Information processing at the *foxa* node of the sea urchin endomesoderm specification network

Smadar Ben-Tabou de-Leon[1] and Eric H. Davidson[1]

Division of Biology 156-29, California Institute of Technology, Pasadena, CA 91125

Contributed by Eric H. Davidson, April 20, 2010 (sent for review March 10, 2010)

The *foxa* regulatory gene is of central importance for endoderm specification across Bilateria, and this gene lies at an essential node of the well-characterized sea urchin endomesoderm gene regulatory network (GRN). Here we experimentally dissect the *cis*-regulatory system that controls the complex pattern of *foxa* expression in these embryos. Four separate *cis*-regulatory modules (CRMs) cooperate to control *foxa* expression in different spatial domains of the endomesoderm, and at different times. A detailed mutational analysis revealed the inputs to each of these *cis*-regulatory modules. The complex and dynamic expression of *foxa* is regulated by a combination of repressors, a permissive switch, and multiple activators. A mathematical kinetic model was applied to study the dynamic response of *foxa cis*-regulatory modules to transient inputs. This study shed light on the mesoderm–endoderm fate decision and provides a functional explanation, in terms of the genomic regulatory code, for the spatial and temporal expression of a key developmental control gene.

cis-regulatory analysis | gene regulatory networks | mathematical modeling | embryonic development

The sea urchin *foxa* gene encodes a forkhead transcription factor ortholog, which is used in the process of endoderm specification in many bilaterians and also in cnidarians (1–6). In *Xenopus* (6) and the sea urchin (1) a major function of the *foxa* gene during embryonic development is maintenance of the endoderm–mesoderm boundary, by repression of mesoderm fate in the endoderm. Interference with translation of *foxa* mRNA in sea urchin causes loss of gut formation and specification of excess mesodermal derivatives (1). The spatial expression pattern of *foxa* in the sea urchin embryo is dynamic. It is initially expressed broadly in the endomesoderm progenitor field, but later the *foxa* gene is silenced in the mesoderm, continuing to be transcribed only in the veg2 endoderm lineage (Fig. 1*A*) (1, 7, 8).

The *foxa* gene is a canonical member of an evolutionarily ancient subcircuit of the endoderm specification gene regulatory network (GRN) in echinoderms (1, 7, 9). Perturbation data predict multiple regulatory inputs into the *foxa* node of the endomesoderm specification network, both positive and negative (1, 7). To learn how the genomic control system integrates these inputs, producing the *foxa* expression pattern as an output, we undertook a comprehensive *cis*-regulatory analysis. Embryonic expression of *foxa* turns out to be controlled by four separate *cis*-regulatory modules (CRMs). The CRMs receive repressing and activating inputs and function combinatorially to drive the *foxa* spatiotemporal expression pattern.

Results

Expression and Regulatory Structure of the *foxa* Gene. Transcriptional expression of the *foxa* gene starts at about 11 h postfertilization (hpf). At 15 hpf *foxa* is expressed in all of the descendants of the veg2 ring of cells and is absent from the skeletogenic mesoderm lineage (SM), from veg1 and from the ectoderm (Fig. 1*A*). The inner veg2 descendants give rise to the nonskeletogenic mesoderm lineage (NSM), and the outer veg2 descendants give rise to endoderm (Fig. 1*A*). From 15 to 18 hpf, *foxa* expression continues in both NSM and endoderm progenitors, but at 18–20 hpf, *foxa* expression is shut down in the NSM progenitors and transcription continues thereafter only in the endoderm (Fig. 1 *A* and *C*). The clearance of *foxa* transcripts from the NSM marks the physical separation in different cells of the distinct regulatory states underlying the respective specification of the NSM and the endoderm (7). Starting at about 22 hpf, *foxa* is also expressed in a patch of cells in the oral ectoderm where the mouth will form (Fig. 1 *A* and *C*).

To provide a standard of comparison for the ensuing *cis*-regulatory dissection, we constructed recombinant *Strongylocentrotus purpuratus foxa* BACs that contain either GFP or RFP coding sequences knocked into the *foxa* translation start site. These BACs contain the remainder of the *foxa* gene, plus 98 kb of upstream and 50 kb of downstream noncoding sequence (Fig. 1*B*). In sea urchin embryos exogenous DNA is stably integrated in a clone of cells early in cleavage, resulting in a random mosaic pattern of incorporation. When the recombinant BAC is injected into sea urchin eggs, it faithfully expresses its fluorophore in the same embryonic spatial domains where endogenous *foxa* transcripts are found. For example, Fig. 1 *D* and *E* show embryos expressing the *foxa:GFP* BAC; in each embryo shown here a clone of cells displays activity within the appropriate endodermal domain (Fig. 1*D*) or the oral ectoderm domain (Fig.1*E*). The BAC constructs thus contain all of the regulatory information necessary to generate normal *foxa* spatial expression.

Sequence patches in the vicinity of the *foxa* gene that displayed significant interspecific conservation were identified by comparison of the *foxa* BAC of *S. purpuratus* with the *Lytechinus variegatus* BAC-containing orthologous sequence (10) (Fig. S1). A series of gene transfer experiments using deletion and fusion expression constructs, guided by the results of Fig. S1, then resulted in identification of four conserved *cis*-regulatory regions, such that when all are included in the construct, a correct pattern of expression is generated. Thus together these four regions contain all of the regulatory information necessary for *foxa* expression. These are the regions *F*, *I*, *J*, and *K* of Fig. 1*B* (their sequences are given in *SI Text 1*).

To study the regulation of spatial gene expression, various versions of *foxa* CRMs driving expression of GFP were coinjected with the complete *foxa:RFP* BAC (11). Coinjected constructs are ligated together in the egg cytoplasm before the concatenates are taken up into blastomere nuclei. Thus the GFP and RFP constructs are incorporated into the same cells of the embryo (12). The reporter constructs are coinjected together with about 3-fold excess carrier DNA to preclude cross-regulation (11). The advantage of this method is that the experiment and the control are observed in the same embryos. This is illustrated with a construct including the three CRMs *F*, *I*, and *J*, which in combination suffice to drive correct spatial expression. When the reporter construct *FIJ:GFP* (i.e., *F*, *I*, and *J* fused together in front of the GFP coding sequence) is coinjected

Author contributions: S.B.-T.d.-L. and E.H.D. designed research; S.B.-T.d.-L. performed research; S.B.-T.d.-L. and E.H.D. analyzed data; and S.B.-T.d.-L. and E.H.D. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. E-mail: smadar@caltech.edu or davidson@caltech.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1004824107/-/DCSupplemental.

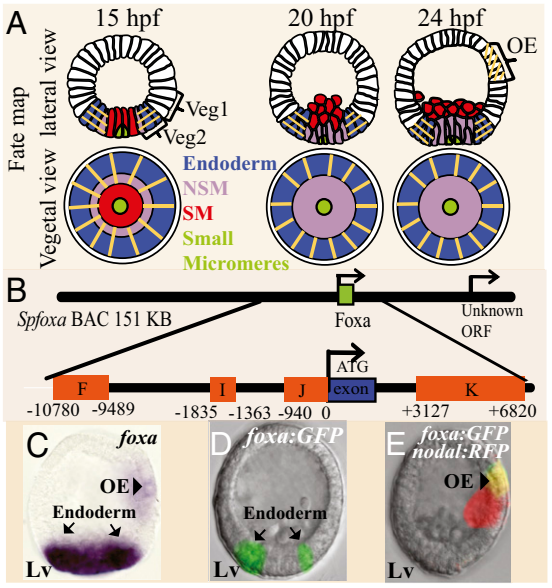

Fig. 1. Expression and regulatory structure of the *foxA* gene. (*A*) Lineage fate map showing lateral and vegetal views at 15, 20, and 24 hpf. The orange lines indicate the domains where *foxa* is expressed. Green, small micromeres; red, SM; purple, veg2 NSM; blue, veg2 endoderm; white, veg1 and the ectoderm; OA, oral ectoderm. (*B*) Diagram of *foxA* CRMs (orange boxes) in *SpfoxA* BAC. The numbers on the CRMs edges indicate their distance from *foxA* start of translation. *foxA* single exon is marked as a blue box. (*C*) Whole mount in situ hybridization of *foxa* at 24 hpf. *foxa* is expressed in the endoderm (arrows) and in the oral ectoderm (arrowhead). (*D* and *E*) Expression of *foxA:GFP* BAC at 24 hpf. The reporter is expressed in the endoderm (*D*) and in the oral ectoderm (*E*). Because the oral and aboral sides of the embryo are indistinguishable at 24 hpf, we coinjected the *foxa:GFP* BAC with *Nodal:RFP BAC*, which is expressed in the oral ectoderm. The figure shows the overlay of the two flourophores. Lv, lateral view.

with *foxa:RFP* BAC, the two constructs express in exactly the same cells as shown in Fig. 2 *A–C*. Quantitative comparison between the spatial expression driven by the *foxa:RPF* BAC and by the construct *FIJ:GFP* injected alone (Fig. S2) (example, Fig. 2*G*), further demonstrates that the two reporters contain the same spatial regulatory information.

Spatial Restriction of *foxa* Expression Pattern by the Tcf-Groucho/β-Catenin Switch. Further experiments showed that the regulatory information necessary for correct spatial expression is encoded in a single Tcf (T cell factor) site in the most 5′ end of module *F*. A mutation of this site caused the mutated *FIJ:GFP* construct to express ectopically in any cells of the embryo. For example, Fig. 2 *D–F* shows an embryo coinjected with the mutated *FIJ:GFP* construct plus the *foxa:RFP* BAC. Whereas the RFP is expressed correctly in veg2 endoderm, the GFP is expressed in the endoderm but also ectopically in veg1 ectoderm. Fig. 2 *H* and *I* show that when the Tcf site is mutated, the construct produces ectopic expression in the SM and in the NSM. The ubiquitous factor that drives expression in all of the cells of the embryo probably binds to module *J*, as the reporter construct *J:GFP* drives ubiquitous expression.

Tcf acts as a toggle switch, depending on its cofactors (13, 14). Early in development, Tcf coactivator β-catenin is stabilized and nuclearized in the vegetal plate of the sea urchin embryo by a maternal cytoplasmic system (15, 16). At this time Tcf acts to permit expression of target genes such as *foxa* in the entire endomesoderm territory, while repressing expression of the same genes in the ectoderm. By midblastula stage nuclear β-catenin has cleared from the SM nuclei and is localized in the veg2

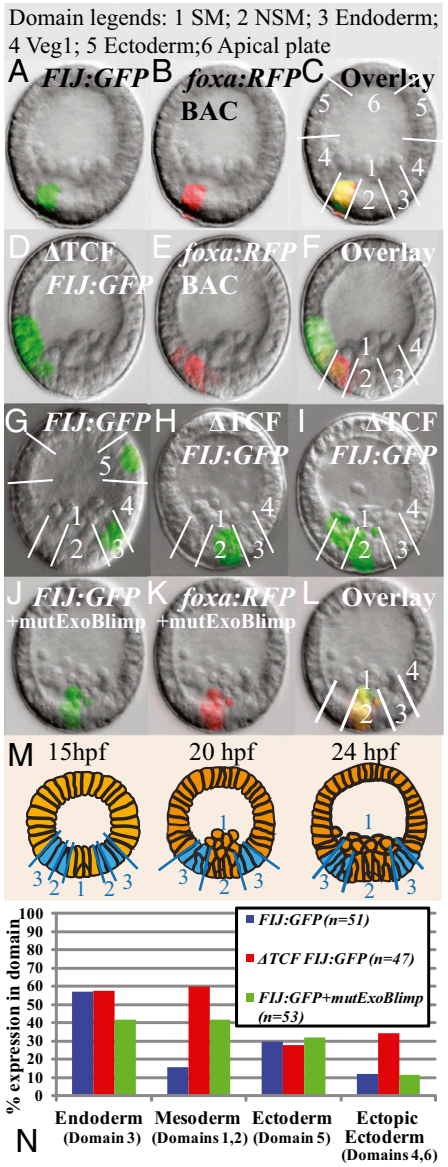

Fig. 2. Spatial restriction of *foxa* expression. Throughout the figure embryonic domains are enumerated as follows: 1 SM, 2 NSM, 3 veg2 endoderm, 4 veg1, 5 ectoderm, and 6 apical plate. (*A–L*) Expression patterns of embryos injected with different reporter constructs at 23–24 hpf. (*A–C*) *FIJ:GFP* coinjected with *foxA:RFP* BAC: (*A*) GFP, (*B*) RFP, and (*C*) overlay. (*D–F*) Coinjection of *foxA:RFP* BAC and a version of *FIJ:GFP* where a Tcf site in *F* was mutated. (*D*) GFP, (*E*) RFP, and (*F*) overlay. (*G*) Intact *FIJ:GFP* injected alone drives correct expression in the endoderm and in the ectoderm. (*H* and *I*) Tcf mutated version of *FIJ:GFP* drives GFP in the NSM (*H*) and in the SM, NSM, and endoderm (*I*). (*J–L*) *FIJ:GFP* coinjected with *foxA:RFP* BAC and mutExoBlimp1 construct that drives *blimp1* in the NSM: (*J*) GFP, (*K*) RFP, and (*L*) overlay. (*M*) Diagrams of β-catenin nuclearization at 15, 20, and 24 hpf. Cells where β-catenin is nuclearized are in cyan, the rest of the cells are in orange. (*N*) Quantification of spatial expression at 23–24 hpf. Percentages sum to more than 100% as some embryos express in two or more tissue types. At 24 hpf the oral and aboral sides of the embryo are indistinguishable, therefore, we considered only domains 4 and 6 to be ectopic ectoderm expression. Data are based on three replicate experiments totaling ≈120 embryos for each construct. Number of expressing embryos for each construct is indicated in the graph key.

lineage nuclei, i.e., in cells that will give rise to NSM plus endoderm (Fig. 2*M*) (16), and at this time *foxa* expression is likewise restricted to these cells (Figs. 1*A* and 2*M*) (7). At mesenchyme blastula stage (>20 hpf) detectable nuclearized β-catenin clears

from the NSM as well and remains visible only in veg2 plus veg1 endoderm (16). This leads to silencing of *foxa* expression in the NSM due to Tcf-Groucho repression. Thus the same factor that initially enables broad *foxa* expression throughout the veg2 endomesoderm later restricts it to the endodermal domain of this lineage (cf. Figs. 1*A* and 2*M*).

Zygotic β-catenin nuclearization is maintained in the endomesodermal territories by Wnt signaling. Among the Wnt ligands that execute this function is Wnt8 (17). The transcription factor Blimp1 is one of the activators of the *wnt8* gene (18). In normal development Blimp1 turns down its own expression first in the SM, then in the NSM (19). If *blimp1* autorepression is prevented by mutation of the *blimp1 cis*-regulatory target sites in a *Blimp1* expression construct, ectopic expression of both *blimp1* and *wnt8* genes follows in the NSM at times when both genes are normally shut off there (19). If indeed the Tcf-Groucho complex is the NSM repressor of *foxa*, then if by the same means we force a lingering *wnt8* expression and hence β-catenin nuclearization in the NSM, repression of *foxa* in this domain should not occur, and *foxa* transcripts should fail to clear from the NSM. To test this prediction we used the same mutated *blimp1* expression construct lacking Blimp1 target sites as in the earlier work (19). This construct was coinjected together with *FIJ:GFP* and *foxa:*RFP BAC. The clear result was that both the GFP and RFP are now expressed ectopically in the SM and NSM (Fig. 2 *J–L*).

Quantification of spatial expression for the experiments of Fig. 2 are shown in Fig. 2*N*. The percentage of embryos expressing ectopically in the mesoderm (SM and NSM) and in the ectoderm has increased significantly in the Tcf mutant. When *FIJ:GFP* was coinjected with the *blimp1* expression construct lacking autorepression sites, the ectoderm expression is unchanged, but the mesoderm expression is sharply increased. This is the predicted result, on the basis that mesoderm repression of *foxa* depends on the Tcf toggle switch. Taken together, these results indicate that Tcf is the input responsible for keeping the *foxa* gene off outside of the endoderm, at least in embryos <22 hpf, and that this regulatory transaction is mediated by the genomic Tcf site in module *F*.

Combinatorial Control of *foxa* Activation and Expression Level. To study the regulation of *foxa* expression level we used quantitative PCR (QPCR) to measure the expression levels generated by different reporter constructs over time. Construct expression kinetics were compared with that of the *foxa* recombinant BAC (Fig. 3*A*). Maps and genomic coordinates of each construct are presented in Fig. S3. In the absolute terms of molecules of reporter transcript per incorporated construct DNA molecule, the kinetic profile generated by the *foxa* recombinant BAC is significantly higher than that of the construct *FIJ:GFP* (Fig. 3*A*). The missing information is not encoded in sequences between modules *F*, *I*, and *J*, because the construct *F* to *J*, which spans the whole genomic region from module *F* through modules *I* and *J* to the reporter knockin (Fig. S3), generates the same expression level as does *FIJ:GFP* (Fig. 3*A*). However, a significant amplification of expression level occurs when the 3′ module *K* is included (*F to K*, Fig. 3*A*). Modules *F*, *I*, *J*, and *K* together generate a kinetic profile identical to that of the *foxa:RFP* BAC, indicating that together they contain all of the information necessary for quantitative as well as spatial *foxa* expression.

We conducted detailed mutation and perturbation analyses to determine the binding sites that transduce functional *foxa* inputs in the four CRMs. A map of the sites is shown in Fig. 3*B*. The binding site sequences and their exact locations within the CRMs are to be found in *SI Text 1*, and a list of the sites per se plus their mutated forms are provided in Table S1.

Delta-Notch and Suppressor of Hairless [Su(H)]. Perturbation analysis shows that interference with Notch signaling prevents *foxa* clearance from the NSM (8) but does not affect *foxa* expression level (7). When the construct *FIJ:GFP* is coinjected with Notch MO, the

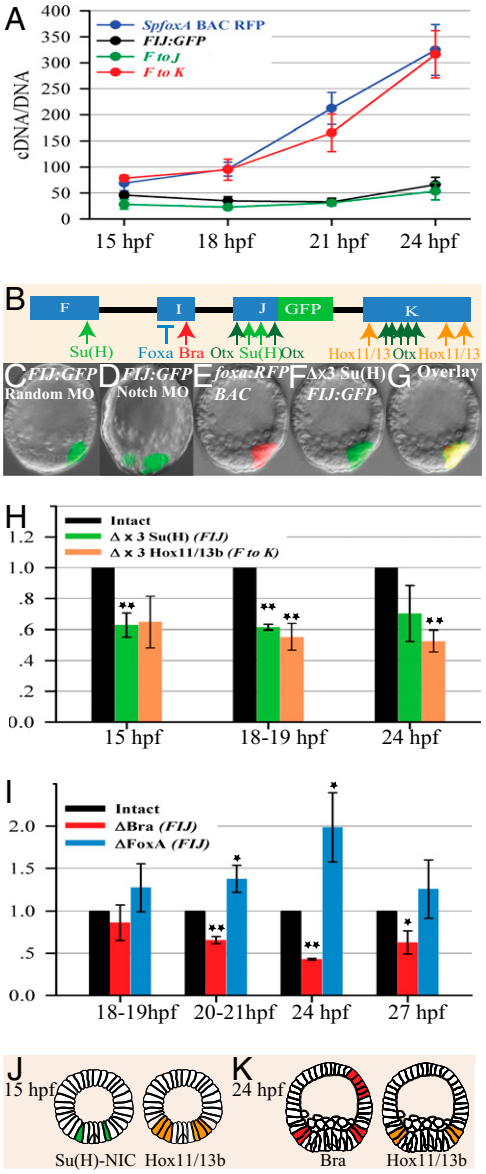

Fig. 3. Quantitative analysis of *foxa* inputs. Throughout the figure error bars show ±1 SE. (*A*) QPCR time courses of *SpfoxA* BAC:RFP and the constructs: *F* to *K*, *F* to *J*, and *FIJ:*GFP. The values are cDNA (mRNA) copies per injected DNA copy. Values are based on at least three independent batches of 150 injected embryos per time point. (*B*) Diagram of *foxA* CRMs and the binding sites of the different inputs. Color codes for all inputs are consistent throughout. (*C*) The construct *FIJ:GFP* coinjected with Random MO shows normal expression in the endoderm at 25 hpf. (*D*) The construct *FIJ:GFP* coinjected with Notch MO, the GFP fails to clear from the NSM. (*E–G*) Coinjection of the reporter *FIJ:GFP* where the three Su(H) binding sites are mutated with *foxA:RFP* BAC. The GFP expression is unaffected by the mutation and is identical to the RFP expression. (*E*) RFP, (*F*) GFP, (*G*) overlay. (*H* and *I*) QPCR measurement of the effect of site mutations on the level of reporter constructs. The results are the ratio between the expression of the mutated construct and the intact construct. The construct that was used to measure the effect of each mutation is indicated in the figure key. Significance was calculated by one-tailed *Z*-test. *, $P < 0.01$; **, $P < 0.001$. Numerical values of means, number of repeats, and *P* values are presented in Table S2. (*J*) Spatial expression of *foxa* inputs at 15 hpf. Delta signal from the SM is inducing Su(H)-NIC activity in the NSM. Hox11/13b is expressed in all veg2 descendants. (*K*) Spatial expression of *foxa* inputs at 24 hpf. Brachyury and Hox11/13b have progressed toward veg1 descendants and only partially overlap with *foxA* expressing cells.

GFP lingers in the NSM at 24–25 hpf, (Fig. 3 *C* and *D* and quantification in Fig. S4*A*) and the GFP level is unchanged, similarly to the behavior of the endogenous *foxa*. To study the direct effect of Notch signaling we mutated the three putative binding site of Su(H) in the construct *FIJ:GFP*. The mutations did not prevent the reporter clearance from the NSM and at 24–25 hpf it expresses correctly in the endoderm (Fig. 3 *E–G* and quantification in Fig. S4*B*). The mutations reduced the expression level of the reporter at 15 hpf and 18 hpf (Fig. 3*H*). At 10–18 hpf, Notch signaling is occurring in veg2 cells due to reception of the Delta ligand produced by the adjacent SM cells (Fig. 3*J*) (20). Our analysis shows that in its initial phase of activity, Notch signaling helps to boost *foxa* expression in the veg2 domain; however, the clearance from the NSM is an indirect effect, probably due to the interaction with the Tcf toggle-switch mechanism, as discussed below.

Hox11/13b. New perturbation studies revealed that *hox11/13b* MO decreases *foxa* levels starting from 18 hpf (7). There are three putative Hox11/13b binding sites in region *K*. We mutated these sites in a reporter construct spanning the whole genomic region from *F* to *K* and observed a 50% reduction in the expression level at 18 hpf and 24 hpf (Fig. 3*H*). At 18 hpf this reduction brings the expression level of *F* to *K* down to the average expression level of *FIJ:GFP*, indicating that Hox11/13b is the main input into *K*. However, a comparison of the results in Fig. 3 *A* and *H* shows that at 24 hpf the difference in expression level due to Hox11/13b site mutations accounts for only about half of the input from module *K*. This implies additional late activating input into module *K*, probably Otx, as explained below. There is a paradoxical aspect to the late (i.e., 24 hpf) effect of *hox11/13b* site mutations on expression levels of the *foxa* construct: although until about 21 hpf, *foxa* and *hox11/13b* expression overlap in the veg2 endoderm (Fig. 3*J*) (7), after this time, *hox11/13b* expression clears progressively from veg2 and becomes active in veg1, where *foxa* is not expressed (Fig. 3*K*) (7, 21). We address this apparent paradox below.

Otx. Perturbation analysis shows that a splice blocking MO targeted to the homeodomain exon of the *otx* gene [i.e., one that will affect all Otx isoforms (7)] affects *foxa* level at 24 hpf (Fig. S5*A*). At 24 hpf this MO reduces the level of the reporter *F* to *K* by about 80%, similarly to the reduction of the endogenous *foxa* expression level (Fig. S5*A*). The reporter construct *FIJ:GFP* is unaffected by this treatment (Fig. S5*A*) and the mutations of the Otx sites in *FIJ* do not reduce the reporter level at this time (Fig. S5). These results show that at 24 hpf the response to Otx is encoded in module *K*, probably through the 10 putative Otx sites identified in this module (*SI Text 1*). In addition, at 27 hpf the mutation of two putative Otx sites in the proximal module *J* reduces the expression of the constructs *FIJ:GFP* and *J:GFP* by 30% (Fig. S5*B*) indicating that at this time Otx binds to module *J* as well. Different Otx isoforms are expressed everywhere in the embryo through development. Because *foxa* expression is always more restricted spatially than is *otx* expression, the Otx input probably acts to boost *foxa* level, providing no spatial information.

Brachyury. We identified a putative Brachyury site in module *I*. Mutation of this site in the *FIJ:GFP* reporter construct, and in the *F* to *J* construct, significantly decreased reporter expression level at 20–27 hpf (Fig. 3*I*). The *brachyury* gene turns on at about 15 hpf and is initially expressed in the same domain as *foxa* (7). At about 21 hpf *brachyury* expression begins to fade in veg2 and becomes active in veg1 and also in a patch in the oral ectoderm (Fig. 3*K*). Thus after 24 hpf the spatial overlap of the endodermal expression of *foxa* and *brachyury* decreases significantly. Our results show, however, that Brachyury input is required for the normal level of *foxa* expression at 20–27 hpf. We address the delayed and lingering effect of Brachyury mutation below.

FoxA. Perturbation analysis showed that the *foxa* gene represses its own transcriptional activity (1). At 24 hpf the injection of Foxa MO increased the level of the construct *FIJ:GFP* by 2-fold, similar to the increase of the level of the endogenous *foxa* gene in the same injections. A mutation of a single putative FoxA site in module *I* increased the level of the *FIJ:GFP* reporter transcript at 20–24 hpf (Fig. 3*I*). This result verifies that FoxA is an autorepressor that reduces, but does not eliminate, its gene product level.

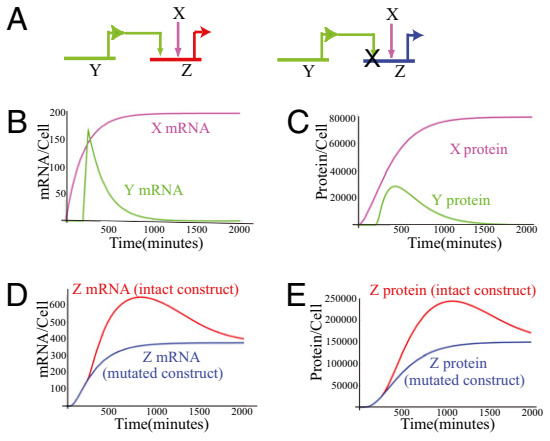

Fig. 4. Kinetic model of a site mutation of a transient input in a reporter construct. (*A*) Schematic diagrams of the inputs to the intact (*Left*) and the mutated (*Right*) constructs. (*B*) mRNA expression kinetics of the inputs, *X* (fuchsia) and *Y* (green). (*C*) Protein kinetics of the inputs, *X* and *Y*, same color code as in *B*. (*D*) mRNA kinetics of the intact construct (red) and the mutated construct (blue). (*E*) Protein kinetics of intact and mutated constructs, same color code as in *D*. Kinetic parameters: mRNA and protein turnover rate $K_{dm} = K_{dp} = 0.005\ min^{-1}$. Transcriptional delay, $T_m = 20$ min. See *SI Text 2* for additional details.

Lagging Kinetics of Target Gene Expression in Response to a Transient Input. The mutational analysis shows significant delays in the response of *foxa* reporter constructs to some of their activating inputs. For example, Hox11/13b appears to contribute significantly to the boost in *foxa* expression at 24 hpf (Fig. 3*H*) although the spatial overlap between *hox11/13b* and *foxa* is reduced at this time (cf. Figs. 1*A* and 3*K*). The gene *brachyury* turns on at 15 hpf, yet its effect on *foxa* reporter construct is detectable only at 20 hpf and lingers until 27 hpf, 3 h after these gene products have segregated spatially (cf. Figs. 1*A* and 3*K*).

To understand this kind of phenomenon, it is necessary to consider how long after a positively acting input becomes available do the effects on its target gene become detectable; and how long after the input becomes unavailable, do detectable effects on its target gene outputs linger. We applied a mathematical model developed earlier (22–24) to simulate the expression dynamics of an intact construct and a construct where a binding site of a transient input is mutated. The set of differential equations derived for this purpose is given in *SI Text 2*, and the results of the simulations are shown in Fig. 4. Here the target gene, *Z*, has two inputs, the transcription factors produced by genes *X* and *Y* (Fig. 4*A*); i.e., gene *Z* would represent the *foxa* gene or *foxa cis*-regulatory constructs. Consideration of the logic function executed by the *cis*-regulatory element of *Z* on its inputs is essential for a realistic simulation. Mutations of Hox11/13b (Fig. 3*H*) and Brachyury (Fig. 3*I*) binding sites each resulted in decreases of up to 50% of the total expression level of the *foxa* constructs: thus the contribution of each activator is additive with respect to the others. Therefore we simulate the logic executed by the CRM of *Z* as (*X* additive OR *Y*) (22–24). We assumed a half-life of 2 h for all mRNAs and proteins. It is important to note that these half-lives are short relative to the average values observed in sea urchin embryos (23); a priori, the extent of the lingering effect will be greater the longer the half-life.

Kinetic profiles for the mRNA and transcription factor proteins produced by genes *X* and *Y* are shown in Fig. 4 *B* and *C*. Kinetic profiles for mRNA and protein produced by gene *Z* in the intact condition and under *Y* site mutation are shown in Fig. 4 *D* and *E*. In this simulation, the difference between the outputs of the intact construct and of the mutated one reaches 2-fold about 5 h after gene *Y* transcription is shut off (Fig. 4 *B* and *D*). The reason for the delayed onset effect is that it requires some hours for the transient driver input to produce sufficient target gene output that the difference is detectable over the background provided by the nontransient driver. The significant difference in transcript levels generated by the control vs. the mutated construct lingers for about 8 h more, long after *Y* mRNA is essentially gone. Qualitatively, the basic reason for the lag in effects is that the mRNA and protein products of gene *Y* remain present and functionally active after the gene ceases activity, until they stochastically decay, and likewise the transcripts of the target gene *Z* accumulate as well, until they also stochastically decay. The import of the exercise in Fig. 4 is that it provides a realistic explanation for the delayed difference between control construct output and output of constructs in which the Brachyury and Hox11/13b target sites were mutated (Fig. 3 *H* and *I*). Furthermore, the delayed *cis*-regulatory response reported here is relevant to any case of additive OR logic applied on transient inputs.

Discussion

Here we identify the genomic regulatory information that determines the complex spatial and temporal expression of a gene, which is key to the process of endoderm specification. The regulatory system is encompassed in four modules, the inputs, functions, and interactions of which we now understand. We also solve a problem that occurs frequently in kinetic perturbation data sets used for GRN analysis. The *foxa* gene participates in the network functions that result in separation of endodermal and mesodermal fates, and its immediate roles in this process are controlled by the *cis*-regulatory system, which is the subject of this work.

***Cis*-Regulatory Genomic Code.** The *foxa* gene describes a complex pattern of expression (Fig. 1) (7). In the early blastula its transcription is activated and confined to the veg2 endomesodermal cell lineage. But then *foxa* expression is silenced in the exact portion of this lineage that will assume a mesodermal (NSM) fate, while the gene continues to be expressed in the endodermal portion of the veg2 descendants. For these endoderm cells, *foxa* later provides canonical regulatory functions that are essential to specification of the foregut endoderm (7). This general description of *foxa* expression pattern can be decomposed into a set of specific regulatory operations, which the *cis*-regulatory analysis causally resolves, as follows: (*i*) Initiation of transcription: the gene begins to be expressed in the embryo because modules *F*, *J*, and *K* contain target sites for three early positive transcription factors, namely. the permissive Tcf-β-catenin complex (module *F*), the Delta-Notch signal transducer Su(H) (module *J*), and Hox11/13b (module *K*). (*ii*) Expression in veg2 endomesoderm but absence of expression in SM or ectoderm: the genomic basis for this is the Tcf site of module *F*, which subjects this gene to toggle-switch spatial expression control, dependent on whether a given nucleus contains Tcf-β-catenin or the repressive Tcf-Groucho (Figs. 2*M* and 5*A*, 15 hpf). In addition, the Su(H) sites are functional only in veg2 cells that are in direct contact with the Delta ligand (Figs. 3*J* and 5*A*). Also the prominent driver of *foxa*, Hox11/13b, is expressed only in veg2 at this time (Fig. 3*J*). (*iii*) Silencing of foxa transcription in future NSM: the genomic basis for this silencing is also the Tcf site (module *F*) as shown in Fig. 2 *D–I*. The underlying molecular mechanism is the clearance of β-catenin from nucleus of this territory and the establishment of Tcf-Groucho repression there.

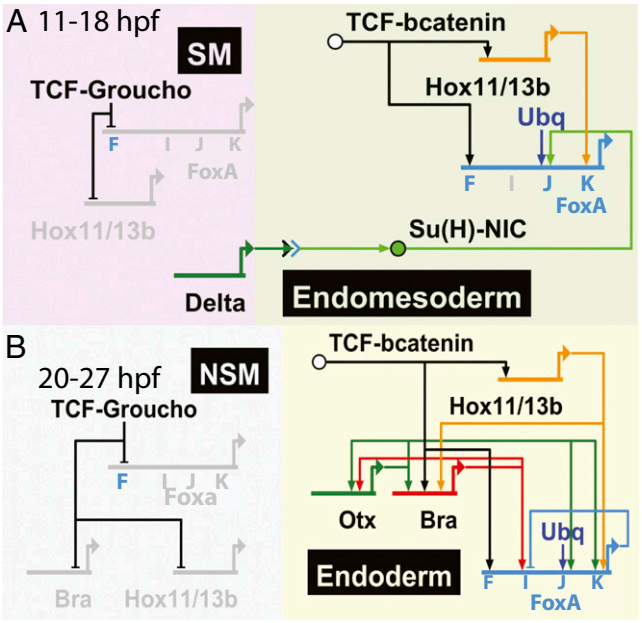

Fig. 5. Endomesoderm GRN and the active inputs into *foxA* CRMs. Active modules are marked in light blue; inactive modules and genes are marked in gray. (*A*) At 11–18 hpf *foxA* is activated by Hox11/13b, the Delta-Notch signal from the SM cells and unknown ubiquitous activator. *foxA* is repressed in the ectoderm and in the SM by Tcf-Groucho. (*B*) At 20–27 hpf *foxA* is activated by Hox11/13b, Otx, Brachyury, and unknown ubiquitous activator. Starting from 20 hpf, FoxA autorepresses its own gene expression. *foxA* is restricted to the endoderm by Tcf-Groucho repression elsewhere.

The mechanisms responsible for β-catenin clearance are still not fully resolved and there may be several contributory events. Notch signaling is clearly required, as interference with Notch signaling in the NSM precursors blocks clearance of *foxa* transcripts from these cells (Fig. 3 *C* and *D* and Fig. S4A) (8). This could work via the Nemo-like kinase, which reduces positive Tcf activity (25), presumably leading to institution of the repressive Tcf-Groucho complex. Note that from the local vantage point afforded by the *foxa* gene per se, this mechanism could reflect the means by which endodermal vs. NSM fate is established, as many other endodermal genes are controlled by the Tcf switch (Fig. 5) (7). (*iv*) Increase in rate of expression after 18 hpf and control of amplitude: the increased rate after 18 hpf is caused by the boost received from the Otx (module *K*) and the spatially confined Brachyury (module *I*) and Hox11/13b (module *K*). From 20 hpf on the level of transcription is modulated by *foxa* autorepression (module *I*, Fig. 3*I*).

There are two missing pieces of information. First, the identity of the ubiquitous activator that binds in module *J* is unknown. Second, it remains unclear which factors prevent *foxa* expression in veg1 endoderm in the 24–30 hpf period. Both Brachyury and Hox11/13b are by then being expressed in veg1, but *foxa* transcription remains strictly confined to the veg2 lineage.

With these exceptions, our results demonstrate that the *cis*-regulatory analysis provides a unique explanation, in terms of specific DNA sequence features, for each aspect of blastula stage *foxa* expression.

Combinatorial Modular Function. It has been clear for some time that in bilaterian animals, genes typically use multiple CRMs to mediate expression in diverse regulatory states (for review, ref. 26). Recently, however, a new aspect of module choice has arisen, the existence of CRMs, which in combination contribute to a given phase of developmental expression, rather than generating completely distinct phases (e.g., ref. 27). The use of BAC

reporter constructs enables the study of the totality of the *cis*-regulatory system and has demonstrated the function of multiple modules in the same cells at given phases of sea urchin development (28, 29). The several noncontiguous DNA segments harboring these modules interact combinatorially with the basal transcription apparatus (11). The analysis presented here provides a unique view of CRMs cooperation: each of *foxa*'s four CRMs executes a particular task, but their specific interactions are required to generate the overall pattern of expression. The proximal module *J* integrates ubiquitous and local [Su(H)-NIC] drivers, but its output depends on module *F*. In the absence of the spatial repression function mediated by module *F*, module *J* drives expression ubiquitously. The Hox11/13b sites of module *K* contribute to *foxa* expression level, early and late, interacting with the proximal module *J* in cells where module *F* is permissive. Module *I* is active only after 20 hpf and it together with module *K* mediates later quantitative control of output. These interactions imply that formation of intermodule complexes following occupation of transcription factor target sites is an essential aspect of the regulatory control mechanism. This is unlikely to be a peculiarity of the *foxa* gene. The sequence basis underlying control of functional interactions among CRMs is emerging as a new frontier in understanding the genomic regulatory code.

Materials and Methods

Microinjection and QPCR Measurement of GFP mRNA in Eggs Expressing GFP Constructs. PCR products were purified with the Qiagen Qiaquick PCR purification Kit or Zymo DNA Clean and Concentrator kit. The fragments were then microinjected into fertilized *S. purpuratus* eggs as described (30). Linearized BAC constructs were desalted by drop dialysis into TE buffer on a 0.025-μm VSWP filter (Millipore). Approximately 500–800 molecules of the desired reporter construct were injected, along with a 3- to 6-fold molar excess of HindIII-digested carrier sea urchin DNA per egg, in a 4-pL volume of 0.12 M KCl. A similar injection solution was made for BAC reporters but with 100–200 copies of the BAC per 4 pL. Embryos were collected at different stages for assessment of spatial activity by fluorescence microscopy or for quantitative analysis of transcript prevalence by QPCR. For QPCR measurements RNA and DNA of about 150 injected embryos per time point was isolated using Qiagen Allprep kit. The RNA was then reverse transcribed to cDNA using iScript cDNA synthesis kit from Bio-Rad. QPCR for both DNA and cDNA was performed as described (31). The level of cDNA was computed by comparison with an internal standard (ubiquitin) cDNA, and the level of injected DNA was computed in comparison with nodal DNA.

Constructs. Standard PCR and fusion PCR techniques using the High Fidelity PCR kit (Roche) and Long Template High Fidelity PCR kit (Roche) were used to build reporter constructs. Binding-site sequences were mutated by PCR, and the resulting constructs were checked by sequencing. The mutation PCR primers were designed with about 20-bp sequences surrounding both ends of the target site. The target site was changed into a mutant form of the candidate transcription factor binding sites. Whole mount in situ hybridization was performed as described in ref. 1.

ACKNOWLEDGMENTS. We thank Isabelle Peter, Jongmin Nam, and Joel Smith for valuable comments on the manuscript. We thank Sharon Kuo for help with constructs preparation. This research was supported by National Institutes of Health Grants GM 61005 and HD 37105. S.B.-T.d.-L was a fellow of the Human Frontier Science Program Organization.

1. Oliveri P, Walton KD, Davidson EH, McClay DR (2006) Repression of mesodermal fate by foxa, a key endoderm regulator of the sea urchin embryo. *Development* 133: 4173–4181.
2. Friedman JR, Kaestner KH (2006) The Foxa family of transcription factors in development and metabolism. *Cell Mol Life Sci* 63:2317–2328.
3. Boyle MJ, Seaver EC (2008) Developmental expression of foxA and gata genes during gut formation in the polychaete annelid, Capitella sp. I. *Evol Dev* 10:89–105.
4. Hiruta J, Mazet F, Yasui K, Zhang P, Ogasawara M (2005) Comparative expression analysis of transcription factor genes in the endostyle of invertebrate chordates. *Dev Dyn* 233:1031–1037.
5. Burtscher I, Lickert H (2009) Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development* 136:1029–1038.
6. Suri C, Haremaki T, Weinstein DC (2004) Inhibition of mesodermal fate by Xenopus HNF3beta/FoxA2. *Dev Biol* 265:90–104.
7. Peter IS, Davidson EH (2010) The endoderm gene regulatory network in sea urchin embryos up to mid-blastula stage. *Dev Biol* 340:188–199.
8. Croce JC, McClay DR (2010) (Dynamics of Delta/Notch signaling on endomesoderm segregation in the sea urchin embryo. *Development* 137:83–91.
9. Hinman VF, Nguyen AT, Cameron RA, Davidson EH (2003) Developmental gene regulatory network architecture across 500 million years of echinoderm evolution. *Proc Natl Acad Sci USA* 100:13356–13361.
10. Brown CT, Xie Y, Davidson EH, Cameron RA (2005) Paircomp, FamilyRelationsII and Cartwheel: Tools for interspecific sequence comparison. *BMC Bioinformatics* 6:70.
11. Nam J, et al. (2007) Cis-regulatory control of the nodal gene, initiator of the sea urchin oral ectoderm gene network. *Dev Biol* 306:860–869.
12. Livant DL, Hough-Evans BR, Moore JG, Britten RJ, Davidson EH (1991) Differential stability of expression of similarly specified endogenous and exogenous genes in the sea urchin embryo. *Development* 113:385–398.
13. Range RC, Venuti JM, McClay DR (2005) LvGroucho and nuclear beta-catenin functionally compete for Tcf binding to influence activation of the endomesoderm gene regulatory network in the sea urchin embryo. *Dev Biol* 279:252–267.
14. Kim CH, et al. (2000) Repressor activity of Headless/Tcf3 is essential for vertebrate head formation. *Nature* 407:913–916.
15. Weitzel HE, et al. (2004) Differential stability of beta-catenin along the animal-vegetal axis of the sea urchin embryo mediated by dishevelled. *Development* 131: 2947–2956.
16. Logan CY, Miller JR, Ferkowicz MJ, McClay DR (1999) Nuclear beta-catenin is required to specify vegetal cell fates in the sea urchin embryo. *Development* 126:345–357.
17. Wikramanayake AH, et al. (2004) Nuclear beta-catenin-dependent Wnt8 signaling in vegetal cells of the early sea urchin embryo regulates gastrulation and differentiation of endoderm and mesodermal cell lineages. *Genesis* 39:194–205.
18. Minokawa T, Wikramanayake AH, Davidson EH (2005) cis-Regulatory inputs of the wnt8 gene in the sea urchin endomesoderm network. *Dev Biol* 288:545–558.
19. Smith J, Theodoris C, Davidson EH (2007) A gene regulatory network subcircuit drives a dynamic pattern of gene expression. *Science* 318:794–797.
20. Sweet HC, Gehring M, Ettensohn CA (2002) LvDelta is a mesoderm-inducing signal in the sea urchin embryo and can endow blastomeres with organizer-like properties. *Development* 129:1945–1955.
21. Arenas-Mena C, Cameron RA, Davidson EH (2006) Hindgut specification and cell-adhesion functions of Sphox11/13b in the endoderm of the sea urchin embryo. *Dev Growth Differ* 48:463–472.
22. Davidson EH (1986) *Gene Activity in Early Development* (Academic, Orlando), 3rd Ed.
23. Bolouri H, Davidson EH (2003) Transcriptional regulatory cascades in development: Initial rates, not steady state, determine network kinetics. *Proc Natl Acad Sci USA* 100: 9371–9376.
24. Ben-Tabou de-Leon S, Davidson EH (2009) Modeling the dynamics of transcriptional gene regulatory networks for animal development. *Dev Biol* 325:317–328.
25. Rottinger E, et al. (2006) Nemo-like kinase (NLK) acts downstream of Notch/Delta signalling to downregulate TCF during mesoderm induction in the sea urchin embryo. *Development* 133:4341–4353.
26. Davidson EH (2006) *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution.* (Academic, San Diego).
27. Hong JW, Hendrix DA, Levine MS (2008) Shadow enhancers as a source of evolutionary novelty. *Science* 321:1314.
28. Lee PY, Nam J, Davidson EH (2007) Exclusive developmental functions of gatae cis-regulatory modules in the Strongylocentrorus purpuratus embryo. *Dev Biol* 307: 434–445.
29. Wahl ME, Hahn J, Gora K, Davidson EH, Oliveri P (2009) The cis-regulatory system of the tbrain gene: Alternative use of multiple modules to promote skeletogenic expression in the sea urchin embryo. *Dev Biol* 335:428–441.
30. Rast JP (2000) Transgenic manipulation of the sea urchin embryo. *Methods Mol Biol* 136:365–373.
31. Revilla-i-Domingo R, Minokawa T, Davidson EH (2004) R11: A cis-regulatory node of the sea urchin embryo gene network that controls early expression of SpDelta in micromeres. *Dev Biol* 274:438–451.