

RESEARCH ARTICLE

A master regulatory loop that activates genes in a temporally coordinated manner in muscle cells of ascidian embryos

Izumi Oda and Yutaka Satou*

ABSTRACT

Ascidian larval muscle cells present a classic example of autonomous development. A regulatory mechanism for these cells has been extensively investigated, and the regulatory gene circuit has been documented from maternal factors to a muscle-specific gene. In the present study, we comprehensively identified genes expressed specifically in ascidian muscle cells, and found that all of them are under control of a positive regulatory loop of *Tbx6-r.b* and *Mrf*, the core circuit identified previously. We also found that several transcription factors under control of the *Tbx6-r.b/Mrf* regulatory loop exhibited various temporal expression profiles, which are probably important for creating functional muscle cells. These results, together with results of previous studies, provide an exhaustive view of the regulatory system enabling autonomous development of ascidian larval muscle cells. It shows that the *Tbx6-r.b/Mrf* regulatory loop, but not a single gene, serves a 'master' regulatory function. This master regulatory loop not only controls spatial gene expression patterns, but also governs temporal expression patterns in ascidian muscle cells.

KEY WORDS: Muscle cells, *Tbx6*, Myogenic factor, Ascidian, Regulatory loop

INTRODUCTION

Tadpole larvae of ascidians, which belong to the sister group of vertebrates, differentiate 36 muscle cells in their tails. The anterior 28 muscle cells (B-line muscle cells) have been investigated as a classical example of autonomous development. The core circuit of the genetic pathway behind this autonomous differentiation is a feedback loop consisting of two transcription factor genes, *Tbx6-r.b* and *Mrf* (Brown et al., 2007; Imai et al., 2006; Meedel et al., 2007; Yagi et al., 2005; Yu et al., 2019) (Fig. 1A).

The muscle determinant, *Zic-r.a* or *Macho-1*, is primarily responsible for initiation of the program that specifies muscle cells (Nishida and Sawada, 2001; Satou et al., 2002), while another *Zic*-related transcription factor gene, *Zic-r.b*, which positively regulates *Tbx6-r.b* and *Mrf*, is activated by a pathway under control of β -catenin and *Gata.a*, but independently of *Zic-r.a* (Imai et al., 2016, 2002; Tokuoka et al., 2021). In addition, *Tbx6-r.b* and *Mrf* regulate each other (Yu et al., 2019) (Fig. 1A).

Mrf is the sole ortholog for the vertebrate myogenic factor genes *MyoD*, *Myf5*, myogenin, and *Mrf4* (Araki et al., 1994;

Meedel et al., 1997). Among these myogenic transcription factors, *MyoD* was first identified as a factor that can convert mouse fibroblast cells to myoblast cells (Lassar et al., 1986). This factor has been regarded as a 'master control gene', positioned at the top of a genetic pathway specifying muscle fate. In fact, knockout of myogenic factor genes greatly impairs myogenesis in vertebrate embryos (Hasty et al., 1993; Kassir-Duchossoy et al., 2004; Nabeshima et al., 1993; Rudnicki et al., 1993). However, the myogenetic factor gene plays a relatively minor role in muscle differentiation of flies, nematodes, and sea urchin (Balagopalan et al., 2001; Chen et al., 1994; Venuti et al., 1993). Similarly, knockdown of *Mrf* in ascidian embryos using specific morpholino oligonucleotides does not completely eliminate muscle cells (Imai et al., 2006; Meedel et al., 2007). Indeed, *Myl.c*, which encodes myosin light chain, is expressed in early embryos independently of *Mrf*, but under control of *Tbx6-r.b* (Fig. 1A) (Yu et al., 2019). As mentioned above, *Mrf* constitutes a positive feedback loop with *Tbx6-r.b*, and many muscle structural genes may be under control of this feedback loop. The first question in the present study is whether this regulatory loop has a 'master regulatory' role to control all muscle structural genes.

The activity of this feedback loop subsides gradually, and *Tbx6-r.b* expression largely disappears before the tailbud stage (Takatori et al., 2004). Instead of *Tbx6-r.b*, *Tbx15/18/22* begins to be expressed under control of the regulatory loop and maintains expression of *Myl.c* in late embryos and larvae through a shared binding site for *Tbx6-r.b* and *Tbx15/18/22* in the late enhancer of the upstream regulatory region of *Myl.c* (Fig. 1A). In addition to this shared binding site, *Myl.c* has a site that binds *Tbx6-r.b*, but not *Tbx15/18/22*, and *Myl.c* is activated through this site in the early enhancer before gastrulation. In this way, *Myl.c* is regulated by *Tbx6-r.b* in early embryos, then by *Tbx6-r.b* and *Mrf*, and finally by *Tbx15/18/22* (Yu et al., 2019). This observation indicates that expression patterns of genes expressed in muscle cells can be regulated temporally by a combination of transcription factors. A previous comprehensive gene expression assay showed various temporal expression patterns of muscle-specific genes (Satou et al., 2001). On the basis of these observations, the next question is how expression of muscle-specific genes is temporally regulated. *Snai* (also known as *Snail*), *Meox* (*Mox*), and *Tcf15-r* (*Paraxis-like*), which encode transcription factors, are expressed in the muscle lineage (Cao et al., 2019; Fujiwara et al., 1998; Imai et al., 2004; Satou et al., 2001); therefore, they may contribute to temporal regulation of muscle-specific genes as *Tbx15/18/22* does.

Together with results of previous studies, the present study answers these questions by uncovering the entire genetic pathway from maternal factors to expression of muscle-specific genes through the lifetime of muscle cells of ascidian larvae, and illuminates how the gene regulatory system coordinates gene expression to differentiate functional muscle cells.

Department of Zoology, Graduate School of Science, Kyoto University, Sakyo, Kyoto 606-8502, Japan.

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

DOI: 10.1093/dev/204382

Handling Editor: Cassandra Extavour
Received 9 September 2024; Accepted 13 December 2024

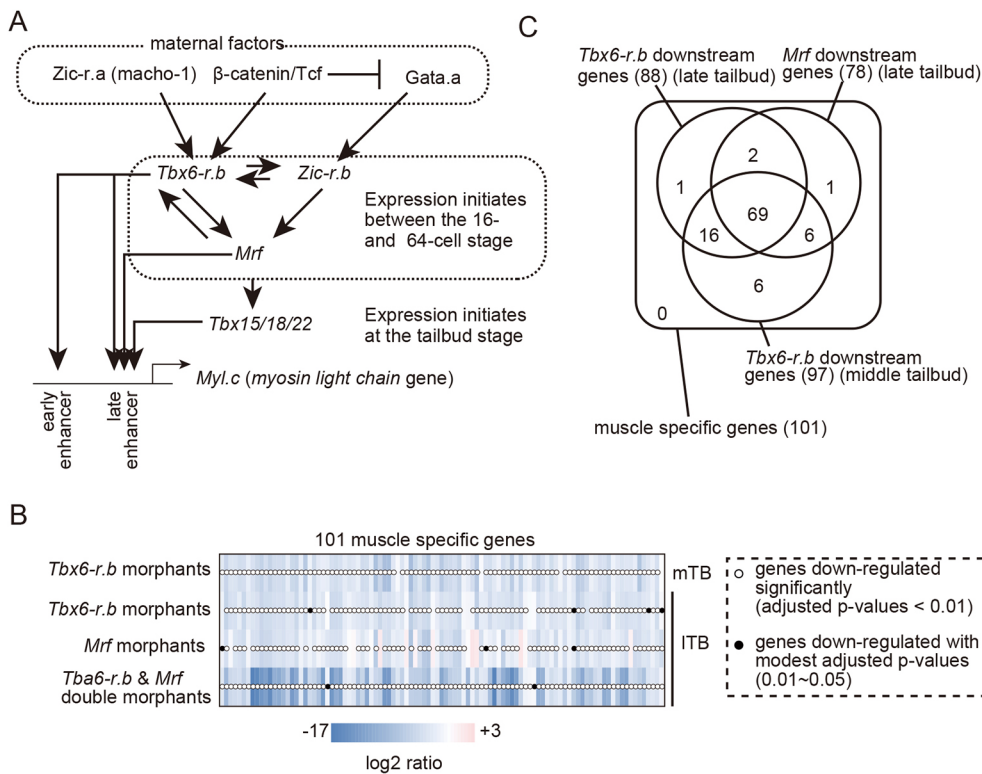


Fig. 1. RNA-sequencing experiments revealed that all 101 muscle-specific genes are under control of the regulatory loop of *Tbx6-r.b* and *Mrf*. (A) The gene regulatory circuit for differentiation of muscle cells. The core regulatory loop is enclosed by a dotted line. (B) A heatmap of changes in expression of 101 muscle-specific genes between control and morphant embryos as designated on the left. RNA-sequencing results were analyzed with DESeq2 (Love et al., 2014). Genes with significantly downregulated expression are shown by circles. ITB, late tailbud embryos; mTB, middle tailbud embryos. (C) A Venn diagram that shows overlaps among three RNA-sequencing experiments (threshold *P*-value: 0.01).

RESULTS

All muscle-specific genes are regulated under control of *Tbx6-r.b* and *Mrf*

First, we used publicly available single-cell transcriptome data for late tailbud embryos to identify muscle-specific genes (Cao et al., 2019). We mapped the raw data to a gene model set (KY21 set) that was recently published (Satou et al., 2022). We found that 275 genes were significantly enriched in a cluster representing muscle cells ($P < 0.01$; mean UMI count > 1), but this set included genes that are expressed not only in muscle, but also in other tissues. To exclude such genes, we calculated mean expression counts per gene per cluster, and found that expression counts of 102 genes in the muscle cluster were at least ten times as abundant as those in all other clusters. This set contained nine genes encoding muscle actin, six genes encoding myosin heavy chain, six genes encoding myosin light chain, one gene encoding troponin I, one gene encoding troponin T, two genes encoding troponin C, two genes encoding tropomyosin, and one gene encoding muscle-type creatine kinase; thus, muscle-specific genes were successfully identified using this method. In addition, we randomly chose six genes from this list, and confirmed that these genes are expressed specifically in muscle cells (Fig. S1). We excluded one gene (KY21.Chr1.1843) from the list, because it was a single-exon gene encoded in an intron of another gene. Specifically, it encoded a short protein (30 amino acids) with no clear similarity to proteins of other animals. Moreover, its GC content (25%) was much lower than the genome average (35%) (Dehal et al., 2002), and this low GC content potentially causes mis-mapping of sequencing tags and makes downstream assays difficult. Thus, in subsequent analyses, we used the remaining 101 genes (Table S1).

We then performed RNA sequencing of tailbud embryos injected with a control morpholino oligonucleotide (MO), a MO against *Tbx6-r.b* or a MO against *Mrf*. We used late tailbud embryos for both morphants, and we also used middle tailbud embryos for

Tbx6-r.b morphants, because *Tbx6-r.b* expression ceases earlier than *Mrf* expression (Yu et al., 2019). While the *Mrf* MO has been used in previous studies (Imai et al., 2006; Yu et al., 2019), the *Tbx6-r.b* MO employed in the present study was different from the MO we used previously. For this reason, we first examined expression of two known targets (muscle actin genes and *Myl.c*; Yagi et al., 2005) of *Tbx6-r.b* in embryos injected with the new MO, and found that this new MO downregulated expression of both genes in the same way as the old one, while expression levels of control genes (*Cdk5* and *Slk/Stk10*) hardly changed (Fig. S2).

The above 101 genes were all significantly downregulated in *Tbx6-r.b* morphant embryos and/or *Mrf* morphant embryos (adjusted $P < 0.01$; Fig. 1B; Table S2). Most of them (69 genes) were downregulated in all three experiments (Fig. 1C). This is not surprising because *Tbx6-r.b* and *Mrf* constitute a positive feedback loop (Yu et al., 2019). As further confirmation, we performed an RNA-sequencing analysis using late tailbud embryos injected simultaneously with MOs of *Tbx6-r.b* and *Mrf*, and found that 99 genes were downregulated significantly (adjusted $P < 0.01$; Fig. 1B; Table S2). Expression of the remaining two genes (*Slc38a3* and *Obscn*) was also downregulated with modest adjusted *P*-values (0.017 and 0.024). We quantified their mRNA levels in double morphants of *Tbx6-r.b* and *Mrf* with reverse transcription and quantitative PCR (RT-qPCR), and found that these genes were indeed downregulated in double morphants of *Tbx6-r.b* and *Mrf* (Fig. S3). These data indicate that the 101 muscle-specific genes are under control of the regulatory loop of *Tbx6-r.b* and *Mrf*.

Temporal expression profiles of genes specifically expressed in the muscle lineage

To examine temporal expression profiles of the above 101 muscle-specific genes, we used publicly available RNA-sequencing data that quantified RNA levels at various stages (Brozovic et al., 2018). We mapped data of early gastrula, middle gastrula, middle neurula,

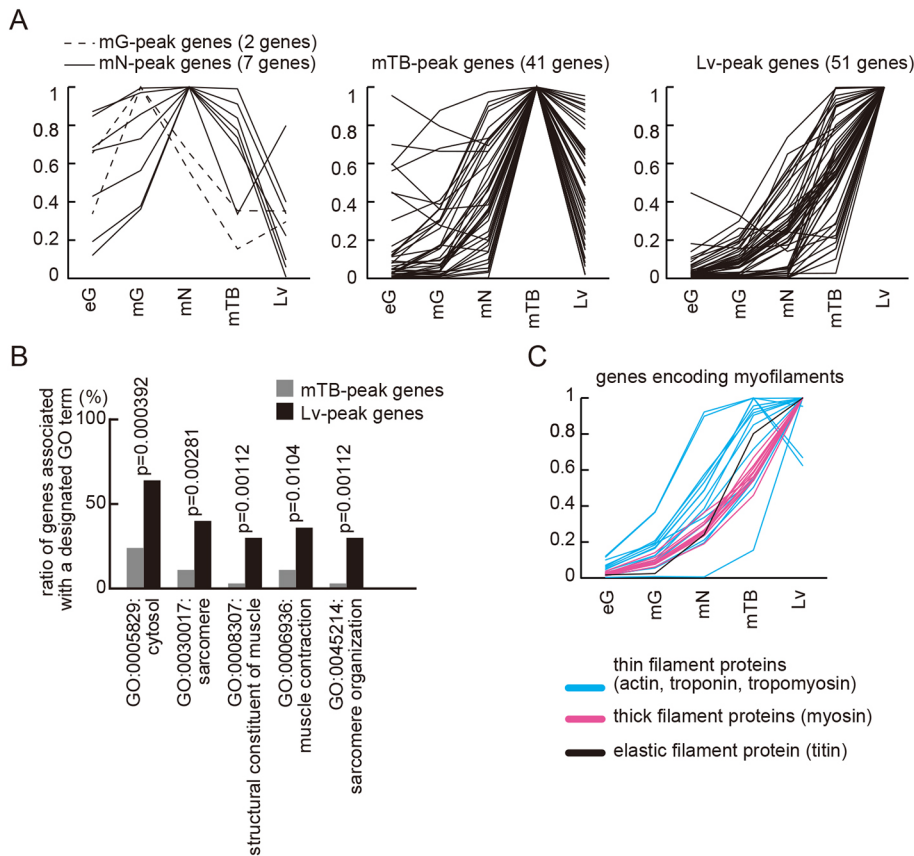


Fig. 2. Genes encoding myofiber proteins tend to be expressed abundantly in larvae.

(A) Temporal expression patterns of the 101 muscle-specific genes calculated using publicly available RNA-sequencing data (Brozovic et al., 2018). TPM values were converted to relative values against the highest value for each gene. Genes with peak expression at the middle gastrula (mG; dashed lines) or middle neurula (mN; solid lines) stages are shown in the left panel (mG-peak genes and mN-peak genes). Genes with peak expression at the middle tailbud (mTB) and larval (Lv) stages are shown in the middle and right panels (mTB-peak genes and Lv-peak genes). No genes showed peak expression at the early gastrula (eG) stage. (B) mTB-peak genes and Lv-peak genes were analyzed using GO terms (Ashburner et al., 2000). Differences between the two populations were examined using Fisher's exact tests. GO terms that passed the multiple testing correction using the Benjamini–Hochberg method (threshold value=0.1) are shown with *P*-values of Fisher's exact tests. (C) Temporal expression profiles of genes encoding myofiber proteins.

middle tailbud, and larval stages to the KY21 gene model set, and found that 51 genes are most abundantly expressed in larvae (Lv-peak genes), 41 genes in middle tailbud embryos (mTB-peak genes), seven genes in middle neurula embryos (mN-peak genes), and two genes in middle gastrula embryos (mG-peak genes) (Fig. 2A; Table S3).

Among the 101 genes, 94 encoded proteins with clear homologs in the human proteome (BLASTP E-value, $\leq 1 \times 10^{-5}$). Using gene ontology (GO) terms associated with these homologous human proteins, we analyzed differences in GO terms between the mTB-peak genes ($n=38$) and Lv-peak genes ($n=47$). The remaining two gene groups were not analyzed because gene numbers were too small for statistical comparisons. Eight cellular component GO terms, four molecular function GO terms, and two biological process GO terms were associated with 15 or more proteins (Tables S4 and S5).

Among the eight cellular-component GO-terms, 'cytosol' and 'sarcomere' were significantly more frequently associated with Lv-peak genes than with mTB-peak genes ($P=3.92 \times 10^{-4}$ and $P=2.81 \times 10^{-3}$, respectively; Fisher's exact test). Accordingly, one molecular function GO term, 'structural constituent of muscle' ($P=1.12 \times 10^{-3}$), and two biological process terms, 'muscle contraction' ($P=1.04 \times 10^{-2}$) and 'sarcomere organization' ($P=1.12 \times 10^{-3}$), were significantly more frequently associated with Lv-peak genes (Fig. 2B; Table S5). Consistent with this, 25 of 28 genes encoding myofilament proteins (myosin, actin, tropomyosin, troponin, and titin) were expressed most abundantly at the larval stage (Fig. 2C). By contrast, 'plasma membrane' was more frequently associated with mTB-peak genes ($P=2.44 \times 10^{-2}$), although multiple corrections using the Benjamini–Hochberg method (threshold=0.1) did not result in statistical significance

(Table S5). Indeed, two genes encoding cholinergic receptors (KY21.Chr7.717 and KY21.Chr7.961) and three of four genes encoding solute carrier proteins (KY21.Chr1.1444, KY21.Chr3.8, and KY21.Chr6.592) were most abundant at the middle tailbud stage. These observations suggest that populations of mTB-peak genes and Lv-peak genes are qualitatively different, and that genes expressed in muscle cells are temporally regulated according to protein functions.

Temporal expression profiles of transcription factor genes in the muscle lineage

A mechanism that temporally regulates muscle-specific genes has partly been revealed. *Myf.c*, which encodes a myosin light chain, has two cis-regulatory elements. The first of these ('early enhancer' in Fig. 1A) contains a Tbx6-r.b-binding site and regulates expression in early embryos; *Tbx6-r.b* initiates expression earlier than *Mrf*. The second ('late enhancer' in Fig. 1A) contains binding sites for Tbx6-r.b and *Mrf* and regulates expression in late embryos (Yu et al., 2019). The second also binds Tbx15/18/22, which is expressed in late embryos, and maintains expression in late embryos and larvae (Yu et al., 2021). In addition, *Meox*, *Tcf15-r*, and *Snai* are expressed in the muscle lineage (Cao et al., 2019; Fujiwara et al., 1998; Imai et al., 2004; Satou et al., 2001); therefore, they likely also contribute to regulation of the temporal expression profiles of muscle-specific genes. Notably, *Tbx15/18/22*, *Meox* and *Tcf15-r* are the only transcription factor genes among the above 101 muscle-specific genes (*Snai* is not included in the 101 muscle-specific genes because it is rarely expressed at the late tailbud stage).

Meox expression becomes detectable specifically in the muscle lineage by *in situ* hybridization at the late gastrula stage, and continues to be expressed at the middle tailbud stage

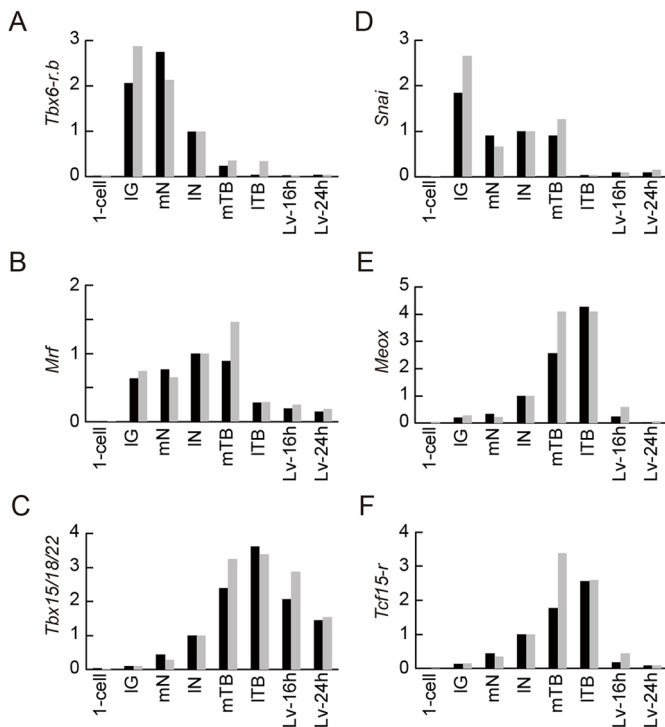


Fig. 3. Temporal gene expression profiles of six transcription factor genes. (A–F) Expression levels per embryo were quantified by RT-qPCR assays at 1-cell, late gastrula (IG), middle neurula (mN), late neurula (IN), middle tailbud (mTB), late tailbud II (ITB), 16h-larva (Lv-16 h), and 24 h-larva stages (Lv-24 h), and calculated as relative values against expression levels at the late neurula stage. Results from two batches of embryos are represented with different colors.

(Imai et al., 2004) (Fig. S4). *Tcf15-r* is also expressed specifically in the muscle lineage from late gastrula to tailbud embryos (Satou et al., 2001) (Fig. S4). *Snai* expression in the muscle lineage begins at the 32-cell stage, and is hardly detectable before the middle tailbud stage (Fujiwara et al., 1998; Imai et al., 2004).

To understand more precisely the temporal expression profiles of these transcription factor genes, which are expressed in the muscle lineage, we examined their expression profiles by RT-qPCR. As we showed previously (Yu et al., 2021), *Tbx6-r.b* expression was decreased and hardly undetectable before the late tailbud stage, whereas expression of *Mrf* and *Tbx15/18/22* was readily detected in larvae (Fig. 3A–C). *Snai* expression was detected at the middle tailbud stage, but rapidly decreased before the late tailbud stage (Fig. 3D). Expression profiles of *Meox* and *Tcf15-r* were similar; these genes were expressed most abundantly at the tailbud stage, and were barely detectable in larvae (Fig. 3E,F). *Tbx15/18/22*, *Meox*, and *Tcf15-r* belonged to the 101 muscle-specific genes, and their temporal expression profiles were also included in Table S3 and Fig. 2A as mTB-peak genes. Their overall expression profiles agreed with those obtained from RT-qPCR. These observations indicate that *Meox*, *Tcf15-r*, and *Snai* may be involved in temporal regulation of muscle structural genes, as in the case of *Tbx15/18/22*.

Meox and Tcf15-r downstream genes are highly expressed in larvae

Because the above RNA-sequencing experiments showed that *Meox* and *Tcf15-r* were downregulated in double morphants of *Tbx6-r.b* and *Mrf* (Fig. 4A; Table S2), *Meox* and *Tcf15-r* may regulate gene expression in late embryos under control of the regulatory loop

of *Tbx6-r.b* and *Mrf*. To test this hypothesis, we first performed RNA-sequencing experiments using late tailbud embryos developed from eggs injected with an MO against *Meox* or *Tcf15-r*. We took biological duplicates for each experiment and found that six and nine genes were downregulated significantly in *Meox* and *Tcf15-r* morphants, respectively (Fig. 4B; Table S2). To compare these results with our previous result for *Tbx15/18/22*, we re-analyzed *Tbx15/18/22* data using the latest gene model set (Satou et al., 2022), which we used in the present study. There were overlaps among downstream genes of these three factors (Fig. 4C).

The *Meox* MO used in the above experiment was designed to block translation. To confirm its specificity, we designed another MO, which was expected to block splicing. This splicing-block MO did block splicing, although its efficiency was not very high (Fig. S5A). Next, we examined the expression level of *Ckmt*, because it was the most prominently downregulated gene in the RNA-sequencing experiment using embryos injected with the *Meox* translation-block MO (Fig. 4B). RT-qPCR experiments showed that this gene was downregulated in embryos injected with the splicing-block MO and in embryos injected with the translation-block MO. Likewise, a splicing-block MO against *Tcf15-r* did indeed block splicing (Fig. S5C), and *Acta.h*, which encodes muscle actin, was downregulated in embryos injected with the translation-block MO and in embryos injected with the splicing MO against *Tcf15-r* (Fig. S5D). For this experiment, we used primers designed to detect an intron of *Acta.h* and measured expression levels of its nascent transcript, because the ascidian genome contains several muscle actin genes and their exons are highly similar. These observations indicated that all these MOs specifically knocked down *Meox* and *Tcf15-r*.

According to temporal expression profiles based on RNA-sequencing data (Brozovic et al., 2018), five of six *Meox*-downstream genes were Lv-peak genes and the remaining one was an mTB-peak gene (Fig. S6). Similarly, all nine *Tcf15-r*-downstream genes were Lv-peak genes (Fig. S6). Among the 42 genes regulated by *Tbx15/18/22*, 33 genes were Lv-peak genes, and nine genes were mTB-peak genes (Fig. S6). By contrast, the remaining 54 genes, which were not under control of *Meox*, *Tcf15-r*, or *Tbx15/18/22*, showed a different pattern from those under control of *Meox*, *Tcf15-r*, and/or *Tbx15/18/22* (Fig. 4D; $P=5.5 \times 10^{-7}$, Fisher's exact test). This observation indicates that *Meox*, *Tcf15-r*, and *Tbx15/18/22* individually or together maintain expression levels in late embryos and larvae.

Snai regulates muscle-specific genes in earlier stages

Next, we examined *Snai* morphants at the middle tailbud stage by RNA sequencing. Because the *Snai* MO was used in previous studies (Imai et al., 2006, 2009; Tokuoka et al., 2018), we did not examine specificity of this MO in the present study. We found that 43 genes were significantly downregulated (adjusted $P < 0.01$) (Fig. 5A; Table S2). The temporal expression profile of *Snai* shown in Fig. 3 suggests that *Snai* may be involved in regulation of genes at earlier stages than *Meox*, *Tcf15-r*, and *Tbx15/18/22*. Indeed, *Tcf15-r* and *Meox*, which are expressed most abundantly at the middle tailbud stage (see Fig. 3), were downregulated in *Snai* morphants (Fig. 5A). To test this hypothesis further, we compared expression profiles of genes under control of *Snai* with those of genes that are not. Temporal expression profiles of genes regulated by *Snai*, but not by *Meox*, *Tcf15-r*, or *Tbx15/18/22* were markedly different from those of genes regulated by *Meox*, *Tcf15-r*, or *Tbx15/18/22* but not by *Snai* (Fig. 5B,C). Their profiles also differed from those of genes not under control of *Snai*, *Meox*, *Tcf15-r*, or *Tbx15/18/22* (Fig. 5C; Fig. S7).

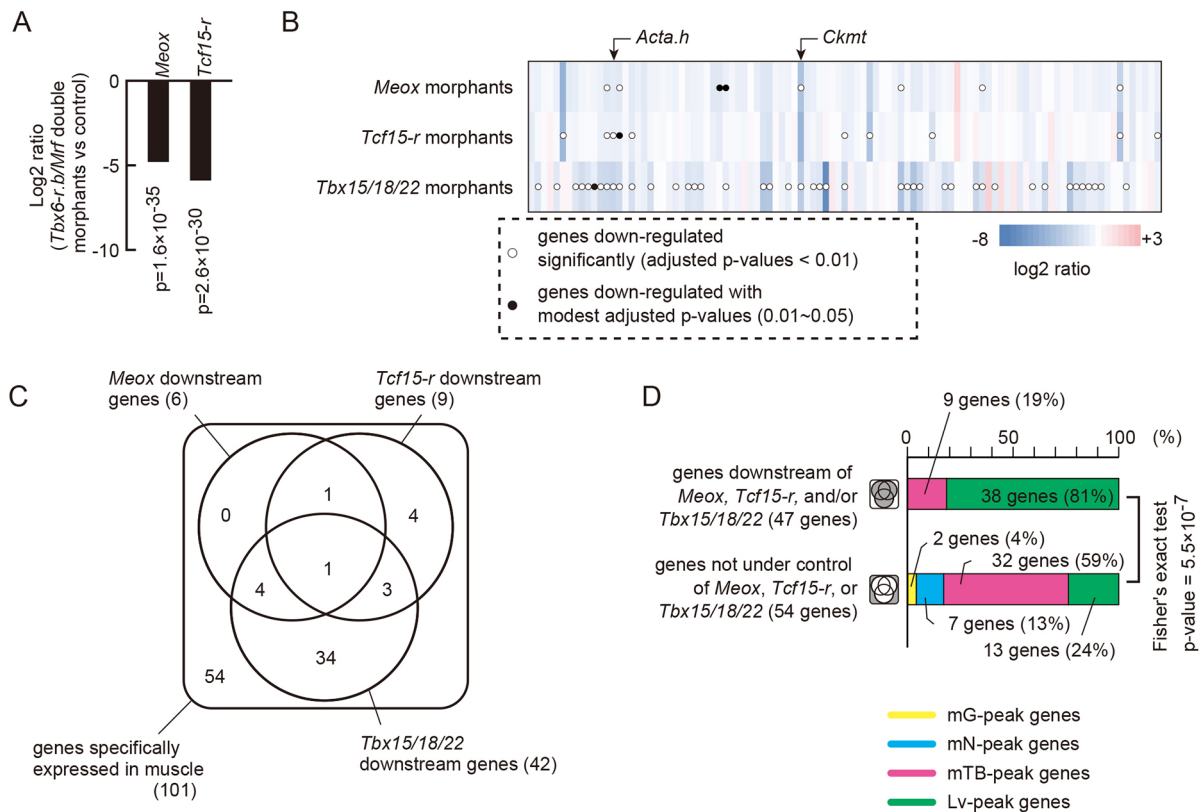


Fig. 4. *Meox* and *Tcf15-r* regulate subsets of muscle-specific genes, most of which are expressed most abundantly in larvae. (A) RNA-sequencing experiments using double morphants of *Tbx6-r.b* and *Mrf* (see Fig. 1) showed that *Meox* and *Tcf15-r* are under control of *Tbx6-r.b* and *Mrf*. Adjusted *P*-values, which are shown in the panel, were calculated using DESeq2 (Love et al., 2014). (B) A heatmap of changes in expression of the 101 muscle-specific genes between control and morphant embryos of *Meox*, *Tcf15-r*, or *Tbx15/18/22*. RNA-sequencing results were analyzed with DESeq2 (Love et al., 2014). Genes with significantly downregulated expression are shown with circles. We used RNA-sequencing data for *Tbx15/18/22* morphants that were published previously (Yu et al., 2021) and re-mapped them to the latest gene model set (Satou et al., 2022). (C) A Venn diagram that shows overlaps among genes under control of *Meox*, *Tcf15-r*, and/or *Tbx15/18/22*. (D) Percentages of mG-peak, mN-peak, mTB-peak, and Lv-peak genes in two gene populations; genes under control of *Meox*, *Tcf15-r*, and/or *Tbx15/18/22*; and genes under control of none of them.

Muscle-specific transcription factors combinatorically regulate temporal expression patterns of downstream genes

The above analysis indicated that temporal expression profiles of muscle-specific genes are determined by combinations of upstream regulatory factors and encoded in cis-regulatory regions of downstream genes. Because *Snai* promoted expression predominantly around the tailbud stage, we next set out to confirm that *Snai*-binding sites can promote expression around the tailbud stage using reporter constructs. For this purpose, we first examined the upstream regulatory region of *Meox* and *Tcf15-r*. The aforementioned RNA-sequencing analyses showed that *Meox* and *Tcf15-r* are under control of *Snai* (Fig. 5A; Table S2), and we identified two putative *Snai*-binding sites for each of these genes using JASPAR scan (Sandelin et al., 2004) and the *Ciona* *Snai* position weight matrix (Nitta et al., 2019) (Fig. S8). Upstream regions of *Meox* and *Tcf15-r* containing these putative *Snai*-binding sites promoted reporter expression in muscle cells of middle tailbud embryos (Fig. 6A,B). Introduction of mutations in these *Snai*-binding sites impaired expression of the reporters in muscle cells (Fig. 6C,D), and these putative *Snai*-binding sites bound *Snai* *in vitro* (Fig. 6E).

A previous study demonstrated that *Myl.c*, the expression peak of which was at the larval stage (Table S3), has two cis-regulatory modules in its 146-bp upstream region that can promote reporter expression at the tailbud stage (Fig. 6F; Fig. S8) (Yu et al., 2021, 2019). Deletion of the shared binding site for *Tbx6-r.b* and *Tbx15/*

18/22 greatly reduced the activity of this regulatory region (Fig. 6G). We next inserted 51-bp and 45-bp upstream regions of *Meox* and *Tcf15-r* that contained *Snai*-binding sites instead of the shared binding site for *Tbx6-r.b* and *Tbx15/18/22*. These reporter constructs were expressed in muscle cells of middle tailbud embryos (Fig. 6H,I), while introduction of mutations into *Snai*-binding sites greatly reduced expression (Fig. 6J,K). Therefore, the inserted *Snai* sites promoted expression in middle tailbud stage.

Finally, we compared expression levels of the reporter genes in middle tailbud embryos with those in larvae using RT-qPCR (Fig. 6L). The expression level of the reporter with the wild-type upstream region of *Myl.c* increased greatly from the middle tailbud stage to the larval stage, whereas those of the reporters with wild-type upstream regions of *Meox* and *Tcf15-r* rarely changed. Expression levels of the chimera construct, which were tested in the above experiments (Fig. 6H,I), rarely changed between the middle tailbud and larval stages. Therefore, replacements of shared binding sites for *Tbx6-r.b* and *Tbx15/18/22* with *Snai*-binding sites of the upstream regions of *Meox* and *Tcf15-r* shifted the expression peak earlier, as expected.

DISCUSSION

The regulatory circuit for differentiation of muscle cells of ascidian larvae

Fig. 7 summarizes the genetic pathway for differentiation of B-line muscle cells of ascidian larvae. All 101 muscle-specific genes

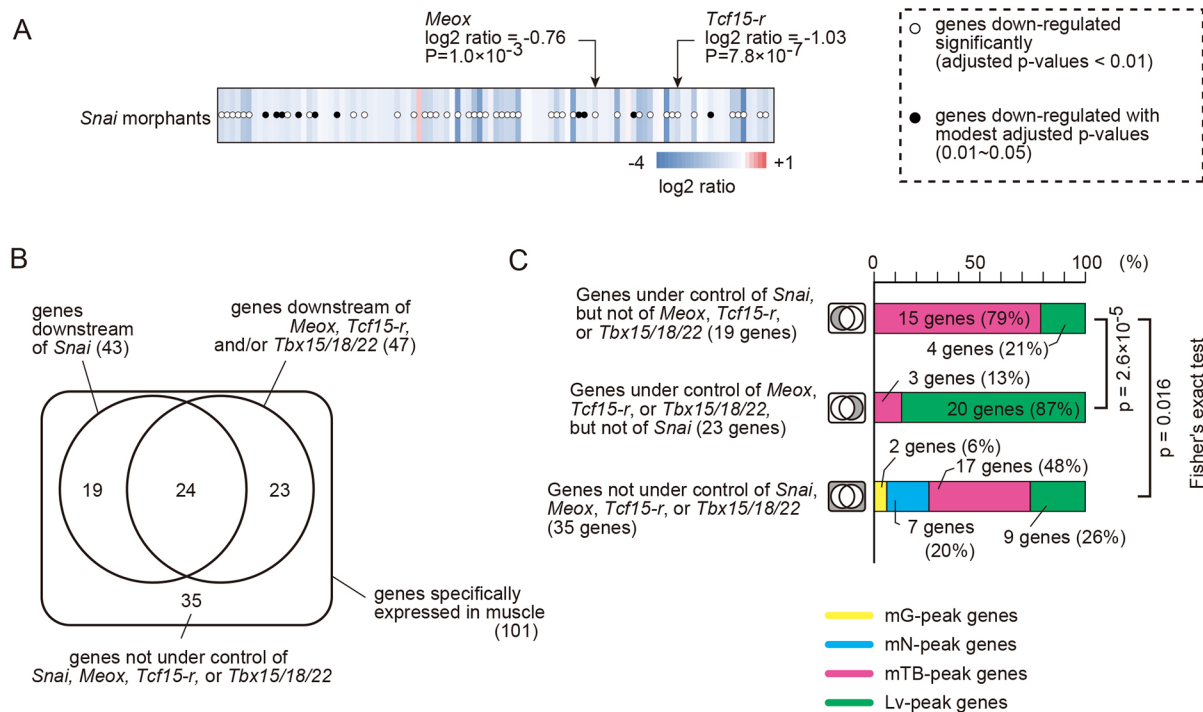


Fig. 5. *Snai* regulates a subset of muscle-specific genes, most of which are expressed most abundantly in middle tailbud embryos. (A) A heatmap of changes in expression of the 101 muscle-specific genes between control and *Snai* morphant embryos. RNA-sequencing results were analyzed with DESeq2 (Love et al., 2014). Genes with significantly downregulated expression are shown with circles. (B) A Venn diagram that shows overlaps between genes under control of *Snai* and genes under control of *Meox*, *Tcf15-r*, and/or *Tbx15/18/22*. (C) Percentages of genes with peak expression at four developmental stages in three gene populations: genes under control of *Snai* but not of *Meox*, *Tcf15-r*, or *Tbx15/18/22*; genes under control of *Meox*, *Tcf15-r*, and/or *Tbx15/18/22* but not of *Snai*; and genes under control of none of them.

examined in the present study were under control of this genetic pathway.

In this pathway, *Tbx6-r.b* and *Mrf* constitute a central regulatory loop, and *Tbx15/18/22* is activated under control of this regulatory loop. *Tbx15/18/22* maintains expression of *Myl.c* by the larval stage (Yu et al., 2021). Here, we provide evidence that *Meox* and *Tcf15-r* also maintain expression of a subset of muscle-specific genes, which partly overlaps with target genes of *Tbx15/18/22*.

In addition, many muscle-specific genes are abundantly expressed in embryos and rarely expressed in larvae (see Fig. 2). *Meox* and *Tcf15-r* are examples of such genes, and are controlled by *Snai*, which is responsible for expression of muscle-specific genes at the tailbud stage. Thus, in the muscle lineage of ascidian embryos, a combination of *Tbx6-r.b*, *Mrf*, *Snai*, *Meox*, *Tcf15-r*, and *Tbx15/18/22* coordinates temporal expression patterns of muscle-specific genes.

Although *Snai* encodes a well-known transcriptional repressor (Chopra and Levine, 2009), which was originally identified in *Drosophila* (Simpson, 1983), this protein also activates targets in a context-dependent manner (Rembold et al., 2014). We showed that *Snai* participates in activation of its targets in muscle cells of ascidian embryos, whereas a previous study suggests that *Snai* represses *Brachyury*, an essential transcription factor for specification of the notochord (Fujiwara et al., 1998).

In addition, *Mrf* protein levels are controlled by cereblon, a substrate receptor of an E3 ubiquitin ligase, and this helps to control temporal expression profiles of muscle-specific genes (Long et al., 2023). Importantly, *cereblon* is also under control of *Mrf* (Long et al., 2023).

Previous studies have also identified a cis-regulatory motif that resembles the cyclic AMP response element (CRE) motif

(Brown et al., 2007; Johnson et al., 2005; Kusakabe et al., 2004). According to a comprehensive *in situ* hybridization assay for transcription factor genes (Imai et al., 2004), no homologs for genes encoding cAMP responsive element binding proteins (Yamada et al., 2003) are expressed specifically in muscle cells. Although the same study also revealed that two additional transcription factor genes, *Mnx* and *Otp*, are expressed specifically in muscle cells transiently in early embryos, binding motifs of these homeodomain proteins differ from the CRE motif. In other words, there are no candidate transcription factors to bind to the CRE motif in muscle cells. These observations suggest that a transcription factor that binds to the CRE motif is expressed non-specifically in muscle cells and does not confer spatial specificity to target genes. Instead, it may have a permissive role to help other transcription factors to activate targets sufficiently. It is also possible that these sites bind *Tbx6-r.b*, *Mrf*, *Snai*, *Tbx15/18/22*, *Meox*, or *Tcf15-r*, because the M2d site of *Myl.h* (*Ci-MLC2*), which is important for expression in muscle (Kusakabe et al., 2004), resembles not only the CRE motif but also the binding motif for *Tbx6-r.b* and *Tbx15/18/22*.

Temporal regulation of muscle-specific genes

We have shown that various temporal expression patterns are reproducible and can be designed. Reporters with only *Tbx6-r.b*-binding sites are expressed in early embryos, and initiation timing is controllable via the number of *Tbx6-r.b*-binding sites (Yu et al., 2019). Reporters with a *Tbx6-r.b*-binding site and a *Snai*-binding site are expressed in early embryos and tailbud embryos. Reporters with a *Tbx15/18/22* site are expressed even in larvae (Yu et al., 2021). In other words, after being activated by the core regulatory

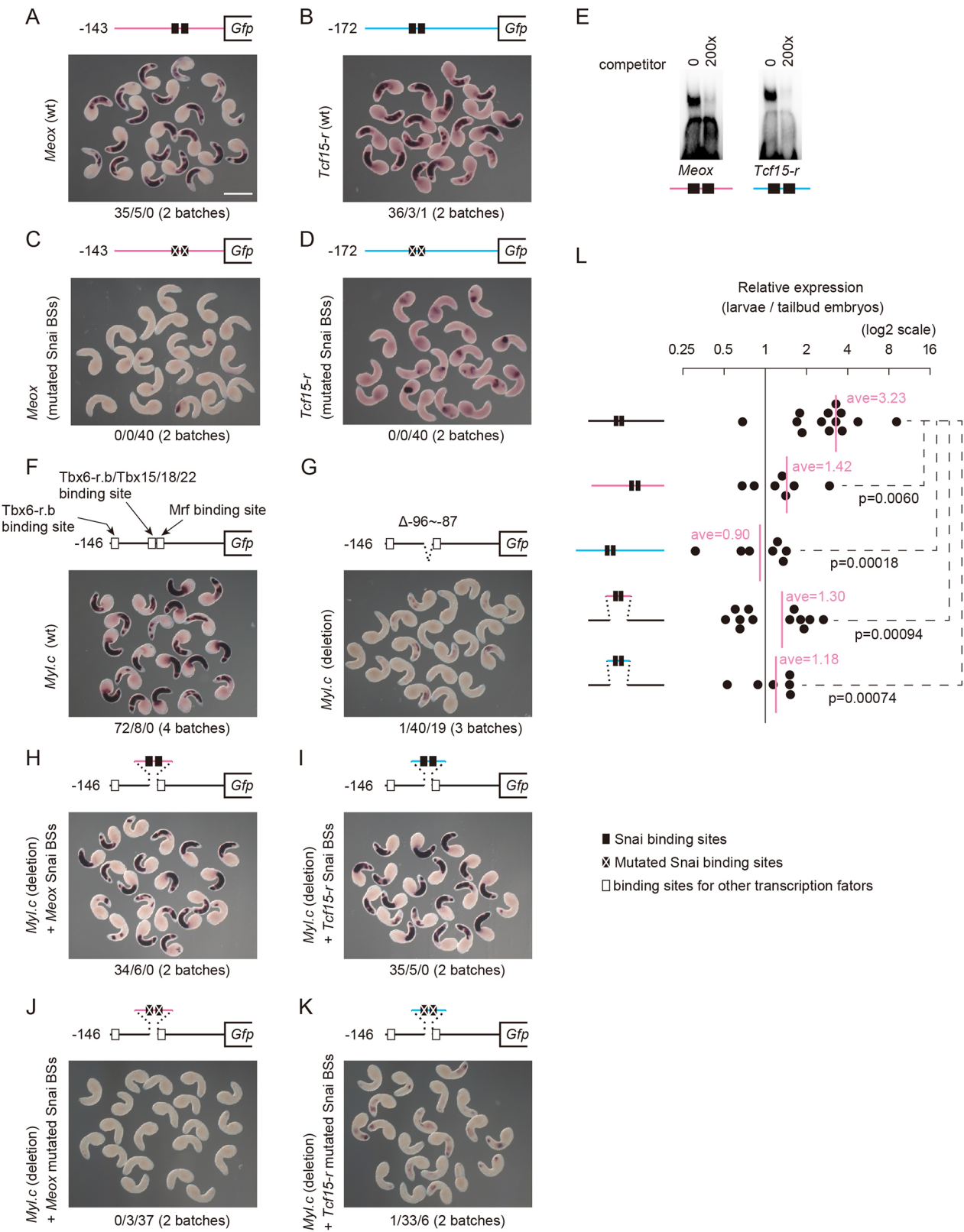


Fig. 6. See next page for legend.

circuit of *Tbx6-r.b* and *Mrf*, muscle-specific genes under control of *Snai* are expressed in tailbud embryos, and muscle-specific genes under control of *Tbx15/18/22*, *Meox*, and/or *Tcf15-r* are expressed even in larvae.

Importantly, these expression profiles reproduce those of endogenous genes. Indeed, genes encoding myofiber proteins tend to be expressed later and other genes tend to be expressed earlier, which suggests that the above temporal regulatory

Fig. 6. Snai-binding sites promote expression around the middle tailbud stage. (A-D,F-K) Expression of a reporter gene (*GFP*) was examined by *in situ* hybridization at the middle tailbud stage. Numbers indicate embryos with strong expression (expression area is over 50% of the area of muscle cells), embryos with weak expression (expression area is under 50% of the area of muscle cells), and embryos with no expression. (A-D) The regions 143-bp and 172-bp upstream of *Meox* (A) and *Tcf15-r* (B), which contain two *Snai*-binding sites, are sufficient to promote expression of the reporters in muscle cells, while mutations introduced to *Snai*-binding sites impair expression of the reporters (C,D). (E) Gel shift assays show that upstream regions of *Meox* and *Tcf15-r* can bind *Snai* *in vitro*. Shifted bands disappear by addition of specific competitors, indicating specific binding of *Snai* to these regions. (F) The 146-bp upstream region of *Myl.c*, which contains a *Tbx6-r.b*-binding site, a shared binding site for *Tbx6-r.b* and *Tbx15/18/22*, and an E-box motif (*Mrf* binding site), is sufficient to promote the reporter in muscle cells, as previously reported (Yu et al., 2019). (G) Deletion of the shared binding site for *Tbx6-r.b* and *Tbx15/18/22*, which is important for expression in tailbud embryos and larvae (Yu et al., 2021, 2019), greatly reduced expression in muscle cells at the middle tailbud stage, although some embryos expressed the reporter very weakly. (H-K) Insertions of 51-bp and 45-bp upstream regions of *Meox* and *Tcf15-r* containing two *Snai*-binding sites into the *Myl.c* deletion construct recovered expression of the reporter in muscle cells (H,I), whereas introduction of mutations into *Snai*-binding sites greatly reduced expression (J,K). (L) Expression levels of the reporters shown in F,A,B,H,I (top to bottom) were compared between the middle tailbud and larval stages, and relative expression levels are shown. Each circle represents an independent experiment, and pink bars show averages. Differences were tested using Wilcoxon's rank sum test, and *P*-values are shown. Scale bar: 200 μ m.

mechanisms are essential for differentiation of functional muscle cells.

This temporal regulatory mechanism is similar to that in the notochord lineage, in which *Brachyury* is specifically expressed (Corbo et al., 1997; Yasuo and Satoh, 1993). Genes with multiple *Brachyury*-binding sites that can act synergistically are expressed earlier. Genes with a single *Brachyury*-binding site are expressed next, and then genes that are indirectly controlled by *Brachyury* are expressed (Katikala et al., 2013). Temporal regulation of gene expression by combinations of transcription factors, and numbers and probably qualities of their binding sites is likely a common mechanism in various tissues of ascidian larvae.

Secondary muscle cell lineages

In addition to autonomously differentiating B-line muscle cells, ascidian larvae have four pairs of muscle cells that are not differentiated autonomously, and these cells are called A-line and b-line muscle cells, because they are derived from A4.1 and b4.2 blastomeres of 8-cell embryos, respectively. In A-line cells, *Tbx6-r.b* and *Mrf* are expressed under control of FGF, Delta/Notch, and Nodal signaling pathways (Hudson et al., 2007). In the b-line, *Tbx6-r.b* is also expressed soon after fate restriction at the early neurula stage (Ishida and Satou, 2024). Although it has not been determined precisely when *Mrf* begins to be expressed in this lineage, *Mrf* is clearly expressed in b-line muscle cells of tailbud embryos (Meedel et al., 1997). These observations indicate that *Tbx6-r.b* and *Mrf* are also responsible for differentiation of A-line and b-line muscle cells, and our RNA-sequencing analysis using morphants of *Tbx6-r.b* and *Mrf* supports this hypothesis.

This means that all three lineages of ascidian larval muscle cells share the gene circuit that begins with the regulatory loop of *Tbx6-r.b* and *Mrf*. Although it is currently unknown which lineage represents an ancestral pathway, it is likely that the regulatory loop of *Tbx6-r.b* and *Mrf* and its downstream circuit has been co-opted from an ancestral lineage to two other lineages, as in the case of

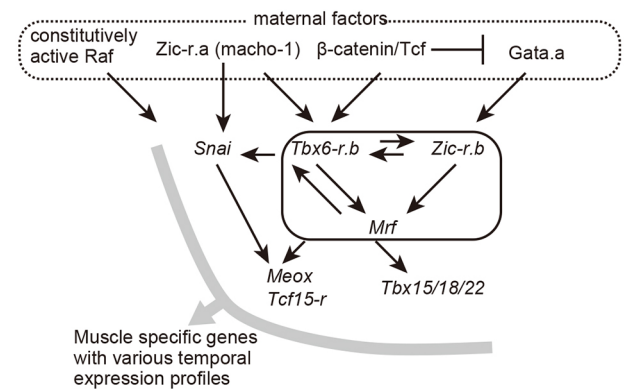


Fig. 7. The gene regulatory circuit from maternal factors to muscle structural genes in ascidian larval muscle development. The present study indicates that all muscle-specific genes are under control of this gene circuit. Muscle-specific genes have various temporal expression profiles that are likely important for muscle development. These temporal expression profiles are created by this gene regulatory circuit.

dorsal and ventral epidermal sensory neurons of ascidian larvae (Waki et al., 2015).

Tbx6-r.b and Mrf constitute the master regulatory loop for muscle differentiation

We identified 101 muscle-specific genes using single-cell transcriptome data, and almost all known genes expressed in muscle cells are included in this list. Therefore, it is likely that almost all genes specifically expressed in muscle cells were examined in the present study. Even if a small number of genes failed to be identified for technical reasons, it is likely that all muscle-specific genes, including such omitted genes, are under control of the regulatory loop of *Tbx6-r.b* and *Mrf*.

In addition to these muscle-specific genes, there are genes that are expressed in multiple cell types, including muscle cells (Satou et al., 2001). Although it is not easy to analyze these genes exhaustively, they often have multiple cis-regulatory modules that direct tissue-specific gene expression. Therefore, it is likely that cis-regulatory modules that induce muscle-specific expression are similarly under control of *Tbx6-r.b* and *Mrf*.

As we showed previously (Yu et al., 2019), *Myl.c* is directly controlled by *Tbx6-r.b* and *Mrf*. Similar cis-regulatory sequences, which are expected to bind *Tbx6-r.b* and *Mrf*, are found in many other muscle-specific genes (Satou and Satoh, 1996; Yuasa et al., 2002). Therefore, these studies altogether indicate that the regulatory loop of *Tbx6-r.b* and *Mrf*, but not *Mrf* alone, acts as the master regulator for muscle differentiation in ascidian embryos.

This mechanism contrasts with that in vertebrate embryos, in which myogenic factors serve the master regulatory function. Interestingly, four myogenic factor genes in vertebrate genomes regulate one another (Tapscott, 2005). For example, Myod and myogenin can activate each other (Brennan et al., 1990; Thayer et al., 1989). Such cross-regulatory circuits may function in a manner equivalent to that of the regulatory loop we found in ascidian embryos, which specifies the muscle fate. If so, acquisition of paralogous genes in the vertebrate lineage may have promoted a change in the structure of the gene regulatory network controlling muscle development, and these paralogs acquired a genuine master regulatory function in the vertebrate lineage.

MATERIALS AND METHODS

Animals and cDNAs

Ciona robusta (or *C. intestinalis*, type A) adults were obtained from the National Bio-Resource Project for *C. intestinalis*. cDNA clones were obtained from our EST clone collection (Satou et al., 2005). Identifiers for genes examined in the present study are listed in Table S6.

Whole-mount *in situ* hybridization and RT-qPCR to detect expression of endogenous genes

In situ hybridization was performed as described previously (Ikuta and Saiga, 2007; Satou et al., 1995). For the experiments shown in Fig. 3, we extracted RNA from 50 embryos using an RNeasy Micro kit (QIAGEN, 74004). Before extraction, we mixed synthetic *Gfp* mRNA as a spike-in control, and used it to normalize differences in extraction efficiency among samples, because no genes expressed at a constant level from eggs to larvae have been identified. Normalized expression values per embryo were compared with those at the late neurula stage. We performed the experiments shown in Figs S2, S3, and S5 in a similar manner, but we did not use spike-in controls for these experiments. Instead, we used *Raf* as an internal control for relative quantification. *Cdk5* and *Slk/Stk10* (Tokuoka et al., 2022) were used as additional controls. SYBR Green chemistry was used for qPCR.

Primers for the experiment are shown in Table S7. There are ten genes encoding muscle actin in the genome (*Acta.a* to *Acta.j*). Primers for the experiment shown in Fig. S2 were expected to detect expression of all muscle actin genes, whereas primers for the experiment shown in Fig. S5 were designed to specifically amplify an intronic region of *Acta.h* to specifically detect nascent *Acta.h* transcripts. We also used primers to amplify an intronic region of *Myl.c* to detect nascent *Myl.c* transcripts.

Gene knockdown assays and reporter assays

MOs (Gene Tools, LLC) used in the present study are listed in Table S8. The *Mrf* MO and a *Tbx6-r.b* MO (5'-TTACAATTCCTCTCTTTTCGATT-3') have been used in previous studies (Imai et al., 2006; Yagi et al., 2005; Yu et al., 2019); therefore, we did not perform experiments to confirm their specificity in the present study. Specificity of a newly designed *Tbx6-r.b* MO (5'-TCATATTCGCCATAGTCTTGCTGG-3'), and MOs against *Meox* and *Tcf15-r* were tested as described in the Results section (Figs S2 and S5). We also used an MO against *lacZ* of *Escherichia coli* as a control. MOs were introduced by microinjection under a microscope. All knockdown phenotypes were confirmed in at least two independent series of experiments.

Genomic positions of the upstream regions of *Meox*, *Tcf15-r*, and *Myl.c* are 6,019,222-6,019,487 on chromosome 8, 250,326-250,558 on chromosome 11, and 5,405,563-5,405,768 on chromosome 8 of the HT assembly (Satou et al., 2019), respectively. The reporter constructs contain a genomic region of the first and second exons and the first intron of *Foxtn2* (chromosome 14: 6,279,296-6,279,458) before GFP, to avoid amplification from introduced DNAs in qPCR (Fig. S9; see below). Reporter constructs were introduced into fertilized eggs by electroporation. Expression of the reporter was examined by *in situ* hybridization and by RT-qPCR.

For RT-qPCR to detect reporter expression, we used a Cell-to-Ct kit (Thermo Fisher Scientific, A35377) to obtain cDNA from 20-30 embryos. Because we did not have good controls expressed at the same level between the tailbud and larval stages, we did not use internal controls, but directly compared expression levels between the same number of tailbud embryos and larvae from the same batch. We used the TaqMan method for qPCR. Primers and a probe were designed to amplify cDNA derived from spliced mRNA (Forward primer: 5'-CAGTTCGTTTAAAGGG-3'; Reverse primer: 5'-GCCGGACACGCTGAACCTT-3'; Probe: 5'-FAM-CTGGTCGAGCTGGACGGCGACTAMRA-3'). To ensure preferential amplification of cDNA from spliced mRNA, we changed the reaction conditions as follows: primer concentrations, 300 nM each; probe concentration, 250 nM; extension duration, 10 s.

Gel shift assay

Recombinant Snai protein was produced as a fusion protein of the Snai DNA-binding domain (amino acids 436-581) and glutathione S-transferase

in *E. coli* (DH5 α strain). The protein was purified under native conditions using MagneGST Glutathione Particles (Promega, V861A). After annealing two complementary oligonucleotides (5'-AAAATACAGGTCAGTCGT-CACCTTTGCTTGCCCAGTTGTTTACTTCGTTTAAA-3' and 5'-AAA-TTTAAACGAAGTAAACAACCTGGGCAAGCAAAGGTGACGACTGA-CCTCGTAT-3' for the *Meox* upstream region; 5'-AAACGTGTAGTT-CAAGACAGCTGACTCACGTGATATCAGCATGAATGG-3' and 5'-AA-ACCATTATGCTGATATCACGTGAGTCAGCTGTCTTGAACCTACA-CG-3' for the *Tcf15* upstream region), both protruding ends of the double-stranded oligonucleotides were filled with biotin-11-dUTP, and these biotin-labeled oligonucleotides were used as probes. Proteins and biotin-labeled probes were mixed in 25 mM HEPES (pH 7.5), 50 mM KCl, 1 mM DTT, 25 ng/ μ l poly(dIdC), 2.5% glycerol, 0.05% NP40, and 100 ng/ μ l bovine serum albumin with or without competitor double-stranded DNAs (200-fold molar excess). Protein amounts were determined empirically. Protein-DNA complexes were detected using a Chemiluminescent Nucleic Acid Detection Module Kit (Thermo Fisher Scientific, 89880).

RNA sequencing

For RNA sequencing, we prepared embryos injected with the *lacZ* MO (0.5 mM), *Tbx6-r.b* MO (0.5 mM), *Mrf* MO (0.5 mM), *Snai* MO (0.5 mM), *Meox* translation-block MO (0.5 mM), or *Tcf15-r* translation-block MO (0.5 mM). We also prepared embryos injected with *Tbx6-r.b* and *Mrf* simultaneously (0.5 mM each), and their controls injected with a double dose of the *lacZ* MO (1 mM). Total RNA was extracted using a Dynabeads mRNA DIRECT Micro Kit (Thermo Fisher Scientific, 61011). For *Mrf* morphants, we made libraries with an Ion Total RNA-Seq Kit Version 2 (Thermo Fisher Scientific, 4475936), and used Ion PGM/S5XL instruments (DRA accession numbers: DRR585068, DRR585069). These embryos were obtained when we collected *Tbx15/18/22* morphants and control embryos in our previous study (Yu et al., 2021). Therefore, controls for *Mrf* morphants were those that we previously used to analyze *Tbx15/18/22* morphants (DRA accession numbers: DRR315181 and DRR315182; Yu et al., 2021). For the remaining samples, libraries were made with an NEBNext Ultra II Directional RNA Library Prep Kit (New England Biolabs, 7765). Obtained reads were deposited in the DRA/SRA database under accession numbers DRR585933-DRR585950. The experiment included two biological replicates. Sequencing reads were mapped to the latest gene model set (Satou et al., 2022) using bowtie2 (Langmead and Salzberg, 2012) with default parameters, and DESeq2 (Love et al., 2014) was used to identify differentially expressed genes with default parameters. RNA-sequencing data of wild-type embryos (Brozovic et al., 2018) were mapped in the same way. Single-cell transcriptome data (Cao et al., 2019) were mapped to the same set of the genome and gene models using Cell Ranger (10x Genomics) with default parameters.

Ciona muscle-specific proteins were mapped to the human proteome (Uniprot release 2021_01, file name: UP000005640_9606.fasta) (UniProt-Consortium, 2019) using BLASTP with default parameters (Altschul et al., 1990). Best-hit proteins with E-values $<1 \times 10^{-5}$ were regarded as homologs. GO annotations were downloaded from the Gene Ontology web site (file name: goa_human.gaf, downloaded on 10 June 2024) (Ashburner et al., 2000).

Acknowledgements

We thank Chikako Imaizumi, Manabu Yoshida and other members working under the National Bio-Resource project (MEXT, Japan) in our group and the University of Tokyo for providing experimental animals.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: Y.S.; Formal analysis: I.O., Y.S.; Investigation: I.O.; Visualization: I.O., Y.S.; Writing – original draft: I.O., Y.S.

Funding

This research was supported by grants from the Japan Society for the Promotion of Science (21H02486, 24K02032, 24K21274 to Y.S.; 21K06017 to I.O.) and from the Takeda Science Foundation to Y.S.

Data availability

RNA-sequencing data are available in the DRA/SRA database under accession numbers DRR585068, DRR585069, and DRR585933–DRR585950 (BIOPROJECT number [PRJDB18571](https://www.ncbi.nlm.nih.gov/bioproject/18571)).

Peer review history

The peer review history is available online at <https://journals.biologists.com/dev/lookup/doi/10.1242/dev.204382.reviewer-comments.pdf>

References

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410. doi:10.1016/S0022-2836(05)80360-2
- 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. doi:10.1007/BF00457803
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T. et al. (2000). Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29. doi:10.1038/75556
- Balogopalan, L., Keller, C. A. and Abmayr, S. M. (2001). Loss-of-function mutations reveal that the *Drosophila nautilus* gene is not essential for embryonic myogenesis or viability. *Dev. Biol.* **231**, 374–382. doi:10.1006/dbio.2001.0162
- Brennan, T. J., Edmondson, D. G. and Olson, E. N. (1990). Aberrant Regulation of MyoD1 contributes to the partially defective myogenic phenotype of Bc3h1 cells. *J. Cell Biol.* **110**, 929–937. doi:10.1083/jcb.110.4.929
- 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. doi:10.1126/science.1145893
- Brozovic, M., Dantec, C., Dardailon, J., Dauga, D., Faure, E., Gineste, M., Louis, A., Naville, M., Nitta, K. R., Piette, J. et al. (2018). ANISEED 2017: extending the integrated ascidian database to the exploration and evolutionary comparison of genome-scale datasets. *Nucleic Acids Res.* **46**, D718–D725. doi:10.1093/nar/gkx1108
- Cao, C., Lemaire, L. A., Wang, W., Yoon, P. H., Choi, Y. A., Parsons, L. R., Matese, J. C., Wang, W., Levine, M. and Chen, K. (2019). Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature* **571**, 349–354. doi:10.1038/s41586-019-1385-y
- Chen, L. S., Krause, M., Sepanski, M. and Fire, A. (1994). The *Caenorhabditis elegans* MyoD homolog Hlh-1 is essential for proper muscle function and complete morphogenesis. *Development* **120**, 1631–1641. doi:10.1242/dev.120.6.1631
- Chopra, V. S. and Levine, M. (2009). Combinatorial patterning mechanisms in the *Drosophila* embryo. *Brief. Funct. Genomic. Proteomic.* **8**, 243–249. doi:10.1093/bfpg/elp026
- Corbo, J. C., Levine, M. and Zeller, R. W. (1997). Characterization of a notochord-specific enhancer from the Brachyury promoter region of the ascidian, *Ciona intestinalis*. *Development* **124**, 589–602. doi:10.1242/dev.124.3.589
- Dehal, P., Satou, Y., Campbell, R. K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D. M. et al. (2002). The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* **298**, 2157–2167. doi:10.1126/science.1080049
- Fujiwara, S., Corbo, J. C. and Levine, M. (1998). The snail repressor establishes a muscle/notochord boundary in the *Ciona* embryo. *Development* **125**, 2511–2520. doi:10.1242/dev.125.13.2511
- Hasty, P., Bradley, A., Morris, J. H., Edmondson, D. G., Venuti, J. M., Olson, E. N. and Klein, W. H. (1993). Muscle deficiency and neonatal death in mice with a targeted mutation in the *myogenin* gene. *Nature* **364**, 501–506. doi:10.1038/364501a0
- Hudson, C., Lotito, S. and Yasuo, H. (2007). Sequential and combinatorial inputs from Nodal, Delta2/Notch and FGF/MEK/ERK signalling pathways establish a grid-like organisation of distinct cell identities in the ascidian neural plate. *Development* **134**, 3527–3537. doi:10.1242/dev.002352
- 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. doi:10.1016/j.ydbio.2007.10.005
- 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. doi:10.1242/dev.129.11.2723
- 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. doi:10.1242/dev.01270
- Imai, K. S., Levine, M., Satoh, N. and Satou, Y. (2006). Regulatory blueprint for a chordate embryo. *Science* **312**, 1183–1187. doi:10.1126/science.1123404
- Imai, K. S., Stolfi, A., Levine, M. and Satou, Y. (2009). Gene regulatory networks underlying the compartmentalization of the *Ciona* central nervous system. *Development* **136**, 285–293. doi:10.1242/dev.026419
- Imai, K. S., Hudson, C., Oda-Ishii, I., Yasuo, H. and Satou, Y. (2016). Antagonism between β -catenin and Gata. a sequentially segregates the germ layers of ascidian embryos. *Development* **143**, 4167–4172. doi:10.1242/dev.141481
- Ishida, T. and Satou, Y. (2024). Ascidian embryonic cells with properties of neural-crest cells and neuromesodermal progenitors of vertebrates. *Nat. Ecol. Evol.* **8**, 1154–1164. doi:10.1038/s41559-024-02387-8
- Johnson, D. S., Zhou, Q., Yagi, K., Satoh, N., Wong, W. and Sidow, A. (2005). De novo discovery of a tissue-specific gene regulatory module in a chordate. *Genome Res.* **15**, 1315–1324. doi:10.1101/gr.4062605
- Kassar-Duchossoy, L., Gayraud-Morel, B., Gomès, D., Rocancourt, D., Buckingham, M., Shinin, V. and Tajbakhsh, S. (2004). Mrf4 determines skeletal muscle identity in *Myf5*:MyoD double-mutant mice. *Nature* **431**, 466–471. doi:10.1038/nature02876
- Katikal, L., Aihara, H., Passamaneck, Y. J., Gazdoui, S., Jose-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. doi:10.1371/journal.pbio.1001697
- 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. doi:10.1016/j.ydbio.2004.09.037
- Langmead, B. and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359. doi:10.1038/nmeth.1923
- Lassar, A. B., Paterson, B. M. and Weintraub, H. (1986). Transfection of a DNA locus that mediates the conversion of 10T1/2 fibroblasts to myoblasts. *Cell* **47**, 649–656. doi:10.1016/0092-8674(86)90507-6
- Long, J. J., Mariossi, A., Cao, C., Mo, Z. Y., Thompson, J. W., Levine, M. S. and Lemaire, L. A. (2023). Cereblon influences the timing of muscle differentiation in *Ciona* tadpoles. *Proc. Natl. Acad. Sci. USA* **120**, e2309989120. doi:10.1073/pnas.2309989120
- Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. doi:10.1186/s13059-014-0550-8
- 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. doi:10.1242/dev.124.9.1711
- 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. doi:10.1016/j.ydbio.2006.09.043
- Nabeshima, Y., Hanaoka, K., Hayasaka, M., Esumi, E., Li, S. W., Nonaka, I. and Nabeshima, Y. (1993). *Myogenin* gene disruption results in perinatal lethality because of severe muscle defect. *Nature* **364**, 532–535. doi:10.1038/364532a0
- 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. doi:10.1038/35055568
- Nitta, K. R., Vincentelli, R., Jacox, E., Cimino, A., Ohtsuka, Y., Sobral, D., Satou, Y., Cambillau, C. and Lemaire, P. (2019). High-throughput protein production combined with high-throughput SELEX identifies an extensive atlas of *Ciona robusta* transcription factor DNA-binding specificities. *Methods Mol. Biol.* **2025**, 487–517. doi:10.1007/978-1-4939-9624-7_23
- Rembold, M., Ciglar, L., Yáñez-Cuna, J. O., Zinnen, R. P., Girardot, C., Jain, A., Welte, M. A., Stark, A., Leptin, M. and Furlong, E. E. M. (2014). A conserved role for Snail as a potentiator of active transcription. *Genes Dev.* **28**, 167–181. doi:10.1101/gad.230953.113
- Rudnicki, M. A., Schneegelsberg, P. N. J., Stead, R. H., Braun, T., Arnold, H.-H. and Jaenisch, R. (1993). MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell* **75**, 1351–1359. doi:10.1016/0092-8674(93)90621-V
- 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. doi:10.1093/nar/gkh012
- 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. doi:10.1046/j.1440-169X.1996.t01-1-00013.x
- Satou, Y., Kusakabe, T., Araki, L. 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. doi:10.1046/j.1440-169X.1995.t01-2-00010.x
- 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. doi:10.1242/dev.128.15.2893
- 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. doi:10.1007/s00427-002-0218-3
- 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. doi:10.2108/zsj.22.837
- Satou, Y., Nakamura, R., Yu, D., Yoshida, R., Hamada, M., Fujie, M., Hisata, K., Takeda, H. and Satoh, N. (2019). A nearly complete genome of *Ciona intestinalis*

- type A (*C. robusta*) reveals the contribution of inversion to chromosomal evolution in the genus *Ciona*. *Genome Biol. Evol.* **11**, 3144–3157. doi:10.1093/gbe/evz228
- Satou, Y., Tokuoka, M., Oda-Ishii, I., Tokuhito, S., Ishida, T., Liu, B. and Iwamura, Y. (2022). A manually curated gene model set for an ascidian, *Ciona robusta* (*Ciona intestinalis* Type A). *Zool. Sci.* **39**, 253–260. doi:10.2108/zs210102
- Simpson, P. (1983). Maternal-zygotic gene interactions during formation of the dorsoventral pattern in *Drosophila* embryos. *Genetics* **105**, 615–632. doi:10.1093/genetics/105.3.615
- 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. doi:10.1002/dvdy.20082
- Tapscott, S. J. (2005). The circuitry of a master switch: MyoD and the regulation of skeletal muscle gene transcription. *Development* **132**, 2685–2695. doi:10.1242/dev.01874
- Thayer, M. J., Tapscott, S. J., Davis, R. L., Wright, W. E., Lassar, A. B. and Weintraub, H. (1989). Positive auto-regulation of the myogenic determination gene MyoD1. *Cell* **58**, 241–248. doi:10.1016/0092-8674(89)90838-6
- Tokuoka, M., Kobayashi, K. and Satou, Y. (2018). Distinct regulation of *Snail* in two muscle lineages of the ascidian embryo achieves temporal coordination of muscle development. *Development* **145**, dev163915. doi:10.1242/dev.163915
- Tokuoka, M., Maeda, K., Kobayashi, K., Mochizuki, A. and Satou, Y. (2021). The gene regulatory system for specifying germ layers in early embryos of the simple chordate. *Sci. Adv.* **7**, eabf8210. doi:10.1126/sciadv.abf8210
- Tokuoka, M., Kobayashi, K., Lemaire, P. and Satou, Y. (2022). Protein kinases and protein phosphatases encoded in the *Ciona robusta* genome. *Genesis* **60**, e23471. doi:10.1002/dvg.23471
- UniProt-Consortium. (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515. doi:10.1093/nar/gky1049
- Venuti, J. M., Gan, L., Kozlowski, M. T. and Klein, W. H. (1993). Developmental potential of muscle cell progenitors and the myogenic factor SUM-1 in the sea urchin embryo. *Mech. Dev.* **41**, 3–14. doi:10.1016/0925-4773(93)90051-X
- Waki, K., Imai, K. S. and Satou, Y. (2015). Genetic pathways for differentiation of the peripheral nervous system in ascidians. *Nat. Commun.* **6**, 8719. doi:10.1038/ncomms9719
- 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. doi:10.1016/j.ydbio.2005.03.029
- Yamada, L., Kobayashi, K., Degnan, B., Satoh, N. and Satou, Y. (2003). A genomewide survey of developmentally relevant genes in *Ciona intestinalis*. IV. Genes for HMG transcriptional regulators, bZip and GATA/Gli/Zic/Snail. *Dev. Genes Evol.* **213**, 245–253. doi:10.1007/s00427-003-0316-x
- Yasuo, H. and Satoh, N. (1993). Function of vertebrate T gene. *Nature* **364**, 582–583. doi:10.1038/364582b0
- Yu, D., Oda-Ishii, I., Kubo, A. and Satou, Y. (2019). The regulatory pathway from genes directly activated by maternal factors to muscle structural genes in ascidian embryos. *Development* **146**, dev173104. doi:10.1242/dev.173104
- Yu, D., Iwamura, Y., Satou, Y. and Oda-Ishii, I. (2021). *Tbx15/18/22* shares a binding site with *Tbx6-r.b* to maintain expression of a muscle structural gene in ascidian late embryos. *Dev. Biol.* **483**, 1–12. doi:10.1016/j.ydbio.2021.12.012
- 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. doi:10.1093/oxfordjournals.jbchem.a003191