Development 137, 1613-1623 (2010) doi:10.1242/dev.046789

Genomic cis-regulatory networks in the early *Ciona intestinalis* embryo

Atsushi Kubo[1], Nobuhiro Suzuki[1], Xuyang Yuan[2], Kenta Nakai[2], Nori Satoh[1,*], Kaoru S. Imai[3] and Yutaka Satou[1,†]

SUMMARY

Precise spatiotemporal gene expression during animal development is achieved through gene regulatory networks, in which sequence-specific transcription factors (TFs) bind to cis-regulatory elements of target genes. Although numerous cis-regulatory elements have been identified in a variety of systems, their global architecture in the gene networks that regulate animal development is not well understood. Here, we determined the structure of the core networks at the cis-regulatory level in early embryos of the chordate *Ciona intestinalis* by chromatin immunoprecipitation (ChIP) of 11 TFs. The regulatory systems of the 11 TF genes examined were tightly interconnected with one another. By combining analysis of the ChIP data with the results of previous comprehensive analyses of expression profiles and knockdown of regulatory genes, we found that most of the previously determined interactions are direct. We focused on cis-regulatory networks responsible for the *Ciona* mesodermal tissues by examining how the networks specify these tissues at the level of their cis-regulatory architecture. We also found many interactions that had not been predicted by simple gene knockdown experiments, and we showed that a significant fraction of TF-DNA interactions make major contributions to the regulatory control of target gene expression.

KEY WORDS: *Ciona intestinalis*, Gene regulatory network, Chromatin immunoprecipitation

INTRODUCTION

During animal development, gene expression is highly regulated, both spatially and temporally. This regulation is primarily achieved at the transcriptional level by the sequence-specific binding of transcription factors (TFs). Such interactions are governed by gene regulatory networks, in which genes exert their effects in a sequential and combinatorial manner (Levine and Davidson, 2005). Remarkable progress has been made in reconstructing the gene regulatory networks responsible for developmental processes. For example, the expression profiles and regulatory networks involved in dorsoventral patterning and mesoderm specification of the fly embryo (Stathopoulos and Levine, 2005; Jakobsen et al., 2007; Li et al., 2008; Sandmann et al., 2006; Sandmann et al., 2007; Liu et al., 2009; Zinzen et al., 2009), endomesoderm specification in *Xenopus* and sea urchin embryos (Koide et al., 2005; Loose and Patient, 2004; Oliveri et al., 2008) and vulva development in *Caenorhabditis elegans* (Inoue et al., 2005; Ririe et al., 2008) have been reported. Significant progress has also been made towards understanding how tissues are specified in the early embryo of the simple chordate *Ciona intestinalis* (Imai et al., 2006). Since the fates of most cells are determined by the early gastrula stage in *Ciona* and most cells in the early embryo can be identified, gene regulatory networks can be analyzed at single-cell resolution.

However, our knowledge of most of these networks is incomplete or provisional, particularly in chordates. Among the regulatory interactions in the *Ciona* embryo that have been revealed by gene knockdown and overexpression studies, only a small fraction have been examined at the cis-regulatory level (Bertrand et al., 2003; Brown et al., 2007; Corbo et al., 1997b; Di Gregorio et al., 2001; Oda-Ishii et al., 2005; Yagi et al., 2004). This is partly because an examination of cis-regulatory elements is very labor intensive. In addition, although most studies of cis-regulatory elements have identified elements that are qualitatively necessary and sufficient for recapitulating the endogenous gene expression patterns, they did not reveal the comprehensive binding profiles of TFs on cis-regulatory regions. Recent studies have revealed that weak or secondary cis-regulatory modules that have often been ignored in such analyses may or may not contribute to driving downstream genes in a quantitative respect (Segal et al., 2008; Zinzen et al., 2009). It therefore remains important to determine the complete binding profiles of TFs on the genome for an integrative understanding of gene regulatory networks in the *Ciona* embryo.

Ciona tadpole larvae have a basic chordate body plan, with the notochord located in the center of the tail and flanked laterally by muscle, ventrally by endoderm and dorsally by the nerve cord (Satoh et al., 2003). The mesenchyme fills the space between the epidermal and endodermal tissues in the trunk. The fate of most cells is already specified before gastrulation is complete. The *Ciona* genome contains ~670 TF genes and their expression profiles during embryonic development have almost all been described at single-cell resolution (Imai et al., 2004; Miwata et al., 2006). Comprehensive functional assays examining regulatory genes expressed in the early embryo have revealed, at single-cell resolution, the molecular pathways that regulate early embryonic specification events (Imai et al., 2006). These studies identified 11 TF genes that play a central role in early regulatory networks. Here, we have examined their in vivo occupancies on genomic DNA and determined the architecture of their networks.

[1]Department of Zoology, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto, 606-8502, Japan. [2]The Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. [3]Department of Biodiversity, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto, 606-8502, Japan.

*Present address: Marine Genomics Unit, Okinawa Institute of Science and Technology Promotion Corporation, Uruma Okinawa, 904-2234, Japan
†Author for correspondence (yutaka@ascidian.zool.kyoto-u.ac.jp)

Accepted 9 March 2010

MATERIALS AND METHODS

Ascidian embryos

Ciona adults were obtained from the National BioResource Project for *Ciona* (Japan) and were maintained in aquaria. Eggs and sperm were surgically obtained from gonoducts. Following insemination, eggs were reared at 18°C in filtered sea water.

Chromatin immunoprecipitation

We made the following DNA constructs encoding GFP-tagged TFs using Multisite Gateway Technology (Invitrogen): Brachyury-promoter(2.2k)-attB1-Brachyury coding region-attB2-GFP; FoxA-a-promoter(3.3k)-attB1-GFP-attB2-FoxA-a coding region; FoxD-promoter(3.1k)-attB1-FoxD coding region-attB2-GFP; MyoD-promoter(3.5k)-attB1-MyoD coding region-attB2-GFP; Neurogenin-promoter(3.1k)-attB1-Neurogenin coding region-attB2-GFP; Otx-promoter(2.9k)-attB1-Otx coding region-attB2-GFP; Snail-promoter(3.1k)-attB1-Snail coding region-attB2-GFP; SoxC-promoter(2.0k)-attB1-SoxC coding region-attB2-GFP; Tbx6b-promoter(3.1k)-attB1-Tbx6b coding region-attB2-GFP; Twist-like1-promoter(4.9k)-attB1-Twist-like1 coding region-attB2-GFP; and ZicL-promoter(1.8k)-attB1-ZicL coding region-attB2-GFP. Each of these constructs was introduced into fertilized eggs by electroporation (Corbo et al., 1997a). The embryos were fixed with 1% formaldehyde for 15 minutes. The developmental stages at which these embryos were fixed are shown in Fig. 1. Chromatin extracted from 0.5-1.0 g of embryos was sonicated to obtain DNA fragments with an average size of 1 kb. The DNA fragments were enriched by immunoprecipitation with a polyclonal antibody specific for GFP (Invitrogen). After reversal of the fixation, the immunoprecipitated and input DNAs were analyzed by quantitative PCR (qPCR) or amplified ligation-mediated PCR for the microarray analysis.

qPCR was performed using the SYBR Green method (Applied Biosystems). The ChIP enrichment ratio for a given region was determined as the ratio of the region in the precipitated DNA to that in the whole genome relative to the average for five genomic regions that were not expected to be bound by ZicL and Brachyury.

For the microarray analysis, immunoprecipitated DNA and genomic DNA were randomly labeled with Cy3 and Cy5 and hybridized to whole-genome *Ciona* arrays using the CGH protocol provided by Agilent Technologies. Two or three independent experiments were performed to reduce experimental error. Microarray data were submitted to the GEO public database (for accession numbers, see Table 1).

Design of the whole-genome *Ciona* arrays

We used a set of two 244k arrays (version 1) for 16 experiments and one 1M array (version 2) for seven experiments. These arrays were custom made by Agilent Technologies. The arrays used in each experiment are listed in Table 1. Each of these two sets of arrays covers the entire euchromatic region of the *Ciona* genome (Dehal et al., 2002) (GEO accession numbers for the array design: GPL6555/GPL6556 for the version 1 array; GPL8993 for the version 2 array). The version 1 array contains 393,266 unique probes separated by an average of 292 bp. In addition to the uniformly dispersed probes, 56,113 probes separated by an average of 79 bp were designed to densely cover the region spanning from 5 kb upstream of the transcription start site to the first intron of each regulatory gene. The version 2 array contains 971,042 unique probes separated by an average of 60 bp to densely cover the same genomic region as the version 1 arrays.

Data analysis

The raw data from the tiling arrays were normalized using Feature Extraction software (version 9.0, Agilent Technologies) with Linear and Lowess normalization. To identify significant probes and regions, we used two different programs called ChIPmix (Martin-Magniette et al., 2008) and Chipotle (Buck et al., 2005), which are based on totally different algorithms: the former is based on a linear regression mixture model and identifies significant probes, whereas the latter uses a sliding window approach and identifies significant regions. The Chipotle program (window size, 1 kb; step size, 250 bp) was first used to determine candidate significant regions. Among the DNA segments identified by Chipotle, we chose those containing one or more probes that the ChIPmix program identified as significant. To reduce noise, we performed two or three independent experiments for each of the 11 TFs (biological replicates). Only the segments identified as positive in both experiments were considered significant in the present study.

Table 1. Microarrays used in each experiment and GEO accession numbers

Experiment	Microarray version	Accession number
Brachyury	1	GSM300069/GSM300071
Brachyury	1	GSM300471/GSM300475
FoxA-a	1	GSM441213/GSM441214
FoxA-a	2	GSM441215
FoxD	1	GSM441225/GSM441226
FoxD	2	GSM441227
MyoD	1	GSM325773/GSM325774
MyoD	1	GSM326036/GSM326037
Neurogenin	1	GSM441228/GSM441229
Neurogenin	2	GSM441230
Otx	1	GSM441231/GSM441232
Otx	2	GSM441233
Snail	1	GSM441234/GSM441235
Snail	2	GSM441236
SoxC	1	GSM441237/GSM441238
SoxC	2	GSM441239
Tbx6b	1	GSM441272/GSM441273
Tbx6b	2	GSM441274
Twist-like1	1	GSM441275/GSM441276
Twist-like1	1	GSM441277/GSM441278
Twist-like1	1	GSM441279/GSM441280
ZicL	1	GSM271902/GSM271903
ZicL	1	GSM271904/GSM271905

For mapping of the identified segments to the genes, we used the newest gene model set (Satou et al., 2008). If a portion of or the entire DNA segment was mapped within a genomic region containing a gene, the DNA segments were considered to be associated with this gene. If a given DNA segment was located in a region between two genes that were in head-to-tail orientation, and the distance between the segment and the second gene was less than 5 kb, the segment was considered to be associated with the second gene. If the two genes were aligned in head-to-head orientation, the segment was considered to be associated with the closer gene provided that the distance was less than 5 kb. If the two genes were aligned in tail-to-tail orientation, the segment was not considered to be associated with either gene.

The enrichment of TF binding sequences in the identified regions was examined as previously described (Li et al., 2008). The background regions were randomly selected non-coding regions of the *Ciona* genome that did not overlap with the binding regions of the TF in question. The position weight matrices (PWMs) of ZicL and Tbx6b were constructed from the results of Selex (systematic evolution of ligands by exponential enrichment) experiments (Yagi et al., 2004; Yagi et al., 2005). For the remaining nine TFs of unknown PWM, we searched for their best homologs, the PWMs of which were found in the TRANSFAC database (Matys et al., 2003) using the BLASTP program. Patser (Hertz and Stormo, 1999) was used to search for matches to the PWMs. The enrichment was calculated by dividing the density of matches in the bound regions by the density of matches in the background regions inside 125 bp non-overlapping sliding windows across the regions 4 kb upstream to 4 kb downstream of the peak positions of identified binding regions.

Gene knockdown and overexpression experiments

In the present study, we used morpholino antisense oligonucleotides (MOs) (Gene Tools) for *MyoD*, *Twist-like1* and *Brachyury*, the specificities of which were confirmed in a previous study (Imai et al., 2006). Synthetic capped mRNAs were synthesized from cDNAs cloned into pBluescript RN3 vector (Lemaire et al., 1995) using the mMESSAGE mMACHINE Kit (Ambion). Fifteen fmoles of MOs or 1.5 pg (Brachyury) or 9 pg (MyoD and Twist-like1a) of synthetic capped mRNAs in 30 pl of solution were injected into dechorionated eggs after fertilization, as described previously (Imai et al., 2006). Cleavage of some embryos was arrested at the 110-cell stage with 2.5 μg/ml cytochalasin B, and the arrested embryos were cultured further

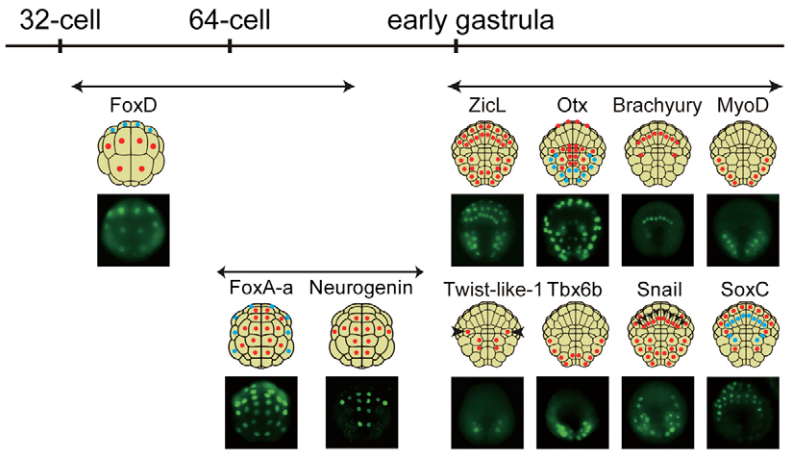

Fig. 1. An overview of the ChIP experiments. *Ciona intestinalis* embryos electroporated with DNA constructs encoding GFP-tagged transcription factors (TFs) under the control of their own promoters reproduced the endogenous expression patterns, which are depicted above the images showing GFP expression. Red circles indicate the blastomeres with mRNA expression at the stage indicated. Blastomeres indicated by light blue circles do not express the TF genes indicated, but their ancestral cells do. Samples for each experiment were collected at the stages indicated by double-headed arrows. Arrowheads indicate endogenous expression that was not reproduced with our constructs.

until the control embryos reached the early tailbud stage (10 hours after fertilization). Relative quantification of mRNA levels (by qRT-PCR) was performed as described (Imai et al., 2004).

RESULTS

Identification of in vivo binding of eleven TFs in the early *Ciona* embryo

The developmental fates of blastomeres in the *Ciona* embryo have been determined by the gastrula stage (Nishida, 1987). A comprehensive study has revealed that 53 TF genes are zygotically expressed and regulate one another in complex networks before gastrulation begins (Imai et al., 2006). To dissect the architecture of these networks at the level of protein-DNA interactions, we focused on 11 TF genes that play core roles in gene regulatory networks for endomesoderm specification: *Brachyury*, *FoxA-a*, *FoxD*, *MyoD*, *Neurogenin*, *Otx*, *Snail*, *SoxC*, *Tbx6b*, *Twist-like1* and *ZicL*. Because the *Ciona* genome contains multiple copies of *FoxD*, *Tbx6b* and *ZicL* as gene clusters and their precise copy numbers have not yet been determined, these genes are collectively referred to *FoxD*, *Tbx6b* and *ZicL* in this paper. Likewise, there are two copies of *Twist-like1*, which are highly similar to each other, and we collectively refer to these as *Twist-like1*.

We prepared 11 gene-fusion constructs that encode GFP-tagged TFs expressed under the control of their own promoters (e.g. a fusion gene that encodes GFP-tagged Brachyury driven by the *Brachyury* promoter). When these constructs were introduced into eggs, the resultant embryos expressed the fusion genes at the same time and in the same blastomeres as the endogenous genes (Fig. 1). Exceptions were the *Twist-like1* and the *Snail* constructs. *Twist-like1* is normally expressed in three cell lineages (A7.6, B7.7 and B8.5), but our construct drove Twist-like1-GFP expression only in the B7.7 and B8.5 lines. *Snail* expression in the notochord lineage is normally very weak. Our *Snail* construct did not recapitulate this expression in the notochord lineage but did drive Snail-GFP expression in the remaining lineages.

Expression of these genes did not affect embryonic morphology at the stage when the embryos were fixed (Fig. 1). The fixed embryos were subjected to ChIP using anti-GFP antibodies, and subsequently to microarray analysis. To define significant regions, we used two programs employing totally different algorithms. DNA segments regarded as positive by both programs were defined as significant. To confirm that our approach successfully identified TF binding sites, we analyzed the sequences of ZicL and Tbx6b binding regions defined with three different false discovery rates (FDRs), as the consensus binding motifs of these two TFs are known (Yagi et al., 2004; Yagi et al., 2005). The frequencies of matches to the consensus binding sequences for ZicL and Tbx6b around peaks in 0.1% FDR were generally better than in 0.01% and 1% FDRs (Fig. 2A). As expected, the frequencies of the consensus binding sequences for ZicL and Tbx6b were markedly higher around peaks in the identified regions (Fig. 2A), suggesting that our method was able to successfully identify the TF binding regions.

Brachyury and *Ci-tropomyosin-like* are the only known direct targets of ZicL and Brachyury, respectively (Yagi et al., 2004; Di Gregorio and Levine, 1999). As an independent confirmation, we inspected the TF binding sites of these genes. Our ZicL ChIP profile showed a sharp peak around two known strong ZicL binding sites (Fig. 2B). The Brachyury ChIP profile also showed a peak around the known Brachyury binding site in the *Ci-tropomyosin-like* promoter (Fig. 2C). These peaks were included in significant regions identified with all the FDRs described above. We also performed ChIP-qPCRs for these two known interactions. The ChIP-qPCR results showed excellent agreement with the ChIP-chip results (Fig. 2B,C, orange dots).

Next, we examined the promoters of genes that were identified in previous studies as likely direct targets of one of the 11 TFs on the basis of expression assays (Imai et al., 2004) and gene knockdown assays (Imai et al., 2006). Among 29 interactions that had been found in the gene knockdown assays and for which both the source and target genes are expressed in the same cells, 28, 23 and 19 interactions were indicated to be direct under the FDRs of 1%, 0.1% and 0.01%, respectively. The remainder of the interactions were not regarded as direct. *Otx* expression in the A-line lineage requires a cis-regulatory module that includes Fox binding sites (Oda-Ishii et al., 2005) and is suppressed in *FoxA-a* morphants (Imai et al., 2006). The FoxA-a binding to this cis-regulatory element was counted with FDRs of 1% and 0.1%, but not with the most stringent FDR (0.01%) (Fig. 3A and see Fig. S7 in the supplementary material). Similarly, several lines of evidence have suggested that *MyoD* is directly regulated by ZicL. First, *MyoD* expression is suppressed in *ZicL* morphants (Imai et al., 2006). Second, *MyoD* and *ZicL* are both expressed in presumptive muscle cells and the time windows of their expression overlap (Imai et al., 2004). Lastly, there is a putative ZicL binding site near to the peaks found in the *MyoD* upstream region (P=1.2×10^{-5}; Fig. 2D). This putative binding was observed under the FDRs of 1% and 0.1%, but not under the most stringent FDR of 0.01%. On the basis of the above observations, in the following sections we generally

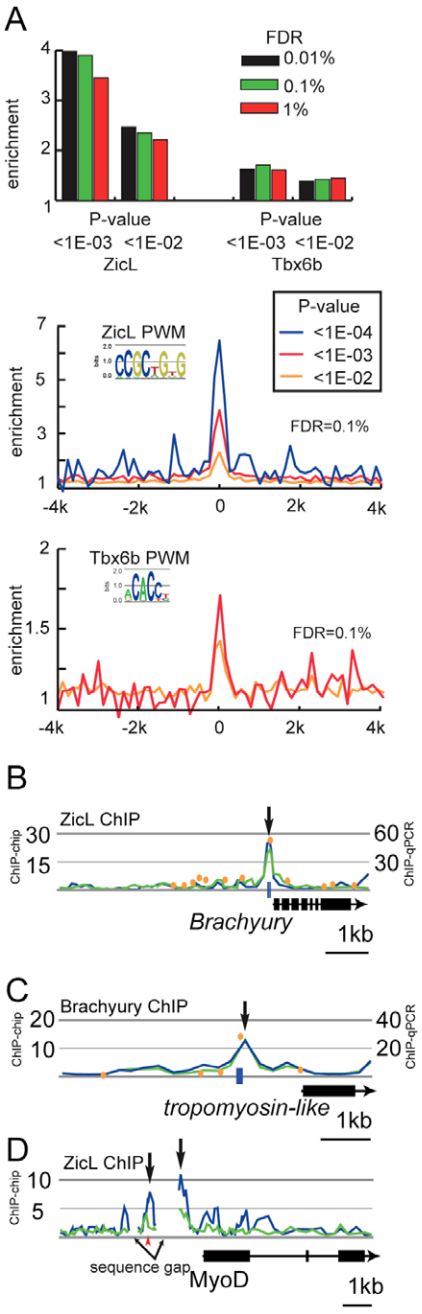

Fig. 2. Confirmation that TF binding sites were successfully identified. (**A**) Fold enrichment of matches to the position weight matrices (PWMs) for ZicL and Tbx6b in 125 bp windows. (Upper panel) Fold enrichment at the peak positions at three different false discovery rates (FDRs). (Middle and lower panels) Fold enrichment across the 0.1% FDR peaks at three different thresholds. The zero position on the *x*-axis represents the peak position. Note that no matches to the Tbx6b PWM were found at the threshold of 1×10^{-4}. (**B-D**) Mapping of the ZicL (B,D) and Brachyury (C) ChIP data onto genomic regions consisting of the exons and 5 kb upstream of *Brachyury* (B), *Ci-tropomyosin-like* (C) and *MyoD* (D). Each graph shows fold enrichment (*y*-axis, a non-log linear scale) for the chromosomal regions (*x*-axis). Blue boxes on the *x*-axis represent experimentally confirmed ZicL and Brachyury binding sites. The red arrowhead in D indicates a putative ZicL binding site. ChIP-qPCR results are shown with orange dots and ChIP-Chip results are denoted by lines; blue and green lines show the results of two independently performed experiments. Scales for ChIP-qPCR and ChIP-Chip results are shown on the right and left of each graph, respectively.

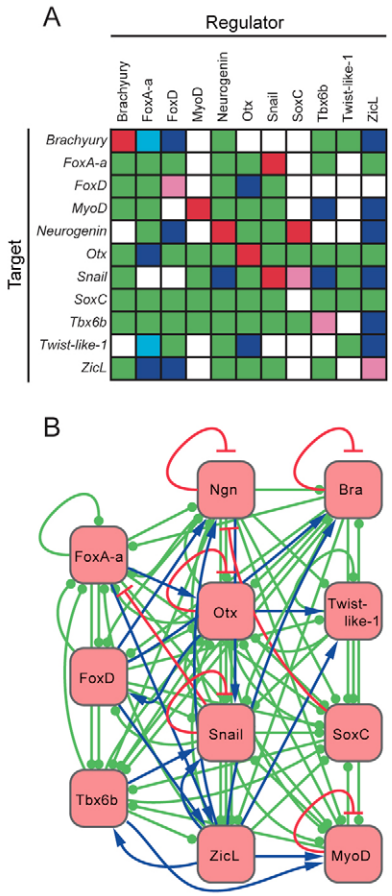

Fig. 3. The core regulatory network in the early *Ciona intestinalis* embryo. (**A**) Comparison between results of the present ChIP assays and the gene knockdown assays of a previous study (Imai et al., 2006). Among negative interactions revealed by the gene knockdown assays, interactions that were and were not confirmed with the ChIP assays (FDR, 0.1%) are shown by red and pink boxes, respectively. Among positive interactions revealed by the gene knockdown assays, interactions that were and were not confirmed with the ChIP assays are shown by blue and light-blue boxes, respectively. Interactions identified only with the ChIP assays are shown by green boxes. White boxes indicate that no interactions were observed. (**B**) Protein-DNA Interactions between the 11 core TFs and the DNA segments encoding these TFs [drawn with Cytoscape (Shannon et al., 2003)]. Blue arrows indicate positive interactions and red T-bars indicate negative interactions. Lines ending with a green circle indicate interactions of undetermined function.

describe the results obtained at an FDR of 0.1%. The number of significant regions and the putative target genes associated with the 11 TFs are summarized in Table 2 and listed in Table S1 in the supplementary material.

As shown in Fig. 2A, the frequencies of the consensus sequences for ZicL and Tbx6b binding were markedly higher around peaks in the identified regions. Since the consensus binding motifs of the other nine TFs had not been determined previously, we performed similar analyses with motifs of homologs in other animals (see Fig. S1A-I in the supplementary material). The frequencies of the consensus binding motifs for six of the TFs, but not FoxD, SoxC or Twist-like1, were markedly higher around peaks. Because the position weight matrices (PWMs) for FoxD, SoxC and Twist-like1

Table 2. An overview of the ChIP results

Transcription factor	Significant regions	Putative target genes
Brachyury	2924	2251
FoxA-a	5265	3885
FoxD	342	316
MyoD	386	331
Neurogenin	4696	3646
Otx	6022	4104
Snail	2436	1922
SoxC	254	234
Tbx6b	978	864
Twist-like1	435	373
ZicL	1707	1178

gave higher background, no significant changes were seen. However, the number of matches to the motifs was markedly higher around peaks than in flanking regions and the background (see Fig. S1J-L in the supplementary material). These observations suggested that our method was able to successfully identify the TF binding regions.

As previously reported in other animals (Li et al., 2008), we found that the regions bound by Brachyury, MyoD, Neurogenin, Snail, Tbx6b, Twist-like1 and ZicL, especially around the peaks, showed a marked GC bias (see Fig. S2 in the supplementary material). This bias is likely to be related to the consensus sequences, because the consensus sequences for these TFs are generally more GC-rich than those of the remaining TFs (see Fig. S4 in the supplementary material). The observed enrichment of recognition sequences was unlikely to be an artifact of GC bias because even if we picked up background sequences with a base composition comparable to the averaged GC content of the bound regions (the difference between the average GC content of the bound and background regions was less than 0.8%), matches to the PWMs were enriched around peaks versus each of the GC-adjusted backgrounds (see Fig. S3 in the supplementary material).

Next, we attempted to discover overrepresented motifs in the regions (360 bp) around the peaks identified by each ChIP experiment using the Trawler program (Ettwiller et al., 2007). We found that overrepresented motifs were similar to the PWMs that were determined experimentally (Tbx6b and ZicL) or to those of homologs in other animals (the remaining nine TFs) (see Fig. S4 in the supplementary material). This further supported the conclusion that the results of our ChIP experiments were of high quality.

It is generally believed that TFs tend to bind near promoters, although many examples are known in which TFs bind to enhancers far from promoters. The distributions of peaks in all experiments, except Snail ChIP, were higher around transcription start sites (see Fig. S5 in the supplementary material). The reason why Snail binding sites were not enriched around transcription start sites is unclear, but this does not necessarily indicate that the results of the Snail ChIP were of low quality. Altogether, these observations support the conclusion that all of our ChIP experiments revealed in vivo occupancies of the TFs.

Tightly interconnected regulatory networks in early embryogenesis

TF genes were significantly enriched among the target genes of the 11 TFs. Among 670 potential TF genes in the *Ciona* genome, at least 607 encode proteins with known TF motifs or proteins with two or more zinc-finger motifs that potentially bind to DNA (Imai et al., 2004; Miwata et al., 2006). As shown in Table 3, we found a significantly greater number of TF genes among the targets than would be expected from random sampling (>99% confidence in χ^2 tests). This enrichment indicates that the TFs examined bind targets selectively and not randomly.

Table 3. Regulatory genes among the target genes of the eleven TFs

	Target genes encoding TFs	
	Number	Percentage
Whole genome	607	4.0
ChIP		
Brachyury	194	8.6
FoxA-a	245	6.3
FoxD	43	13.6
MyoD	77	23.3
Neurogenin	259	7.1
Otx	266	6.5
Snail	168	8.7
SoxC	28	12.0
Tbx6b	126	14.6
Twist-like1	88	23.6
ZicL	184	15.6

We compared the ChIP data with the results of the comprehensive gene knockdown experiments of a previously study (Imai et al., 2006). Among 76 interactions previously found in the early embryo, the ChIP assays indicated that 58 are direct (see Fig. S6 in the supplementary material). In addition, we found 251 novel interconnections. Fig. 3A shows interconnections among the 11 TFs that were subjected to the present ChIP experiments (for results with the mild and stringent FDRs, see Fig. S7 in the supplementary material). Among 121 (11×11) possible interconnections, 84 were observed in the present study (Fig. 3B). Our data indicate that these genes are tightly interconnected with one another.

Regulatory networks for mesodermal tissue specification

Because the gene regulatory network model previously constructed from comprehensive expression profiles and comprehensive knockdowns of regulatory genes is of single-cell resolution (Imai et al., 2006; Imai et al., 2009), we simply integrated the ChIP data into this network by assuming that the examined TFs bind to the targets wherever their mRNAs are expressed. The reconstructed networks had a complex architecture.

The reconstructed regulatory networks allow us to trace development at the single-cell level. Figs S8 and S9 in the supplementary material show the interconnections among the core 11 TFs in A-line and B-line blastomeres, which give rise to endomesodermal tissues, from the 8-cell to the early gastrula stage. Two of the three mesenchymal lineages (B-line mesenchymal cells) and 28 out of 36 muscle cells (B-line muscle cells) in the tadpole larvae are derived from B4.1 blastomeres at the 8-cell stage (Fig. 4A). Thirty-two and eight notochord cells are derived from A4.1 and B4.1 blastomeres, respectively. Previous studies demonstrated that *Twist-like1*, *MyoD* and *Brachyury* are essential for specification of the mesenchyme, muscle and notochord, respectively (Corbo et al., 1997b; Imai et al., 2003; Imai et al., 2006; Meedel et al., 2007; Tokuoka et al., 2004; Yasuo and Satoh, 1998).

Mesenchyme

Twist-like1 is expressed exclusively in the mesenchymal lineage and is regulated by *FoxA-a*, *Otx* and *ZicL*, as indicated by the fact that knockdown of any of these three genes results in loss or reduction

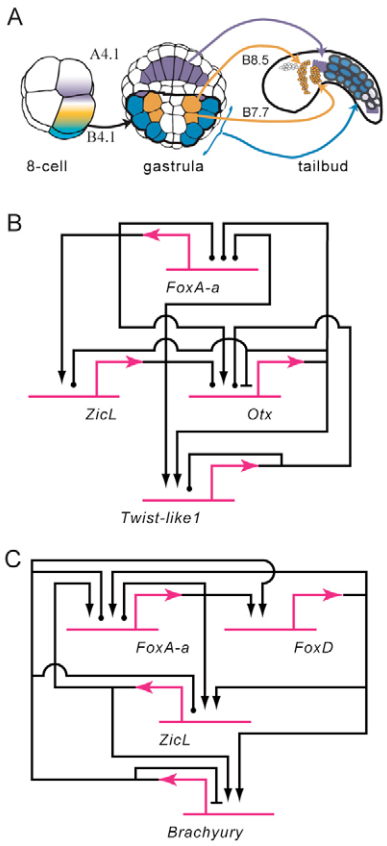

Fig. 4. Cis-regulatory networks for specification of mesodermal tissues. (**A**) Depiction of the B-line mesenchymal (orange), muscle (blue) and notochord (purple) cells at the gastrula and tailbud stages. (**B**,**C**) Overviews of the cis-regulatory network for specification of (B) mesenchyme and (C) notochord. Black arrows indicate interactions that activate transcription and T-bars indicate interactions that repress transcription. Lines ending with a circle indicate interactions of undetermined function. These depictions were drawn with Biotapestry (Longabaugh et al., 2005) and then modified.

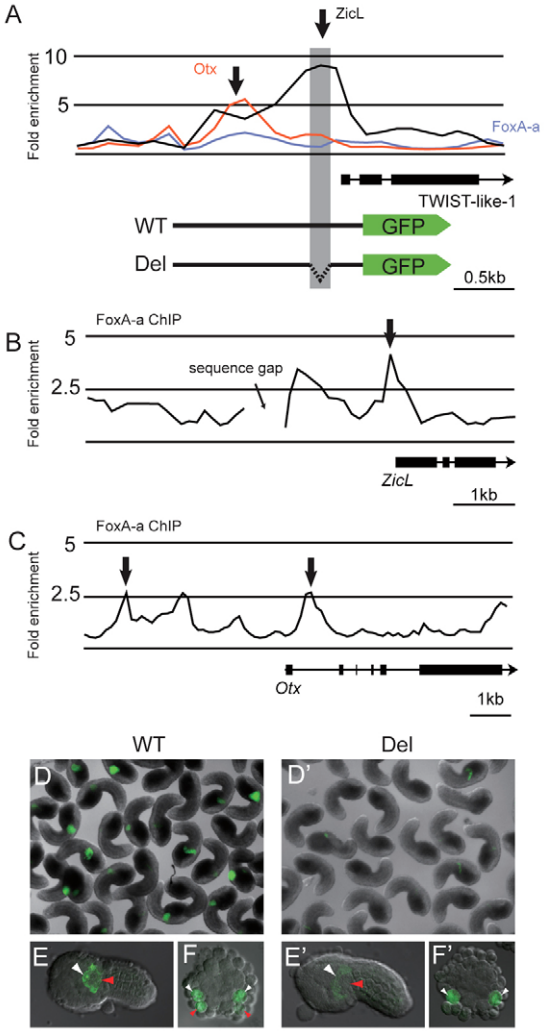

Fig. 5. Regulation of Twist-like1 expression. (**A**) Mapping of the ZicL ChIP (black), Otx ChIP (red) and FoxA-a ChIP (purple) data onto a genomic region consisting of the *Twist-like1* region and 5 kb of upstream sequence. The *y*-axis shows fold enrichment of ChIP DNA relative to input DNA in 125 bp non-overlapping windows (the enrichment ratio was averaged in each window, and medians of the averaged ratio for two replicates are shown). DNA constructs used to examine ZicL binding to the *Twist-like1* promoter are shown beneath. WT, wild-type *Twist-like1* promoter sequence. Del contains a 150 bp deletion that includes the ZicL binding site. (**B**,**C**) Mapping of the FoxA-a ChIP data onto genomic regions consisting of the exons and 5 kb upstream sequences of ZicL (B) and *Otx* (C). (**D-F'**) GFP expression in *Ciona intestinalis* embryos electroporated with fusion genes of GFP and the wild-type (D-F) and deletion mutant (D′-F′) promoters. (D,D′) The wild-type promoter drives GFP expression more strongly than does the mutant promoter, and the number of GFP-positive embryos is greater in D than in D′. (E-F′) The wild-type promoter activated GFP in both B7.7 (red arrowheads) and B8.5 lineages (white arrowheads) of the tailbud embryo (E) and the cleavage-arrested embryo (F), whereas the mutant promoter failed to activate GFP in the B7.7 lineage of the tailbud embryo (E′) and the cleavage-arrested embryo (F′). The embryos shown in E-F′ were subjected to indirect immunofluorescence using an anti-GFP primary antibody and an Alexa 488-labeled secondary antibody to enhance the signal.

of *Twist-like1* expression (Imai et al., 2006). As summarized in Fig. 4B (see also Fig. S11A in the supplementary material), we could not detect direct binding of FoxA-a to the *Twist-like1* promoter (Fig. 5A) even with 1% FDR, but we did find that FoxA-a binds to the upstream regions of *Otx* and *ZicL* (Fig. 5B,C), and that ZicL and Otx bind to the promoter of *Twist-like1* (Fig. 5A). Therefore, it is highly likely that FoxA-a mainly activates *Twist-like1* indirectly through activating *Otx* and *ZicL*.

Twist-like1 expression begins in B7.7 (the posterior B-line mesenchyme) at the 64-cell stage and in B8.5 (the anterior B-line mesenchyme) at the early gastrula stage. These two mesenchymal lines contribute to distinct adult tissues after metamorphosis (Tokuoka et al., 2005). ZicL might be associated with the differences between these two lineages because the contribution of *ZicL* to *Twist-like1* activation is weaker than that of *Otx* (Imai et al., 2006). To confirm this idea, we tested a mutant *Twist-like1* promoter, from which a 150 bp segment containing the identified ZicL binding region was deleted. Because the Otx ChIP result indicated that the Otx binding region is distinct from the ZicL binding region, Otx was expected to bind to this mutant promoter. When introduced into fertilized eggs by electroporation, the wild-type promoter (1550 bp) drove reporter expression in 65% of the embryos (Fig. 5D; n=368, three independent experiments), whereas the mutant promoter drove reporter expression in 36% of the embryos (Fig. 5D′; n=281, three independent experiments). In addition to the significant decrease in

the number of embryos expressing the reporter (Student's *t*-test, $P<0.001$), the overall fluorescence was weaker and the posterior B-line mesenchyme did not appear to express the reporter in the mutant construct (Fig. 5E,E′). To confirm this observation, the experimental embryos were cleavage-arrested at the 110-cell stage. Cells in the arrested embryos cannot divide further, but the developmental programs proceed as in normal embryos. The mutant construct failed to drive reporter expression in the posterior B-line mesenchyme (Fig. 5F,F′). These results suggest that ZicL contributes to the difference between these two lineages.

A previous study showed that nine mesenchyme-specific non-regulatory genes are under the control of *Twist-like1*. None of these genes was identified as a direct target in the present study (Fig. S10A in the supplementary material). Even when applied with an FDR of 1%, only one gene was identified as a direct target. Therefore, it is likely that *Twist-like1* regulates the expression of mesenchyme-specific genes through its downstream regulatory gene circuit, although there is a possibility that Twist-like1 binds to the regulatory elements of these genes at later stages.

Muscle

The B6.2 and B6.4 cell pairs in the 32-cell embryo have the potential to give rise to mesenchyme and muscle (Fig. S8 in the supplementary material). At the 64-cell stage, these cells divide, and one of the daughter cells becomes specified to give rise to the muscle cells. Previous functional assays showed that *ZicL*, *Tbx6b* and *MyoD* are essential for specification of muscle cells (Imai et al., 2002c; Imai et al., 2006; Meedel et al., 2007; Yagi et al., 2005). *Tbx6b* begins to be expressed at the 16-cell stage, and cells expressing *Tbx6b* give rise not only to muscle cells but also to mesenchyme cells. *Tbx6b* expression declines to undetectable levels before the tailbud stage. *ZicL* starts to be expressed at the 32-cell stage in a variety of cells, including those with developmental fates of muscle, mesenchyme, notochord and neurons. *ZicL* expression in the muscle lineage disappears before the late gastrula stage. *MyoD* expression begins at the 44-cell stage exclusively in the muscle lineage under the control of *Tbx6b* and *ZicL*. The present study showed that *ZicL*, *Tbx6b* and *MyoD* constituted a tightly interconnected gene circuit that is responsible for this specification (see Fig. S11B in the supplementary material): (1) ZicL bound to the promoters of *MyoD* and *Tbx6b*; (2) Tbx6b bound to the promoters of *MyoD* and *ZicL*; and (3) MyoD bound to the promoter of *Tbx6b* and to its own promoter. All of these interactions, except MyoD binding to the *Tbx6b* promoter, have been confirmed by functional assays (Yagi et al., 2005; Imai et al., 2006).

To understand how this gene circuit regulates downstream muscle-specific genes, we examined the promoters of 13 muscle structural genes that are well annotated and known to be expressed in the larval tail muscle (see Fig. S10 in the supplementary material). Of these, ten were directly bound by MyoD and Tbx6, one by MyoD and ZicL, one by Tbx6b and ZicL, and one by MyoD alone.

As expected from a previous study (Brown et al., 2007), both MyoD and Tbx6 bound to the promoters of more than three-quarters of the muscle genes examined. To test the action of this feed-forward loop comprising MyoD and Tbx6b in the regulation of muscle-specific gene expression, we examined the expression patterns of genes under the control of this circuit. Of the 155 genes under the direct control of MyoD and Tbx6b, 50 (including the above ten) were already known to be expressed in muscle cells (Fujiwara et al., 2002; Imai et al., 2004; Kusakabe et al., 2002; Miwata et al., 2006; Satou et al., 2001). From the remaining 105 genes, we randomly chose 20, and found that 15 are expressed in muscle cells (Fig. 6), suggesting that this circuit is widely used for the regulation of genes expressed in muscle cells, and also that this circuit might not necessarily be sufficient for driving expression of the target.

Notochord

Brachyury is activated at the 64-cell stage exclusively in the notochord lineage, and this expression specifies the notochord fate (Corbo et al., 1998; Yasuo and Satoh, 1998). As described above, ZicL directly binds to the *Brachyury* promoter and activates its expression (Yagi et al., 2004). It has also been shown that *FoxD* and *FoxA-a* are required for *Brachyury* expression, probably through regulating *ZicL* expression, and that FGF signaling is also required for *Brachyury* expression (Imai et al., 2002a; Imai et al., 2002b; Imai et al., 2006). The present assays showed that not only ZicL, but also FoxD binds to the *Brachyury* promoter (Fig. 4C; see Fig. S11C in the supplementary material). Although *FoxD* mRNA is not present in the notochord lineage at the 32-cell and 64-cell stages, when *ZicL* and *Brachyury* are activated, respectively (*FoxD* is expressed in the ancestors of cells in which *ZicL* and *Brachyury* are expressed; see Fig. S9 in the supplementary material), the ChIP assay indicated that FoxD binds to the promoters of *ZicL* and *Brachyury*. Because knockdown of *FoxD* eliminates *ZicL* and *Brachyury* expression (Imai et al., 2002b; Imai et al., 2006) and because the FoxD-GFP fusion protein exists in the notochord lineage at the 32-cell stage (Fig. 1), it is likely that FoxD protein exists in these cells and binds to the promoters of *ZicL* and *Brachyury* when these two genes begin to be expressed.

FoxA-a binding to the *Brachyury* promoter was not identified under 0.1% FDR. There was, however, a small peak that was counted as significant under 1% FDR. We could not rule out the possibility that FoxA-a binds weakly to the *Brachyury* promoter. It is also possible that FoxA-a could bind weakly to a FoxD binding site because the FoxA-a binding peak coincided with that of FoxD (see Fig. S11 in the supplementary material). Even if this weak binding occurs in vivo, the regulation of *Brachyury* by FoxA-a would largely be achieved indirectly through FoxD and ZicL, as we found strong binding of FoxA-a to the promoters of *FoxD* and *ZicL*.

Next, we examined 14 non-regulatory genes that are known to be expressed in the notochord under the control of *Brachyury* (Hotta et al., 2000; Takahashi et al., 1999). Among them, 11 were identified here as direct targets of Brachyury (see Fig. S10 in the supplementary material). The present results suggest that the remaining three genes are regulated indirectly through a gene circuit under the control of *Brachyury*, although it cannot be ruled out that Brachyury binds to the regulatory elements of these three genes at later stages.

The contribution of TF-DNA binding to regulatory controls of target gene expression

In the present study, we found many interactions between TFs and genomic DNA that were unexpected from preceding gene knockdown assays. Similar observations were also reported in preceding ChIP studies (Jakobsen et al., 2007; Li et al., 2008; Sandmann et al., 2006). To estimate what proportion of the binding makes a major contribution to gene regulation in *Ciona* embryos, we injected *MyoD* mRNA or an MO against *MyoD* into eggs and examined their effects on the expression of the same targets that we analyzed above at the gastrula stage or at the tailbud stage, respectively. The mRNA levels of 14 targets, ten of which were expressed in muscle, were significantly increased (>2-fold) in embryos injected with *MyoD* mRNA (Fig. 6), and *MyoD* MO injection significantly reduced the mRNA levels of three of these

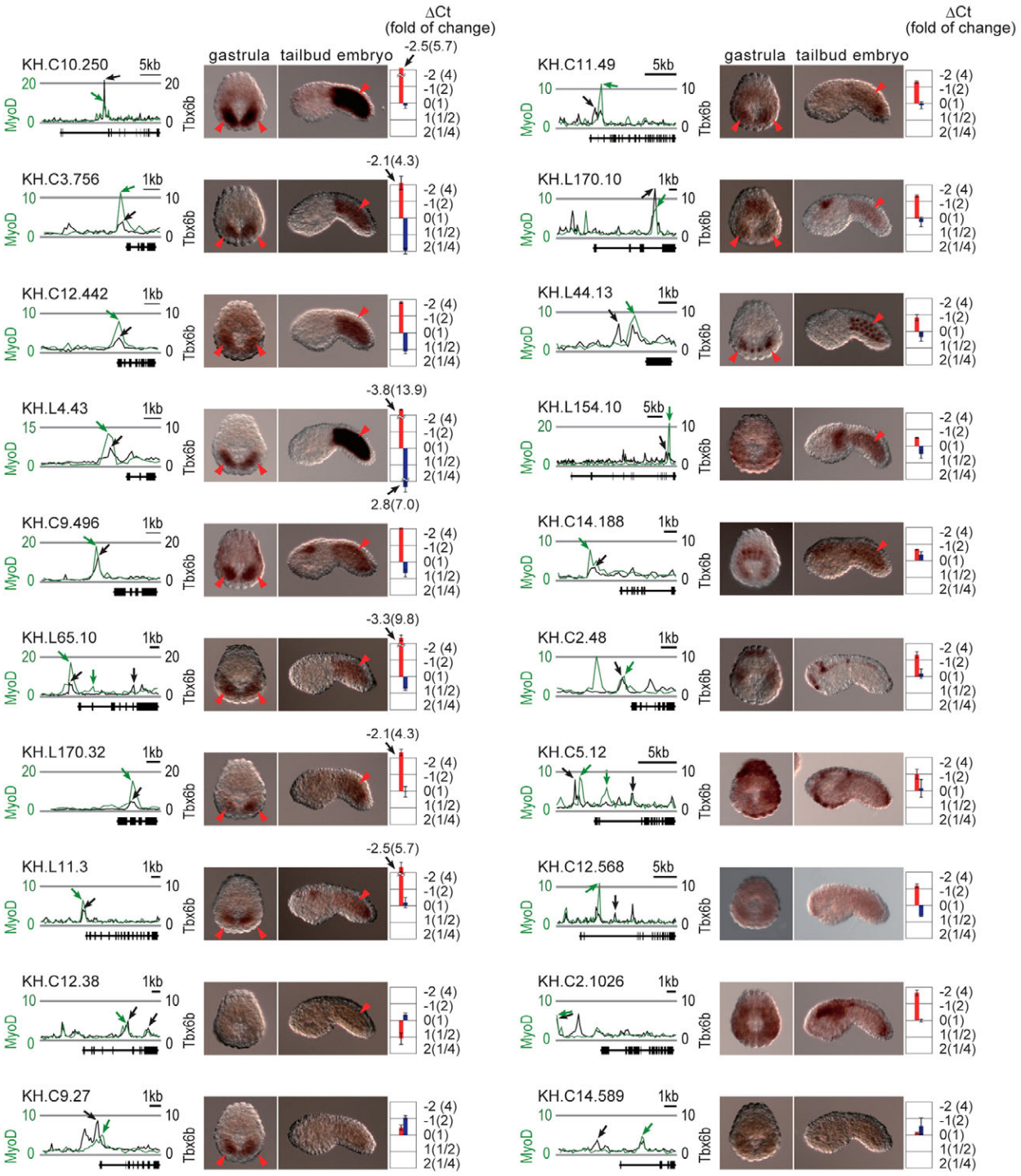

Fig. 6. Expression patterns of common targets of MyoD and Tbx6b in gastrula and tailbud *Ciona intestinalis* embryos. Gene model names are listed at the upper left of each panel. Red arrowheads indicate expression in muscle cells. To the left are shown maps of the MyoD (green lines) and Tbx6b (black lines) ChIP data on genomic regions consisting of the exons and 5 kb upstream sequence of the genes. Each graph shows fold enrichment of ChIP DNA relative to input DNA in 125 bp non-overlapping windows (the enrichment ratio was averaged in each window, and medians of the averaged ratio for two replicates are shown). The left and right scales of each graph are for MyoD and Tbx6b, respectively. Black boxes indicate exons (shown 5′ to 3′ from left to right). Green and black arrows indicate the highest peaks in significant regions. To the right are shown the changes in the mRNA levels of target genes in embryos injected with synthetic mRNAs (red bars) or MOs (blue bars) for MyoD. Error bars represent s.e.m. of duplicate experiments.

targets. The mRNA level of one target (KH.C12.38), which was weakly expressed in muscle at the tailbud stage (Fig. 6), was significantly decreased in embryos injected with *MyoD* mRNA, whereas the mRNA level of one target (KH.C9.27), which was expressed in muscle at the gastrula stage (Fig. 6), was significantly increased in embryos injected with the *MyoD* MO. In total, the

mRNA levels of 16 targets were significantly altered by *MyoD* mRNA overexpression or gene suppression (Fig. 6). The remaining four were not significantly affected, although three of these were expressed in muscle, implying that MyoD binding makes a relatively small contribution to activating these target genes. We also found that eight of 15 Brachyury targets and seven of 12 Twist-like1 targets were significantly affected in the embryos by overexpression or knockdown of *Brachyury* or *Twist-like1*, respectively (Fig. 7 and see Fig. S12 in the supplementary material). Therefore, we estimated that more than half of TF binding makes a major contribution to the regulatory control of gene expression.

DISCUSSION

Generally, high-quality antibodies are required for ChIP experiments. In our experiments, TF-GFP fusion genes were introduced exogenously into eggs and a polyclonal antibody specific for GFP was used for all experiments. This approach successfully identified TF-DNA interactions in the *Ciona* embryos. In this regard, our approach offers a great advantage. A possible drawback is that the introduced TFs might be somewhat overexpressed, although they were expressed at the same time and in the same blastomeres as the endogenous genes, being under the control of their own promoters (Fig. 1). It is possible that some of the interactions we observed in the present study were stronger, and thus more prevalent, than would normally occur in wild-type embryos. However, as the introduced TFs did not disrupt the morphology of the embryos at the time when we collected the samples, it seems likely that any effect of overexpression was minimal. Even though we might have overestimated the binding strength, our experiments indicate that in early embryos these TFs have access to the DNA segments we identified. However, because the introduced gene is not always expressed in all cells and the expressed GFP-tagged TF would compete for binding with the endogenous TFs, sensitivity could be reduced, and therefore some weak interactions could have been missed. Nonetheless, the number of identified regions was comparable to that reported in previous ChIP studies in other animals, in which ChIP assays were performed with TF-specific antibodies and normal embryos (Jakobsen et al., 2007; Li et al., 2008; Sandmann et al., 2006; Sandmann et al., 2007; Zeitlinger et al., 2007), suggesting that the majority of binding sites were identified in our experiments.

We found more interactions than would be predicted from previous functional assays. There were some target genes that are not expressed zygotically in early embryos. However, this observation is not necessarily contradictory because the binding of TFs to cis elements may repress target gene expression. In addition, the binding of these TFs might be insufficient for regulating the targets and they might instead need to work cooperatively with other TFs that are expressed at other time points.

A large proportion of the previously identified interactions (Imai et al., 2006) were revealed to be direct in the present study, whereas a small number of interactions were suggested to be indirect. Most of the latter interactions can be explained, as shown, for example, by the regulation of *Twist-like1* (Fig. 4B) and *Brachyury* (Fig. 4C) by FoxA-a. Indirect interactions that cannot be explained based on the current knowledge are the auto-repressive interactions of *FoxD*, *Tbx6b* and *ZicL*. Since these auto-interactions were identified as significant under a milder FDR (1%), weak interactions might exist.

The present study also showed that tissue-specific genes are regulated in a different way in each of three ascidian mesodermal tissues, after key regulatory genes are expressed. In muscle, most muscle structural genes were found to be under the direct control of MyoD and/or a feed-forward loop of MyoD and Tbx6b. In the mesenchyme lineage, Twist-like1 did not seem to regulate the mesenchyme-specific non-regulatory genes directly. In the notochord, 11 of 14 notochord-specific non-regulatory genes were under the direct control of Brachyury.

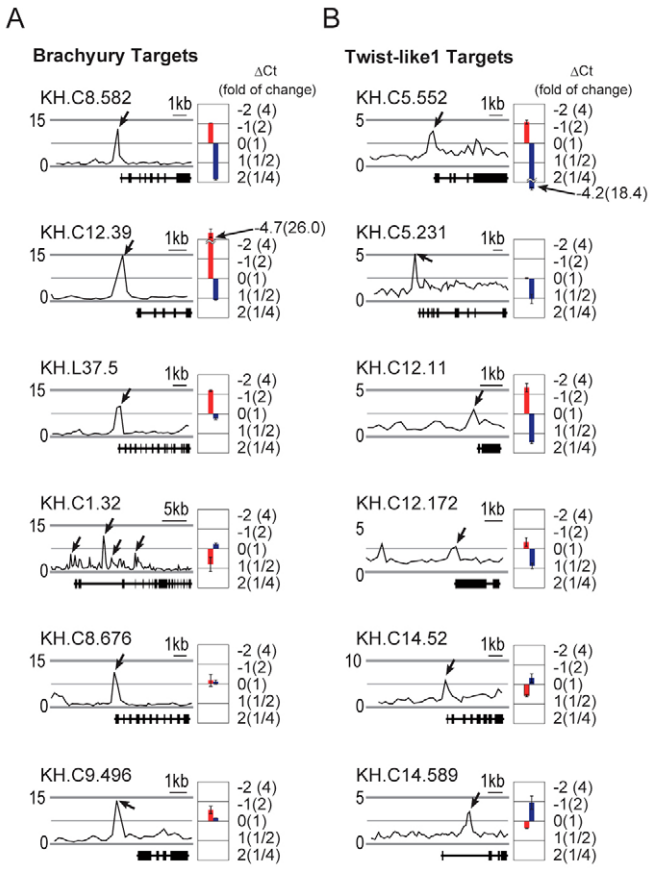

Fig. 7. Targets of Brachyury and Twist-like1. (**A**,**B**) To the left are shown maps of the Brachyury (A) and Twist-like1 (B) ChIP data on genomic regions consisting of the exons and 5 kb upstream sequence of the genes indicated in each panel. Each graph shows the fold enrichment of ChIP DNA relative to input DNA in 125 bp non-overlapping windows [the enrichment ratio was averaged in each window, and medians of the average ratio for two (Brachyury ChIP) or three (Twist-like1 ChIP) replicates are shown]. Black boxes indicate exons (shown 5′ to 3′ from left to right). Black arrows indicate the highest peaks in significant regions. To the right are shown the changes in the mRNA levels of target genes in embryos injected with synthetic mRNAs (red bars) or MOs (blue bars) for *Brachyury* (A) and *Twist-like1* (B). Error bars represent s.e.m. of duplicate experiments.

Information inherited from parental cells and signals from nearby cells are integrated at the cis-regulatory elements of target genes to control their transcription. Because changes in the cis-regulatory regions of TFs are thought to have made major contributions to animal evolution (Stern and Orgogozo, 2008), these elements should carry traces of evolutionary processes. Deciphering the logic of how these elements can be changed will lead to a better understanding of how developmental systems have evolved.

Acknowledgements

We thank all the staff members of the Field Science Center (Onagawa) of Tohoku University and the Maizuru Fisheries Research Station of Kyoto University for collecting *Ciona intestinalis*. This research was supported mainly

by Grants-in-Aid from the MEXT, Japan, to Y.S. (17687022, 20687014, 21671004) and also by gCOE program A06 of Kyoto University. A.K. was supported by a JSPS Research Fellowship (194010).

Competing interests statement
The authors declare no competing financial interests.

Supplementary material
Supplementary material for this article is available at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.046789/-/DC1

References

Bertrand, V., Hudson, C., Caillol, D., Popovici, C. and Lemaire, P. (2003). Neural tissue in ascidian embryos is induced by FGF9/16/20, acting via a combination of maternal GATA and Ets transcription factors. *Cell* **115**, 615-627.
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.
Buck, M. J., Nobel, A. B. and Lieb, J. D. (2005). ChIPOTle: a user-friendly tool for the analysis of ChIP-chip data. *Genome Biol.* **6**, R97.
Corbo, J. C., Erives, A., Di Gregorio, A., Chang, A. and Levine, M. (1997a). Dorsoventral patterning of the vertebrate neural tube is conserved in a protochordate. *Development* **124**, 2335-2344.
Corbo, J. C., Levine, M. and Zeller, R. W. (1997b). Characterization of a notochord-specific enhancer from the Brachyury promoter region of the ascidian, Ciona intestinalis. *Development* **124**, 589-602.
Corbo, J. C., Fujiwara, S., Levine, M. and Di Gregorio, A. (1998). Suppressor of hairless activates brachyury expression in the Ciona embryo. *Dev. Biol.* **203**, 358-368.
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.
Di Gregorio, A. and Levine, M. (1999). Regulation of Ci-tropomyosin-like, a Brachyury target gene in the ascidian, Ciona intestinalis. *Development* **126**, 5599-5609.
Di Gregorio, A., Corbo, J. C. and Levine, M. (2001). The regulation of forkhead/HNF-3beta expression in the Ciona embryo. *Dev. Biol.* **229**, 31-43.
Ettwiller, L., Paten, B., Ramialison, M., Birney, E. and Wittbrodt, J. (2007). Trawler: de novo regulatory motif discovery pipeline for chromatin immunoprecipitation. *Nat. Methods* **4**, 563-565.
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.
Hertz, G. Z. and Stormo, G. D. (1999). Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics* **15**, 563-577.
Hotta, K., Takahashi, H., Asakura, T., Saitoh, B., Takatori, N., Satou, Y. and Satoh, N. (2000). Characterization of Brachyury-downstream notochord genes in the Ciona intestinalis embryo. *Dev. Biol.* **224**, 69-80.
Imai, K. S., Satoh, N. and Satou, Y. (2002a). Early embryonic expression of FGF4/6/9 gene and its role in the induction of mesenchyme and notochord in Ciona savignyi embryos. *Development* **129**, 1729-1738.
Imai, K. S., Satoh, N. and Satou, Y. (2002b). An essential role of a FoxD gene in notochord induction in Ciona embryos. *Development* **129**, 3441-3453.
Imai, K. S., Satou, Y. and Satoh, N. (2002c). 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., Satoh, N. and Satou, Y. (2003). A Twist-like bHLH gene is a downstream factor of an endogenous FGF and determines mesenchymal fate in the ascidian embryos. *Development* **130**, 4461-4472.
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., Stolfi, A., Levine, M. and Satou, Y. (2009). Gene regulatory networks underlying the compartmentalization of the Ciona central nervous system. *Development* **136**, 285-293.
Inoue, T., Wang, M., Ririe, T. O., Fernandes, J. S. and Sternberg, P. W. (2005). Transcriptional network underlying *Caenorhabditis elegans* vulval development. *Proc. Natl. Acad. Sci. USA* **102**, 4972-4977.
Jakobsen, J. S., Braun, M., Astorga, J., Gustafson, E. H., Sandmann, T., Karzynski, M., Carlsson, P. and Furlong, E. E. (2007). Temporal ChIP-on-chip reveals Biniou as a universal regulator of the visceral muscle transcriptional network. *Genes Dev.* **21**, 2448-2460.
Koide, T., Hayata, T. and Cho, K. W. (2005). *Xenopus* as a model system to study transcriptional regulatory networks. *Proc. Natl. Acad. Sci. USA* **102**, 4943-4948.
Kusakabe, T., Yoshida, R., Kawakami, I., Kusakabe, R., Mochizuki, Y., Yamada, L., Shin-i, T., Kohara, Y., Satoh, N., Tsuda, M. et al. (2002). Gene expression profiles in tadpole larvae of Ciona intestinalis. *Dev. Biol.* **242**, 188-203.
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.
Levine, M. and Davidson, E. H. (2005). Gene regulatory networks for development. *Proc. Natl. Acad. Sci. USA* **102**, 4936-4942.
Li, X. Y., MacArthur, S., Bourgon, R., Nix, D., Pollard, D. A., Iyer, V. N., Hechmer, A., Simirenko, L., Stapleton, M., Luengo Hendriks, C. L. et al. (2008). Transcription factors bind thousands of active and inactive regions in the Drosophila blastoderm. *PLoS Biol.* **6**, e27.
Liu, Y. H., Jakobsen, J. S., Valentin, G., Amarantos, I., Gilmour, D. T. and Furlong, E. E. (2009). A systematic analysis of Tinman function reveals Eya and JAK-STAT signaling as essential regulators of muscle development. *Dev. Cell* **16**, 280-291.
Longabaugh, W. J., Davidson, E. H. and Bolouri, H. (2005). Computational representation of developmental genetic regulatory networks. *Dev. Biol.* **283**, 1-16.
Loose, M. and Patient, R. (2004). A genetic regulatory network for *Xenopus* mesendoderm formation. *Dev. Biol.* **271**, 467-478.
Martin-Magniette, M. L., Mary-Huard, T., Berard, C. and Robin, S. (2008). ChIPmix: mixture model of regressions for two-color ChIP-chip analysis. *Bioinformatics* **24**, i181-186.
Matys, V., Fricke, E., Geffers, R., Gossling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A. E., Kel-Margoulis, O. V. et al. (2003). TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* **31**, 374-378.
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.
Miwata, K., Chiba, T., Horii, R., Yamada, L., Kubo, A., Miyamura, D., Satoh, N. and Satou, Y. (2006). Systematic analysis of embryonic expression profiles of zinc finger genes in Ciona intestinalis. *Dev. Biol.* **292**, 546-554.
Nishida, H. (1987). Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. III. Up to the tissue restricted stage. *Dev. Biol.* **121**, 526-541.
Oda-Ishii, I., Bertrand, V., Matsuo, I., Lemaire, P. and Saiga, H. (2005). Making very similar embryos with divergent genomes: conservation of regulatory mechanisms of Otx between the ascidians Halocynthia roretzi and Ciona intestinalis. *Development* **132**, 1663-1674.
Oliveri, P., Tu, Q. and Davidson, E. H. (2008). Global regulatory logic for specification of an embryonic cell lineage. *Proc. Natl. Acad. Sci. USA* **105**, 5955-5962.
Ririe, T. O., Fernandes, J. S. and Sternberg, P. W. (2008). The *Caenorhabditis elegans* vulva: a post-embryonic gene regulatory network controlling organogenesis. *Proc. Natl. Acad. Sci. USA* **105**, 20095-20099.
Sandmann, T., Jensen, L. J., Jakobsen, J. S., Karzynski, M. M., Eichenlaub, M. P., Bork, P. and Furlong, E. E. (2006). A temporal map of transcription factor activity: mef2 directly regulates target genes at all stages of muscle development. *Dev. Cell* **10**, 797-807.
Sandmann, T., Girardot, C., Brehme, M., Tongprasit, W., Stolc, V. and Furlong, E. E. (2007). A core transcriptional network for early mesoderm development in Drosophila melanogaster. *Genes Dev.* **21**, 436-449.
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., 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., 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.
Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U. and Gaul, U. (2008). Predicting expression patterns from regulatory sequence in Drosophila segmentation. *Nature* **451**, 535-540.
Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498-2504.
Stathopoulos, A. and Levine, M. (2005). Genomic regulatory networks and animal development. *Dev. Cell* **9**, 449-462.
Stern, D. L. and Orgogozo, V. (2008). The loci of evolution: how predictable is genetic evolution? *Evolution* **62**, 2155-2177.

Takahashi, H., Hotta, K., Erives, A., Di Gregorio, A., Zeller, R. W., Levine, M. and Satoh, N. (1999). Brachyury downstream notochord differentiation in the ascidian embryo. *Genes Dev.* **13**, 1519-1523.

Tokuoka, M., Imai, K. S., Satou, Y. and Satoh, N. (2004). Three distinct lineages of mesenchymal cells in Ciona intestinalis embryos demonstrated by specific gene expression. *Dev. Biol.* **274**, 211-224.

Tokuoka, M., Satoh, N. and Satou, Y. (2005). A bHLH transcription factor gene, Twist-like 1, is essential for the formation of mesodermal tissues of Ciona juveniles. *Dev. Biol.* **288**, 387-396.

Yagi, K., Satou, Y. and Satoh, N. (2004). A zinc finger transcription factor, ZicL, is a direct activator of Brachyury in the notochord specification of Ciona intestinalis. *Development* **131**, 1279-1288.

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.

Yasuo, H. and Satoh, N. (1998). Conservation of the developmental role of Brachyury in notochord formation in a urochordate, the ascidian Halocynthia roretzi. *Dev. Biol.* **200**, 158-170.

Zeitlinger, J., Zinzen, R. P., Stark, A., Kellis, M., Zhang, H., Young, R. A. and Levine, M. (2007). Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the Drosophila embryo. *Genes Dev.* **21**, 385-390.

Zinzen, R. P., Girardot, C., Gagneur, J., Braun, M. and Furlong, E. E. (2009). Combinatorial binding predicts spatio-temporal cis-regulatory activity. *Nature* **462**, 65-70.