Temporal regulation of the muscle gene cascade by Macho1 and Tbx6 transcription factors in Ciona intestinalis

Jamie E. Kugler1, Stefan Gazdoiu1, Izumi Oda-Ishii1, Yale J. Passamaneck1, Albert J. Erives2 and Anna Di Gregorio1,*

1Department of Cell and Developmental Biology, Weill Medical College of Cornell University, 1300 York Avenue, Box 60, New York, NY 10065, USA
2Department of Biological Sciences, Dartmouth College, Hanover, NH 03755, USA
*Author for correspondence (and2015@med.cornell.edu)

Accepted 9 April 2010
Journal of Cell Science 123, 2453-2463
© 2010. Published by The Company of Biologists Ltd
doi:10.1242/jcs.066910

Summary
For over a century, muscle formation in the ascidian embryo has been representative of ‘mosaic’ development. The molecular basis of muscle-fate predetermination has been partly elucidated with the discovery of Macho1, a maternal zinc-finger transcription factor necessary and sufficient for primary muscle development, and of its transcriptional intermediaries Tbx6b and Tbx6c. However, the molecular mechanisms by which the maternal information is decoded by cis-regulatory modules (CRMs) associated with muscle transcription factor and structural genes, and the ways by which a seamless transition from maternal to zygotic transcription is ensured, are still mostly unclear. By combining misexpression assays with CRM analyses, we have identified the mechanisms through which Ciona Macho1 (Ci-Macho1) initiates expression of Ci-Tbx6b and Ci-Tbx6c, and we have unveiled the cross-regulatory interactions between the latter transcription factors. Knowledge acquired from the analysis of the Ci-Tbx6b CRM facilitated both the identification of a related CRM in the Ci-Tbx6c locus and the characterization of two CRMs associated with the structural muscle gene fibrillar collagen 1 (CiFCol1). We use these representative examples to reconstruct how compact CRMs orchestrate the muscle developmental program from pre-localized ooplasmic determinants to differentiated larval muscle in ascidian embryos.

Key words: Ciona intestinalis, Ascidian, Brachyury, Tbx6, Collagen, Macho, Muscle, Notochord, T-box, Transcription

Introduction
Ascidians, or tunicates, provide an essential genetic and evolutionary reference for studies of chordate development because of their well-characterized modes of development, their experimental amenability, their phylogenetic proximity to vertebrates and the organization of their larval body plan (Davidson and Christiana, 2006; Deluc et al., 2006; Passamaneck and Di Gregorio, 2005). Ascidian larvae are characterized by the presence of an axial notochord flanked by two rows of muscle cells, classified as primary (B-lineage) and secondary (A- and b-lineage) (Meedel et al., 1987), and by a rudimentary dorsal nervous system (Satoh, 1994). As shown by classical embryological studies, the development of the primary larval muscle cells in solitary ascidians proceeds cell-autonomously from the first cleavages (Meedel et al., 1987; Nishida, 1992; Ortolani, 1955). Around the neurula stage, muscle precursors begin to position themselves paraxially (e.g. Rhee et al., 2005) and continue to divide synchronously until the early tailbud stage; after which they elongate ~fourfold, in the absence of cell division and evident positional reorganization, to allow tail extension (Passamaneck et al., 2007).

Recent molecular investigations have shown that primary muscle development is initially orchestrated by a maternally deposited muscle determinant, the zinc-finger transcription factor Macho-1 (Nishida, 2002; Nishida and Sawada, 2001). In Ciona intestinalis, overexpression of Ci-macho1 and its morpholino-mediated knockdown cause ectopic expression and silencing, respectively, of the T-box transcription factor genes Ci-Tbx6b and Ci-Tbx6c, indicating that Ci-Macho1 somehow activates their early expression (Yagi et al., 2004a). In turn, Ci-Tbx6b and Ci-Tbx6c, which are considered the result of a recent lineage-specific duplication event (Dehal et al., 2002), function as mediators of Ci-Macho1, because their morpholino-induced knockdown results in the downregulation of muscle-specific structural genes, such as those encoding muscle actin, myosin chains and creatine kinase, among others (Yagi et al., 2005). Similarly, in Holocynthia roretzi, another solitary ascidian distantly related to Ciona, Hr-Macho1 has been shown to activate the expression of Hr-Tbx6, as well as that of structural muscle genes, such as Hr-muscle actin (Sawada et al., 2005). In addition to controlling primary muscle formation, Ci-Macho1 acts cooperatively with β-catenin to induce the formation of the heart field by activating Ci-Mesp (Christiaen et al., 2009; Davidson et al., 2005; Satou et al., 2004); the function of Ci-Macho1 in heart specification is also mediated by the Ci-Tbx6-related transcription factors (Christiaen et al., 2009; Davidson et al., 2005).

Despite the wealth of information on the gene regulatory network that initiates and sustains muscle development in Ciona and other ascidian embryos (Hudson and Yasuo, 2008; Imai et al., 2006; Meedel et al., 2007), the cis-regulatory mechanisms that integrate maternal and zygotic information along this complex gene cascade, from egg to swimming larva, are largely unexplored, although some common logic has been identified in a limited subset of cis-regulatory modules (CRMs) (Erives, 2009). For example, it is still unclear how Ci-Macho1 activates its transcriptional intermediaries, and information on the structure and function of the CRMs controlling expression of direct Ci-Macho1 targets, and of their downstream structural muscle genes, is limited and fragmentary.
By combining misexpression assays and CRM analyses, we have begun to address these points and to gain a mechanistic understanding of how the activity of maternal and zygotic transcription factors is coordinated at the cis-regulatory level. Through misexpression assays, we have probed the transcriptional plasticity of the *Ciona mesoderm* to assess the ability of temporally and spatially misexpressed Ci-Macho1 to ectopically activate muscle genes in notochord cells, and we have identified a cross-regulatory interaction between Ci-Tbx6b and Ci-Tbx6c. Through a combination of sequence inspection, point-mutation analyses and electrophoretic mobility assays, we have characterized CRMs from the upstream genomic regions of these two genes, which occupy important nodes in the muscle gene regulatory network, as well as two muscle CRMs from a representative structural gene, *Fibrillar Collagen-1 (CiFCol1)*, which is robustly expressed in muscle cells from late gastrulation until the late tailbud stages.

We present evidence that the expression patterns of *Ci-Tbx6b* and *CiFCol1* are recapitulated by compact bipartite CRMs, consisting of ‘early’ and ‘late’ moieties that are essential for initiation and maintenance of gene expression, respectively, and that the coordinated activity of these CRMs ensures continuity in the expression of genes necessary for muscle development and differentiation.

### Results

#### Effects of the ectopic expression of Ci-Macho1 on the transcription of *Ci-Tbx6b* and *Ci-Tbx6c*

To gain insights on the mechanisms used by Ci-Macho1 to control expression of *Ci-Tbx6b* and *Ci-Tbx6c*, we misexpressed this transcription factor spatially and temporally in notochord cells using the *Ciona Brachyury (Ci-Bra)* promoter region (Corbo et al., 1997). *Ci-macho1* is normally expressed in germ-cell precursors in 110-cell-stage embryos, and in sensory vesicle, nerve cord and germ cells in mid-tailbud embryos (Fig. 1A,C,D) (Satou et al., 2002a); no detectable expression in notochord cells was seen at any stage. For this reason, the *Ci-Bra* promoter, which is active predominantly in notochord cells and mesenchyme, and only sporadically in a few muscle cells, was chosen for the misexpression experiments. When the misexpression experiments were attempted

---

**Fig. 1. Transcriptional changes in response to ectopically expressed Ci-Macho1, Ci-Tbx6b and Ci-Tbx6c.** (A–O,Q) Microphotographs of control and transgenic *Ciona intestinalis* embryos hybridized in situ with digoxigenin-labeled antisense RNA probes. Note that in some cases, because of mosaic incorporation of the transgene(s), not all the cells of the same lineage show staining. The percentages of transgenic embryos displaying the patterns shown, and the total number of embryos scored, are indicated at the bottom of each panel. (A,B,D,F,H,J,L) Mid-tailbud embryos; individual embryos are oriented with anterior to the left, dorsal up. (A,B,G,H,M) Low-magnification microphotographs of embryos from representative experiments. (C,E,G,H,M-O,Q) 110-cell-stage embryos, vegetal views. (A–P) Effects of the misexpression of *Ci-macho1* on *Ci-Tbx6b* and *Ci-Tbx6c* expression. (A,C,D) Control embryos expressing zygotic *Ci-macho1* in germ-line precursors (yellow arrowhead), sensory vesicle and nerve cord (blue arrowheads). (B,E,F) *Bra>macho* transgenics efficiently misexpress *Ci-macho1* in notochord (red arrowheads) and mesenchyme cells (pink arrowheads) at early (E) and late stages (F). (G,H) *Bra>macho* transgenics hybridized with the *Ci-Tbx6b* probe. (I,J) *Ci-Tbx6b* is expressed at the 110-cell stage in muscle precursors (I; orange arrowheads) but is no longer expressed by the early tailbud stage (J; see also supplementary material Fig. S1A–C). (K) Ectopic expression of *Ci-Tbx6b* is detected in the notochord of 110-cell-stage *Bra>macho* transgenics (red arrowhead). (L) Mid-tailbud *Bra>macho* embryos show ectopic expression of *Ci-Tbx6b* only in mesenchyme (pink arrowhead). (M) *Bra>macho* transgenics hybridized with the *Ci-Tbx6c* probe. (N) Expression of *Ci-Tbx6c* in control embryos is detected in a subset of muscle precursors (orange arrowheads). (O) In 110-cell *Bra>macho* embryos, ectopic expression of *Ci-Tbx6c* is detected in notochord precursors (red arrowheads). (P) Changes in gene expression in 110-cell stage *Bra>macho* embryos versus control embryos, as monitored by qRT-PCR. (Q) Left panel shows a 110-cell-stage *Bra>Tbx6b* embryo hybridized with the *Ci-Tbx6b* probe. Right panel shows a 110-cell-stage *Bra>Tbx6c* embryo hybridized with the *Ci-Tbx6c* probe. Ectopic notochord staining is indicated by red arrowheads. Scale bars: 50 μm.
using the \textit{Ci-FoxA-a} promoter, which encompasses a wider expression territory (Di Gregorio et al., 2001), the effects on embryogenesis were too extensive to allow meaningful interpretation of the results (data not shown).

Transgenic embryos carrying the \textit{Bra>macho} construct (Fig. 1B) efficiently express \textit{Ci-macho1} in notochord cells and their precursors (Fig. 1E,F) and display a short, stubby tail (Fig. 1B,F,L). The 110-cell-stage transgenic \textit{Bra>macho} embryos hybridized with a probe for \textit{Ci-Tbx6b} (Fig. 1G), which is normally expressed in muscle precursors (Fig. 1I) (Takatori et al., 2004), showed ectopic expression of this gene in notochord precursors (Fig. 1K); however, at the early tailbud stage (Fig. 1H), when \textit{Ci-Tbx6b} is normally undetectable (Fig. 1J), a consistent ectopic signal was observed only in a subpopulation of trunk mesenchyme cells (Fig. 1H,L). This result can be explained by previous observations of the activity of the \textit{Ci-Bra} promoter region in mesenchyme cells (Corbo et al., 1997).

\textit{Ci-Tbx6c} (Fig. 1M-O), which is normally expressed only in a subset of muscle precursors at the 110-cell stage (Fig. 1N), was ectopically activated in transgenic \textit{Bra>macho} embryos in notochord precursors only at this stage (Fig. 1O), and was not detected at later stages (data not shown). The transcriptional changes induced by the misexpression of \textit{Ci-macho1} in 110-cell transgenic embryos were validated and quantified by qRT-PCR (Fig. 1P). Together, these results suggest that notochord precursors are competent to respond to \textit{Ci-Macho1}, although only at early stages of their development.

**Cross-regulatory interactions between \textit{Ci-Tbx6b} and \textit{Ci-Tbx6c}**

We next sought to assess whether \textit{Ci-Tbx6b} and \textit{Ci-Tbx6c}, which possess identical binding affinities in vitro (Yagi et al., 2005), are each capable of activating transcription of the other gene. To test this hypothesis, we misexpressed \textit{Ci-Tbx6b} and \textit{Ci-Tbx6c} in the notochord by cloning the respective cDNAs downstream of the \textit{Ci-Bra} promoter. As a first step, we ascertained whether these genes were efficiently transcribed in the notochord of transgenic embryos; a representative control experiment is presented in supplementary material Fig. S1. Subsequently, we hybridized embryos carrying the \textit{Bra>Tbx6b} transgene with the \textit{Ci-Tbx6c} probe, and embryos carrying the \textit{Bra>Tbx6c} transgene with the \textit{Ci-Tbx6b} probe (Fig. 1Q). We found ectopic expression of \textit{Ci-Tbx6c} in notochord precursors in embryos bearing the \textit{Bra>Tbx6b} transgene (Fig. 1Q, left panel; compare with Fig. 1N) and similarly, transcription of \textit{Ci-Tbx6b} was ectopically activated in embryos carrying the \textit{Bra>Tbx6c} transgene (Fig. 1Q, right panel; compare with Fig. 1I). Under the experimental conditions that we used, no cross-hybridization between the two probes was observed. These results indicate that these related transcription factors possess the potential for reciprocal activation.

**A bipartite CRM recapitulates spatial and temporal muscle expression of \textit{Ci-Tbx6b}**

Beginning at the 16-cell stage, \textit{Ci-Tbx6b} is specifically expressed in the precursors of the larval muscles derived from the B5.1 blastomere pair (Takatori et al., 2004). In 110-cell embryos, \textit{Ci-Tbx6b} expression expands to include numerous muscle precursors (Fig. 1I), whereas by the neurula stage it becomes confined to the posterior-most muscle cells (Takatori et al., 2004). \textit{Ci-Tbx6b} (gene model ci0100144249) is clustered in a tandem arrangement with \textit{Ci-Tbx6c} (ci0100144293) on sequence scaffold 126 (JGI v2.0; http://genome.jgi-psf.org/Cioin2/Cioin2.home.html) (Dehal et al., 2002) (Fig. 2A,S; supplementary material Fig. S2A). A genomic fragment spanning 2.4 kb was PCR-amplified from the 5'-flanking region of \textit{Ci-Tbx6b} and its various truncations were tested in vivo by parallel electroporations (e.g. Di Gregorio and Levine, 2002) to identify the minimal sequences necessary for its function. Truncated versions of the 2.4 kb fragment were cloned upstream of both the endogenous \textit{Ci-Tbx6b} promoter and the \textit{Ci-FoxA-a} basal promoter (supplementary material Fig. S2B). The 2.4 kb CRM directed strong expression in muscle and mesenchyme cells (Fig. 2B-D; supplementary material Fig. S2D-F and Table S1); occasional staining was also observed in trunk ventral cells, which are the heart precursors (data not shown). The use of the \textit{Ci-FoxA-a} basal promoter (Oda-Ishii and Di Gregorio, 2007) allowed the identification of a minimal 112 bp fragment sufficient for activity in a heterologous context (shaded area in supplementary material Fig. S2B).

Sequence analysis of the 2.4 kb \textit{Ci-Tbx6b} CRM identified three bona fide binding sites for \textit{Ci-Macho1} (green squares in Fig. 2A), all of which partially matched the consensus sequence previously identified (Yagi et al., 2004a) (Fig. 2A, bottom). We performed an electrophoretic mobility-shift assay (EMSA) to verify the functionality of these sites, and found that all three sequences were bound in vitro by a Histidine-tagged \textit{Ci-Macho1} fusion protein (supplementary material Fig. S3A), as well as by a GST-tagged \textit{Ci-Macho1} protein (data not shown), with affinities that reflected their different homologies with the published consensus sequence (Fig. 2A, bottom). The three \textit{Ci-Macho1}-binding sites were found within a 463 bp region. To verify the involvement of the \textit{Ci-Macho1}-binding sites in the temporal control of the CRM activity, we compared the developmental windows of transcriptional activity of the three \textit{Ci-Tbx6b} constructs shown in Fig. 2A. The first construct contains the 2.4 kb wild-type \textit{Ci-Tbx6b} CRM fused to \textit{lacz}, the second construct contains an 862 bp 5'-truncated version of the previous fragment, which lacks all \textit{Ci-Macho1}-binding sites, and the third construct consists of the 2.4 kb \textit{Ci-Tbx6b} CRM carrying mutations in all three \textit{Ci-Macho1}-binding sites. To accurately determine the respective windows of activity, one-cell stage \textit{Ciona} embryos were electroporated separately with each construct, cultured until the 32-cell (Fig. 2B,E,H), 110-cell (Fig. 2C,F,I) and early neurula stages (Fig. 2D,G,J), then fixed and hybridized in situ with a \textit{lacz} RNA probe. In embryos electroporated with the 2.4 kb wild-type \textit{Ci-Tbx6b} CRM, a strong hybridization signal was detected in the B6.4 pair of muscle precursors starting from the 32-cell stage (Fig. 2B). At the 110-cell stage, \textit{lacz} expression was detected in descendants of B6.4 blastomeres and in the B8.7 and B8.8 pairs, which derive from the B6.2 pair (Fig. 2C). In early neurulae, \textit{lacz} expression was detected only in the posterior-most muscle cells (Fig. 2D). This pattern faithfully recapitulates the expression of \textit{Ci-Tbx6b} (Takatori et al., 2004). Embryos electroporated with the 862 bp CRM, which does not contain the \textit{Ci-Macho1}-binding sites, showed no detectable signal at either the 32-cell stage (Fig. 2E) or at the 110-cell stage (Fig. 2F). \textit{lacz} expression became detectable by the early neurula stage in only the posterior-most muscle cells (Fig. 2G; note that the asymmetrical expression is due to mosaic incorporation of the transgene). Finally, the vast majority of embryos carrying the 2.4 kb CRM with mutations in all \textit{Ci-Macho1}-binding sites showed no expression of the reporter at either the 32-cell (Fig. 2H) or the 110-cell stage (Fig. 2I), and displayed active transcription in the posterior-most muscle cells only at the neurula stage (Fig. 2J).
These results indicate that mutations in the Ci-Macho1-binding sites do not affect the spatial (i.e. lineage-specific) activity of the 2.4 kb Ci-Tbx6b CRM, but are sufficient to cause a considerable delay in the onset of transcription driven by this fragment.

Together, these observations suggest that: (1) the temporal information required for the early activity of the 2.4 kb Ci-Tbx6b CRM is stored in its distal region; (2) this temporal information is encoded by the Ci-Macho1-binding sites; and (3) the 862 bp proximal region contains the cis-regulatory elements responsible for the late muscle activity observed in neurulae.

The proximal region of the Ci-Tbx6b CRM recapitulates the late muscle expression of Ci-Tbx6b

The proximal 112 bp region of the Ci-Tbx6b CRM (shaded area in supplementary material Fig. S2B), which was first identified using the Ci-FoxA-a basal promoter, was further tested using the endogenous 154 bp Ci-Tbx6b promoter. The resultant 266 bp construct (Fig. 2K; dark orange area in the schematics in Fig. 2A) was used for lacZ time-course experiments. The majority of the embryos electroporated with the 266 bp construct did not show staining at the 32-cell (Fig. 2L) or at the 110-cell stage (Fig. 2M);
however, by the early neurula stage, 87% of the embryos showed strong staining in muscle cells (Fig. 2N). These results suggest that this 266 bp region is sufficient to recapitulate the late activity of the 862 bp fragment previously analyzed (Fig. 2H-J).

Sequence inspection identified at least three putative binding sites within this region: a distal sequence showing an incomplete match with the CREB-binding sites previously identified in other Ciona muscle CRMs (TGACG core; blue diamond in Fig. 2K) (Kusakabe et al., 2004; Brown et al., 2007), a T-box-binding site matching the core consensus sequence identified for Ci-Tbx6b and Ci-Tbx6c via SELEX assays (7 out of 10 matches; red oval in Fig. 2K) (Yagi et al., 2005), and an ‘AC’-core E-box (pink triangle in Fig. 2K) (Erives et al., 1998). In vivo analysis of progressive truncations and specific point mutations showed that removal of the imperfect CREB-binding site did not noticeably affect either the intensity of the muscle staining or the percentage of embryos showing activity (data not shown). However, when the T-box-binding site was mutated, a considerable reduction was observed not only in the intensity of the muscle staining but also in the percentage of embryos showing activation of the reporter gene (Fig. 2O,P,R). Finally, a mutation of the ‘AC’-core E-box left the muscle staining unaffected (Fig. 2Q,R). The results of three independent experiments are quantified in the graph in Fig. 2R.

Comparable results were obtained when transgenic embryos were muscle CRMs (TGACG core; blue diamond in Fig. 2K) (Erives et al., 1998). In vivo analysis of progressive truncations and specific point mutations showed that removal of the imperfect CREB-binding site did not noticeably affect either the intensity of the muscle staining or the percentage of embryos showing activity (data not shown). However, when the T-box-binding site was mutated, a considerable reduction was observed not only in the intensity of the muscle staining but also in the percentage of embryos showing activity (Fig. 2O,P,R). An additional mutation of the ‘AC’-core E-box left the muscle staining unaffected (Fig. 2Q,R). The results of three independent experiments are quantified in the graph in Fig. 2R.

Together, these findings suggest that the proximal region of the Ci-Tbx6b CRM is mainly controlled by a T-box transcription factor. To further investigate this point, we carried out EMSA using bacterially expressed Ci-Tbx6b and Ci-Tbx6c proteins and a radiolabeled oligonucleotide probe containing the putative T-box-binding site (supplementary material Fig. S3B). For these experiments we synthesized full-length proteins to identify possible differences in binding affinities (see Materials and Methods); nevertheless, we found that this sequence was bound with similar intensity by both GST-Ci-Tbx6b and GST-Ci-Tbx6c fusion proteins (supplementary material Fig. S3B). These results suggest that the late-acting Ci-Tbx6b CRM functions as an autoregulatory and/or a cross-regulatory enhancer sequence.

Lastly, we used the information gathered from the analysis of the Ci-Tbx6b CRM to predict the location of a related muscle CRM in the Ci-Tbx6c locus. Ci-Tbx6c is clustered with Ci-Tbx6b within a 8 kb genomic region (Fig. 2S; supplementary material Fig. S2A). Within the 2 kb sequence directly upstream of Ci-Tbx6c, we identified a region spanning 293 bp which contained bona fide binding sites for Ci-Macho1 and Ci-Tbx6b/c, as well as an ‘AC’-core E-box (Fig. 2S). When we tested this fragment in vivo we found that, as expected, it was able to direct expression in most muscle cells (Fig. 2T), recapitulating the expanded expression pattern seen for Ci-Tbx6c at stages later than 110-cell (Takatori et al., 2004). This result suggests that the muscle activity of the highly related Ciona Tbx6 genes is controlled by the same basic set of structural cis-regulatory elements.

The composite cis-regulatory region of CiFcol1 harbors two muscle CRMs with different temporal onsets

We next sought to investigate whether the structural and functional criteria that we had identified for the Tbx6-related transcription factor genes applied to the cis-regulatory region of a structural muscle gene. The Ciona Fibrillar Collagen-I gene (CiFcol1; JGI gene model ci0100150759) (Wada et al., 2006) encodes a member of the fibrillar collagen family related to the vertebrate clade A collagen genes (Wada et al., 2006). During Ciona early embryogenesis, this gene is first expressed mostly in muscle precursors and by the neural plate stage its expression has expanded to mesenchyme, notochord, endoderm and CNS (Satou et al., 2001) (our unpublished results). Expression of CiFcol1 is maintained in all these tissues at very high levels throughout the mid-tailbud stage (Fig. 3A) until hatching (Satou et al., 2001) (our unpublished results).
The sustained expression of *CiFColl* in muscle cells suggested that transcription of this gene might rely upon a combination of early and late activators. Sequence inspection of its 5′-flanking region indicated that *CiFColl* might be controlled, at least at early developmental stages, by Ci-Macho1 and/or its intermediaries. To verify this hypothesis, we first used the misexpression assays described in Fig. 1. Since *CiFColl* is normally expressed in notochord precursors starting from the late gastrula stage, and is not present at detectable levels in these cells in early gastrulae (Fig. 3B), we harvested embryos at this developmental stage which had been electroporated with either *Bra>macho*, *Bra>Tbx6b* or *Bra>Tbx6c* and monitored the effects on the activation of *CiFColl* transcription in notochord precursors by WMISH (Fig. 3C-E). These experiments show that Ci-Macho1, Ci-Tbx6b and Ci-Tbx6c are all able to induce precocious onset of *CiFColl* expression in the notochord lineage.

We then investigated the molecular mechanisms mediating the response of this structural gene to the misexpression of the three muscle transcription factors by characterizing the *CiFColl cis*-regulatory region. We found that a 2.2 kb fragment from the 5′-flanking region of *CiFColl* (yellow rectangle in Fig. 3F) was able to fully recapitulate its expression pattern (inset in Fig. 3F). Within this main fragment, three different CRMs were identified, including one 400 bp distal muscle CRM (blue rectangle in Fig. 3F), and two adjacent proximal CRMs, one directly expression in notochord (data not shown) and the other, 230 bp long, which was active in muscle cells (orange rectangle in Fig. 3F). The distal muscle CRM contains two generic E-boxes (Fig. 3G, pink triangles) and directs expression in muscle cells (Fig. 3G, orange arrowheads; see lineage map in supplementary material Fig. S4A) and in the two rows of lateral ependymal cells of the nerve cord (Fig. 3G, blue arrowheads). The 230 bp proximal muscle CRM contains four putative T-box-binding sites (generic sequence: TNNCAC; Fig. 3H, red ovals) and one putative low-affinity Ci-Macho1-binding site (Yagi et al., 2004a) (Fig. 3H, green rectangle) and directs expression in most of the primary muscle cells and in trunk mesenchyme.

To accurately define the window of activity of each muscle CRM, we performed time-course experiments using the *lacZ* probe (Fig. 3I,J,L-Q,R). In addition, we monitored the accumulation of β-galactosidase in the muscle cells where each CRM was active by performing X-Gal staining (Fig. 3K,P). Embryos carrying the distal 400 bp muscle CRM began accumulating *lacZ* transcripts between the 32-cell and the 110-cell stage (Fig. 3L). From these observations, we conclude that of the two muscle CRMs identified within the *CiFColl* locus were mediating the response of the endogenous gene to the ectopic expression of *Ci-Tbx6b* and/or *Ci-Tbx6c*, we co-electroporated each CRM with either *Bra>Tbx6b* or *Bra>Tbx6c*. As shown in Fig. 4, misexpression of *Ci-Tbx6b* was sufficient to induce ectopic expression of the *CiFColl* proximal 230 bp muscle CRM in notochord cells (Fig. 4A,B), consistent with the possibility that Ci-Tbx6b activates this enhancer. While performing controls for these experiments, we noticed that in some batches of embryos the pFBASP6 vector was itself slightly responsive to the

**Selective activation of the *CiFColl* CRMs by ectopically expressed Tbx6 transcription factors**

To test whether the muscle CRMs identified within the *CiFColl* locus were mediating the response of the endogenous gene to the ectopic expression of *Ci-Tbx6b* and/or *Ci-Tbx6c*, we co-electroporated each CRM with either *Bra>Tbx6b* or *Bra>Tbx6c*. As shown in Fig. 4, misexpression of *Ci-Tbx6b* was sufficient to induce ectopic expression of the *CiFColl* proximal 230 bp muscle CRM in notochord cells (Fig. 4A,B), consistent with the possibility that Ci-Tbx6b activates this enhancer. While performing controls for these experiments, we noticed that in some batches of embryos the pFBASP6 vector was itself slightly responsive to the
Bra>Tbx6b construct; however, the ectopic activation observed in the notochord when the 230 bp CiFCol1 CRM was used in these co-electroporations instead of the empty vector was at least 3.5-times higher (Fig. 4D). No ectopic activation in the notochord was observed when the 230 bp CiFCol1 CRM was co-electroporated with Bra>Tbx6c (Fig. 4C,D). In addition, we tested the response of these CRMs to another Tbx6-related transcription factor, Ci-Tbx6a (Takatori et al., 2004), and to a more distantly related T-box factor, Ci-Tbx15/18/22 (Erives and Levine, 2000; Takatori et al., 2004), and we found no response above background to either factor (data not shown). Finally, no ectopic activation was observed when the distal CiFCol1 CRM was co-electroporated with either Bra>Tbx6b or Bra>Tbx6c (data not shown); this result is consistent with the lack of evident Tbx6b/c-binding sites in this sequence (Fig. 3G).

The flow of maternal and zygotic developmental information in the muscle cells of Ciona

Our findings are summarized by the model in Fig. 5, which reconstructs part of the cis-regulatory hierarchy underlying the cell-autonomous developmental program that characterizes primary muscle formation in Ciona. Arrows in Fig. 5 indicate the interactions between transcription factors and cis-regulatory sequences, either demonstrated by the work presented here (black) or inferred (gray).

After fertilization, binding of maternally encoded Ci-Macho1 protein to the early Ci-Tbx6b CRM activates transcription of this gene in muscle precursors. By the 16-cell stage, when maternal Ci-macho1 transcripts starts declining (Satou et al., 2002a), transcription of Ci-Tbx6b begins. At later stages, it is likely that the Ci-Tbx6b early CRM is progressively vacated and is first aided and then progressively replaced by the late CRM, which is activated by Ci-Tbx6b and/or Ci-Tbx6c. Given the similarities that we have identified in the Ci-Tbx6b and Ci-Tbx6c cis-regulatory sequences, it is plausible that similar mechanisms activate the Ci-Tbx6c CRM. Ci-Tbx6b in turn activates the proximal CRM of CiFCol1, which recapitulates the early transcription of this gene. By the neurula stage, as expression of Ci-Tbx6b fades, different muscle activators begin binding the distal, late-acting CiFCol1 CRM and allow transcription of this gene to proceed without interruptions until the late tailbud stage.

Discussion

Spatial and temporal heterogeneity of the embryonic territories responsive to Ci-Macho1

For over a century, it has been known that upon fertilization and first cleavages of the ascidian egg, maternally loaded cytoplasmic determinants are differentially segregated into the resulting blastomeres, thus giving rise to the early determination and invariance that characterize their developmental fates (Conklin, 1905; Deno et al., 1984; Ortolani, 1955). Although recent molecular evidence has highlighted the major role of inductive processes in the formation of most embryonic tissues (reviewed by Lemaire, 2009), the cell-autonomous development of the primary muscle of the ascidian larva is largely attributable to the maternal determinant Ci-Macho1 (Nishida and Sawada, 2001). Ci-macho1 postplasmic mRNA is relocated after fertilization by the cortical centrosome-attracting body (CAB) (Sardet et al., 2003). As cleavage proceeds, in both Halocynthia (Nishida and Sawada, 2001) and Ciona (Satou et al., 2002a) Ci-macho1 mRNA becomes progressively restricted to a narrow region of the embryo, the B7.6 blastomeres; however, the Macho1 protein is generally believed to persist in an unlocalized form, and to be distributed to all descendants of the B4.1 cells (Kondoh et al., 2003). Studies in Halocynthia show that for the proper formation of other lineages that also derive from the B4.1 cells, such as mesenchyme and endoderm, the function of Macho1 needs to be actively suppressed by FGF and BMP signaling pathways (Kim et al., 2000; Kim and Nishida, 1999; Kondoh et al., 2003). Similar mechanisms are also likely responsible for the functional suppression of zygotically expressed Ci-Macho1 in the Ciona CNS, considering that Ci-FGF16/19/20 is expressed in the Ciona CNS through tailbud stages and is required for neural development (Imai et al., 2004; Bertrand et al., 2004).

The misexpression experiments described here suggest that no such restraining mechanism is present in notochord cells before the early tailbud stage. In fact, at early developmental stages the ectopic activation of both Ci-Tbx6b and Ci-Tbx6c was seen in notochord precursors of both lineages in Bra>macho embryos, whereas at the mid-tailbud stage only the ectopic activation of Ci-Tbx6b was observed, and it was confined to a subset of mesenchyme cells. These cells are most likely descendants of the B7.3 blastomere, a 64-cell stage precursor of both secondary

![Fig. 5. A working model for the muscle cis-regulatory hierarchy in Ciona.](Image 338x147 to 542x449)

The regulatory interactions identified by this study are summarized and plotted against the developmental time-course of Ciona embryogenesis at 18°C (black vertical bar on left). Horizontal green bar represents the maternally-deposited Ci-macho1 transcripts, green squares depict the Ci-Macho1 protein. Horizontal bars symbolize the cis-regulatory regions of Ci-Tbx6b and Ci-Tbx6c (yellow) and CiFCol1 (orange); transcription start sites are indicated by horizontal arrows. Transcripts are shown in scheme as vertical bars on the right. Transcription factors are depicted as described in Fig. 2.
notochord and mesenchyme cells (Satoh, 1994). The differential competence of the notochord to respond to Ci-Macho1 might be explained by the requirement for temporally and spatially localized co-factors and/or transcriptional intermediaries. Alternatively, as in the case of the CNS, Ci-Macho1 might be functionally suppressed in the notochord of tailbud embryos by the activation of the FGF signaling pathway, as suggested by the observation that Ci-FGFR is expressed in the notochord beginning at the early tailbud stage (Imai et al., 2004; Shi et al., 2009). These mechanisms might also account for the relatively mild phenotype that we observed in embryos carrying the Bra>macho transgene, whereby the notochord is still able to form, even in transgenic embryos where mosaic incorporation is minimal.

**The spatio-temporal expression pattern of Ci-Tbx6b is recapitulated by early and late cis-regulatory sequences**

Using in vivo transient transgenic assays, we have identified a 2.4 kb CRM upstream of Ci-Tbx6b that is able to faithfully recapitulate the muscle expression of this gene. The temporal muscle activity of the 2.4 kb CRM represents the composite read-out of early- and late-acting cis-regulatory sequences, which interpret maternal and zygotic information. The Ci-Tbx6b CRM contains a distal region which functions as the repository of the temporal information necessary to recapitulate the early expression pattern previously reported for Ci-Tbx6b (Takatori et al., 2004). When this distal region is deleted, muscle activity is not lost, but its onset is considerably delayed. Sequence inspection and point-mutation analyses suggested that this early-acting distal region might be controlled by maternal Ci-Macho1, because three putative binding sites for this factor are present in this sequence. We found that these sites are bound in vitro by Macho1 and that their concomitant mutation is sufficient to cause the same delay in the onset of transcriptional activity that was observed when the entire fragment encompassing them was deleted. Together, these observations provide a mechanistic cis-regulatory explanation to the results of our misexpression assay, as well as to previous results showing that overexpression of Ci-Macho1 is sufficient to induce ectopic expression of Ci-Tbx6b (Yagi et al., 2004a) and that, likewise, Hr-Macho1 is able to ectopically induce Hr-Tbx6, among other muscle genes (Sawada et al., 2005). It is noteworthy that in Ciona, Ci-ZicL cooperates with Ci-Macho1 to promote muscle development (Yagi et al., 2005); this zygotic zinc-finger transcription factor is related to Ci-Macho1 and recognizes a similar consensus binding site in vitro (Yagi et al., 2004b). Interestingly, one of the three Ci-Macho1-binding sites that we have characterized in the Ci-Tbx6b CRM, namely site ‘C’, contains permutations of the published ZicL consensus site that are compatible with binding in vitro (Yagi et al., 2004b). If this site is bound in vivo by either transcription factor, then this would explain the observation that Ci-Tbx6b is still weakly expressed in Ci-Macho1 morphant embryos, whereas its expression is no longer detectable in Ci-Macho1 and Ci-ZicL double-morphants (Yagi et al., 2005).

Within the 2.4 kb CRM, a 266 bp proximal region is able to direct transcription only from neurulation onwards, thus acting as a late muscle enhancer. Sequence analysis of this region revealed the presence of an imperfect CREB-binding site, a T-box-binding site (generic sequence: TNNCAC) partly matching the core consensus sequence previously reported for Ci-Tbx6b/c (Yagi et al., 2005), and an ‘AC’-core E-box. Both CREB-binding sites and AC-core E-boxes have been previously shown to be necessary for muscle activity of other muscle CRMs (Brown et al., 2007; Erives et al., 1998; Kusakabe et al., 2004); however, in this case, only the T-box site substantially contributes to the muscle activity, qualitatively and quantitatively. Through EMSA, we have shown that this T-box site is bound in vitro by both Ci-Tbx6b and Ci-Tbx6c.

**Modular organization and temporal properties of the cis-regulatory region of CiFC01**

Originally isolated in a subtractive screen aimed to identify genes downstream of Ci-Bra (Takahashi et al., 1999), the CiFC01 gene attracted our interest because of its sustained muscle expression, which begins around mid-gastrulation, and because its upstream region is enriched in T-box-binding sites.

Dissection of a 2.2 kb genomic fragment located upstream of the transcription start site of CiFC01 revealed the presence of discrete CRMs active in all the tissues where CiFC01 is expressed. In particular, this 2.2 kb fragment harbors two distinct muscle CRMs: a distal CRM containing two generic E-boxes and depleted of T-box-binding sites and Ci-Macho1-binding sites, and a proximal CRM containing four clustered T-box-binding sites, some of which are bound weakly in vitro by the Ci-Tbx6b protein (data not shown), and a low-affinity Ci-Macho1-binding site. The heterogeneity of these sequences is reflected by the temporal activity of the two CRMs, because the distal one, which does not contain any apparent T-box-binding sites, is activated later than the proximal one, which is enriched in these motifs. In particular, the distal CiFC01 muscle CRM is active in a small subset of muscle precursors from the 110-cell stage to the neurula stage, and only by the early tailbud stage does its territory expand to encompass all muscle cells. Afterwards, it remains active in the majority of muscle cells. Therefore, the spatial range of action of this CRM in the muscle seems to be controlled by an activator(s) functioning from neurulation onwards. The presence of two E-boxes in this sequence prompted us to investigate the possible involvement of transcription factors of the bHLH family in the regulation of this CRM. We found that neither mutation of the E-boxes nor misexpression, individual or combined, of two bHLH transcription factors, Ci-MRF (Meedel et al., 2007) and Ci-paraxis (Erives, 2009; Imai et al., 2004) had any detectable effect (data not shown), thus leaving the identification of the late activator(s) to future investigations.

Conversely, the proximal CiFC01 muscle CRM is ignited early in most muscle cell precursors, starting from the 32-cell stage, but its activity fades by the mid-tailbud stage. We conclude that the additive activity of the two CRMs is probably responsible for the sustained expression of CiFC01 in muscle cells.

Interestingly, misexpression of Ci-Macho1, Ci-Tbx6b or Ci-Tbx6c in notochord cells all result in ectopic activation of CiFC01 in this territory. Although we cannot rule out that this might be attributable to the low-affinity Ci-Macho1-binding site in the CiFC01 early CRM, given the late onset of CiFC01 muscle expression it seems more likely that Ci-Macho1 activates expression of CiFC01 indirectly, through Ci-Tbx6b. To test this hypothesis we monitored the response of the CiFC01 proximal muscle CRM to the misexpression of Ci-Tbx6b in notochord cells. We found that misexpression of Ci-Tbx6b caused the ectopic activation of the CiFC01 proximal muscle CRM in the notochord, whereas misexpression of Ci-Tbx6c did not have any effect. We conclude that the ectopic activation of CiFC01 seen in notochord cells of embryos carrying the Bra>Tex6c construct
might occur indirectly, via the activation of \textit{Ci-Tbx6b} expression by Ci-Tbx6c.

Finally, no ectopic activation was observed when the distant \textit{CiFCol1} muscle CRM was co-electroporated with either construct (data not shown), consistent with the lack of Tbx6b/c-binding sites in its sequence.

Muscle gene regulation in ascidians versus vertebrates: lineage-specific innovations and evolutionarily conserved mechanisms

By analyzing the \textit{cis}-regulatory sequences that mediate the response to Ci-Macho1 and its mediators, this study has begun to provide sharper insights into the molecular mechanisms controlling cell-autonomous muscle development in the ascidian embryo. Given the large number of genes that respond to Ci-Tbx6b and Ci-Tbx6c (Yagi et al., 2005) (this study), it is conceivable that the mechanisms of transcriptional regulation that control the CRMs presented here might be shared by several other muscle genes. This hypothesis is supported by the abundance of putative Tbx6b/c-binding sites in muscle CRMs identified by other groups (Brown et al., 2007; Kusakabe et al., 2004) and by our laboratory (data not shown).

Although the early cell-fate determination mediated by Macho-like proteins in muscle cells has been described so far as an ascidian-specific mechanism, transcription factors of the Zic family, of which Macho, Zicl and related proteins represent a diverged branch (Aruga et al., 2006), are known to be required for shaping the body plan of widely different animals (Merzdorf, 2007). In addition, Tbx6-related proteins in \textit{Ciona} appear to be part of an evolutionarily conserved kernel that is employed for the specification and differentiation of paraxial mesoderm in several other chordates, including mouse (Chapman and Papaioannou, 1998; White et al., 2003), \textit{Xenopus} (Tazumi et al., 2008; Uchiyama et al., 2001) and zebrafish (Goering et al., 2003). Hence, the elucidation of the \textit{cis}-regulatory mechanisms used by these transcription factors to modulate expression of their target genes should provide insights on the inner workings of other model systems in which \textit{cis}-regulatory elements are less tractable, including higher chordates.

Materials and Methods

Ascidians and electroporation

Adult \textit{Ciona intestinalis} were purchased from Marine Research and Educational Products (M-REP; Carlsbad, CA). Fertilization, dechorionation, electroporation and Adult were used as a template, in the presence of either \textit{Hi-III Taq} (Invitrogen, Carlsbad, CA, USA) or Turbo \textit{Pfu} (Stratagene, La Jolla, CA, USA) DNA polymerase. All plasmids were checked for accuracy either manually (Sambrook et al., 1989) or through automated sequencing facilities.

Whole-mount in situ hybridization (WISH)

This was performed as previously described (Oda-Ishii and Di Gregorio, 2007) using digoxigenin-labeled antisense RNA probes. The \textit{CiFCol1} probe was prepared from EST GC01e23, the \textit{Ci-macho1} probe from EST GC18f07, the \textit{Ci-Tbx6b} probe from EST GC99f19, and the \textit{Ci-Tbx6c} probe from EST GC43g03 (Satou et al., 2002b). The lacZ probe was kindly provided by Yukata Nibu (Weill Cornell Medical College, New York, NY).

Quantitative RT-PCR (qRT-PCR)

\textit{Ciona} zygotes from a single round of fertilization were divided into two groups and were either electroporated with \textit{Ci-Bra>gfp} or co-electroporated with \textit{Ci-Bra>gfp} and \textit{Ci-macho1 constructs}. Once they reached the 4-cell stage, the embryos showing fluorescence in the notochord were selected for RNA extraction using an epifluorescence microscope. RNA was extracted using the RNeasy Protect Mini kit (Qiagen, Valencia, CA, USA), and treated with RNase-free DNase (Qiagen) using the on-column digestion protocol to remove possible genomic DNA contamination. cDNA was synthesized using the Superscript III kit (Invitrogen). This process was repeated twice on different batches of embryos to generate two independent samples of \textit{Bra-macho} cDNA and corresponding wild-type control cDNA.

qRT-PCR samples were prepared using 2× SYBR Green Hotstart Mix (Applied Biosystems, Foster City, CA, USA) according to the manufacturer’s instructions. The resulting data were analyzed using the SDS2.3 software (Applied Biosystems). Both sets of cDNAs were loaded with the appropriate primers on a single plate, to eliminate variation between PCR amplifications. Samples were run in triplicate, using \textit{Ci-GAPDH} as a control. Error bars represent the variation seen between the two sets of cDNAs. The relative standard curve method was used to calculate the fold change in gene expression, as described by the manufacturer (Taqman Applied Biosystems, CA). \textit{Bra-macho} cDNA was used as a control for the fold increase in gene expression relevant in consideration of the low number of notochord precursors at the 110-cell stage (ten), the electrophoresis efficiency, and the high variability in the number of cells that efficiently incorporate the construct(s) within each transgenic embryo.

Recombinant protein synthesis and purification

The full-length \textit{Ci-macho1} coding region was PCR-amplified from the \textit{Bra-macho} plasmid described above and cloned into the \textit{BamHI} and \textit{SacI} sites of the pRSET-B vector (Invitrogen). Full-length \textit{Ci-Tbx6b} and \textit{Ci-Tbx6c} coding regions were PCR-amplified and cloned into the pET3c vector (GE Healthcare, Piscataway, NJ) in frame with the GST tag. The resulting plasmids were transformed into \textit{E. coli} BL21 (DE3) cells harboring the pP32 plasmid (Affiniti Research Products, Exeter, UK) and the fusion proteins were purified as previously described (Gazdou et al., 2005).
The 6XHis-tagged Ci-Machol protein was expressed at 15°C in the presence of 0.1 mM IPTG.

Electrophoretic mobility shift assays (EMSA)
The following double-stranded oligonucleotides were used (only the 5’-3’ strand is reported, mutations are underlined): Macho site-A-wt, CAAACATGC-AGCAGATGAGGACAGGAGACA; Macho site-A-MUT, CAAACATGC-CAGGAGATGAGGACAGGAGACA; Macho site-B-wt, GACAAAGCATCTGGAGGAGCAGTGGGAGATA; Macho site-B-MUT, GACAAAGCATCTGGAGGAGCAGTGGGAGATA; Macho site-C-wt, AAGTAATGAGGACCCCGTGCTGAACCT; Macho site-C-MUT, AAGTAATGAGGACCTGGTGGAATA; T-box site wt, 5’-CGCATTGCTGACATTGTTG-3’; T-box site MUT, 5’-CGCATTGCTGAGTAAGTTTGTGATTTTGAAT-3’. Radioactive labeling and subsequent procedures were performed as previously described (Dunn and Di Gregorio, 2009).

We thank Yutaka Nibu and the members of the Di Gregorio and Nibu labs for valuable comments on the manuscript and Hiroki Takahashi (NIIBB, Japan) for the Ci-Thbx6 cDNA clone. We are grateful to Mami Takeda and Gary Esses for technical help. This work was supported by grant 1-FY08-430 from the March of Dimes Birth Defects Foundation and by NIH/NICHD grant R01HD50704 to A.D.G. J.E.K. was supported in part by a Jacques Cohenca pre-doctoral fellowship. A.K., H. Sato, N. Satoh, and Y. Nibu were supported by a grant from the Uehara Memorial Foundation (Japan). A.D.G. thanks the Ira T. Hirsch Scholar. Deposited in PMC for release after 12 months.

Supplementary material available online at http://jcs.biologists.org/cgi/content/full/123/4/2453/DC1

References


