinction of muscle-specific properties in somatic cell heterokaryons. Dev. Biol. *101*, 463–476.

Linial, M. (1987). Creation of a processed pseudogene by retroviral infection. Cell *49*, 93–102.

Linkhart, T. A., Clegg, C. H., and Hauschka, S. D. (1981). Myogenic differentiation in permanent clonal mouse myoblast cell lines: regulation by macromolecular growth factors in the culture medium. Dev. Biol. *86*, 19–30.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Pachnis, V., Belayew, A., and Tilghman, S. M. (1984). Locus unlinked to alpha-fetoprotein under control of the murine *raf* and *Rif* genes. Proc. Natl. Acad. Sci. USA *81*, 5523–5527.

Palmer, T. D., Hock, R. A., Osborne, W. R. A., and Miller, A. D. (1987). Efficient retrovirus-mediated transfer and expression of a human adenosine deaminase gene in diploid skin fibroblasts from an adenosine deaminase-deficient human. Proc. Natl. Acad. Sci. USA *84*, 1055–1059.

Prochownik, E. V., and Kukowska, J. (1986). Deregulated expression of c-*myc* by murine erythroleukaemia cells prevents differentiation. Nature *322*, 848–850.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463–5467.

Sasse, J., Horwitz, A., Pacifici, M., and Holtzer, H. (1984). Separation of precursor myogenic and chondrogenic cells in early limb bud mesenchyme by a monoclonal antibody. J. Cell Biol. *99*, 1856–1866.

Schneider, M. D., Payne, P. A., Ueno, H., Perryman, M. B., and Roberts, R. (1986). Dissociated expression of c-*myc* and a *fos*-related competence gene during cardiac myogenesis. Mol. Cell. Biol. *6*, 4140–4143.

Sejersen, T., Sumegi, J., and Ringertz, N. R. (1985). Density-dependent arrest of DNA replication is accompanied by decreased levels of c-*myc* mRNA in myogenic but not in differentiation-defective myoblasts. J. Cell. Physiol. *125*, 465–470.

Sen, R., and Baltimore, D. (1986). Inducibility of κ immunoglobulin enhancer-binding protein NF-κB by a posttranslational mechanism. Cell *47*, 921–928.

Stone, J., De Lange, T., Ramsay, G., Jakobovits, E., Bishop, J. M., Varmus, H., and Lee, W. (1987). Definition of regions in human c-*myc* that are involved in transformation and nuclear localization. Mol. Cell. Biol. *7*, 1697–1709.

Taylor, S. M., and Jones, P. A. (1979). Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine. Cell *17*, 771–779.

Taylor, S. M., and Jones, P. A. (1982). Changes in phenotypic expression in embryonic and adult cells treated with 5-azacytidine. J. Cell. Physiol. *111*, 187–194.

Thompson, C. B., Challoner, P. B., Neiman, P. E., and Groudine, M. (1985). Levels of c-*myc* oncogene mRNA are invariant throughout the cell cycle. Nature *314*, 363–366.

Villares, R., and Cabrera, C. V. (1987). The *achaete-scute* gene complex of D. melanogaster: conserved domains in a subset of genes required for neurogenesis and their homology to *myc*. Cell *50*, 415–424.

Walsh, F. S., and Ritter, M. A. (1981). Surface antigen differentiation during human myogenesis in culture. Nature *289*, 60–64.

Westin, E. H., Wong-Staal, F., Gelmann, E. P., Dalla-Favera, R., Papas, T. S., Lautenberger, J. A., Eva, A., Reddy, E. P., Tronick, S. R., Aaronson, S. A., and Gallo, R. C. (1982). Expression of cellular homologues of retroviral *onc* genes in human hematopoietic cells. Proc. Natl. Acad. Sci. USA *79*, 2490–2494.

Wigler, M., Pellicer, A., Silverstein, S., Axel, R., Urlaub, G., and Chasin, L. (1979). DNA-mediated transfer of the adenine phosphoribosyltransferase locus into mammalian cells. Proc. Natl. Acad. Sci. USA *76*, 1373–1376.

Wolpert, L., and Stein, W. (1984). Positional information and pattern formation. In Pattern Formation, G. Malacinski and S. Bryant, eds. (New York: MacMillan), p. 1.

Wright, W. E. (1984a). Induction of muscle genes in neural cells. J. Cell Biol. *98*, 427–435.

Wright, W. E. (1984b). Control of differentiation in heterokaryons and hybrids involving differentiation-defective myoblast variants. J. Cell Biol. *98*, 436–443.

Wright, W. E. (1985). The amplified expression of factors regulating myogenesis in L6 myoblasts. J. Cell Biol. *100*, 311–316.

Wright, W. E., and Aronoff, J. (1983). The suppression of myogenic function in heterokaryons formed by fusing chick myocytes to diploid rat fibroblasts. Cell Differ. *12*, 299–306.

Note Added in Proof

We have found that transfection of the MyoD1 expression vehicle into nondifferentiating aza-myoblast variants (1 and 2) induces myogenesis in these cell lines. A recent publication of the murine L-*myc* gene sequence indicates that this gene also shares sequence similarity (a. a. 318–339) with MyoD1 (Legouy et al., 1987, EMBO J. *6*, 3359–3366).