The gene regulatory system for specifying germ layers in early embryos of the simple chordate

Tokuoka, Miki; Maeda, Kazuki; Kobayashi, Kenji; Mochizuki, Atsushi; Satou, Yutaka

DEVELOPMENTAL BIOLOGY

The gene regulatory system for specifying germ layers in early embryos of the simple chordate

Miki Tokuoka[1], Kazuki Maeda[2], Kenji Kobayashi[1], Atsushi Mochizuki[3], Yutaka Satou[1]*

In animal embryos, gene regulatory networks control the dynamics of gene expression in cells and coordinate such dynamics among cells. In ascidian embryos, gene expression dynamics have been dissected at the single-cell resolution. Here, we revealed mathematical functions that represent the regulatory logics of all regulatory genes expressed at the 32-cell stage when the germ layers are largely specified. These functions collectively explain the entire mechanism by which gene expression dynamics are controlled coordinately in early embryos. We found that regulatory functions for genes expressed in each of the specific lineages contain a common core regulatory mechanism. Last, we showed that the expression of the regulatory genes became reproducible by calculation and controllable by experimental manipulations. Thus, these regulatory functions represent an architectural design for the germ layer specification of this chordate and provide a platform for simulations and experiments to understand the operating principles of gene regulatory networks.

INTRODUCTION

The principles by which gene regulatory networks (GRNs) control gene expression in individual cells remain incompletely understood. In particular, in multicellular organisms, GRNs in individual cells are connected with one another through cell-cell interactions and constitute a large GRN, which governs gene expression in individual cells of an organism. For this reason, a platform making it possible to simulate the operating principles of a whole system at the single-cell resolution is necessary to understand how GRN dynamics are regulated.

Early embryos of ascidians, which are invertebrate chordates, are simple and provide a unique opportunity for understanding such principles of GRN dynamics at the whole-embryo level. At the 32-cell stage, germ layers are largely specified and neural induction occurs (Fig. 1A) (*1*, *2*). A comprehensive in situ hybridization assay (*3*) has revealed that 13 genes encoding regulatory factors (transcription factors and signaling molecules) begin to be expressed zygotically in nine different patterns at the 32-cell stage (hereafter called as downstream genes). Because seven maternal transcription factors begin to regulate 14 genes between the 8- and 16-cell stages (*3*–*8*), the regulatory factors encoded by these 14 regulatory genes, together with seven maternal factors, regulate gene expression at the 32-cell stage (see Fig. 1, B and C; note that one possible upstream factor, Wnttun5, is not included in Fig. 1B). Because all regulatory genes have been examined in these comprehensive assays, we can expect that the whole regulatory system in early embryos will be clarified by analyzing these regulatory genes. However, regulatory interactions among upstream factors and downstream genes, which have been revealed by previous studies (*2*, *4*, *9*–*13*), do not provide sufficient information to reproduce GRN dynamics in every cell of 32-cell embryos.

In the present study, we reveal the regulatory logics of the 13 downstream genes that are expressed in 32-cell embryos of the ascidian (*Ciona robusta*, also called as *C. intestinalis* type A) as mathematical functions to predict and control gene expression. Empirically, in early ascidian embryos, gene regulatory mechanisms can be explained in a qualitative manner (*14*), and therefore, regulatory functions may be represented as Boolean functions. We succeeded in representing mathematical functions describing how each of the 13 downstream genes is regulated by 21 upstream factors as Boolean functions.

RESULTS

Regulatory functions that govern gene expression at the 32-cell stage

The distribution of the upstream factors shown in Fig. 1B is based on observations in previous studies (*5*–*7*, *9*, *12*, *13*, *15*–*21*) (Supplementary Text), and expression patterns of the downstream genes shown in Fig. 1C are based on in situ hybridization in a previous study (*3*). Our purpose here is to obtain Boolean functions describing how 13 downstream genes are regulated by 21 upstream factors. Regulatory functions are directly given in the form of truth tables T^n of expression patterns of 21 upstream factors (7 maternal factors and 14 zygotic factors) and each of the 13 downstream genes (Fig. 2A). However, each T^n contains 2,097,152 ($=2^{21}$) conditions, and it is practically difficult to completely fill the truth tables. To overcome this problem, we first excluded Ets1/2, T cell factor 7 (Tcf7), GATA binding protein a (Gata.a), and posterior end mark 1 (Pem1) from our analysis because we did not need to consider these upstream factors explicitly; Ets1/2 acts as an effector of the mitogen-activated protein kinase (MAPK) pathway, which is regulated positively by fibroblast growth factor 9/16/20 (Fgf9/16/20) signaling and CA-Raf (constitutively active Raf) and negatively by Ephrin A-d (Efna.d) signaling (*4*, *15*, *22*); Tcf7 acts as a positive regulator with nuclear β-catenin; Gata.a acts as a positive regulator in cells in which the nuclear β-catenin is not present (*7*, *13*); and Pem1 always represses transcription by suppressing the function of RNA polymerase II (*16*, *17*) (we did not consider the most posterior germline blastomeres where Pem1 is localized and transcriptionally silent, either). Wnttun5, a tunicate-specific Wnt ligand, was also excluded; this signaling ligand controls the orientation of cell divisions, and it is not likely that this ligand directly regulates gene expression (*18*). Consequently, the number of upstream factors was reduced to 16.

[1]Department of Zoology, Graduate School of Science, Kyoto University, Sakyo, Kyoto 606-8502, Japan. [2]Faculty of Informatics, University of Fukuchiyama, 3370 Hori, Fukuchiyama, Kyoto 620-0886, Japan. [3]Institute for Frontier Life and Medical Sciences, Kyoto University, Sakyo, Kyoto 606-8507, Japan.
*Corresponding author. Email: yutaka@ascidian.zool.kyoto-u.ac.jp

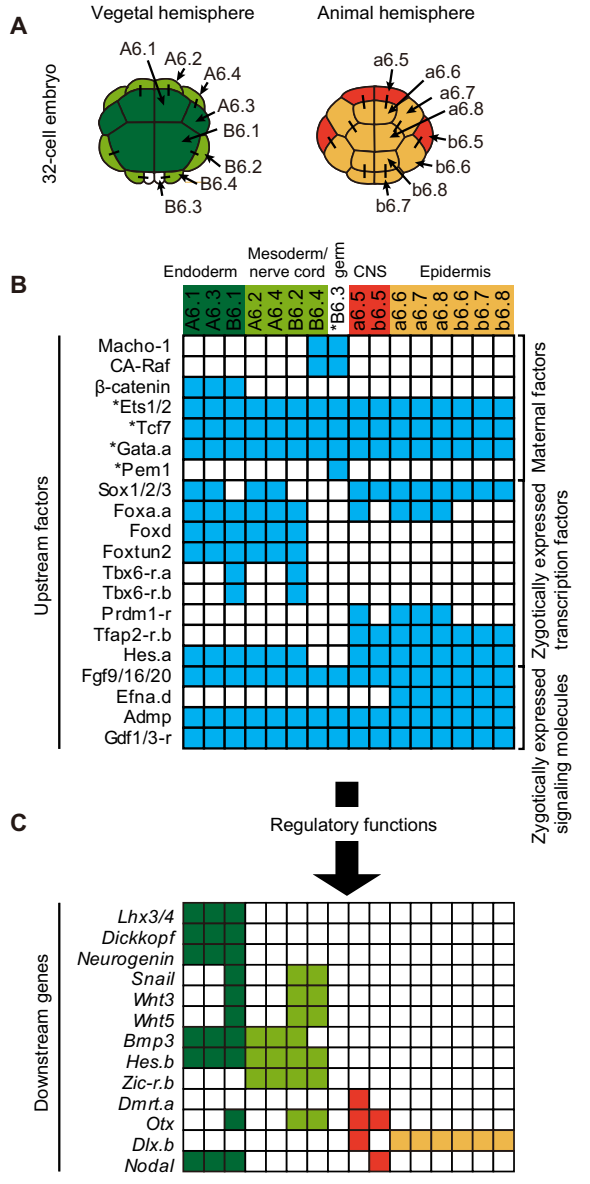

Fig. 1. Gene expression patterns in early *Ciona* embryos. (**A**) Schematics of a bisymmetrical 32-cell embryo. Cells are largely specified to two ectodermal fates, vegetal marginal (largely mesodermal) fate and vegetal central (largely endodermal) fate, which are shown by different colors. (**B**) Possible upstream factors of genes expressed in 32-cell embryos. Cells where each upstream factor is expected to act are shown in cyan. We do not explicitly consider Ets1/2, Tcf7, Gata.a, Pem1, or the most posterior germline cells where Pem1 is localized (asterisks; see main text and Supplementary Text for details). Nine transcription factors encoded by genes expressed at the 16-cell stage are expected to act at the 32-cell stage in daughters of cells with their mRNAs. In addition, five signaling ligands act in 32-cell embryos. Fgf9/16/20 and Efna.d regulate the MAPK pathway positively and negatively, respectively. Cells in which they act have been determined by staining of doubly phosphorylated extracellular signal–regulated kinase in wild-type and *Efna.d* morphant embryos (*18*). Cells in which anti-dorsalizing morphogenetic protein (Admp) and growth differentiation factor 1/3-related (Gdf1/3-r) act have been inferred from a combination of experiments and mathematical analysis (*19*). Note that Efna.d is a cell membrane–bound protein transmitting the signal in a contact-dependent manner and that all cells are in direct contact with cells expressing Fgf9/16/20 and Gdf1/3-r (*21*). See main text and Supplementary Text for details. (**C**) Expression of downstream genes that initiate expression at the 32-cell stage. Cells where each gene is expressed are colored.

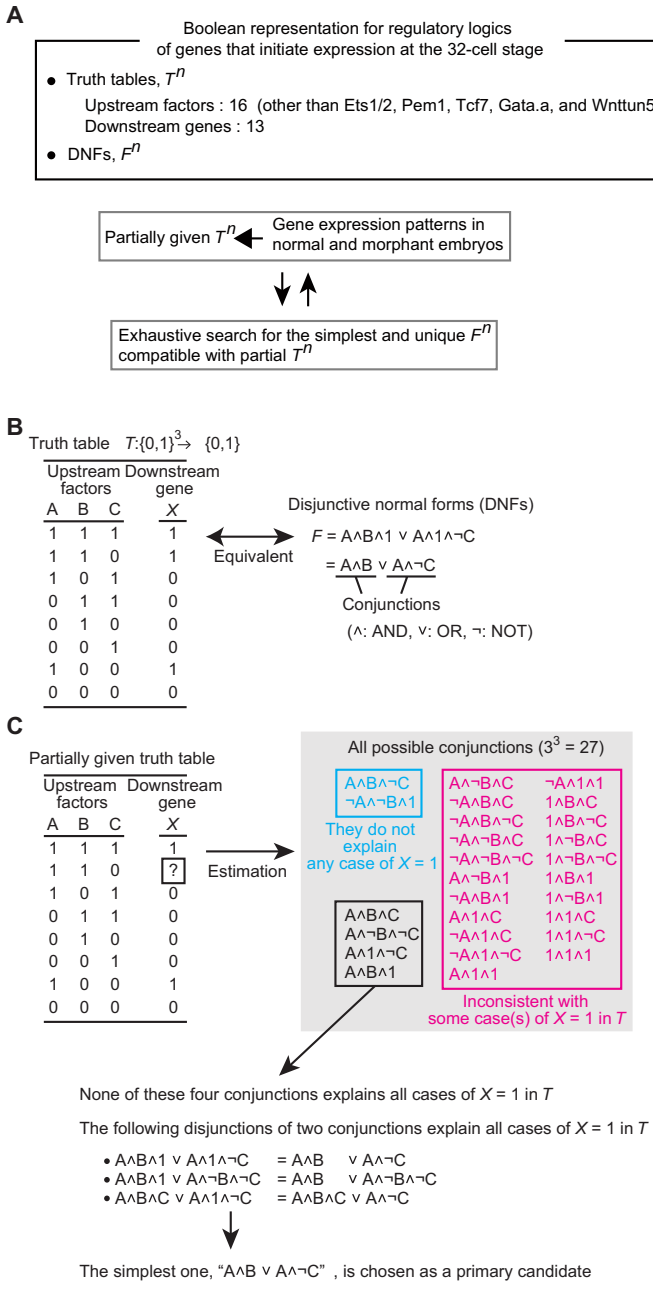

Fig. 2. Boolean representation of regulatory logics of genes that initiate expression at the 32-cell stage. (**A**) From partial truth tables T^n with missing values, disjunctive normal forms (DNFs) are inferred on the basis of experiments. We repeated experiments, in which one or more upstream factors were down-regulated, until an exhaustive search successfully identified a unique candidate DNF. (**B** and **C**) Intuitive explanation of the method to infer DNFs. In this hypothetical system, three upstream factors, A, B, and C, are expressed and possibly regulate the gene *X*. (B) Given a complete truth table, the corresponding regulatory function in the DNF can be determined. Conversely, given a regulatory function in the DNF, the corresponding truth table can be determined uniquely. (C) Estimation of DNFs corresponding to a partially given truth table. First, among all 27 possible conjunctions, conjunctions inconsistent with one or more cases of $X = 1$ are excluded (magenta). Second, conjunctions that do not explain any case of $X = 1$ in the partial truth table are also excluded (cyan). If any single conjunction in the remaining set does not fully explain the partial truth table, then disjunctions of multiple conjunctions are examined. In this case, three DNFs fully explain the truth table. We consider the simplest one as a primary candidate.

Then, we attempted to find disjunctive normal forms (DNFs) compatible with partial T^n, describing gene expression patterns in normal embryos (see Fig. 1) and experimental embryos. The DNF is one of the canonical Boolean logic forms represented as the sum (disjunction; OR; ∨) of products (conjunction; AND; ∧) like A∧B∨A∧¬C (¬: NOT). Namely, in this example, "A∧B" and "A∧¬C" are called conjunctions, and "A," "B," and "C" are called literals symbolizing the effects of upstream factors on the focal downstream gene. The DNF was chosen because this form is more easily interpreted from a biological viewpoint; each conjunction probably represents one regulatory module that requires the simultaneous binding of multiple upstream factors.

There are three possible effects of an upstream factor on a downstream gene; upstream factors up-regulate, down-regulate, or do not regulate a downstream gene. Because of 16 upstream factors, there are 43,046,721 ($=3^{16}$) possible conjunctions. In the present study, we tried to find minimal DNF F^n compatible with partial T^n, which represents gene expression patterns under normal conditions and a limited number of experimental conditions. First, we removed conjunctions incompatible with T^n from the first candidates. Conjunctions that do not explain any expression of the focal downstream gene are also removed. Then, we tried to find combinations of the remaining conjunctions that fully explain the expression pattern of the focal gene in normal and experimental embryos. In most cases, multiple DNFs are compatible with partial T^n. We consider DNFs with the smallest number of conjunctions and the smallest number of literals as primary candidates. An intuitive explanation for this method is shown in Fig. 2 and the details are described in Materials and Methods. The computer code that we used is available at https://github.com/kmaed/mindnf.

When multiple candidates were obtained after an exhaustive search, we performed additional experiments (Fig. 2A). In this way, we attempted to determine the "simplest" DNF for each target gene from partial T^n (table S1), which represents the expression patterns under normal conditions and 4 to 17 experimental conditions (Fig. 3 and fig. S1), and succeeded in determining the F^n for 12 targets basically through knockdown experiments (Table 1).

In the simplest case of *Bmp3*, knockdown of either *Foxa.a* or *Foxd* using specific antisense morpholino oligonucleotides (MOs) resulted in complete loss of *Bmp3* expression, which is normally found in all vegetal cells except for the germline and its sister cells (B6.3 and B6.4) (Figs. 1C and 3, A to C). Meanwhile, knockdown of *Fgf9/16/20* or *Macho-1* or treatment with U0126, which is an inhibitor of the MAPK pathway, did not affect *Bmp3* expression (fig. S1A). F^{Bmp3} was therefore formulated as Foxa.a∧Foxd.

F^{Bmp3} is satisfied in the vegetal cells except B6.3 and B6.4 of normal embryos, because only these vegetal cells express Foxa.a and Foxd simultaneously (Fig. 1B). Thus, F^{Bmp3} and the distribution patterns of the upstream factors explain the specific expression pattern of *Bmp3*. Regulatory functions for the other downstream genes were similarly determined (Supplementary Text).

The only exception was *Nodal*; two candidates explained the expression of *Nodal* after 13 experiments, as shown in fig. S1M. These two models for *Nodal* consisted of four conjunctions, and three of them were shared between these two models. Because the unshared conjunctions represented expression observed only under experimental conditions (Table 1), we did not determine which of these closely resembling DNFs was correct.

Thus, we succeeded in recapitulating all gene expression patterns of the 32-cell embryo of this simple chordate using the mathematical functions, which are represented in the Boolean formula. As shown in the next sections, these functions allow us to overview the embryo-wide genetic program of germ layer formation and to control embryo-wide gene expression patterns.

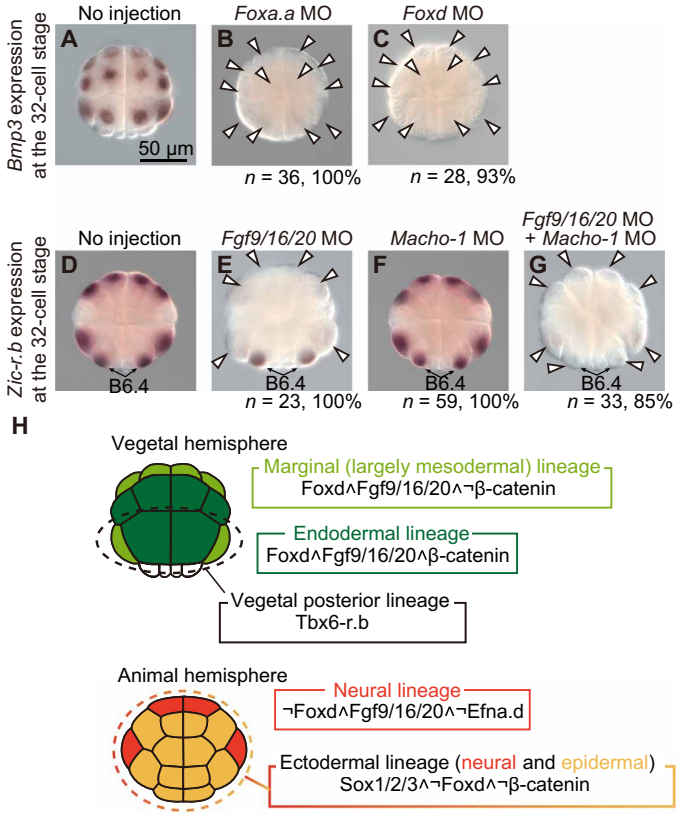

Fig. 3. Regulatory logics for gene expression in 32-cell embryo. (**A** to **G**) Expression of *Bmp3* (A to C) and *Zic-r.b* (D to G) in normal embryos (A and D) and embryos injected with MOs indicated above the photographs (B, C, and E to G). Number of embryos examined and the proportion of embryos that each panel represents are shown below the panels. Arrowheads indicate loss of expression. (**H**) Regulatory logics that specify five different lineages at the 32-cell stage. Note that the regulatory mechanisms in B6.4 do not share the mechanisms shown here (see the text).

Regulatory mechanisms in the sister cells of the germline cells are distinct from those in other somatic cells

We noticed that the gene expression in B6.4 is regulated differently from that in the other somatic lineages. This is particularly interesting because the parental cells of B6.4 cells have a germ cell fate, and their somatic daughters, B6.4 cells, initiate zygotic transcription at the 32-cell stage (the other somatic cells initiate zygotic transcription at the 16-cell stage).

At the 32-cell stage, *Zic-r.b* is expressed in four pairs of marginal vegetal cells, which are largely mesodermal (Fig. 3D). The regulation in the anterior three pairs was represented by Foxa.a∧Foxd∧Fgf9/16/20∧¬β-catenin, which is consistent with earlier studies (*5*, *9*, *13*). Meanwhile, *Zic-r.b* was regulated by two different mechanisms redundantly in B6.4 (Table 1); there are a *Macho-1*–dependent mechanism and an *Fgf9/16/20*-dependent mechanism (Fig. 3, E to G, and Supplementary Text). Similarly, *Otx* was also activated in B6.4 by two independent mechanisms (Table 1).

Snail is also activated by different mechanisms between the B6.4 lineage and other lineages (*15*). The DNFs (F^{Wnt3} and F^{Wnt5}) for *Wnt3*

Table 1. Representation of regulatory logics in DNFs for expression of genes that initiate expression at the 32-cell stage.

Downstream genes	DNFs*	Cells where gene expression is activated in normal embryos
Lhx3/4	Foxd∧Fgf9/16/20∧β-catenin	A6.1/A6.3/B6.1
Neurogenin	Foxd∧Fgf9/16/20∧β-catenin	A6.1/A6.3/B6.1
Dickkopf	Foxd∧Fgf9/16/20∧β-catenin	A6.1/A6.3/B6.1
Snail	CA-Raf∧Macho-1∨	B6.4
	Tbx6-r.b	B6.1/B6.2
Wnt3	CA-Raf∧Macho-1∨	B6.4
	Tbx6-r.b	B6.1/B6.2
Wnt5	CA-Raf∧Macho-1∨	B6.4
	Tbx6-r.b	B6.1/B6.2
Bmp3	Foxa.a∧Foxd	A6.1-4/B6.1/B6.2
Hes.b	Foxd∧Fgf9/16/20∨	A6.1-4/B6.1/B6.2
	¬Sox1/2/3∧¬Hes.a ∧Fgf9/16/20∧¬Efna.d∧¬Prdm1-r†	B6.4
Zic-r.b	Foxa.a∧Foxd∧Fgf9/16/20 ∧¬β-catenin∨	A6.2/A6.4/B6.2
	¬β-catenin∧Macho-1∨	B6.4
	¬Sox1/2/3∧¬Hes.a ∧Fgf9/16/20∧¬β-catenin∨	B6.4
	Foxa.a∧¬Hes.a∧Fgf9/16/20 ∧¬β-catenin	None‡
Dmrt.a	Sox1/2/3∧Foxa.a ∧¬Foxd∧Fgf9/16/20∧¬Efna.d ∧¬β-catenin	a6.5
Otx	¬β-catenin∧Macho-1∨	B6.4
	Tbx6-r.a∧Fgf9/16/20∨	B6.1/B6.2
	Tbx6-r.b∧Fgf9/16/20∨	B6.1/B6.2
	¬Foxd∧Fgf9/16/20∧¬Efna.d∨	a6.5/b6.5/B6.4
	Fgf9/16/20∧¬Gdf1/3-r∧¬Admp	None‡
Dlx.b	Sox1/2/3∧Foxa.a∧¬Foxd∧¬β-catenin∨	a6.5-8
	Sox1/2/3∧Efna.d∨	a6.6-8/b6.6-8
	Sox1/2/3∧¬Foxd∧¬Fgf9/16/20	None‡
Nodal	Foxa.a∧Fgf9/16/20∧β-catenin∨	A6.1/A6.3/B6.1
	Tbx6-r.b∧β-catenin∨	B6.1
	Sox1/2/3∧¬Foxa.a ∧¬Foxd∧Fgf9/16/20∧¬Efna.d ∧¬β-catenin∨	b6.5
	Fgf9/16/20∧¬Gdf1/3-r ∧¬Admp∧¬Prdm1-r/¬Foxa.a§	None‡

*Note that β-catenin acts with Tcf7, which is present in all cells, for activating targets, and that β-catenin represses Gata.a (*7*, *13*). Because Gata.a is present in all cells, ¬β-catenin means that Gata.a becomes active. †This conjunction was determined by not only knockdown experiments but also overexpression experiments (fig. S2). ‡These conjunctions represent expression in experimental conditions only. §We were not able to determine this conjunction uniquely. The expression patterns of *Prdm1-r* and *Foxa.a* indicate that either of them or both are involved. However, this conjunction represents expression in experimental conditions only but not expression in normal 32-cell embryos.

and *Wnt5*, which are expressed in the same pattern as *Snail*, were identical to F^{Snail} (Table 1 and Supplementary Text). Similarly, *Hes.b* is activated in B6.4, and its regulatory mechanism was different from those that activate *Zic-r.b*, *Otx*, and *Snail*/*Wnt3*/*Wnt5* in B6.4 and *Hes.b* in the other cells (Table 1). Thus, the regulatory mechanisms of these genes differ between B6.4 and other cells, and the regulatory mechanisms also differ among these genes. In other words, ascidian embryos may have evolved specific genetic programs for gene expression in B6.4, probably because the initiation of zygotic transcription occurs one step later in this lineage than in the other somatic lineages.

Core regulatory logics specifying cell fates including three germ layer fates

In contrast to the case in B6.4, we noticed that there are regulatory logics shared by specific cell groups. One of the mechanisms that activate *Otx* in B6.4, "¬Foxd∧Fgf9/16/20∧¬Efna.d," is also used for the activation of *Otx* in the neural precursors (a6.5 and b6.5). While *Otx* is expressed in the anterior and posterior neural precursors, *Dmrt.a* is specifically expressed in the anterior neural precursors and *Nodal* is expressed in the posterior neural precursors. As shown in Table 1, the responsible conjunctions in their DNFs contained ¬Foxd∧Fgf9/16/20∧¬Efna.d in common. This observation suggests that ¬Foxd∧Fgf9/16/20∧¬Efna.d represents the core regulatory logic for neural induction in ascidian embryos (Fig. 3H).

In contrast, the regulatory logic "Sox1/2/3∧¬Foxd∧¬β-catenin," which was shared between *Dmrt.a* and *Nodal*, was also found in the regulatory function of *Dlx.b*, which is expressed in the epidermal and anterior neural precursors. Therefore, this may represent the regulatory mechanism for specifying ectodermal fate, although F^{Otx} does not contain this logic (Fig. 3H).

The neural and ectodermal core logics included negative regulation by Foxd. Conversely, most genes expressed in the vegetal hemisphere were positively regulated by Foxd (Table 1), as previously indicated (*23*). Similarly, our results also confirmed a previously proposed model that β-catenin plays a key role in discriminating medial vegetal (largely endodermal) cells from marginal vegetal (largely mesodermal) ones (*5*, *9*, *13*) (Table 1). In addition, Fgf9/16/20 is necessary for inducing both groups of genes. Thus, Foxd, β-catenin, and Fgf9/16/20 act together and constitute core mechanisms to specify endodermal and mesodermal fates, Foxd∧Fgf9/16/20∧β-catenin for endoderm and Foxd∧Fgf9/16/20∧¬β-catenin for mesoderm (Fig. 3H).

There was an additional regulatory system that drove expression in the posterior vegetal cells. This system depends on *Tbx6-r.b* (Fig. 3H), which is activated by Macho-1 and β-catenin at the 16-cell stage (*7*, *20*). *Snail*, *Wnt3*, and *Wnt5* were activated by Tbx6-r.b, and *Otx* was activated by Tbx6-r.b and its paralog Tbx6-r.a in the posterior vegetal cells. The above observations suggest that there are five core regulatory logics and that modified versions of them are used for regulating the downstream genes.

Controlling gene expression patterns using the regulatory functions

We next attempted to control gene expression using F^n. For this purpose, we used 16-cell stage embryos, because zygotic gene expression scarcely occurs before the 16-cell stage (*6*), and therefore, we can examine direct effects of upstream factors.

Lhx3/4 is not expressed in normal 16-cell embryos (*3*), which is consistent with the prediction from $F^{Lhx3/4}$ and expression pattern of its upstream factors (Fig. 4A). $F^{Lhx3/4}$ predicts that *Lhx3/4* will be

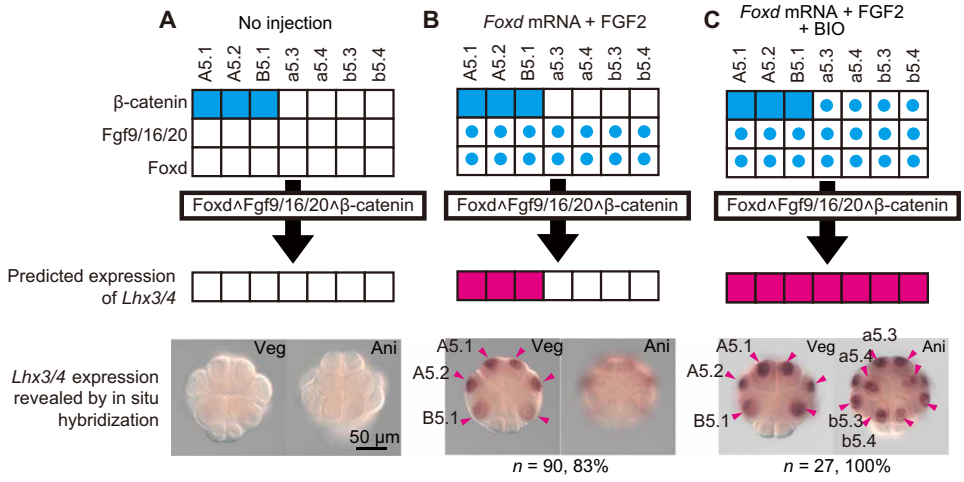

Fig. 4. Regulation of *Lhx3/4* at the 16-cell stage. (**A**) While β-catenin is present in three pairs of vegetal cells (cyan boxes in the top), neither Fgf9/16/20 nor Foxd is expected to act in normal 16-cell embryos (white boxes). Therefore, $F^{Lhx3/4}$ predicts that *Lhx3/4* is not expressed (white boxes in the center) and no expression is detected by in situ hybridization (bottom). (**B**) If embryos injected with *Foxd* mRNA (0.75 pg) are treated with FGF2, then Foxd and Fgf signaling are expected to act in all cells (cyan dots), and $F^{Lhx3/4}$ predicts that *Lhx3/4* is expressed in the vegetal hemisphere where nuclear β-catenin is present (magenta boxes). In situ hybridization shows that this is the case. (**C**) If embryos are additionally treated with BIO, then *Lhx3/4* is predicted to be expressed in all cells except the most posterior germline pair, and in situ hybridization revealed that this was the case.

expressed in the vegetal hemisphere even at the 16-cell stage with Foxd protein and Fgf signaling (Fig. 4B). To test this prediction, we injected a synthetic *Foxd* mRNA into unfertilized eggs and allowed them to develop in seawater containing recombinant FGF2 to mimic Fgf9/16/20 signaling. In such experimental embryos, *Lhx3/4* was expressed in three vegetal cell pairs at the 16-cell stage as predicted (Fig. 4B). We further added BIO, a glycogen synthase kinase 3 inhibitor that upregulates β-catenin activity (*5*, *9*), at the late 8-cell stage. $F^{Lhx3/4}$ predicted that *Lhx3/4* would be expressed in the animal and vegetal hemispheres, and *Lhx3/4* was expressed as predicted (Fig. 4C).

While *Bmp3* is not expressed in normal embryos, overexpression of *Foxa.a* and *Foxd* induced *Bmp3* expression in all cells except the germline cells of 16-cell embryos (fig. S3A), which demonstrates that *Foxa.a*∧*Foxd* represents a condition sufficient for inducing the expression of *Bmp3* and that *Bmp3* expression is controllable with F^{Bmp3}. Note that *Foxa.a* and *Foxd* are expressed under the control of nuclear β-catenin at the 16-cell stage. However, as shown in fig. S3A, overexpression of *Foxa.a* and *Foxd* induced *Bmp3* expression in the animal hemisphere, where nuclear β-catenin is not present. This observation supports the notion that β-catenin is not directly involved in the expression of *Bmp3* and is consistent with the regulatory function that does not contain β-catenin. Similarly, we succeeded in controlling the expression of *Snail*, *Bmp3*, *Hes.b*, *Zic-r.b*, *Dmrt.a*, *Otx*, and *Dlx.b* (Supplementary Text and fig. S3, B to G). Although we did not succeed in completely controlling *Nodal* expression at the 16-cell stage, we succeeded in controlling the expression at the 32-cell stage (Supplementary Text and fig. S3H). Predictions were implemented in our website, http://ghost.zool.kyoto-u.ac.jp/sim32/index.html; it predicts how changes in activities of the upstream factors or regulatory functions affect gene expression patterns.

Evolutionary plasticity in a regulatory function

A previous study (*9*), which probably used a closely related *Ciona* species (*C. intestinalis* or *C. intestinalis* type B) reproductively isolated from the species that we used (*C. robusta* or *C. intestinalis* type A)

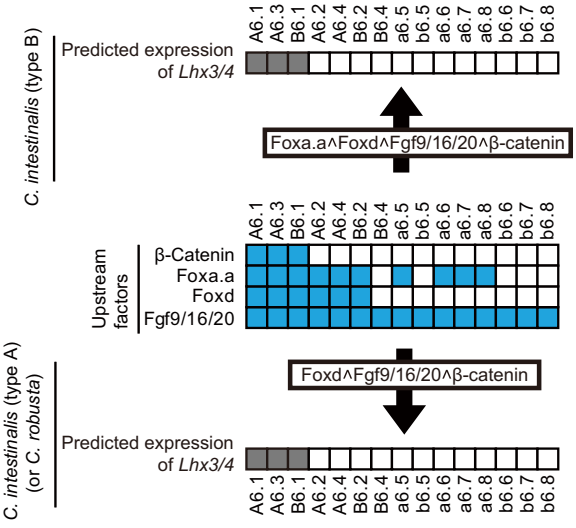

Fig. 5. Evolutionary plasticity of the regulatory mechanism for *Lhx3/4* expression. The present study revealed that $F^{Lhx3/4}$ is Foxd∧Fgf9/16/20∧β-catenin in an ascidian, *C. intestinalis* type A (or *C. robusta*). However, in an ascidian from a different population (*C. intestinalis* type B or simply called *C. intestinalis*), *Foxa.a* is necessary for *Lhx3/4* expression (*9*); therefore, the mechanism is likely to be slightly different. However, even under the assumption that $F^{Lhx3/4}$ is Foxa.a∧Foxd∧Fgf9/16/20∧β-catenin, the expression pattern of *Lhx3/4* is unchanged in normal embryos.

judging from the sampling location of the animals (*24*, *25*), showed that Foxa.a also participates in the regulation of *Lhx3/4*; therefore, $F^{Lhx3/4}$ in this species may be Foxa.a∧Foxd∧Fgf9/16/20∧β-catenin but not Foxd∧Fgf9/16/20∧β-catenin. We calculated a gene expression pattern using the regulatory function Foxa.a∧Foxd∧Fgf9/16/20∧β-catenin and found that *Lhx3/4* was predicted to be expressed in the same cells under normal conditions even with this regulatory function (Fig. 5). Because *Lhx3/4* is a key regulatory gene for specifying the endoderm and adult heart precursors (*9*, *26*, *27*), strong evolutionary

pressure is expected to maintain the expression pattern of *Lhx3/4* between these two species. The above observation may therefore represent the evolutionary plasticity of gene regulation.

DISCUSSION

Our study mathematically formulated the regulatory system for gene expression at the stage when the germ layers are largely specified. This formulation revealed that there are core regulatory mechanisms, each of which is commonly found in F^n for genes expressed in a specific lineage. These may represent ancient regulatory mechanisms, from which individual mechanisms have evolved. It is also possible that these regulatory mechanisms may have evolved independently and convergently; because of the limited number of upstream factors, similar regulatory mechanisms may have evolved.

On the other hand, the regulatory mechanisms for the expression in B6.4 were different from one another. The B6.4 pair is peculiar among cells of the 32-cell embryo, because their parental cells have germ cell fate and are transcriptionally inactive at the 16-cell stage. The regulatory mechanisms used in this B6.4 lineage may be required for developmental programs to proceed coordinately with those in the other somatic cells, in which transcription begins one step earlier at the 16-cell stage. This observation may suggest that each of these mechanisms, which regulate gene expression in the sister cells of the germline cells, has evolved independently. Because these latter somatic lineages have five core regulatory logics, shown in Fig. 3H, in common, this germline sister lineage (B6.4) contrasts with the other somatic lineages.

Our study succeeded in representing regulatory mechanisms for downstream genes as mathematical functions. The only exception is that the regulatory function for *Nodal*. F^{Nodal} contains three conjunctions that represent expression in normal embryos. It is possible that each of these conjunctions does not represent sufficient conditions. However, it is hard to imagine that all conjunctions lack upstream factors, despite extensive theoretical and experimental analyses. In addition, we succeeded in controlling *Nodal* expression at the 32-cell stage with the regulatory function we determined. Therefore, it is more likely that a higher order of transcriptional regulation, including epigenetic regulation, is involved in *Nodal* regulation at the 16-cell stage.

In the present study, each regulatory function was represented in the DNF, which is represented by disjunctions of conjunctions. The conjunction that directs *Zic-r.b* expression in A6.2, A6.4, and B6.2 is represented by Foxa.a∧Foxd∧Fgf9/16/20∧¬β-catenin. The upstream regulatory region of *Zic-r.b* contains a binding site for Ets1/2, which is an effector of Fgf signaling, four Fox sites, and four binding sites for Gata.a, which becomes active in the absence of nuclear β-catenin (*13*, *28*). On the other hand, an earlier study identified a cis-regulatory element that can drive *Otx* in a6.5 (*4*). The conjunction that directs *Otx* expression in a6.5 is represented by ¬Foxd∧Fgf9/16/20∧¬Efna.d, and this cis-element contains binding sites for Ets1/2 (an effector of the signaling pathway that is positively regulated by Fgf9/16/20 and negatively regulated by Efna.d) but not for clear Foxd binding sites. In other words, although the conjunction means that Foxd represses *Otx* expression, no Foxd binding sites are found in this cis-element. However, because ectopic expression in the vegetal cells, where Foxd is expressed, is driven with a construct that contains this cis-element only (*4*), this cis-element may lack Foxd binding sites to suppress ectopic expression. This element also contains Gata.a binding sites. Because Gata.a acts as an effector of the Fgf signaling pathway, in this case (*4*), Gata.a is not included in the conjunction; note that Ets1/2 is not included with the same reason. Thus, these earlier observations are consistent with the idea that each conjunction of the DNFs may represent one regulatory module, although this notion needs to be assessed more extensively in future studies.

Our finding that regulatory mechanisms were represented as Boolean functions indicates that qualitative, but not quantitative, controls are important for regulation in early embryos of this chordate. This property may not be specific to ascidian embryos, as the endomesodermal GRN in early sea urchin embryos is represented by Boolean functions (*29*). This property may be revealed in these two model animals because GRN dynamics have been analyzed extensively. If so, then it is possible that GRN dynamics are similarly regulated in early embryos of other animals. Alternatively, it is also possible that this property is related to features specific to ascidian and sea urchin embryos. These features may include rapid development, relatively small numbers of cells, and highly reproducible embryonic structures (*30*).

The present study made the embryo-wide gene expression patterns predictable at the single-cell resolution and therefore provides a platform for simulations and experiments to understand the operating principles and dynamics of GRNs. Our study made GRN dynamics in early ascidian embryos reproducible by calculation. In experiments, it is more difficult to change regulatory functions than to change expression patterns of upstream factors. However, we can now easily examine in computational simulations how a specific change in the regulatory function of a gene affects its gene expression pattern, as in the case of *Lhx3/4* in two closely related *Ciona* lineages. In this way, our study provides a platform for simulations and experiments to understand the operating principles of GRNs.

MATERIALS AND METHODS

Animals and cDNAs

Adult *C. intestinalis* (type A or *C. robusta*) were obtained from the National BioResource Project for *C. intestinalis* in Japan. Complementary DNA (cDNA) clones were obtained from our expressed sequence tag clone collection (*31*). Identifiers (*32*–*34*) for genes examined in the present study are shown in table S2.

Ciona is excluded from legislation regulating scientific research on animals in Japan. Although there is no scientific evidence that this animal can experience pain, discomfort, or stress, we made our best efforts to minimize the number of animals used for experiments and to minimize potential harm that animals might experience when we obtained the eggs and sperm from them.

Gene knockdown and overexpression assays

Sequences of the MOs (Gene Tools LLC) are as follows: *Sox1/2/3*, 5′-CAGTTTAATGACGTGTGAGACTTTA-3′ (*35*); *Foxa.a*, 5′-ATCCGATTTCAAAAGCTTTCTCAGA-3′ (*11*, *36*); *Foxd*, 5′-GCACACAACACTGCACTGTCATCAT-3′ (*3*, *11*, *36*); *Tbx6-r.b*, 5′-TTGAGCCTCTCACGTCGCCAT-3′ and 5′-TTACAATTT-CCTCTCTCTTTCGATT-3′ (a mixture of these two oligonucleotides was used for simultaneous knockdown of paralogs) (*11*, *37*); *Tfap2-r.b*, 5′-CGGACAGAATTCGAATATCACTCAT-3′ (*35*); *Hes.a*, 5′-TTC-TTCGTTCAACAGGCATGATTGT-3′ (*13*, *38*, *39*); *Fgf9/16/20*, 5′-CATAGACATTTTCAGTATGGAAGGC-3′ (*3*, *11*, *15*, *18*, *19*, *40*); *Efna.d*, 5′-TTGAGTTGCCATTCTTCGTTTTAAT-3′ (*11*, *18*, *19*); *Gdf1/3-r*, 5′-CATCTTTAACCCAACACTTTCAACG-3′ (*18*, *19*); *Admp*,

5′-TATCGTGTAGTTTGCTTTCTATATA-3′ (*11*, *18*, *19*, *41*); β-catenin, 5′-CTGTTCATCATCATTTCAGCCATGC-3′ (*3*, *7*); *Macho-1*, 5′-ATCCCATCGTACCAGTAAAGGCCAT-3′ (*7*, *20*); and *Prdm1-r*, 5′-CGTAACTTTCGCGGTGATTCCTCAT-3′ and 5′-GTCTGAACACACATGATTCCGACAT-3′ (a mixture of these two oligonucleotides was used for simultaneous knockdown of paralogs) (*38*). The above MOs that block translation were used previously, and therefore, additional specificity tests were not performed in the present study. Two MOs for *Tbx6-r.a* that block splicing were designed in the present study: 5′-TTAATCTCATTTCTTACCTCCCTGC-3′ and 5′-ACAAAATACGCCACCAACCTGGAAT-3′. We confirmed that these MOs effectively block splicing by reverse transcription followed by polymerase chain reaction (PCR) (fig. S4AB), in which reverse transcription was performed with the oligo(dT) primer and PCR was performed with the following primers: 5′-TGCACCGCTGCTTGAACA-3′ and 5′-TTCTACCCGGTGATGGACTATCA-3′. Amplified fragments were sequenced to confirm that these fragments were derived from abnormal transcripts (fig. S4C). As shown in fig. S1K, while *Otx* expression was not affected by the knockdown of *Tbx6-r.b* alone, it was affected by coinjection of the *Tbx6-r.b* with either of two MOs for *Tbx6-r.a*. Because both of these *Tbx6-r.a* MOs yielded the same phenotype regarding *Otx* expression (*Otx* is the only gene that was affected by these MOs), these MOs are suggested to act specifically. The MOs were introduced into unfertilized eggs by microinjection. The MOs for *Hes.a* and *Sox1/2/3* were injected at a concentration of 1 mM. Two *Tbx6-r.b* MOs were coinjected to suppress all paralogs (0.25 mM each). All other MOs were injected at a concentration of 0.5 mM.

Synthetic transcripts were prepared from cDNA cloned into the pBluescript RN3 vector (*42*) using the mMESSAGE mMACHINE T3 kit (Thermo Fisher Scientific) and injected into unfertilized eggs. BIO (GSK-3 Inhibitor IX; Merck, no. 361550), human recombinant basic FGF (FGF2; Merck, no. 662005), and U0126 (MAPK kinase inhibitor; Sigma-Aldrich, no. F0291) were added to seawater at concentrations of 2.5 mM, 10 ng/ml, and 2 μM, respectively.

Whole-mount in situ hybridization

In situ hybridization was performed basically following the protocol described previously (*43*). Embryos were fixed in 4% paraformaldehyde in 0.1 M MOPS buffer (pH 7.5) and 0.5 M NaCl at 4°C for over 16 hours. After washing with phosphate-buffered saline containing 0.1% Tween 20 (PBST), the embryos were treated with proteinase K (2 μg/ml) in PBST for 30 min at 37°C and then washed with PBST. The embryos were again fixed with 4% paraformaldehyde in PBST for 1 hour at room temperature, washed with PBST, and immersed in hybridization buffer for 1 hour at 55°C. Then, the hybridization buffer was replaced with fresh hybridization buffer containing a digoxigenin-labeled probe, and the embryos were incubated at 55°C for at least 18 hours. The hybridization buffer contained 50% formamide, 5× SSC, yeast tRNA (100 μg/ml), 5× Denhart's solution, and 1% SDS. After hybridization, the embryos were washed twice at 55°C for 15 min in 50% formamide, 5× SSC, and 1% SDS and twice at 55°C for 15 min in 50% formamide, 2× SSC, and 1% SDS. The embryos were treated for 30 min at 37°C with ribonuclease A (20 μg/ml) in 2× SSC and 0.1% Tween 20. The embryos were further washed twice at 55°C for 15 min in 0.2× SSC and 0.1% Tween 20, and the washing solution was replaced with PBST. The embryos were blocked at room temperature for 30 min with 0.5% blocking reagent (Roche) dissolved in 100 mM tris-HCl (pH 7.5) and 150 mM NaCl and then exposed to alkaline phosphatase–conjugated antidigoxigenin antibody (1:2000 dilution; Roche, no. 11093274910) at 4°C overnight. The embryos were washed with PBST and then with detection buffer [100 mM NaCl, 50 mM $MgCl_2$, and 100 mM tris buffer (pH 8.0)] before signal detection with nitro blue tetrazolium and bromochloroindolyl phosphate.

Boolean representation of regulatory functions

In the 32-cell ascidian embryo, 21 upstream factors are expressed (Fig. 1B; note that Wnttun5 is not included in the figure). However, because Tcf7 acts as a positive regulator with nuclear β-catenin and because Gata.a acts as a positive regulator in which nuclear β-catenin is not present (*7*, *13*), we did not explicitly consider Tcf7 or Gata.a to simplify the calculations. In addition, Pem1 always represses transcription by suppressing the function of RNA polymerase II (*16*, *17*), and it is therefore unnecessary to consider this upstream factor explicitly. Ets1/2 acts as an effector of the MAPK pathway, which is regulated positively by Fgf9/16/20 signaling and CA-Raf and negatively by Efna.d signaling (*4*, *15*, *22*). Therefore, we do not consider Ets1/2 explicitly either. Because Wnttun5 controls the orientation of cell divisions and it is not likely that this ligand directly regulates gene expression (*18*), this was also excluded. For these reasons, we consider the remaining 16 upstream factors.

From gene expression patterns in normal and experimental embryos, regulatory functions of genes are directly given in the form of a truth table T^n

$$T^n:\{0,1\}^{16} \rightarrow \{0,1\},(n = 1,\ldots,13)$$

where the numbers of upstream factors and downstream genes are 16 and 13, respectively. Only some of the elements of T^n are given by experiments because the number of observed expression patterns of upstream factors is much smaller than the size of the binary space (2^{16}).

Our goal is to determine the simplest DNF that is consistent with the given partial truth table T^n; see below for the meaning of the "simplest DNF." The DNF is one of the canonical forms of the Boolean logic. In the case of the ascidian 32-cell embryo, DNFs, F^n, for 13 downstream genes are written in a form of logical disjunction of at least one logical conjunction of literals

$$F^n = \bigvee_{i=1}^{N(n)} C_i^n,\ C_i^n = \bigwedge_{j=1}^{16} l_{i,j}^n$$

where $N(n)$ is the number of conjunctions in the simplest DNF of F^n, and $l_{i,j}^n$ symbolizes the effect of an upstream factor j on the focal downstream gene n by using a literal G_j, that is

$$l_{i,j}^n = \begin{cases} G_j, \text{if gene } j \text{ up-regulates gene } n \text{ in } C_i^n \\ \neg G_j, \text{if gene } j \text{ down-regulates gene } n \text{ in } C_i^n \\ 1, \text{if gene } j \text{ does not regulate gene } n \text{ in } C_i^n \end{cases}$$

To determine the simplest DNF, let us consider all 3^{16} possible logical conjunctions of literals and write each conjunction as C_i (i = 1, …,3^{16}). The problem is then to find the minimum subsets of indexes $S^n \subseteq \{1, \cdots, 3^{16}\}$ such that

$$F^n = \bigvee_{i \in S^n} C_i$$

To reduce the computation time, we find S^n in a step-by-step manner based on the truth table T^n for each downstream gene n. In the following algorithm, $\mathbf{s} \in \{0,1\}^{16}$ denotes binary states of observed expression patterns of 16 upstream factors in normal and experimental embryos.

(i) Initially, set $S' = \{1, \cdots, 3^{16}\}$.

(ii) For each i, if $C_i(\mathbf{s}) = 1$ and $T^n(\mathbf{s}) = 0$ for any $\mathbf{s}$, then remove i from S'.

(iii) For each i, if $C_i(\mathbf{s}) = 0$ for all $\mathbf{s}$ such that $T^n(\mathbf{s}) = 1$, then remove i from S'.

(iv) Find the simplest minimum subset $S^n \subseteq S'$ satisfying the following equation for all $\mathbf{s}$

$$\bigvee_{i \in S^n} C_i(\mathbf{s}) = T^n(\mathbf{s})$$

In other words, under the assumption that the simplest one is most likely, our method fills the missing parts in T^n from limited sets of experiments. The source code for identifying simplest DNFs is available at https://github.com/kmaed/mindnf.

T^n shown in Table 1 were deduced from the expression patterns at the single-cell resolution in normal embryos (Fig. 1, B and C) and experimental embryos (Fig. 3 and fig. S1). When more than 60% of embryos expressed target genes, we considered that target genes were expressed (table S1).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/24/eabf8210/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. H. Nishida, N. Satoh, Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. *Dev. Biol.* **110**, 440–454 (1985).
2. C. Hudson, S. Darras, D. Caillol, H. Yasuo, P. Lemaire, A conserved role for the MEK signalling pathway in neural tissue specification and posteriorisation in the invertebrate chordate, the ascidian *Ciona intestinalis*. *Development* **130**, 147–159 (2003).
3. K. S. Imai, K. Hino, K. Yagi, N. Satoh, Y. Satou, Gene expression profiles of transcription factors and signaling molecules in the ascidian embryo: Towards a comprehensive understanding of gene networks. *Development* **131**, 4047–4058 (2004).
4. V. Bertrand, C. Hudson, D. Caillol, C. Popovici, P. Lemaire, 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 (2003).
5. C. Hudson, N. Kawai, T. Negishi, H. Yasuo, β-catenin-driven binary fate specification segregates germ layers in ascidian embryos. *Curr. Biol.* **23**, 491–495 (2013).
6. I. Oda-Ishii, T. Abe, Y. Satou, Dynamics of two key maternal factors that initiate zygotic regulatory programs in ascidian embryos. *Dev. Biol.* **437**, 50–59 (2018).
7. I. Oda-Ishii, A. Kubo, W. Kari, N. Suzuki, U. Rothbächer, Y. Satou, A maternal system initiating the zygotic developmental program through combinatorial repression in the ascidian embryo. *PLOS Genet.* **12**, e1006045 (2016).
8. K. Miwata, T. Chiba, R. Horii, L. Yamada, A. Kubo, D. Miyamura, N. Satoh, Y. Satou, Systematic analysis of embryonic expression profiles of zinc finger genes in *Ciona intestinalis*. *Dev. Biol.* **292**, 546–554 (2006).
9. C. Hudson, C. Sirour, H. Yasuo, Co-expression of *Foxa.a*, *Foxd* and *Fgf9/16/20* defines a transient mesendoderm regulatory state in ascidian embryos. *eLife* **5**, e14692 (2016).
10. K. S. Imai, K. Kobayashi, W. Kari, U. Rothbächer, N. Ookubo, I. Oda-Ishii, Y. Satou, Gata is ubiquitously required for the earliest zygotic gene transcription in the ascidian embryo. *Dev. Biol.* **458**, 215–227 (2020).
11. K. S. Imai, M. Levine, N. Satoh, Y. Satou, Regulatory blueprint for a chordate embryo. *Science* **312**, 1183–1187 (2006).
12. U. Rothbächer, V. Bertrand, C. Lamy, P. Lemaire, A combinatorial code of maternal GATA, Ets and β-catenin-TCF transcription factors specifies and patterns the early ascidian ectoderm. *Development* **134**, 4023–4032 (2007).
13. K. S. Imai, C. Hudson, I. Oda-Ishii, H. Yasuo, Y. Satou, Antagonism between β-catenin and Gata.a sequentially segregates the germ layers of ascidian embryos. *Development* **143**, 4167–4172 (2016).
14. Y. Satou, K. S. Imai, Gene regulatory systems that control gene expression in the *Ciona* embryo. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **91**, 33–51 (2015).
15. M. Tokuoka, K. Kobayashi, Y. Satou, Distinct regulation of *Snail* in two muscle lineages of the ascidian embryo achieves temporal coordination of muscle development. *Development* **145**, dev163915 (2018).
16. G. Kumano, N. Takatori, T. Negishi, T. Takada, H. Nishida, A maternal factor unique to ascidians silences the germline via binding to P-TEFb and RNAP II regulation. *Curr. Biol.* **21**, 1308–1313 (2011).
17. M. Shirae-Kurabayashi, K. Matsuda, A. Nakamura, Ci-Pem-1 localizes to the nucleus and represses somatic gene transcription in the germline of *Ciona intestinalis* embryos. *Development* **138**, 2871–2881 (2011).
18. N. Ohta, Y. Satou, Multiple signaling pathways coordinate to induce a threshold response in a chordate embryo. *PLOS Genet.* **9**, e1003818 (2013).
19. N. Ohta, K. Waki, A. Mochizuki, Y. Satou, A Boolean function for neural induction reveals a critical role of direct intercellular interactions in patterning the ectoderm of the ascidian embryo. *PLoS Comput. Biol.* **11**, e1004687 (2015).
20. K. Yagi, N. Satoh, Y. Satou, Identification of downstream genes of the ascidian muscle determinant gene *Ci-macho1*. *Dev. Biol.* **274**, 478–489 (2004).
21. O. Tassy, F. Daian, C. Hudson, V. Bertrand, P. Lemaire, A quantitative approach to the study of cell shapes and interactions during early chordate embryogenesis. *Curr. Biol.* **16**, 345–358 (2006).
22. V. Picco, C. Hudson, H. Yasuo, Ephrin-Eph signalling drives the asymmetric division of notochord/neural precursors in *Ciona* embryos. *Development* **134**, 1491–1497 (2007).
23. S. Tokuhiro, M. Tokuoka, K. Kobayashi, A. Kubo, I. Oda-Ishii, Y. Satou, Differential gene expression along the animal-vegetal axis in the ascidian embryo is maintained by a dual functional protein Foxd. *PLOS Genet.* **13**, e1006741 (2017).
24. L. Caputi, N. Andreakis, F. Mastrototaro, P. Cirino, M. Vassillo, P. Sordino, Cryptic speciation in a model invertebrate chordate. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 9364–9369 (2007).
25. M. L. Nydam, R. G. Harrison, Introgression despite substantial divergence in a broadcast spawning marine invertebrate. *Evolution* **65**, 429–442 (2011).
26. Y. Satou, K. S. Imai, N. Satoh, Early embryonic expression of a LIM-homeobox gene *Cs-lhx3* is downstream of β-catenin and responsible for the endoderm differentiation in *Ciona* savignyi embryos. *Development* **128**, 3559–3570 (2001).
27. L. Christiaen, A. Stolfi, B. Davidson, M. Levine, Spatio-temporal intersection of Lhx3 and Tbx6 defines the cardiac field through synergistic activation of *Mesp*. *Dev. Biol.* **328**, 552–560 (2009).
28. C. Anno, A. Satou, S. Fujiwara, Transcriptional regulation of *ZicL* in the *Ciona intestinalis* embryo. *Dev. Genes Evol.* **216**, 597–605 (2006).
29. I. S. Peter, E. Faure, E. H. Davidson, Predictive computation of genomic logic processing functions in embryonic development. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16434–16442 (2012).
30. L. Guignard, U.-M. Fiúza, B. Leggio, J. Laussu, E. Faure, G. Michelin, K. Biasuz, L. Hufnagel, G. Malandain, C. Godin, P. Lemaire, Contact area-dependent cell communication and the morphological invariance of ascidian embryogenesis. *Science* **369**, eaar5663 (2020).
31. Y. Satou, T. Kawashima, E. Shoguchi, A. Nakayama, N. Satoh, An integrated database of the ascidian, *Ciona intestinalis*: Towards functional genomics. *Zool. Sci.* **22**, 837–843 (2005).
32. Y. Satou, K. Mineta, M. Ogasawara, Y. Sasakura, E. Shoguchi, K. Ueno, L. Yamada, J. Matsumoto, J. Wasserscheid, K. Dewar, G. B. Wiley, S. L. Macmil, B. A. Roe, R. W. Zeller, K. E. M. Hastings, P. Lemaire, E. Lindquist, T. Endo, K. Hotta, K. Inaba, 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 (2008).
33. A. Stolfi, Y. Sasakura, D. Chalopin, Y. Satou, L. Christiaen, C. Dantec, T. Endo, M. Naville, H. Nishida, B. J. Swalla, J. N. Volff, A. Voskoboynik, D. Dauga, P. Lemaire, Guidelines for the nomenclature of genetic elements in tunicate genomes. *Genesis* **53**, 1–14 (2015).
34. Y. Satou, R. Nakamura, D. Yu, R. Yoshida, M. Hamada, M. Fujie, K. Hisata, H. Takeda, N. Satoh, 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 (2019).
35. K. S. Imai, H. Hikawa, K. Kobayashi, Y. Satou, *Tfap2* and *Sox1/2/3* cooperatively specify ectodermal fates in ascidian embryos. *Development* **144**, 33–37 (2017).
36. K. Kobayashi, K. Maeda, M. Tokuoka, A. Mochizuki, Y. Satou, Controlling cell fate specification system by key genes determined from network structure. *iScience* **4**, 281–293, (2018).
37. D. Yu, I. Oda-Ishii, A. Kubo, Y. Satou, The regulatory pathway from genes directly activated by maternal factors to muscle structural genes in ascidian embryos. *Development* **146**, dev173104 (2019).
38. T. Ikeda, T. Matsuoka, Y. Satou, A time delay gene circuit is required for palp formation in the ascidian embryo. *Development* **140**, 4703–4708 (2013).

39. T. Ikeda, Y. Satou, Differential temporal control of *Foxa.a* and *Zic-r.b* specifies brain versus notochord fate in the ascidian embryo. *Development* **144**, 38–43 (2017).
40. M. Tokuoka, K. S. Imai, Y. Satou, N. Satoh, Three distinct lineages of mesenchymal cells in *Ciona intestinalis* embryos demonstrated by specific gene expression. *Dev. Biol.* **274**, 211–224 (2004).
41. K. S. Imai, Y. Daido, T. G. Kusakabe, Y. Satou, Cis-acting transcriptional repression establishes a sharp boundary in chordate embryos. *Science* **337**, 964–967 (2012).
42. P. Lemaire, N. Garrett, J. B. Gurdon, 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 (1995).
43. S. Wada, Y. Katsuyama, S. Yasugi, H. Saiga, Spatially and temporally regulated expression of the LIM class homeobox gene *Hrlim* suggests multiple distinct functions in development of the ascidian, *Halocynthia roretzi*. *Mech. Dev.* **51**, 115–126 (1995).

Acknowledgments: We thank R. Yoshida, S. Aratake, M. Yoshida, and other members working under the National BioResource Project (MEXT, Japan) for providing the experimental animals. A draft of the manuscript was edited by a language editing service provided by Edanz Group.

Funding: CREST program of the Japan Science and Technology Agency (JST) JPMJCR13W6 (A.M. and Y.S.) and JPMJCR1922 (A.M.), Japan Society for the Promotion of Science grant 17KT0020 (Y.S.), Japan Society for the Promotion of Science grant 19H05670 (A.M.), and Japan Society for the Promotion of Science grant 20J40280 (M.T.). **Author contributions:** Investigation and data curation: M.T. and K.K. Formal analysis: K.M., A.M., and Y.S. Conceptualization: Y.S. Writing–original draft: Y.S. Writing–review and editing: M.T., K.M., K.K., A.M., and Y.S. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 21 November 2020
Accepted 23 April 2021
Published 9 June 2021
10.1126/sciadv.abf8210

Citation: M. Tokuoka, K. Maeda, K. Kobayashi, A. Mochizuki, Y. Satou, The gene regulatory system for specifying germ layers in early embryos of the simple chordate. *Sci. Adv.* **7**, eabf8210 (2021).

Science Advances

The gene regulatory system for specifying germ layers in early embryos of the simple chordate

Miki Tokuoka, Kazuki Maeda, Kenji Kobayashi, Atsushi Mochizuki and Yutaka Satou

Sci Adv **7** (24), eabf8210.
DOI: 10.1126/sciadv.abf8210

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.