© 2019. Published by The Company of Biologists Ltd | Development (2019) 146, dev173104. doi:10.1242/dev.173104

RESEARCH ARTICLE

The regulatory pathway from genes directly activated by maternal factors to muscle structural genes in ascidian embryos

Deli Yu, Izumi Oda-Ishii, Atsushi Kubo* and Yutaka Satou‡

ABSTRACT

Striated muscle cells in the tail of ascidian tadpole larvae differentiate cell-autonomously. Although several key regulatory factors have been identified, the genetic regulatory pathway is not fully understood; comprehensive understanding of the regulatory pathway is essential for accurate modeling in order to deduce principles for gene regulatory network dynamics, and for comparative analysis on how ascidians have evolved the cell-autonomous gene regulatory mechanism. Here, we reveal regulatory interactions among three key regulatory factors, Zic-r.b, Tbx6-r.b and Mrf, and elucidate the mechanism by which these factors activate muscle structural genes. We reveal a cross-regulatory circuit among these regulatory factors, which maintains the expression of *Tbx6-r.b* and *Mrf* during gastrulation. Although these two factors combinatorially activate muscle structural genes in late-stage embryos, muscle structural genes are activated mainly by Tbx6-r.b before gastrulation. Time points when expression of muscle structural genes become first detectable are strongly correlated with the degree of Tbx6-r.b occupancy. Thus, the genetic pathway, starting with *Tbx6-r.b* and *Zic-r.b*, which are activated by maternal factors, and ending with expression of muscle structural genes, has been revealed.

KEY WORDS: Ascidian, *Ciona*, Muscle, Mrf, Tbx6, Zic, Gene regulatory network

INTRODUCTION

The ascidian larva has a typical chordate body plan (Lemaire, 2011; Satoh et al., 2003) and 36 mononuclear striated muscle cells differentiate in the larval tail of an ascidian *Ciona intestinalis* (type A; also called *C. robusta*) (Fig. 1A). Among them, the anterior 28 cells are derived from the posterior vegetal quadrants, which are called the primary lineage cells. These primary lineage muscle cells differentiate cell-autonomously (Conklin, 1905; Deno et al., 1984). The primary lineage consists of three sub-lineages called the B7.4, B7.5 and B7.8 sub-lineages (Fig. 1B).

Macho-1 (also known as *Zic-r.a*) mRNA is localized in the posterior pole of the embryo and is inherited by most posterior cells: B4.1 in eight-cell embryos, B5.2 in 16-cell embryos, B6.3 in 32-cell embryos and B7.6 in 64- and 112-cell embryos (Nishida and Sawada, 2001; Satou et al., 2002). Although cells with this localized mRNA eventually give rise to germ cells (Shirae-Kurabayashi et al., 2006), their somatic sister cells produced at each of the fourth, fifth and sixth divisions contribute to primary lineage muscle cells. In these somatic lineages, *Tbx6-r.b* is activated under the control of Macho-1 and β-catenin (Kugler et al., 2010; Oda-Ishii et al., 2016; Yagi et al., 2005). In addition, *Zic-r.b* (also known as *Zic-L*) is activated in these lineages (Imai et al., 2002; Wada and Saiga, 2002) by a maternal factor, Gata.a, the activity of which is initially weakened by β-catenin (Imai et al., 2016; Oda-Ishii et al., 2016). Subsequently, a myogenic factor gene, *Mrf*, is activated in the three lineages (Araki et al., 1994; Meedel et al., 1997) under the control of Tbx6-r.b and Zic-r.b (Imai et al., 2006). These key regulatory factors are required to activate muscle structural genes (Imai et al., 2006, 2002; Meedel et al., 2007, 1997; Yagi et al., 2005). However, it is not fully understood how *Tbx6-r.b* and *Mrf* expression is maintained after initial activation and how these regulatory factors activate muscle structural genes.

In embryos of *Halocynthia roretzi*, a muscle actin gene (*HrMA4a*) begins expression as early as the 32-cell stage (Satou et al., 1995). Analysis of the cis-regulatory region of *HrMA4a* has identified two crucial elements, 5′-TCGCACTTC-3′ and 5′-GTGATAACAACTG-3′ (Satou and Satoh, 1996). The former contains a sequence similar to the Tbx6-r.b-binding motif (5′-TCaCACcT-3′, lower case indicates mismatches with the motif that has been determined by an *in vitro* selection method; Yagi et al., 2005). The latter contains an E-box (CANNTG), which is a binding site for Mrf, and a sequence similar to the Tbx6-r.b-binding motif. Upstream regulatory regions of muscle structural genes share several common motifs, which contain the E-box motif and the Tbx6-r.b-binding motif, and such motifs are indeed important for specific expression of these genes (Brown et al., 2007; Kusakabe et al., 2004). Consistently, chromatin-immunoprecipitation (ChIP) assays have indicated that two factors simultaneously bind to the upstream sequences of most known muscle-specific genes (Kubo et al., 2010). However, it has not been demonstrated that the above *Halocynthia* elements indeed bind Tbx6-r.b and Mrf. In addition, not all muscle structural genes begin expression at the 32-cell stage. Therefore, it should be elucidated which factor(s) determines the timing of the initiation and strength of their expression in early stage embryos.

The gene that encodes fibrillar collagen 1 is expressed in multiple tissues including muscle (Satou et al., 2001). Its expression in muscle cells is controlled by two enhancers that control early and late expression (Kugler et al., 2010). The early enhancer is regulated by Tbx6-r.b. Although the late enhancer includes an E-box, it has been suggested that Mrf may not regulate this gene (Kugler et al., 2010). Meanwhile, other studies have shown that Mrf regulates many other muscle-specific genes (Imai et al., 2006; Izzi et al., 2013; Kubo et al., 2010; Meedel et al., 2007).

Despite extensive efforts, understanding of the genetic pathway for muscle differentiation in the ascidian embryo is still incomplete. The complete understanding of the regulatory logic, which explains spatial and temporal gene expression patterns in all cell lineages, is required for accurate modeling of the regulatory process and will be

Department of Zoology, Graduate School of Science, Kyoto University, Sakyo, Kyoto 606-8502, Japan.
*Present address: Department of Developmental Neurobiology, Institute of Development, Aging and Cancer, Tohoku University, 4-1 Seiryo, Aoba, Sendai, Miyagi 980-8575, Japan.

‡Author for correspondence (yutaka@ascidian.zool.kyoto-u.ac.jp)

Y.S., 0000-0001-5193-0708

Received 22 October 2018; Accepted 14 January 2019

Development (2019) 146, dev173104. doi:10.1242/dev.173104

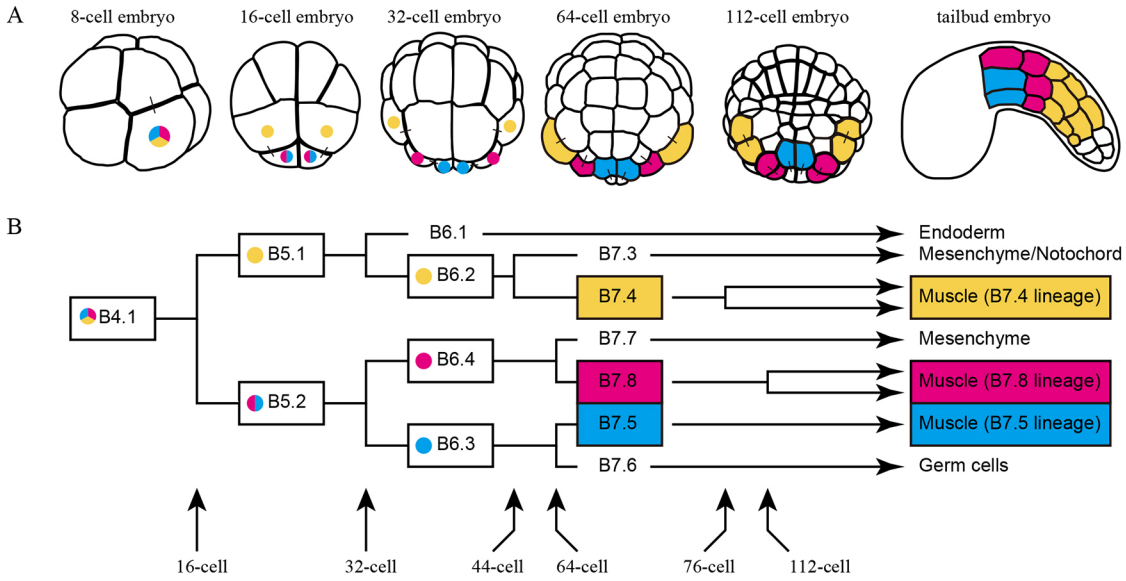

Fig. 1. Three lineages of cells contribute to 28 of the 36 tail muscle cells. (A) Illustrations of the eight-cell embryo to the tailbud embryo. Lateral views are shown for the eight-cell and tailbud embryos, and vegetal views are shown for the other embryos. Cells with muscle fate only are colored, and cells with muscle and other fates are indicated by dots. Sister cell relationships for the muscle lineages are shown by short thin lines. (B) Cell lineages starting with the posterior vegetal cells (B4.1) of the eight-cell embryo. Names for cells with muscle fate only are enclosed and colored with the same colors used in A. Names for cells with muscle and other fates are enclosed by boxes and marked with colored dots.

useful for comparative studies to elucidate principles of network architecture.

In the present study, we addressed the following questions. How is expression of the key regulatory factor genes, *Tbx6-r.b*, *Zic-r.b* and *Mrf*, maintained after initial activation? Do all of these factors directly regulate muscle structural genes? Do previously identified putative Tbx-binding sites and E-boxes actually bind Tbx6-r.b and Mrf? What determines the timing of the initiation and strength of expression of muscle structural genes? By answering these questions, we provide a more complete understanding for the genetic pathway, starting with regulatory genes that are directly activated by maternal factors and ending with expression of muscle structural genes.

RESULTS

Expression profiles of three key transcription factor genes at single cellular resolution

We re-examined expression patterns of *Tbx6-r.b*, *Zic-r.b* and *Mrf* in *Ciona* embryos to determine their precise temporal gene expression profiles. For this purpose, we used the fluorescence *in situ* hybridization method, because nascent transcripts in nuclei were clearly seen with this method (Fig. 2; Fig. S1). As reported previously (Imai et al., 2004, 2002; Meedel et al., 2007; Takatori et al., 2004), in the B7.4 lineage, *Tbx6-r.b*, *Zic-r.b* and *Mrf* began expression at the 16-, 32- and 44-cell stages, respectively. In the B7.8 lineage, *Tbx6-r.b* and *Zic-r.b* expression began simultaneously at the 32-cell stage and *Mrf* expression began at the 64-cell stage. In the B7.5 lineage, *Tbx6-r.b* and *Zic-r.b* expression began at the 64-cell stage and *Mrf* expression began at the 112-cell stage.

The nuclear signal for *Tbx6-r.b* in the B7.4 lineage was not observed between the 32- and 64-cell stages, but it became visible again at the 76-cell stage. This observation indicated that transcription of *Tbx6-r.b* is reactivated at the 76-cell stage in this lineage. Such reactivation of *Tbx6-r.b* was also observed in the B7.8 lineage, because the nuclear signal for *Tbx6-r.b* was not observed at the 64- and 76-cell stages, but was visible at the 112-cell stage. Thus, in these two lineages, two waves of *Tbx6-r.b* expression were observed.

Although *Zic-r.b* expression disappeared rapidly, it was reactivated in the B7.4 lineage (B8.7 and B8.8) at the 76-cell stage. *Mrf* expression became weak after initiation of expression in the B7.4 and B7.8 lineages, but was clearly visible again in the B7.4 lineage at the 76-cell stage and in the B7.8 lineage at the 112-cell stage.

Regulatory interactions among the three key regulatory factors

Regulatory interactions among *Tbx6-r.b*, *Zic-r.b* and *Mrf* have been partially revealed. However, as shown above, their expression patterns are more complex than was previously thought. Therefore, on the basis of the above expression profiles, we re-examined their regulatory interactions.

Tbx6-r.b is initially activated by β-catenin/Tcf7 and Macho-1 at the 16-cell stage (Oda-Ishii et al., 2016). Because initiation of *Tbx6-r.b* expression in the B7.4 and B7.8 lineages is not later than that of *Zic-r.b* and *Mrf* expression (see Fig. 2), it cannot be affected by Zic-r.b or Mrf. Therefore, we examined whether reactivation of *Tbx6-r.b* in these lineages was affected at the 112-cell stage. In *Zic-r.b* and *Mrf* morphants, which were injected with specific antisense morpholino oligonucleotides (MOs) against *Zic-r.b* or *Mrf*, respectively, reactivation of *Tbx6-r.b* was rarely detected in the B7.4 and B7.8 lineages at the 112-cell stage (Fig. 3A-C), whereas expression in B7.5 was unaffected. Thus, *Zic-r.b* and *Mrf* are required for the second wave of *Tbx6-r.b* expression in the B7.4 and B7.8 lineages.

Expression of *Tbx6-r.b* in B7.5 begins immediately after these cells separate from cells with a germ cell fate (Fig. 2), in which transcription is quiescent (Kumano et al., 2011; Shirae-Kurabayashi et al., 2011). Specifically, no zygotic expression occurs in this lineage before the expression of *Tbx6-r.b*. Therefore, we examined whether this *Tbx6-r.b* expression is regulated by a maternal factor, Macho-1, as is the case with expression in the other two lineages. Indeed, knockdown of *Macho-1* resulted in loss of expression of *Tbx6-r.b* in B7.5 at the 64-cell stage (Fig. S2).

Development (2019) 146, dev173104. doi:10.1242/dev.173104

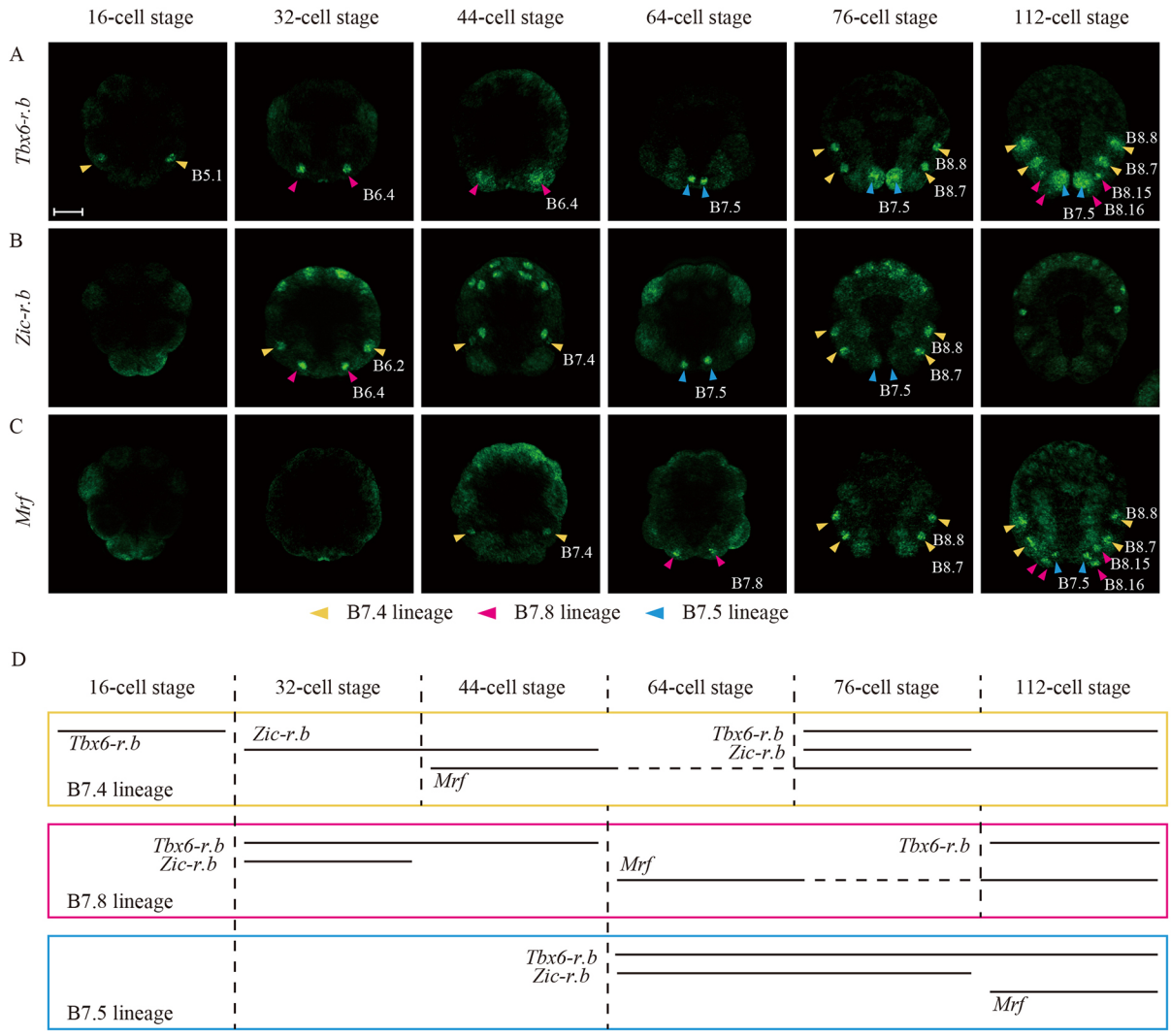

Fig. 2. Expression of three key regulatory genes from the 16-cell to 112-cell stages. (A-C) Expression of *Tbx6-r.b* (A), *Zic-r.b* (B) and *Mrf* (C) in muscle lineages shown by arrowheads with the color code described in Fig. 1. Photomicrographs are *z*-projected image stacks overlaid in pseudocolor. (D) Summary of gene expression profiles of *Tbx6-r.b*, *Zic-r.b* and *Mrf* in each of the three muscle lineages. Vertical dotted lines indicate cell divisions; horizontal lines indicate gene expression; horizontal dotted lines indicate weak gene expression. Scale bar: 50 µm.

Zic-r.b is activated by Gata.a in the B7.4 lineage (Imai et al., 2016), probably in combination with Tbx6-r.b at the 32-cell stage (Yagi et al., 2005). Because *Mrf* expression does not begin at the 32-cell stage (see Fig. 2), it cannot regulate *Zic-r.b* at this stage. Therefore, we examined *Zic-r.b* expression only at the 112-cell stage and found that it was not affected in *Tbx6-r.b* or *Mrf* morphant embryos (Fig. 3D-F). Thus, neither Tbx6-r.b nor Mrf is required for *Zic-r.b* expression in 112-cell embryos.

Mrf expression is regulated by Tbx6-r.b and Zic-r.b at the gastrula stage (Imai et al., 2006). Therefore, we examined whether these regulators are required to initiate *Mrf* expression at earlier stages. As shown in Fig. 3G-L, *Mrf* expression was indeed lost in morphant embryos of *Tbx6-r.b* or *Zic-r.b* at the 64-cell and 112-cell stages. Thus, these three key genes constitute a cross-regulatory circuit.

Timings when expression becomes detectable are different among muscle structural genes

Considering previously published results, it is likely that expression of muscle structural genes begins at various timings in each of the three lineages (Fujiwara et al., 2002; Kusakabe, 1995; Meedel et al., 2007; Nishikata et al., 2001; Satou et al., 1995; Yagi et al., 2005). However, because this information was incomplete, we examined the expression patterns of genes encoding troponin T (Tnt), myosin regulatory light chain (Mrlc), troponin I (Tni), myosin light chain (Mlc) and troponin C (Tnc) using *in situ* hybridization and reverse-transcription followed by quantitative polymerase chain reaction (RT-qPCR) to precisely determine when their expression becomes detectable.

We observed signals for *Tnt* mRNA in the B7.4 and B7.8 lineages of all 64-cell embryos using *in situ* hybridization, although weak signals were observed in the B7.4 lineage of some, but not all, 32-cell embryos (Fig. 4A). The expression in the B7.5 lineage was first seen around the 112-cell stage. Consistently, RT-qPCR assays of four different batches of embryos showed that the expression level of *Tnt* was significantly increased between the 32- and 64-cell stages (Fig. 4A).

The *in situ* hybridization assay showed that *Mrlc* was expressed in the B7.4- and B7.8-lineages of 64-cell embryos, although weak signals were observed in the B7.4 lineage of some, but not all, 32-cell embryos. The expression in the B7.5-lineage was first seen at

Development (2019) 146, dev173104. doi:10.1242/dev.173104

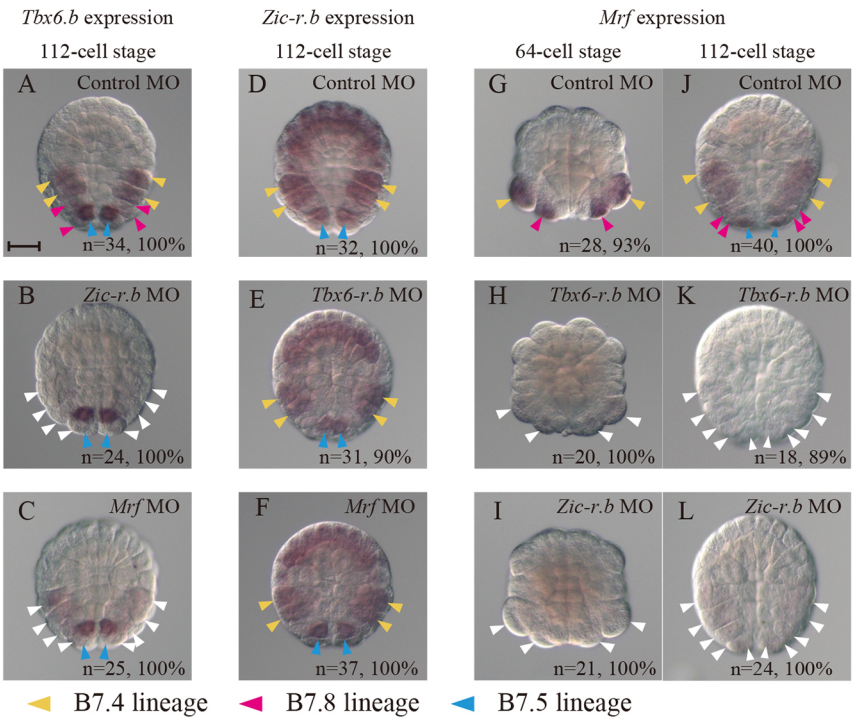

Fig. 3. Regulatory interactions among the three key regulatory genes. (A-L) Expression of *Tbx6-r.b* (A-C), *Zic-r.b* (D-F), and *Mrf* (G-L) in embryos injected with the control MO (A,D,G,J), *Zic-r.b* morphants (B,I,L), *Mrf* morphants (C,F) and *Tbx6-r.b* morphants (E,H,K) at the at the 64- and 112-cell stages. The number of embryos that were examined (*n*) and the percentage of embryos that each panel represents (%) are shown. Expression in the three muscle lineages is indicated by arrowheads with the color code described in Fig. 1. Loss of expression is indicated by white arrowheads. Note that *Mrf* signals are visible in the cytoplasm, but not in the nuclei, of B7.8 of control embryos at the 64-cell stage. Cytoplasmic signals are hardly detected by fluorescence *in situ* hybridization, therefore *Mrf* signals are not seen in B7.8 in Fig. 2C, but are visible in 3G. Scale bar: 50 µm.

the gastrula stage (Fig. 4B). The RT-qPCR assay showed that expression of this gene was significantly increased between the 32- and 64-cell stages (Fig. 4B).

Tni mRNA was first detected in the B7.4 lineage after the 76-cell stage and then in all of the three lineages at the gastrula stage (Fig. 4C), and the RT-qPCR assay consistently showed the increase after the 64-cell stage (Fig. 4C). *In situ* hybridization signals for *Mlc* were first detected in the B7.4 lineage around the 112-cell stage, and the signals became evident in all three lineages by the gastrula stage (Fig. 4D). Consistently, the RT-qPCR assay showed that the expression level of *Mlc* was increased significantly after the 64-cell stage (Fig. 4D). *In situ* hybridization signals for *Tnc* were first detected in all three lineages at the gastrula stage (Fig. 4E), and this increase was confirmed by the RT-qPCR assay (Fig. 4E). Thus, the expression of these genes became detectable at different stages, and therefore it is likely that initiation of expression was different among these muscle structural genes.

Tbx6-r.b, but not *Mrf*, is important for initiation of *Mrlc* expression in early stage embryos

Our previous study showed that Tbx6-r.b and Mrf bind to the upstream regions of most muscle structural genes and that binding of both of these factors is a good predictor for muscle expression (Kubo et al., 2010). However, the above analysis revealed that *Mrf* expression did not necessarily precede expression of the examined muscle structural genes; *Tnt*, *Mrlc* and *Mrf* expression was initiated simultaneously at the 64-cell stage in the B7.8 lineage, whereas in the B7.4 lineage, *Tnt* and *Mrlc* expression preceded *Mrf* expression in some embryos. Therefore, we reasoned that Mrf might be unnecessary for initiation of *Mrlc* expression, but required for its expression in late-stage embryos (gastrula and tailbud embryos).

To test the hypothesis, we knocked down *Mrf* using a specific MO. It has been shown that knockdown of *Mrf* reduces the expression of genes that encode Mlc, Tni and muscle actin (Imai et al., 2006; Meedel et al., 2007). Whereas *Mrlc* expression was similarly reduced by knockdown of *Mrf* at the tailbud stage, at the 64-cell stage the *Mrlc* expression level was unchanged in *Mrf* morphants (Fig. 5A-C). As we have previously shown that *Mrlc* is downregulated in *Tbx6-r.b* morphants at the 112-cell stage (Yagi et al., 2005), we confirmed this finding at the 64-cell stage in the present study. Indeed, the *Mrlc* expression level was reduced at the 64-cell stage of *Tbx6-r.b* morphants (Fig. 5A,C). Thus, Tbx6-r.b, but not Mrf, is responsible for the initiation of *Mrlc* expression, and Mrf is required for *Mrlc* expression in tailbud embryos.

Tbx6-r.b directly activates *Mrlc* expression

To understand how *Mrlc* is activated, we analyzed the upstream region of *Mrlc3*, which is one of the *Mrlc* copies, using reporter constructs. We first prepared a series of four *Gfp* reporter constructs, which contained different lengths of the upstream sequence, and examined whether they were expressed in the muscle lineages at the 64-cell, gastrula and tailbud stages. The construct that contained the upstream 442 bp sequence (−442>*Gfp*) was expressed in 66%-97% of embryos (Fig. 6A-C). The reporter that contained the 146 bp upstream sequence (−146>*Gfp*) was expressed almost as efficiently as −442>*Gfp* (Fig. 6A-C). The construct that contained the 101 bp upstream sequence (−101>*Gfp*) was rarely expressed at the 64-cell stage (Fig. 6A). However, it was expressed almost as efficiently as −442>*Gfp* and −146>*Gfp* at the gastrula and tailbud stages (Fig. 6B,C). Further deletion (−58>*Gfp*) almost completely impaired expression of the reporter at all stages (Fig. 6A-C). These observations indicated a crucial element between −101 and −146 nucleotide positions for expression at the 64-cell stage, and a crucial element between −58 and −101 nucleotide positions for expression in late-stage embryos.

Between −58 and −146 nucleotide positions of the upstream region, two Tbx6-r.b-binding sites and one Mrf-binding site (E-box) were identified using the Patser program (Hertz and Stormo, 1999) with position weight matrices for *Ciona* Tbx6-r.b (Yagi et al., 2005) and mouse Myod1 (MA0499.1 in the JASPAR database; Sandelin et al., 2004). The downstream Tbx6-r.b-binding site (Td site) was next to the Mrf-binding site (E site), and these two sites are therefore designated collectively as the Td/E site hereafter.

Development (2019) 146, dev173104. doi:10.1242/dev.173104

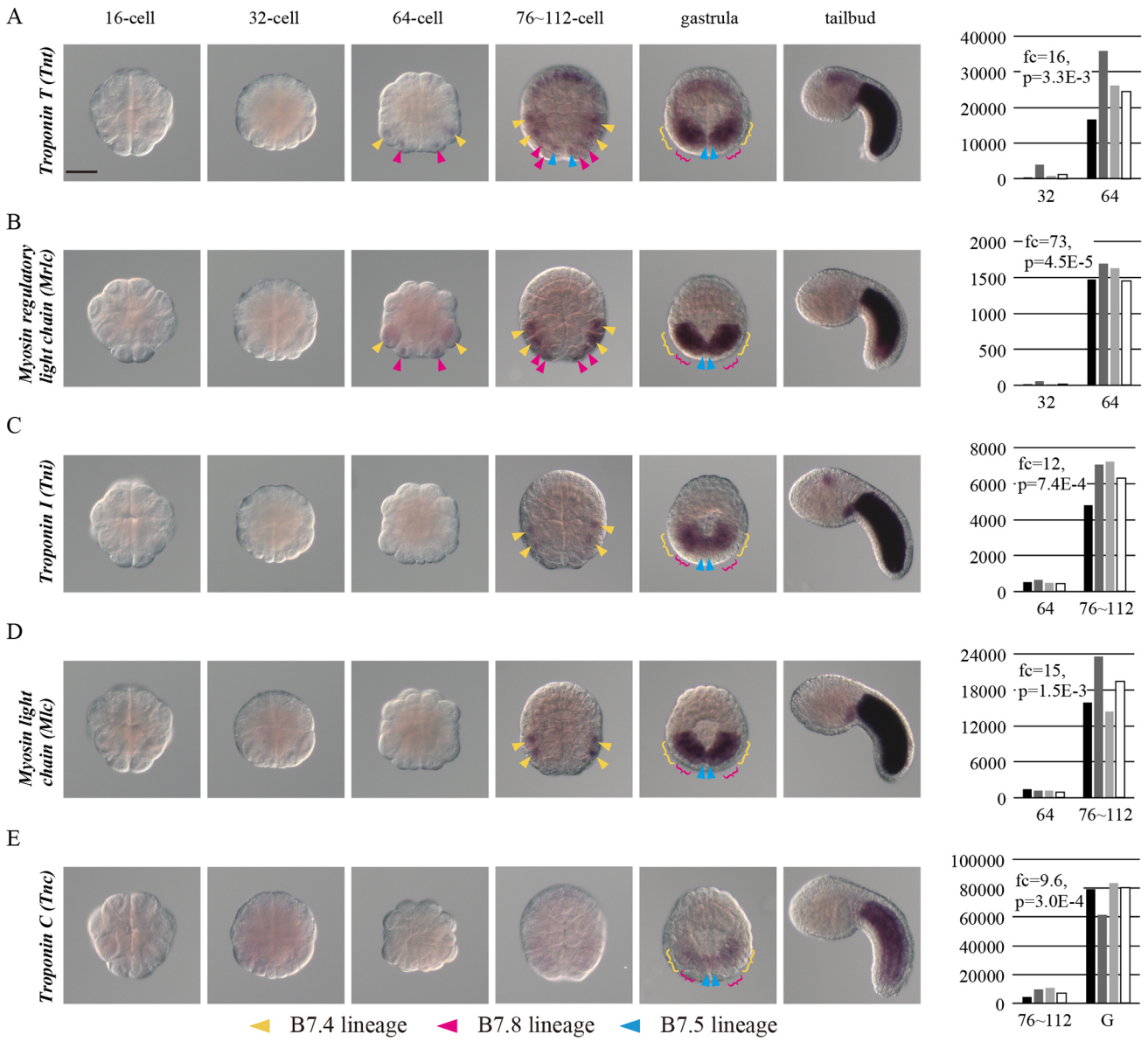

Fig. 4. Temporal and spatial expression patterns of muscle structural genes. (A-E) Expression patterns of *Tnt* (A), *Mrlc* (B), *Tni* (C), *Mlc* (D) and *Tnc* (E) were revealed using *in situ* hybridization (photomicrographs on left) and RT-qPCR (quantification on right). Photomicrograph show signals indicated by arrowheads or lines with the same color code described in Fig. 1 up to the gastrula stage. The *y*-axis in the quantification indicates the number of molecules per embryo. Results from four independent experiments are shown as bars with different colors. One-tailed paired *t*-tests were performed, and *P*-values and fold change values (fc) are shown. Note that *Mrlc* and *Mlc* are multicopy genes. In the analyses shown in this figure, these copies were not discriminated. Scale bar: 50 µm.

This finding was consistent with previously published results that sequences similar to the Tbx protein-binding motif and E-boxes are frequently found in many muscle structural genes (Brown et al., 2007; Kusakabe et al., 2004). We noticed that the upstream Tbx6-r.b-binding site (Tu site) and Td/E site were similar to the two crucial cis-elements for expression of a *Halocynthia* actin gene, *HrMA4a*, which were identified in a previous study (Satou and Satoh, 1996) (Fig. 6D).

Because *Tbx6-r.b* was responsible for initiation of *Mrlc* expression (Fig. 5), we next examined functions of the predicted Tbx6-r.b-binding sites. First, we confirmed that these sites were able to bind Tbx6-r.b *in vitro* using an electrophoresis mobility shift assay. As shown in Fig. 6E, a shifted band was detected for both of the Tu and Td sites, which disappeared with the addition of a specific competitor.

Next, to examine whether these two elements are required for reporter expression in embryos, we introduced mutations into these elements in the reporter that contained the upstream 146-bp sequence. The mutation introduced into the upstream element (µTu) reduced the expression greatly at the 64-cell stage, whereas it did not reduce expression at the gastrula and tailbud stages (Fig. 6F-J). However, the mutation introduced into the downstream element (µTd) reduced the expression greatly in gastrula and tailbud embryos but not in early stage embryos (Fig. 6F-H,K). Thus, these two elements, both of which could bind Tbx6-r.b, played different roles in regulating *Mrlc3* expression.

A mutation introduced into the E-box within the Td/E site did not reduce the number of embryos that expressed the reporter (Fig. 6F-H). However, as shown in Fig. 6L, the number of muscle cells that expressed the reporter was reduced greatly. The peak of Mrf binding,

Development (2019) 146, dev173104. doi:10.1242/dev.173104

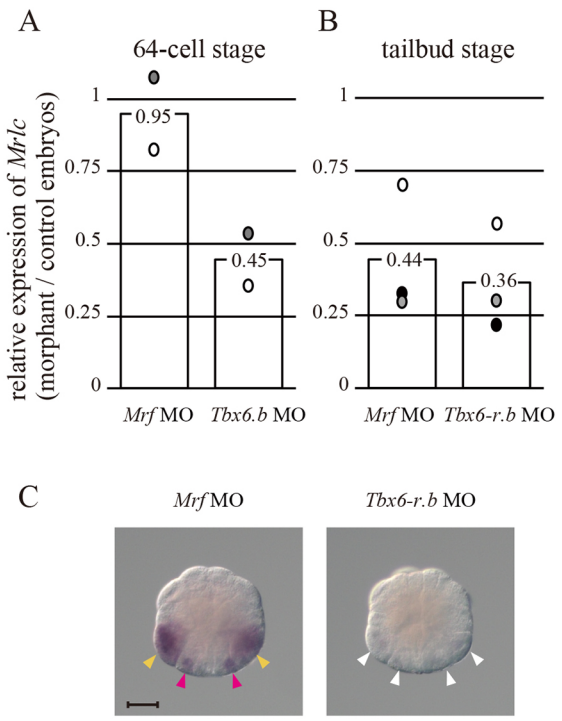

Fig. 5. Tbx6-r.b and Mrf function differently to activate *Mrlc*. (A,B) The expression levels of *Mrlc* in *Mrf* morphants or *Tbx6-r.b* morphants as relative values to the expression level in control embryos was measured by RT-qPCR. Two and three independent experiments were performed at the 64-cell (A) and the tailbud (B) stage, respectively, and shown by dots with different colors. Bars and numbers on them indicate their averages. (C) Expression of *Mrlc* in *Mrf* morphants and *Tbx6-r.b* morphants at the 64-cell stage. Expression signals in the muscle lineages are indicated by arrowheads with the same color code as in Fig. 1. Loss of expression is indicated by white arrowheads. Scale bar: 50 µm.

which was revealed using a ChIP-Chip assay in a previous study (Kubo et al., 2010), was consistently found near the E-box (Fig. S3). Therefore, this E-box is indeed functional, although it may be less important than the Td site.

The degree of Tbx6-r.b occupancy is a good predictor for the timing of when target gene expression becomes detectable and the strength of target gene expression

On the basis of the above observation that Tbx6-r.b is responsible for the expression of *Mrlc* in early stage embryos, we hypothesized that the degree of Tbx6-r.b occupancy to regulatory regions is related to the timing of when expression of Tbx6-r.b targets becomes detectable. In a previous study (Kubo et al., 2010), we analyzed Tbx6-r.b binding using a ChIP assay in which an expression construct that encoded a Tbx6-r.b-Gfp fusion protein under the control of the *Tbx6-r.b* upstream regulatory sequence was electroporated and resultant embryos were subjected to a ChIP assay with an anti-Gfp antibody. However, because this previous study employed microarrays and the resolution was not necessarily sufficient, we similarly prepared the samples and analyzed them using sequencing in the present study (Fig. 7A).

We estimated how much Tbx6-r.b binds to over 1 kb upstream regions of the muscle structural genes that were analyzed in Fig. 4 by calculating summation of fold-change values in the Tbx6-r.b-ChIP samples against the control whole-cell extract samples (ChIP-score), because crucial regulatory elements have been found in 1 kb upstream regions of these genes (Brown et al., 2007; Kusakabe et al., 2004; Satoh et al., 1996; Satou and Satoh, 1996). The calculated scores for *Tnt* and *Mrlc* were higher than those for the three other genes (Fig. 7B). These results suggested that the degree of Tbx6-r.b occupancy may be a good predictor for the timing of when expression of muscle-specific genes becomes detectable. At the same time, this observation indicates that any effect of *Tbx6-r.b* overexpression to specific binding of Tbx6-r.b was minimal, and that the present ChIP assay successfully identified interactions between Tbx6-r.b and DNA.

We previously showed that 15 common target genes of Mrf and Tbx6-r.b are expressed in muscle cells (Kubo et al., 2010). We calculated Tbx6-r.b ChIP scores for these genes, and found that the scores for six genes were similar to or higher than those for *Tnt* and *Mrlc* (Fig. 7C). Next, we compared the expression levels of seven genes in 32-cell, 64-cell, 76- to 112-cell and gastrula embryos with those in 16-cell embryos using RT-qPCR, because the remaining genes were expressed maternally and zygotically and it was difficult to precisely measure their expression levels using RT-qPCR. As shown in Fig. 7D, the gene with the highest score (*KH.C9.27*) was indicated to be expressed strongly as early as the 32-cell stage, and the others with high ChIP scores (*KH.C9.496*, *KH.L4.43* and *KH.C3.756*) were expressed strongly at the 64-cell stage. The expression level of *KH.C11.49*, which showed a lower ChIP score, was increased significantly at the 64-cell stage, but the change was much smaller than the changes observed for the above four genes. The increase in the expression level of *KH.C14.188* was weakly supported at the 64-cell stage and strongly supported at the 112-cell stage by our statistical analysis, and the first significant increase in the expression level of *KH.C10.250* was observed at the 112-cell stage. These two genes had low ChIP scores for Tbx6-r.b, and the changes in their expression levels were small. Specifically, genes with high ChIP scores were expressed strongly in early stage embryos before gastrulation.

To further confirm the hypothesis, we prepared synthetic *Gfp* reporter constructs that contained 2, 4, 8 and 16 copies of the 40 bp region including the Tu site and the basal promoter of *Brachyury* (*2x*, *4x*, *8x* and *16xTu>Gfp*). As shown in Fig. 8A, whereas *2xTu>Gfp* and *4xTu>Gfp* were not expressed at the 32-cell stage, *2xTu>Gfp* was expressed weakly at the 64-cell and gastrula stages, *4xTu>Gfp* was strongly expressed at the 64- and 112-cell stages, and *8xTu>Gfp* and *16xTu>Gfp* were expressed even at the 32-cell stage (Fig. 8B). These results supported our hypothesis. In addition, *4xTu>Gfp* co-injected with synthetic *Tbx6-r.b* mRNA was expressed strongly at the 32-cell stage, whereas the construct that was co-injected with synthetic control *lacZ* mRNA was not (Fig. 8C). Conversely, when the *Tbx6-r.b* MO was co-injected, *8xTu>Gfp* was rarely expressed at the 32-cell stage (Fig. 8C). These data indicate that muscle structural genes that bind more Tbx6-r.b are expressed earlier and more strongly.

DISCUSSION

The genetic program from maternal factors to expression of muscle-specific structural genes

Our results, together with those from previous studies, reveal the genetic pathway for muscle cell differentiation in the ascidian embryo (Fig. 9; Fig. S4). This pathway starts with three maternal factors, Macho-1, β-catenin and Gata.a, and ends with expression of muscle structural genes. A combination of a classical experimental approach and a high-throughput approach revealed this pathway, and we found that the gene expression patterns and regulatory mechanisms were much more complex than was previously expected. First, Macho-1 is maternally expressed and localized in

Development (2019) 146, dev173104. doi:10.1242/dev.173104

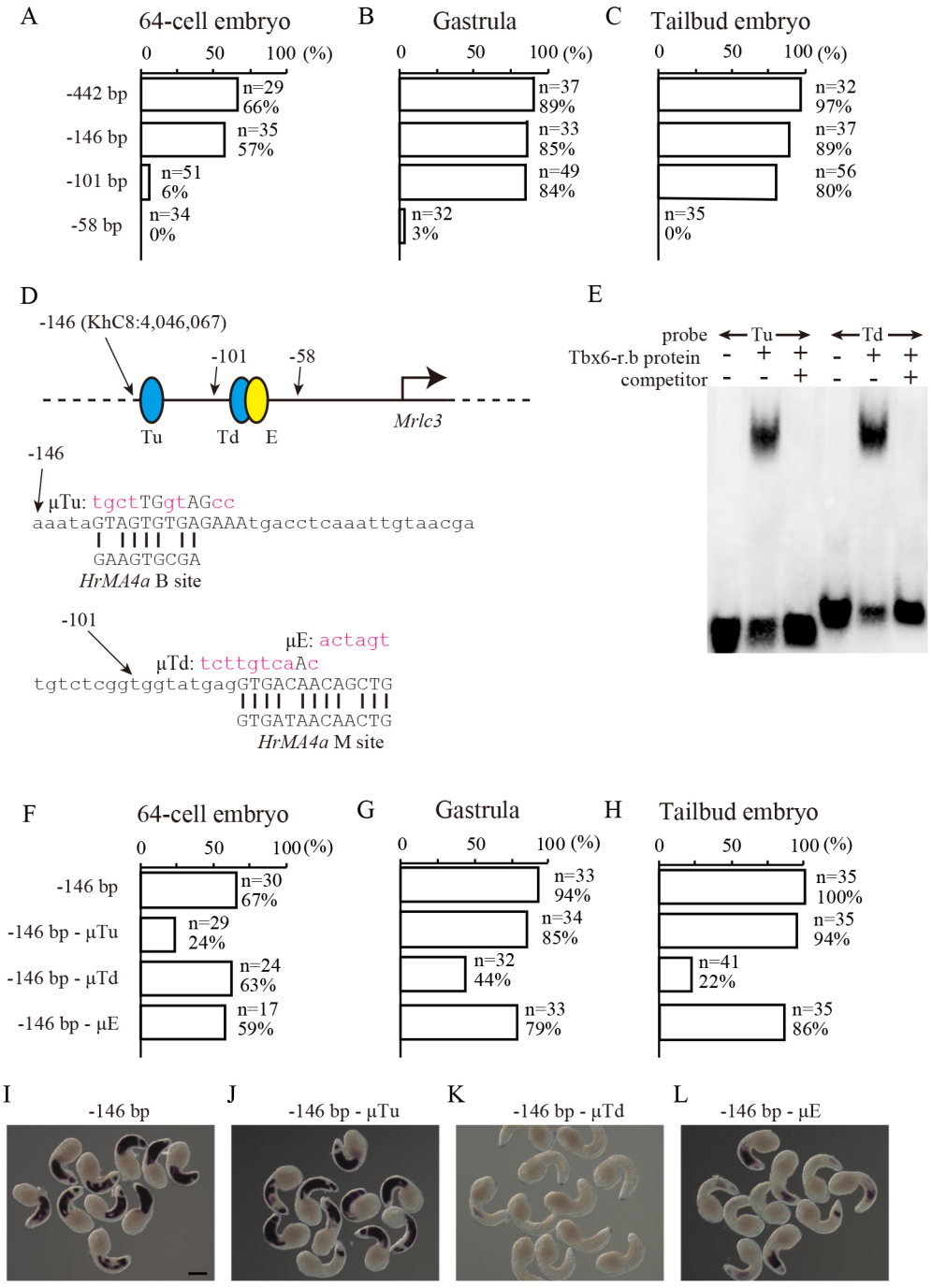

Fig. 6. *Mrlc3* is regulated by Tbx6-r.b. (A-C) Expression of a series of four deletion reporter constructs for *Mrlc3* at the 64-cell (A), gastrula (B) and tailbud (C) stages. The number of embryos (*n*) and the percentage of embryos that expressed the reporter (%) are shown. (D) Depiction of the structure of the upstream region of *Mrlc3*. Two putative Tbx6-r.b-binding sites (Tu and Td sites; blue ovals) and an E-box (a putative Mrf-binding site; yellow oval) are shown. The nucleotide sequence including these binding sites and alignments of Tu and Td/E sites with two essential elements found in *Halocynthia* (Satou and Satoh, 1996) and mutated sequences introduced in μTu, μTd and μE constructs are shown. (E) Electromobility shift (gel-shift) assay showing that Tbx6-r.b specifically bound to the Tu and Td sites *in vitro*. (F-H) Expression of the control and three mutant reporter constructs for *Mrlc3* at the 64-cell (F), gastrula (G) and tailbud (H) stages. The number of embryos (*n*) and the percentage of embryos that expressed the reporter (%) are shown. (I-L) *In situ* hybridization of tailbud embryos with the introduced control and three mutant constructs. The introduced constructs are indicated above each photomicrograph. Note that the percentage of embryos with reporter expression was not greatly different between embryos with the introduced μE and control reporter constructs, but the number of cells with reporter expression was greatly reduced in embryos with the introduced μE construct. Scale bar: 100 μm.

the posterior pole, and acts together with Tcf7/β-catenin to activate *Tbx6-r.b* in the muscle lineages (Nishida and Sawada, 2001; Oda-Ishii et al., 2016; Satou et al., 2002; Yagi et al., 2004). *Zic-r.b* is also activated in muscle lineages under the control of a maternal protein, Gata.a (Imai et al., 2016), probably in combination with Tbx6-r.b at the 32-cell stage (Yagi et al., 2005). Their protein products cooperatively activate *Mrf*, and then *Tbx6-r.b* is activated again by Zic-r.b and Mrf. *Tbx6-r.b*, *Zic-r.b* and *Mrf* are regulated negatively through auto-regulatory loops (Imai et al., 2006), which explains why these genes are downregulated immediately after their initial expression. In addition, although the expression order of these three genes is slightly different among the three muscle lineages, it is persuasive that this cross-activating gene circuit maintains *Tbx6-r.b* and *Mrf* expression during gastrulation.

Expression of *Tbx6-r.b* and *Mrf* decreases gradually after gastrulation. Nevertheless, *Mrf* expression is detected even at the tailbud stage (Araki et al., 1994; Meedel et al., 1997), whereas *Tbx6-r.b* expression is terminated in the primary lineage of muscle cells before the neurula stage (Takatori et al., 2004). It is possible that Tbx6-r.b protein is stable and maintained in muscle cells even in tailbud embryos after degradation of the mRNA. In addition, it is also possible that Tbx6-r.b is not present at the tailbud stage, but is required only for the time when the Td/E site begins to function. Although this remains to be resolved, the ascidian larva undergoes a metamorphosis within a few days, and muscle cells die by apoptosis (Chambon et al., 2002; Lemaire, 2011) and therefore it may not be required to maintain Tbx6-r.b and Mrf expression for a long time.

Development (2019) 146, dev173104. doi:10.1242/dev.173104

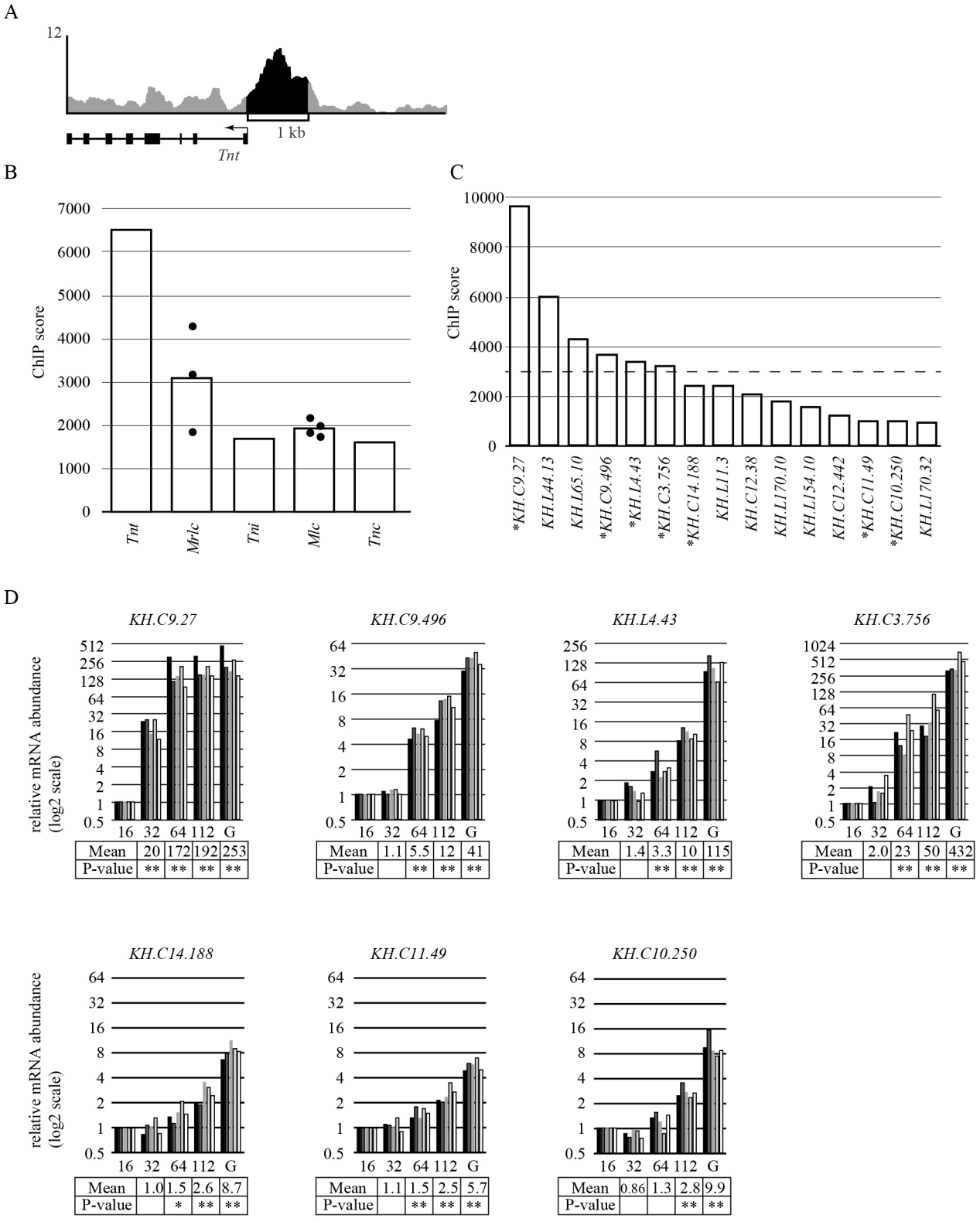

Fig. 7. See next page for legend.

In late-stage embryos (gastrula and later embryos), Mrf plays a role in activating *Mrlc* cooperatively with Tbx6-r.b. A previous study showed that most known muscle structural genes bind Tbx6-r.b and Mrf in their upstream regions (Kubo et al., 2010). Therefore, it is highly likely that these muscle structural genes are regulated similarly to *Mrlc*. However, a small fraction of muscle structural genes, including the gene that encodes muscle creatine kinase, do not bind Mrf strongly (Kubo et al., 2010). This observation suggests that additional regulatory factors may be required to activate such genes. Indeed, previous studies have suggested that another

Development (2019) 146, dev173104. doi:10.1242/dev.173104

Fig. 7. Tbx6-r.b occupancy scores calculated from a ChIP-seq assay correlates with the timing of when expression of muscle genes becomes detectable. (A) A construct driving the expression of a gene that encodes a Tbx6-r.b-Gfp fusion protein was introduced by electroporation, and an anti-Gfp antibody was used for ChIP. Precipitated DNA fragments were analyzed using deep sequencing. A chromosomal region containing the upstream region and exons of *Tnt* is shown. Signal strength within the upstream 1 kb region is shown in black. The graph includes data of two biological duplicates. (B) Tbx6-r.b occupancy scores calculated by a ChIP-seq assay over the 1 kb upstream regions of the five muscle structural genes that were examined in Fig. 4. Scores are based on the summation of fold-change values over the 1 kb regions and are shown by bars. Because *Mrlc* and *Mlc* are multicopy genes, scores for each of these copies are shown by dots and their averages are shown by bars. (C) Results of the same calculation for 15 genes that are expressed in muscle cells (Kubo et al., 2010) are shown by bars. Six of the bars are higher than the level of *Mrlc* (dotted line). (D) Quantification of expression patterns of seven genes, indicated by asterisks in C, as determined using RT-qPCR at the 16-, 32-, 64-, 112-cell and gastrula (G) stages. The *y*-axis indicates relative mRNA abundance compared with mRNA abundance in 16-cell embryos in a log scale. Results from five independent experimental results are shown in the graphs as bars with different colors. *$P \leq 0.01$ (=0.05/5), **$P \leq 0.002$ (=0.01/5) (One-tailed paired *t*-tests with Bonferroni correction method).

transcription factor may be involved in the expression of some muscle genes (Brown et al., 2007; Kusakabe et al., 2004), although its identity is unknown.

A previous study has suggested that such cross-activation among multiple copies of Tbx6-r.b, which are often called Tbx6-r.c and Tbx6-r.d, is important to maintain the expression of *Tbx6-r.b* (Kugler et al., 2010). Although our current data are not necessarily contradictory to this, it is likely that *Tbx6-r.b* expression is mainly maintained by cross-activation between Tbx6-r.b and Mrf. First, the *in situ* hybridization results showed that the initial *Tbx6-r.b* expression was transient and reactivated again in the same lineage of cells. It may be difficult to explain this observation using the previous model. Second, in our previous study, we observed upregulation of *Tbx6-r.b* in *Tbx6-r.b* morphant embryos (Imai et al., 2006), which suggested that Tbx6-r.b represses itself directly or indirectly.

Muscle structural gene expression is quantitatively regulated by Tbx6-r.b in early stage embryos

In the present study, we found a muscle structural gene that began expression at the 32-cell stage of *Ciona* embryos (Fig. 7D). However, not all muscle structural genes are expressed at this stage. Our ChIP data indicated that genes that more strongly bind Tbx6-r.b are expressed more strongly and their expression becomes detectable at earlier stages. We also obtained evidence indicating that the number of Tbx6-r.b-binding sites, which was expected to be related to the extent of Tbx6-r.b binding over the regulatory regions, could be a factor to determine the timing and strength of target gene expression

As discussed previously (Satou and Imai, 2015), most expression patterns of regulatory genes in *Ciona* embryos can be explained simply by combinations of expressed upstream regulatory factors. Namely, the quantitative aspect is less important to control expression of regulatory genes in ascidian embryos. However, our finding in the present study suggests that expression of non-regulatory genes is controlled differently, namely in a quantitative manner. Our data indicate that genes that bind Tbx6-r.b more extensively are activated more strongly, and thereby their expression becomes detectable earlier. This is consistent with a previous finding for genes that are expressed in the notochord (Katikala et al., 2013).

Regulation of muscle structural genes in late-stage embryos

The E-box is a motif for Mrf binding (Blackwell and Weintraub, 1990). Indeed, knockdown of *Mrf* reduced *Mrlc* expression at the tailbud stage. Similarly, it reduces the expression of *Tni* and a muscle actin gene (Imai et al., 2006; Meedel et al., 2007). A mutation introduced into the E-box within the Td/E site of *Mrlc3* also reduced reporter expression. In addition, it has been shown that Mrf binds to the upstream region of *Mrlc3* (Kubo et al., 2010). These data indicate that Mrf activates *Mrlc* directly. Nevertheless, at the 64-cell stage, *Mrlc* expression was not affected by *Mrf* knockdown. Therefore, Mrf plays a role in activating muscle genes in late-stage embryos, but not in early stage embryos before gastrulation. Because *Mrf* expression

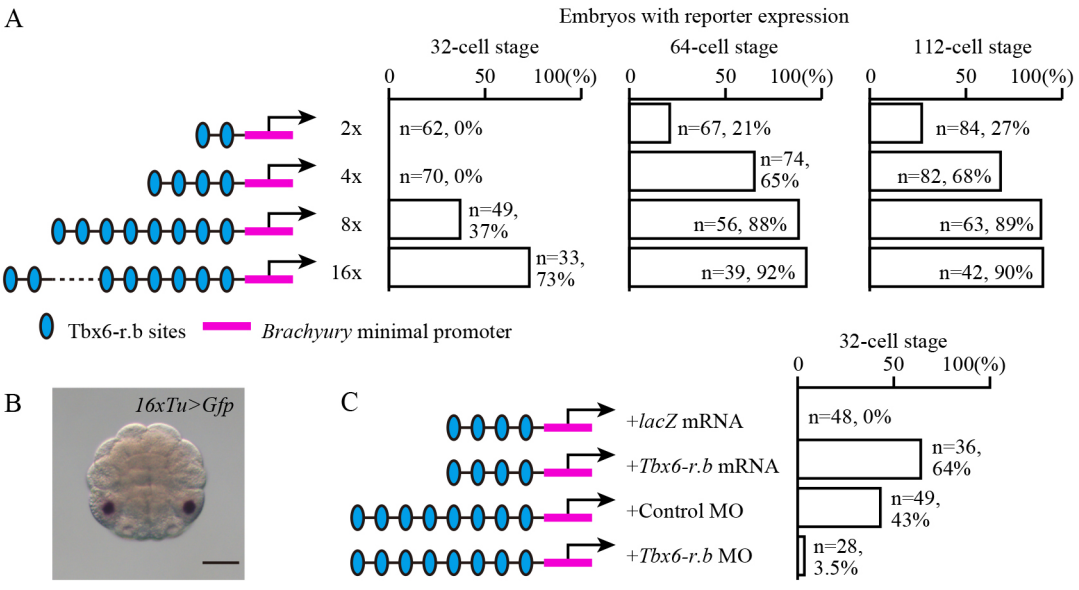

Fig. 8. Reporter gene expression is induced earlier with higher numbers of Tbx6-r.b-binding sites. (A) Expression of four artificial reporter genes, in which two, four, eight and 16 Tbx6-r.b-binding sites and the *Brachyury* minimal promoter sequence were introduced by electroporation. The number of embryos (*n*) and the percentage of embryos that expressed the reporter (%) are shown. (B) A 32-cell embryo that expressed the reporter gene with 16 copies of Tbx6-r.b-binding sites. (C) Expression of the reporter gene with four Tbx6-r.b-binding sites was injected with synthetic *lacZ* or *Tbx6-r.b* mRNA, and the reporter gene with eight Tbx6-r.b-binding sites was injected with the control MO or *Tbx6-r.b* MO. Scale bar: 50 µm.

Development (2019) 146, dev173104. doi:10.1242/dev.173104

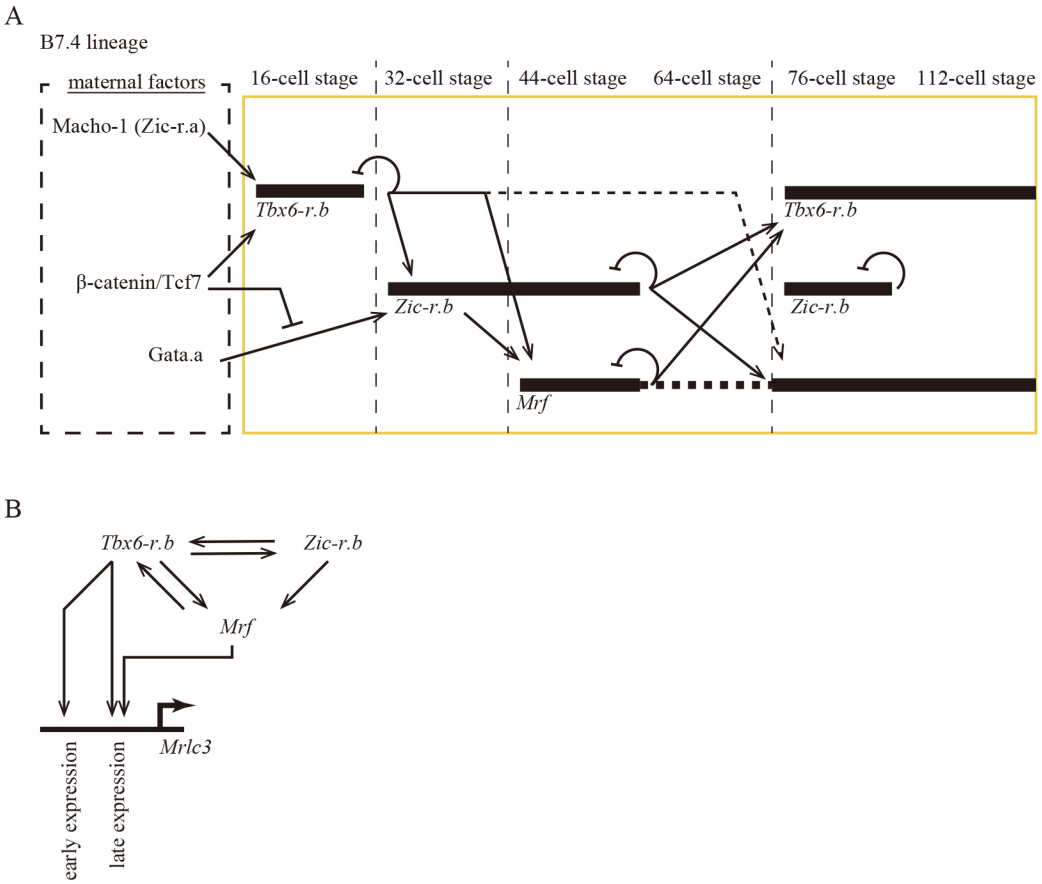

Fig. 9. The genetic pathway from maternal factors to expression of muscle structural genes. (A) Diagram showing regulatory relationships among the key transcription factors in the B7.4 lineage of early stage embryos. Thick lines indicate expression; thick dotted line indicates weak expression (of *Mrf*). Although *Tbx6-r.b* regulates *Mrf* expression at the 76-cell stage, this regulation might be indirect (dashed line). Negative auto-regulatory loops are not shown for expression of *Tbx6-r.b* and *Mrf* at the 76-cell stage and later, because these genes are not inactivated before the 112-cell stage. This diagram was constructed as an assembly of the present and previous studies: see the text for details. Diagrams for the B7.8 and B7.5 lineages are shown in Fig. S4. (B) Schematic showing how the three key regulatory factors regulate *Mrlc3* expression.

begins later than *Tbx6-r.b* expression, a sufficient amount of Mrf may not accumulate before gastrulation.

We also showed that the Td site, which could bind Tbx6-r.b and was adjacent to the E-box, was essential for *Mrlc3* expression in late-stage embryos. It is persuasive that cooperative action between Tbx6-r.b and Mrf is important for the expression in late-stage embryos. This speculation is consistent with the earlier observation that Mrf and Tbx6-r.b bind to the upstream regions of most known muscle structural genes (Kubo et al., 2010). Although it has not been revealed why this Td site functions in late-stage embryos but not in early stage embryos, there may be functional constraints on spacing of Tbx6-r.b-binding sites and E-boxes and their orientation, as has previously been revealed for Gata-binding sites and Ets-binding sites of *Otx* genes (Farley et al., 2015).

As we described above, a nucleotide sequence similar to the Td/E site exists in the upstream region of a *Halocynthia* actin gene, and it is important for the expression of the actin gene (Satou and Satoh, 1996). Similar sequences are found in the upstream region of the three copies of the *Halocynthia Tni* gene (Yuasa et al., 2002). In our preliminary survey, similar sequences were also found in the upstream regions of several, but not all, muscle structural genes in the *Ciona* genome, and it has been revealed that these genes bind Mrf and Tbx6-r.b (Kubo et al., 2010). Therefore, most muscle structural genes may be activated similarly in late-stage embryos.

As shown in Fig. 6, the activity of the μE construct was reduced but not completely abolished. Similarly, mutations introduced into an E-box within the essential regulatory element of a *Halocynthia* muscle actin gene, *HrMA4a*, does not completely abolish, but reduces, reporter expression (Satou and Satoh, 1996). These observations can be explained by the activity of Tbx6-r.b. First, our data indicate that Tbx6-r.b functions through the Td site near the E-box of *Mrlc3*. Second, a potential Tbx6-r.b-binding site was also found near the E-box of *HrMA4a*. Namely, the configuration of the region that contains these two sites is the same as that in the Td/E site (Fig. 6D). It is possible that Tbx6-r.b facilitates binding of Mrf to less conserved sequences, because MEF2 assists binding of Mrf proteins to DNA in vertebrates (Groisman et al., 1996; Molkentin et al., 1995). Indeed, Mrf has been shown to act together with various factors (Berkes et al., 2004; Liu et al., 2010).

Ascidian larval muscle is a tissue that develops autonomously (Conklin, 1905). The genetic pathway that underlies this autonomy begins with Macho-1 and antagonism between two additional maternal factors, Gata.a and β-catenin, and includes the cross-regulatory circuit among Tbx6-r.b, Mrf and Zic-r.b, which is important for activating muscle structural genes. Thus, owing to the previous and present studies, the genetic pathway starting with *Tbx6-r.b* and *Zic-r.b*, which are activated by maternal factors, and ending with expression of muscle structural genes, has been revealed. The comprehensive understanding of this pathway will allow us to

Development (2019) 146, dev173104. doi:10.1242/dev.173104

deduce evolutionary constraints on this pathway using comparative studies among tunicates, and will be useful for revealing how tunicates have evolved the autonomous pathway for muscle cell differentiation. The precise temporal gene expression profiles at the single cell resolution will also facilitate accurate modeling of the dynamics of the gene regulatory network and its further analysis.

MATERIALS AND METHODS

Animals and cDNAs

C. intestinalis (type A) adults were obtained from the National Bio-Resource Project for *Ciona intestinalis*. The cDNA clones were obtained from our EST clone collection (Satou et al., 2005). Identifiers (Satou et al., 2008; Stolfi et al., 2015) for genes that were examined in the present study are as follows: CG.KH2012.S654.1/2/3 for *Tbx6-r.b*, CG.KH2012.C14.307 for *Mrf*, CG.KH2012.L59.1/12/ CG.KH2012.S816.1/2/4 for *Zic-r.b*, CG.KH2012.C4.57 for *Tnt*, CG.KH2012.C8.309 for *Mrlc3*, CG.KH2012.C8.477/859 for *Mrlc1/2*, CG.KH2012.C11.673 for *Tni*, CG.KH2012.C1.1186/20.v1/20.v2/216 for *Mlc* and CG.KH2012.C12.417 for *Tnc*. Transcript models used in the analysis shown in Fig. 7 were CG.KH2012.C4.57.v3.A.SL1-1, CG.KH2012.C8.309.v1.A.nonSL1-1, CG.KH2012.C8.477.v1.A.nonSL1-1, CG.KH2012.C8.859.v1.A.nonSL1-1, CG.KH2012.C11.673.v1.A.SL1-1, CG.KH2012.C1.1186.v1.A.nonSL1-1, CG.KH2012.C1.20.v1.A.nonSL1-2, CG.KH2012.C1.20.v2.A.nonSL2-1 and CG.KH2012.C1.216.v1.A.SL2-1. Because the transcription start site of *Tnc* was not determined precisely, we mapped TSS-seq data (Yokomori et al., 2016) and regarded the nucleotide position 15,401 of chromosome 12 as the transcription start site.

Gene knockdown/overexpression assays and reporter assays

The MOs against *Tbx6-r.b* and *Mrf* (custom-made by Gene Tools), which blocked translation of *Tbx6-r.b* and *Mrf* mRNA, were used for the knockdown experiments (*Tbx6-r.b*, 5′-TTACAATTTCCTCTCTCTTTCGATT-3′; *Mrf*, 5′-GCTCCTCTAGAGAGATACACGTCAT-3′). These MOs have been used in a previous study (Imai et al., 2006), therefore, we did not perform experiments to further confirm their specificity in the present study. We also used the standard control MO purchased from Gene Tools. The MOs were introduced by microinjection under a microscope. Synthetic mRNAs were prepared from cDNA that was cloned into the pBluescript RN3 vector (Lemaire et al., 1995) using the mMESSAGE mMACHINE T3 Kit (Thermo Fisher Scientific) and injected into unfertilized eggs (0.02 mg/ml for *Mrf* and 0.03 mg/ml for *Tbx6-r.b*). All knockdown and overexpression phenotypes were confirmed in at least two independent injections.

The 442 bp upstream sequence used to prepare the reporter construct for *Mrlc3* was obtained from KhC8:4,045,922-4,046,367 of the KH version of the *Ciona* genome sequence (Satou et al., 2008). Various constructs were prepared from this reporter gene. The reporter constructs were introduced by microinjection when they were co-introduced with MOs and/or mRNAs. Otherwise, they were introduced by electroporation. We used *Gfp* as a reporter and examined reporter expression using *in situ* hybridization. The reporter constructs used in the experiment shown in Fig. 8 consisted of 2-16 short DNA fragments containing the Tu site of *Mrlc3* (KhC8:4046071-4046032) and the minimal promoter of *Brachyury* (KhS1404:6203-6275).

Whole-mount *in situ* hybridization and RT-qPCR

In situ hybridization was performed as described previously (Ikuta and Saiga, 2007; Satou et al., 1995). To quantify endogenous gene expression using RT-qPCR we used a Cell-to-Ct kit (Thermo Fisher Scientific). For each measurement, 20 embryos were lysed. For each absolute quantification, we first constructed standard curves using a dilution series of the corresponding cDNA clones. For relative quantification, we used *Pou2f* as an internal control. SYBR Green chemistry was used for qPCR. The primers for the experiment shown in Fig. 7 have been described in a previous study (Kubo et al., 2010). The remaining primers were as follows: *Pou2f*, 5′-AAGATGGTTGCTGGATGCTAATAAT-3′ and 5′-TTGGATTGGAGTGGGAATAACAA-3′; *Tnt*, 5′-GAATACAGCAGCGAGGAGGAGA-3′ and 5′-TCGCTTGCTGGGTTTGCT-3′; *Mrlc* 5′-TGCAATCAACTTTACCGTCTTCCT-3′ and 5′-CTTCGGAGTCGGTTCCATGT-3′; *Tni*, 5′-CGATTTGACATTGAGGCAAGAGTA-3′ and 5′-CGTCTCAGTGGTGGTCGCTTA-3′; *Mlc*, 5′-AGAACGTATGGAAGAAATCAAGGAA-3′ and 5′-TCTCCAACTTGATCGAATCCAA-3′; *Tnc*, 5′-GACGCGAGAGCAAAATTGG-3′ and 5′-TGTCGAACGCTTGCTTGAAT-3′. There are several gene copies that encode Mrlc and Mlc in the genome. Our probes for *in situ* hybridization and primers for RT-qPCR probably detected mRNAs from all of these copies.

Gel-shift assay

Recombinant Tbx6-r.b protein was produced as a fusion protein of the Tbx6-r.b DNA-binding domain and glutathione S-transferase in *Escherichia coli* DH5α strain (Thermo Fisher Scientific). The protein was purified under a native condition using glutathione sepharose 4B (GE Healthcare). After annealing two complementary oligonucleotides (5′-AAAAAATAGTAGTGTGAGAAATGACC-3′ and 5′-AAAGGTCATTTCTCACACTACTATTT-3′ for the Tu site, and 5′-AAAGTGGTATGAGGTGACAACAGCTGATTGGCT-3′ and 5′-AAAAGCCAATCAGCTGTTGTCACCTCATACCAC-3′ for the Td site), both protruding ends of the double-stranded oligonucleotides were filled with biotin-11-dUTP. This biotin-labeled oligonucleotide was used as a probe. Proteins and the biotin-labeled probe were mixed in 10 mM Tris (pH 7.5), 50 mM KCl, 1 mM DTT, 50 ng/µl poly(dIdC), 2.5% glycerol and 0.05% NP40 with or without competitor double-stranded DNAs (100-fold molar excess). Protein amounts that we added in reactions were determined empirically. Protein-DNA complexes were detected using a Chemiluminescent Nucleic Acid Detection Module Kit (Thermo Fisher Scientific).

Chromatin immunoprecipitation

We used a DNA construct that encoded GFP-tagged Tbx6-r.b under the control of the *Tbx6-r.b* promoter (Kubo et al., 2010). Embryos were fixed at the early gastrula stage. The embryos were subjected to ChIP analysis using an anti-GFP antibody (Thermo Fisher Scientific, A6455, 2 µl per 500 mg of embryos), and the immunoprecipitated DNA was amplified using ligation-mediated PCR (Kubo et al., 2010). DNA in whole-cell extracts was used as a control. High-throughput DNA sequencing was performed with an Ion PGM instrument. To calculate fold enrichment for each genomic region, we first called peaks using MACS2 (Zhang et al., 2008) with the option '–nomodel'.

Acknowledgements

We thank Drs Reiko Yoshida, Satoe Aratake and Manabu Yoshida, and other members working under the National Bio-Resource project (Japan Agency for Medical Research and Development) for providing experimental animals.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: Y.S.; Formal analysis: D.Y., I.O.-I., Y.S.; Investigation: D.Y., I.O.-I., A.K.; Writing - original draft: Y.S.; Visualization: D.Y.; Supervision: Y.S.; Funding acquisition: Y.S.

Funding

This research was supported by the Core Research for Evolutional Science and Technology (CREST) program of the Japan Science and Technology Agency (JPMJCR13W6) and a grant from the Japan Society for the Promotion of Science (17KT0020) to Y.S.

Data availability

The ChIP-seq data produced in this study have been deposited in the SRA database under accession number DRA007140.

Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.173104.supplemental

References

Araki, I., Saiga, H., Makabe, K. W. and Satoh, N. (1994). Expression of Amd1, a gene for a Myod1-related factor in the ascidian *Halocynthia roretzi*. *Roux Arch. Dev. Biol.* **203**, 320-327.

Berkes, C. A., Bergstrom, D. A., Penn, B. H., Seaver, K. J., Knoepfler, P. S. and Tapscott, S. J. (2004). Pbx marks genes for activation by MyoD indicating a role

for a homeodomain protein in establishing myogenic potential. *Mol. Cell* **14**, 465-477.
Blackwell, T. K. and Weintraub, H. (1990). Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **250**, 1104-1110.
Brown, C. D., Johnson, D. S. and Sidow, A. (2007). Functional architecture and evolution of transcriptional elements that drive gene coexpression. *Science* **317**, 1557-1560.
Chambon, J. P., Soule, J., Pomies, P., Fort, P., Sahuquet, A., Alexandre, D., Mangeat, P. H. and Baghdiguian, S. (2002). Tail regression in Ciona intestinalis (Prochordate) involves a Caspase-dependent apoptosis event associated with ERK activation. *Development* **129**, 3105-3114.
Conklin, E. G. (1905). Mosaic development in ascidian eggs. *J. Exp. Zool.* **2**, 145-223.
Deno, T., Nishida, H. and Satoh, N. (1984). Autonomous muscle cell differentiation in partial ascidian embryos according to the newly verified cell lineages. *Dev. Biol.* **104**, 322-328.
Farley, E. K., Olson, K. M., Zhang, W., Brandt, A. J., Rokhsar, D. S. and Levine, M. S. (2015). Suboptimization of developmental enhancers. *Science* **350**, 325-328.
Fujiwara, S., Maeda, Y., Shin-I, T., Kohara, Y., Takatori, N., Satou, Y. and Satoh, N. (2002). Gene expression profiles in Ciona intestinalis cleavage-stage embryos. *Mech. Dev.* **112**, 115-127.
Groisman, R., Masutani, H., Leibovitch, M. P., Robin, P., Soudant, I., Trouche, D. and HarelBellan, A. (1996). Physical interaction between the mitogen-responsive serum response factor and myogenic basic-helix-loop-helix proteins. *J. Biol. Chem.* **271**, 5258-5264.
Hertz, G. Z. and Stormo, G. D. (1999). Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics* **15**, 563-577.
Ikuta, T. and Saiga, H. (2007). Dynamic change in the expression of developmental genes in the ascidian central nervous system: revisit to the tripartite model and the origin of the midbrain-hindbrain boundary region. *Dev. Biol.* **312**, 631-643.
Imai, K. S., Satou, Y. and Satoh, N. (2002). Multiple functions of a Zic-like gene in the differentiation of notochord, central nervous system and muscle in Ciona savignyi embryos. *Development* **129**, 2723-2732.
Imai, K. S., Hino, K., Yagi, K., Satoh, N. and Satou, Y. (2004). Gene expression profiles of transcription factors and signaling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks. *Development* **131**, 4047-4058.
Imai, K. S., Levine, M., Satoh, N. and Satou, Y. (2006). Regulatory blueprint for a chordate embryo. *Science* **312**, 1183-1187.
Imai, K. S., Hudson, C., Oda-Ishii, I., Yasuo, H. and Satou, Y. (2016). Antagonism between beta-catenin and Gata.a sequentially segregates the germ layers of ascidian embryos. *Development* **143**, 4167-4172.
Izzi, S. A., Colantuono, B. J., Sullivan, K., Khare, P. and Meedel, T. H. (2013). Functional studies of the Ciona intestinalis myogenic regulatory factor reveal conserved features of chordate myogenesis. *Dev. Biol.* **376**, 213-223.
Katikala, L., Aihara, H., Passamaneck, Y. J., Gazdoiu, S., José-Edwards, D. S., Kugler, J. E., Oda-Ishii, I., Imai, J. H., Nibu, Y. and Di Gregorio, A. (2013). Functional Brachyury binding sites establish a temporal read-out of gene expression in the Ciona notochord. *PLoS Biol.* **11**, e1001697.
Kubo, A., Suzuki, N., Yuan, X., Nakai, K., Satoh, N., Imai, K. S. and Satou, Y. (2010). Genomic cis-regulatory networks in the early Ciona intestinalis embryo. *Development* **137**, 1613-1623.
Kugler, J. E., Gazdoiu, S., Oda-Ishii, I., Passamaneck, Y. J., Erives, A. J. and Di Gregorio, A. (2010). Temporal regulation of the muscle gene cascade by Macho1 and Tbx6 transcription factors in Ciona intestinalis. *J. Cell Sci.* **123**, 2453-2463.
Kumano, G., Takatori, N., Negishi, T., Takada, T. and Nishida, H. (2011). A maternal factor unique to ascidians silences the germline via binding to P-TEFb and RNAP II regulation. *Curr. Biol.* **21**, 1308-1313.
Kusakabe, T. (1995). Expression of larval-type muscle actin-encoding genes in the ascidian Halocynthia roretzi. *Gene* **153**, 215-218.
Kusakabe, T., Yoshida, R., Ikeda, Y. and Tsuda, M. (2004). Computational discovery of DNA motifs associated with cell type-specific gene expression in Ciona. *Dev. Biol.* **276**, 563-580.
Lemaire, P. (2011). Evolutionary crossroads in developmental biology: the tunicates. *Development* **138**, 2143-2152.
Lemaire, P., Garrett, N. and Gurdon, J. B. (1995). Expression cloning of Siamois, a Xenopus homeobox gene expressed in dorsal-vegetal cells of blastulae and able to induce a complete secondary axis. *Cell* **81**, 85-94.
Liu, Y. B., Chu, A., Chakroun, I., Islam, U. and Blais, A. (2010). Cooperation between myogenic regulatory factors and SIX family transcription factors is important for myoblast differentiation. *Nucleic Acids Res.* **38**, 6857-6871.
Meedel, T. H., Farmer, S. C. and Lee, J. J. (1997). The single MyoD family gene of Ciona intestinalis encodes two differentially expressed proteins: Implications for the evolution of chordate muscle gene regulation. *Development* **124**, 1711-1721.
Meedel, T. H., Chang, P. and Yasuo, H. (2007). Muscle development in Ciona intestinalis requires the b-HLH myogenic regulatory factor gene Ci-MRF. *Dev. Biol.* **302**, 333-344.
Molkentin, J. D., Black, B. L., Martin, J. F. and Olson, E. N. (1995). Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell* **83**, 1125-1136.
Nishida, H. and Sawada, K. (2001). macho-1 encodes a localized mRNA in ascidian eggs that specifies muscle fate during embryogenesis. *Nature* **409**, 724-729.
Nishikata, T., Yamada, L., Mochizuki, Y., Satou, Y., Shin-i, T., Kohara, Y. and Satoh, N. (2001). Profiles of maternally expressed genes in fertilized eggs of Ciona intestinalis. *Dev. Biol.* **238**, 315-331.
Oda-Ishii, I., Kubo, A., Kari, W., Suzuki, N., Rothbächer, U. and Satou, Y. (2016). A maternal system initiating the zygotic developmental program through combinatorial repression in the ascidian embryo. *PLoS Genet.* **12**, e1006045.
Sandelin, A., Alkema, W., Engstrom, P., Wasserman, W. W. and Lenhard, B. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* **32**, D91-D94.
Satoh, N., Araki, I. and Satou, Y. (1996). An intrinsic genetic program for autonomous differentiation of muscle cells in the ascidian embryo. *Proc. Natl. Acad. Sci. USA* **93**, 9315-9321.
Satoh, N., Satou, Y., Davidson, B. and Levine, M. (2003). Ciona intestinalis: an emerging model for whole-genome analyses. *Trends Genet.* **19**, 376-381.
Satou, Y. and Imai, K. S. (2015). Gene regulatory systems that control gene expression in the Ciona embryo. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **91**, 33-51.
Satou, Y. and Satoh, N. (1996). Two cis-regulatory elements are essential for the muscle-specific expression of an actin gene in the ascidian embryo. *Dev. Growth Differ.* **38**, 565-573.
Satou, Y., Kusakabe, T., Araki, S. and Satoh, N. (1995). Timing of Initiation of Muscle-Specific Gene-Expression in the Ascidian Embryo Precedes That of Developmental Fate Restriction in Lineage Cells. *Dev. Growth Differ.* **37**, 319-327.
Satou, Y., Takatori, N., Yamada, L., Mochizuki, Y., Hamaguchi, M., Ishikawa, H., Chiba, S., Imai, K., Kano, S., Murakami, S. D. et al. (2001). Gene expression profiles in Ciona intestinalis tailbud embryos. *Development* **128**, 2893-2904.
Satou, Y., Yagi, K., Imai, K. S., Yamada, L., Nishida, H. and Satoh, N. (2002). macho-1-Related genes in Ciona embryos. *Dev. Genes Evol.* **212**, 87-92.
Satou, Y., Kawashima, T., Shoguchi, E., Nakayama, A. and Satoh, N. (2005). An integrated database of the ascidian, Ciona intestinalis: Towards functional genomics. *Zool. Sci.* **22**, 837-843.
Satou, Y., Mineta, K., Ogasawara, M., Sasakura, Y., Shoguchi, E., Ueno, K., Yamada, L., Matsumoto, J., Wasserscheid, J., Dewar, K. et al. (2008). Improved genome assembly and evidence-based global gene model set for the chordate Ciona intestinalis: new insight into intron and operon populations. *Genome Biol.* **9**, R152.
Shirae-Kurabayashi, M., Nishikata, T., Takamura, K., Tanaka, K. J., Nakamoto, C. and Nakamura, A. (2006). Dynamic redistribution of vasa homolog and exclusion of somatic cell determinants during germ cell specification in Ciona intestinalis. *Development* **133**, 2683-2693.
Shirae-Kurabayashi, M., Matsuda, K. and Nakamura, A. (2011). Ci-Pem-1 localizes to the nucleus and represses somatic gene transcription in the germline of Ciona intestinalis embryos. *Development* **138**, 2871-2881.
Stolfi, A., Sasakura, Y., Chalopin, D., Satou, Y., Christiaen, L., Dantec, C., Endo, T., Naville, M., Nishida, H., Swalla, B. J. et al. (2015). Guidelines for the nomenclature of genetic elements in tunicate genomes. *Genesis* **53**, 1-14.
Takatori, N., Hotta, K., Mochizuki, Y., Satoh, G., Mitani, Y., Satoh, N., Satou, Y. and Takahashi, H. (2004). T-box genes in the ascidian Ciona intestinalis: characterization of cDNAs and spatial expression. *Dev. Dyn.* **230**, 743-753.
Wada, S. and Saiga, H. (2002). HrzicN, a new Zic family gene of ascidians, plays essential roles in the neural tube and notochord development. *Development* **129**, 5597-5608.
Yagi, K., Satoh, N. and Satou, Y. (2004). Identification of downstream genes of the ascidian muscle determinant gene Ci-macho1. *Dev. Biol.* **274**, 478-489.
Yagi, K., Takatori, N., Satou, Y. and Satoh, N. (2005). Ci-Tbx6b and Ci-Tbx6c are key mediators of the maternal effect gene Ci-macho1 in muscle cell differentiation in Ciona intestinalis embryos. *Dev. Biol.* **282**, 535-549.
Yokomori, R., Shimai, K., Nishitsuji, K., Suzuki, Y., Kusakabe, T. G. and Nakai, K. (2016). Genome-wide identification and characterization of transcription start sites and promoters in the tunicate Ciona intestinalis. *Genome Res.* **26**, 140-150.
Yuasa, H. J., Kawamura, K., Yamamoto, H. and Takagi, T. (2002). The structural organization of ascidian Halocynthia roretzi troponin I genes. *J. Biochem.* **132**, 135-141.
Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nussbaum, C., Myers, R. M., Brown, M., Li, W. et al. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137.

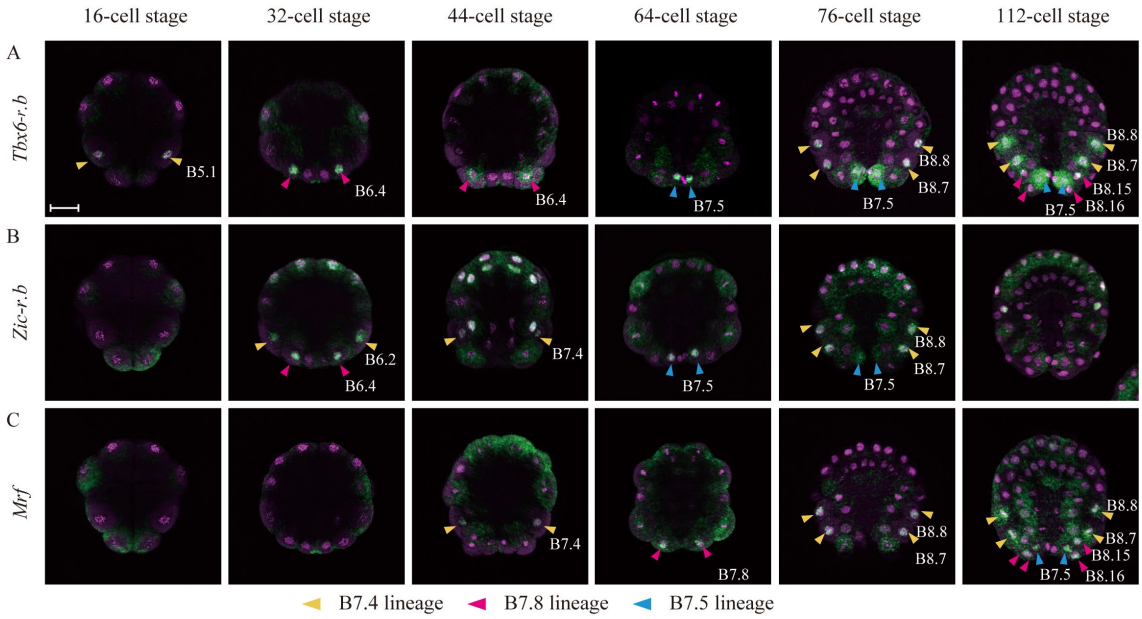

Figure S1. Expression of three key regulatory genes from the 16-cell to 112-cell stages and DAPI staining. Signals for DAPI staining of nuclei (magenta) are superimposed on the images shown in Fig. 2, which show expression of (A) *Tbx6-r.b*, (B) *Zic-r.b*, and (C) *Mrf* in muscle lineages. Arrowheads with the color code described in Fig. 1 indicate expression. The scale bar represents 50 μm.

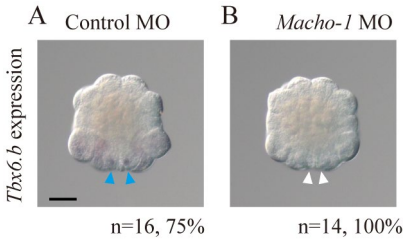

Figure S2. *Tbx6-r.b* expression in B7.5 is regulated by Macho-1. Expression signals of *Tbx6-r.b* are shown in (A) embryos injected with the control MO, and (B) the *Macho-1* MO at the 64-cell stage. The number of embryos examined and the percentage of embryos each panel represents are shown. Expression in B7.5 is indicated by cyan arrowheads. Loss of expression in B7.5 is indicated by white arrowheads. The scale bar represents 50 μm.

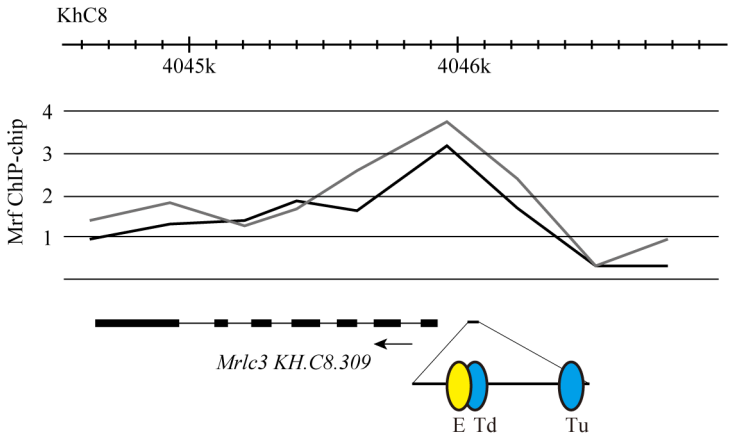

Figure S3. Mrf binds to the upstream region of *Mrlc3*. Gfp-tagged Mrf ChIP-Chip data from a previous study (Kubo et al., 2010) were mapped onto the genomic region containing *Mrlc3* and its upstream region. The graph shows fold enrichment (y-axis) for the chromosomal region (x-axis). The results of two independent experiments are shown by black and gray lines.

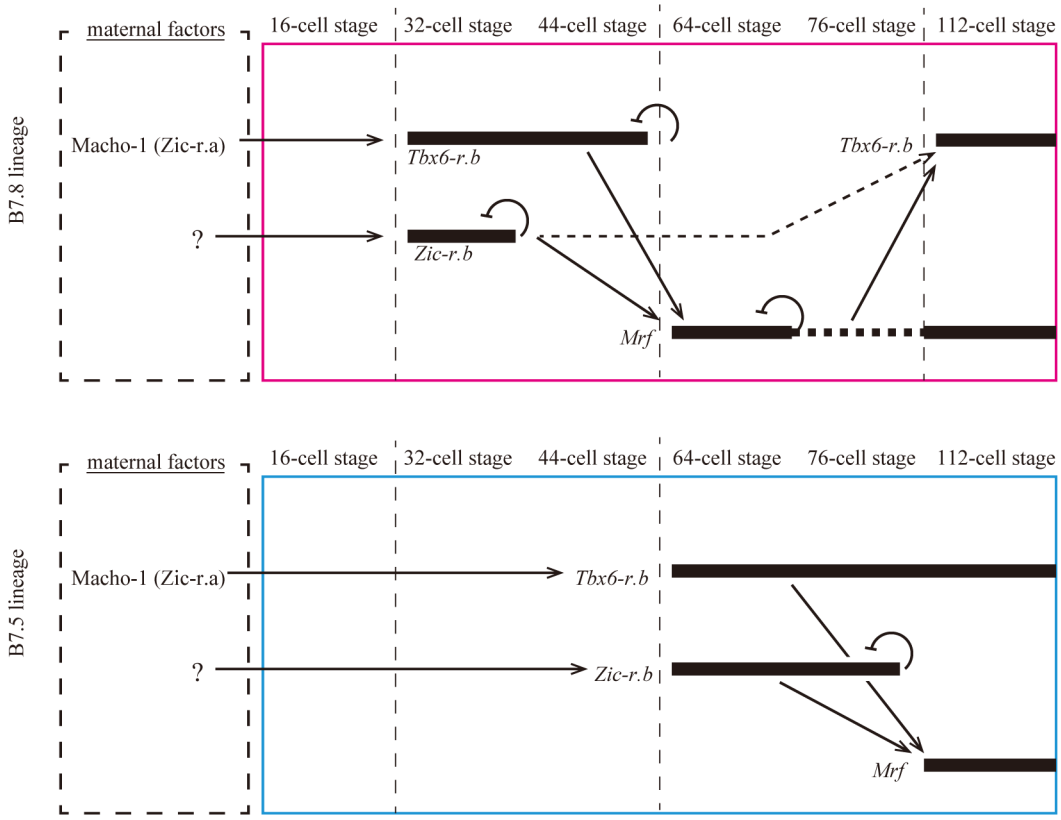

Figure S4. The genetic pathway from maternal factors to expression of muscle structural genes in the B7.8 and B7.5 lineages. Diagrams showing regulatory relationships among the key transcription factors in the (A) B7.8 and (B) B7.5 lineages of early stage embryos. Thick lines indicate expression and a thick dotted line indicates weak expression of *Mrf*. Although *Zic-r.b* regulates *Tbx6-r.b* expression at the 112-cell stage, this regulation might be indirect (dotted line). Negative auto-regulatory loops are not shown for expression of *Tbx6-r.b* and *Mrf* at later stages, because these genes are not inactivated before the 112-cell stage. These diagrams were constructed as an assembly of the present and previous studies. See the text for details.