Assessing regulatory information in developmental gene regulatory networks

Isabelle S. Peter[a,1] and Eric H. Davidson[a,2]

[a]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125

Edited by Douglas H. Erwin, Smithsonian National Museum of Natural History, Washington, DC, and accepted by Editorial Board Member Neil H. Shubin January 24, 2017 (received for review August 17, 2016)

Gene regulatory networks (GRNs) provide a transformation function between the static genomic sequence and the primary spatial specification processes operating development. The regulatory information encompassed in developmental GRNs thus goes far beyond the control of individual genes. We here address regulatory information at different levels of network organization, from single node to subcircuit to large-scale GRNs and discuss how regulatory design features such as network architecture, hierarchical organization, and *cis*-regulatory logic contribute to the developmental function of network circuits. Using specific subcircuits from the sea urchin endomesoderm GRN, for which both circuit design and biological function have been described, we evaluate by Boolean modeling and in silico perturbations the import of given circuit features on developmental function. The examples include subcircuits encoding positive feedback, mutual repression, and coherent feedforward, as well as signaling interaction circuitry. Within the hierarchy of the endomesoderm GRN, these subcircuits are organized in an intertwined and overlapping manner. Thus, we begin to see how regulatory information encoded at individual nodes is integrated at all levels of network organization to control developmental process.

developmental GRN | network topology | circuit function | network hierarchy | Boolean modeling

Developmental process is controlled by gene regulatory networks (GRNs), the regulatory interactions between genes encoding transcription factors and signaling interactions that determine developmental gene expression throughout the genome (1–3). At the level of a single node, the information captured in GRNs is intuitively accessible. However, the information encompassed in GRNs is not just to regulate individual genes. During development, GRNs control the differential specification of cell fates and determine the organization of body parts, organs, and cell types within the animal body plan. As more and more GRNs become experimentally solved, the question arises as to how network features can be recognized that carry information for more complex developmental transactions.

Comparison of different regulatory networks shows that particular constellations of regulatory interactions among a few genes, so-called subcircuits, are recurrently deployed in very different biological contexts (1, 4, 5). These subcircuits are composed of different regulatory genes, but nevertheless encode similar regulatory functions. Specific subcircuit topologies include, for example, positive feedback circuitry, leading to stabilization of gene expression, or mutual-repression circuits that lead to the exclusion of regulatory states (6–9). Several types of network subcircuits have been identified so far, each associated with specific regulatory functions (1, 4, 5). It appears, furthermore, that given types of subcircuits are often found at given positions within the GRN hierarchy, although the number of experimentally solved large-scale networks is so far still small. The existence of these network subcircuits, together with the predictability of their function and position within a GRN, provide evidence that network topology is an important determinant of developmental function.

Beyond individual genes or small subcircuits, we so far know little about the structural regulatory features within entire GRNs that organize and coordinate individual developmental functions at a systems level. It is evident, however, that developmental process is determined by properties at all levels of GRN organization—at the single-node level, the subcircuit level, and the network level. To make explicit those regulatory features that are essential carriers of regulatory information, we use the sea urchin endomesoderm GRN as an example, for which both the architecture of the network as well as the developmental process it controls are well described, and discuss the information provided at each level of network organization.

As demonstrated by a recent Boolean model (10), the endomesoderm GRN for sea urchin embryos is a large-scale developmental GRN that has been experimentally resolved close to completeness (2, 11–13). To assess regulatory information at the single-node level, we can simply compare numbers of regulatory inputs, both activating and repressing, that regulate expression of individual genes. However, even at the level of subcircuits consisting of only few genes, evaluating regulatory information is a complex task. Thus, instead of searching for an absolute measure for information content, we here assess regulatory information in a different way. We assume that the simplest form of regulatory interactions between a set of genes is a linear pathway from the most upstream to the most downstream gene in the circuit. We then examine the regulatory information encoded in a GRN circuit by comparing its regulatory functions to the function of the linear path, thus addressing the gain in regulatory information by network wiring exceeding linear pathways. Using Boolean modeling, we predict specific network topologies, as well as *cis*-regulatory logic processing functions, that determine regulatory function in each type of subcircuit. To assess information encoded at the network level, we investigate the number and types of subcircuits contained within the network, as well as the organization of these subcircuits within network hierarchy, showing that even at this level, subcircuits are not linearly organized, but are heavily intertwined to ensure coordinated function during developmental process.

Regulatory Information at the Single-Node Level

The current model of the sea urchin endomesoderm GRN includes 41 regulatory genes encoding transcription factors, in addition to at least seven signaling interactions. In total, these regulatory genes are connected by 177 regulatory interactions, of

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "Gene Regulatory Networks and Network Models in Development and Evolution," held April 12–14, 2016, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and video recordings of most presentations are available on the NAS website at www.nasonline.org/Gene_Regulatory_Networks.

Author contributions: I.S.P. and E.H.D. designed research; I.S.P. and E.H.D. performed research; and I.S.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. D.H.E. is a guest editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. Email: ipeter@caltech.edu.

[2]Deceased September 1, 2015.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1610616114/-/DCSupplemental.

which 42 mediate repression. To assess the density of regulatory interactions in this network, we can make a simple calculation. Assuming a linear topology of regulatory interactions among a given set of n genes, the number of interactions will be $n - 1$, representing the minimal number of interactions. Conversely, the maximal number of interactions permitted between n genes amounts to n^2. For example, two regulatory genes are connected by at least one interaction (for example $a > b$) and no more than four interactions ($a > b, b > a, a > a$, and $b > b$). Thus, the number of regulatory interactions within a subcircuit of n genes provides the "saturation" of the subcircuit in terms of interactions. With 177 interactions, the sea urchin GRN model is therefore in the lower range of the 40–1,681 permitted interactions.

Although the number of interactions at individual nodes may reflect functional importance in some networks (14), this rule does not apply to developmental GRNs. In GRNs, a high number of regulatory interactions may control the expression of a gene of little regulatory significance, such as a differentiation gene, whereas a regulatory gene upstream in the GRN hierarchy with important regulatory functions may only be connected through few regulatory interactions. In the endomesoderm GRN model, the number of regulatory inputs at individual regulatory nodes ranges from 0 to 7, and the number of outputs ranges from 0 to 15, where the appearance of 0 indicates missing interactions. An example of a regulatory gene with low connectivity and important regulatory function is *pmar1*, a transiently expressed regulatory gene at the top of the skeletogenic GRN that is connected by three regulatory interactions. The expression of this gene is controlled by two regulatory inputs, whereas its only known regulatory function is the repression of *hesc*, thereby controlling the activation of the entire skeletogenic GRN in skeletogenic cells (13, 15). Conversely, Alx1 has many target genes within the same GRN and yet controls only aspects of the functions of the skeletogenic GRN (13, 16, 17). The number of regulatory inputs and outputs at individual network nodes is, therefore, not sufficient to assess regulatory information, and higher levels of network organization need to be considered.

Regulatory Information at the Level of Network Subcircuits

Several types of subcircuits have been defined by comparing frequently occurring constellations of regulatory interactions in GRNs, as discussed extensively in ref. 1. Here, we use specific examples from the same developmental context, the endomesoderm GRN, and evaluate the regulatory information contributed by individual circuit features by Boolean modeling and in silico perturbation. In principle, dynamic ordinary differential equation models could also be used to assess the function of small subcircuits and could perhaps reveal slightly different insights, particularly by providing information on expression levels. Our focus on Boolean modeling stems mostly from its applicability to capturing *cis*-regulatory logic functions, modeling dominant repression, and assessing circuit output simultaneously in time and space.

Positive Feedback Subcircuit. Any topology of regulatory interactions in which a transcription factor either directly or indirectly contributes to the activation of its own gene constitutes a positive feedback circuit. Although the minimum number of genes in a positive feedback circuit is one, a systematic comparison of different GRNs revealed that autoactivation occurs rarely, except in regulatory genes participating in positive feedback subcircuits with other genes (1). The example shown in Fig. 1*A* is a positive feedback subcircuit operating the initial specification of nonskeletogenic mesoderm (NSM) cell fates downstream of Delta/Notch signaling. This subcircuit consists of three genes, glial cell missing (*gcm*), *gatae*, and *six1/2*. The linear hierarchy of this subcircuit is *gcm>gatae>six1/2*. In addition, this subcircuit includes two positive feedback interactions, from *six1/2* into *gcm*, and autoregulatory activation of *gcm* (18). This subcircuit is activated by a Delta/Notch signal that feeds directly into *gcm* and *gatae* (19, 20).

The structure of this subcircuit directly reflects its biological function. It is first activated in precursors of NSM cells by Delta/Notch signaling from adjacent skeletogenic cells. However, shortly after activating this subcircuit, cells divide, and some of the progeny give rise to endodermal cell fates in which the subcircuit has to be inactive in order not to interfere with endoderm specification. Conversely, the Delta/Notch signaling input in the NSM disappears after a few hours, when signal sending skeletogenic cells eventually ingress into the blastocoel. Thus, by this time, the positive feedback circuit has to be active to ensure continued expression of *gcm*. Stabilization of gene expression thus occurs within a narrow time window, not too early and not too late.

A Boolean model of this subcircuit demonstrates this dynamic behavior. Here, the Delta/Notch signaling input (Fig. 1 *A*, ①) acts in OR logic to the other two inputs in the control of *gcm*, such that either Delta/Notch signaling or the two positive feedbacks are sufficient to activate *gcm* expression. On the other hand, the two positive feedbacks operate in AND logic, such that both inputs are necessary for *gcm* expression. Thus the subcircuit is activated by Delta/Notch signaling at first, but continues to be active even after the signaling input disappears (Fig. 1 *A1*). In this constellation, when Delta/Notch signaling occurs for a short time, the feedback circuit is not activated, and *gcm* is expressed only transiently, just as is the case in endodermal cells (Fig. 1 *A2*). Indeed, experimental observations confirm that prolonged exposure to Delta/Notch signaling is necessary to induce stable expression of *gcm* (21). Similarly, if the positive feedbacks are not functional, and *gcm* is exclusively dependent on the Delta/Notch signaling input, expression of all three genes is only supported as long as Delta/Notch signaling is available (Fig. 1 *A3*). If we assume a different *cis*-regulatory logic at the *gcm* node, such that all inputs operate in AND logic, *gcm* expression can not be activated by Delta/Notch signaling alone and will not be expressed (Fig. 1 *A4*). Conversely, if all inputs would operate in OR logic, the autoregulatory feedback mediated by Gcm itself would operate without delay, and this subcircuit would quickly become independent of Delta/Notch signaling (*SI Appendix*).

Thus the relatively simple subcircuit presented in Fig. 1*A* is rich in regulatory information, controlling delayed stabilization of gene expression downstream of a transient signaling input. In this constellation, expression of *six1/2* is turned on with a delay compared with *gcm* expression. Because Gcm and Six1/2 are both required in AND logic to activate *gcm* expression, the positive feedback is active only after the onset of *six1/2* expression, thus ensuring delayed stabilization of gene expression. The information encoded in each component of a subcircuit of this kind is as follows. Transient activation is mediated by interactions ① and ⑤ operating in OR logic to other activating inputs (Fig. 1*A*). Maintenance of gene expression in the absence of the initial input is controlled by the positive feedback in interactions ② and ③, which thus mediate an uncoupling from earlier regulatory events. The temporal delay of input–independent gene expression is operated by interactions ④, ⑥, and ②, together with the AND logic between interactions ② and ③. Interestingly, the organization of the *gcm cis*-regulatory system is in complete agreement with this predicted regulatory logic. Thus, the Delta/Notch signaling input is encoded in a *cis*-regulatory module that is separate from a later-acting module requiring both Gcm and Six1/2 to stabilize gene expression through the positive feedback circuit (18, 19).

Community-Effect Subcircuit. The community-effect subcircuit represents a special form of positive feedback circuit, where the expression of a signaling ligand is controlled downstream of the signaling pathway it activates. The term "community effect" derives from the observation that the maintenance of cell-type–specific gene expression in parenchymal cells depends on intercellular signaling interactions (22). This observation can be explained by relatively simple network circuitry (23).

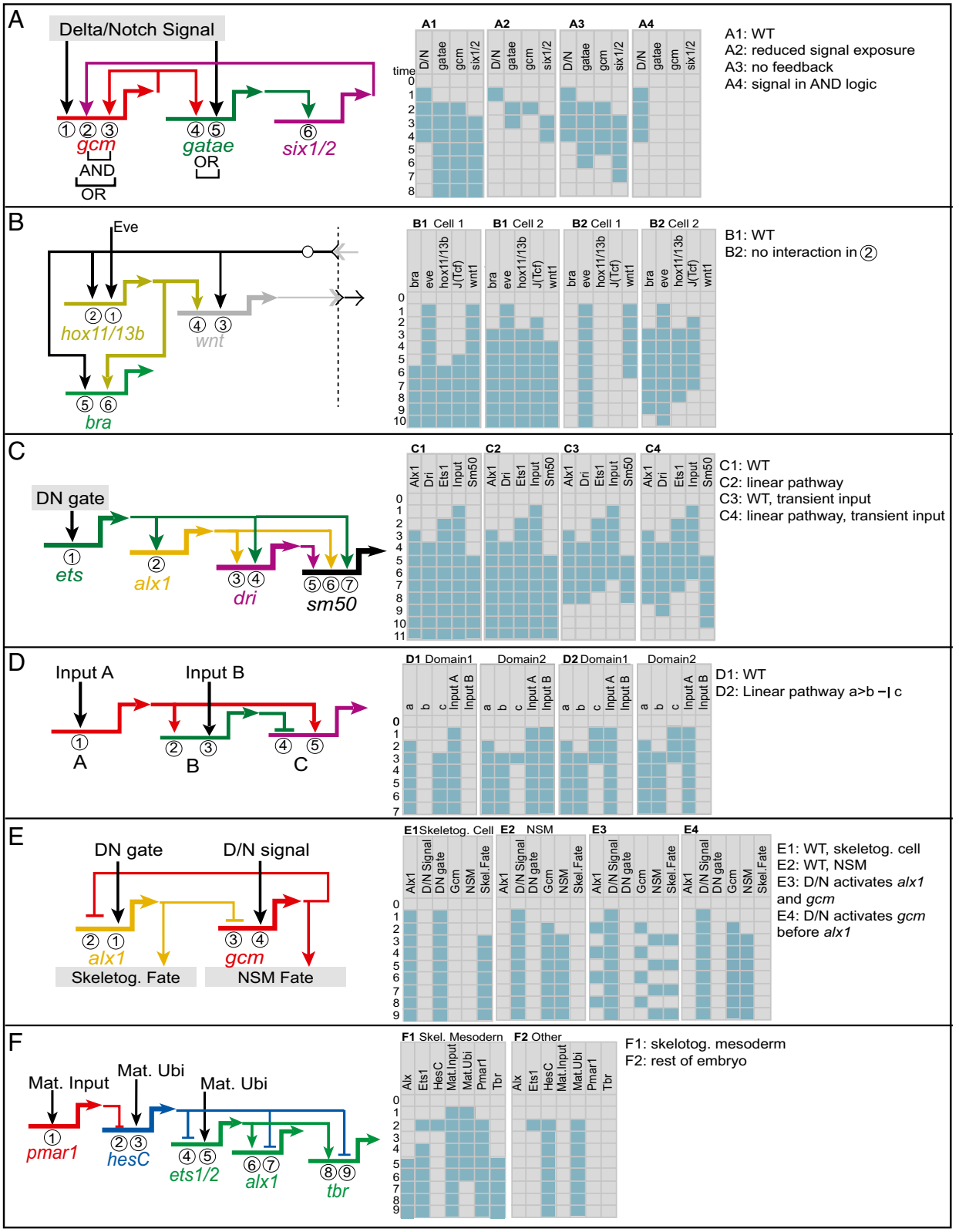

Fig. 1. Structure and function of different types of subcircuit. (*A*) Positive feedback subcircuit. (*B*) Community-effect subcircuit. (*C*) Coherent feedforward subcircuit. (*D*) Incoherent feedforward subcircuit. (*E*) Mutual-repression subcircuit. (*F*) Double-negative gate subcircuit. All except the subcircuit in *D* are examples from the sea urchin endomesoderm GRN. (*Left*) The topologies of regulatory interactions in each subcircuit. (*Right*) The expression of each gene in the subcircuit under each condition, as determined by Boolean modeling. The indicated time steps do not represent real time. Blue, expression; gray, no expression. For equations and additional analyses, see *SI Appendix*. DN, double-negative gate; D/N, Delta/Notch signaling; Mat., maternal; Skel., skeletogenic; Ubi, ubiquitous activator.

In sea urchin embryos, a community-effect subcircuit is activated in the posterior endoderm just before gastrulation. As shown in Fig. 1*B*, the core of this subcircuit consists of *hox11/13b* and *wnt1*. The linear pathway is *hox11/13b*>*wnt1*-Tcf/β-catenin>*hox11/13b*, thus forming a perfect loop. Both Hox11/13b and Wnt signaling furthermore activate *brachyury* (12, 24). This subcircuit is first activated by Wnt signaling from adjacent anterior endoderm cells. During gastrulation, signal-sending and -receiving endodermal cells are physically separated, and gene expression is maintained by the community-effect circuit.

The Boolean model of this community-effect subcircuit in Fig. 1*B* shows that, upon inducing expression of *wnt1* in one cell (Fig. 1 *B1*, cell 1), signaling will induce expression of *hox11/13b* in adjacent cells (Fig. 1 *B1*, cell 2). In turn, Hox11/13b activates expression of *wnt1*, which signals back to other cells, including cell 1. The result is that, eventually, an entire domain of cells will express the same regulatory state downstream of a signaling input, even if the signaling input was initially not received uniformly. Furthermore, as the model shows, the initial signal can be transient, such that if *wnt1* expression in cell 1 is turned on for a limited time interval, gene expression will be sustained thereafter by the activity of the subcircuit. The crucial interactions for these functions are the positive feedback between *hox11/13b* and Wnt signaling. If the subcircuit is perturbed by removing the Wnt signaling input into *hox11/13b*, the activation of this circuit will not spread to nearby cells, nor will gene expression be stabilized once the transient initial input disappears (Fig. 1 *B2*). Similarly, if the same circuit architecture contains instead of *wnt1*, a regulatory gene encoding a transcription factor, the stabilization of gene expression will only occur within cells receiving the initial signal, but no propagation of regulatory states to adjacent cells will occur (*SI Appendix*).

The function of this subcircuit is controlled by the following circuit components. The external input into this circuit is provided by interaction 1, providing context-specific activation limited to the posterior endoderm (Fig. 1*B*). All interactions ①–⑥ are required for replication of regulatory-state expression throughout the domain. Synchronous gene expression within a field of cells is mediated by interactions ②, ③, and ⑤. The intercellular positive feedback loop mediated by interactions ②, ③, and ④ is responsible both for the propagation of gene expression to nearby cells, as well as to ensure maintenance of gene expression once the initial activating input is no longer available. Thus, again, this subcircuit ensures uncoupling from earlier events and stabilization of gene expression in the absence of upstream inputs. The import of this subcircuit in development is evident because, particularly during later development, cells within a given domain will not be at equal distance to the signal source, and will receive varying levels of signal ligands. Where nevertheless cells assume a common fate, community-effect circuits will ensure equal activation of gene expression throughout the domain. The function of this subcircuit is reminiscent of the positive feedback subcircuit, which operates in individual cells such as migrating mesodermal cells, whereas the community effect is active in a domain of cells that remain in close contact with each other. The result of both circuits, however, is to ensure irreversibility and unidirectionality of the developmental process.

Coherent Feedforward Subcircuit. A feedforward circuit is any circuit in which an upstream transcription factor controls the expression of two or more target genes encoding at least one transcription factor controlling the expression of the same genes as the upstream factor ($a > b$, $a > c$, and $b > c$). In a coherent feedforward constellation, the two paths controlling expression of gene c execute the same regulatory functions, activation or repression of the downstream gene. Although this subcircuit can have several configurations, by far the most frequently encountered in developmental GRNs is a subcircuit in which all interactions are activating.

The example shown in Fig. 1*C* is a component of the skeletogenic GRN (13) and represents a typical example for a coherent feedforward subcircuit. A differentiation gene, here *sm50*, is driven by multiple transcription factors, which in addition control each other. Thus, the upstream regulatory factor Ets1 activates the expression of two regulatory genes, *alx1* and *dead ringer* (*dri*), and the differentiation gene *sm50*. In turn, Alx1 activates the expression of *dri* and *sm50*, and Dri also induces expression of *sm50*. This circuit represents a nested feedforward architecture in which all possible feedforward interactions are present.

A function frequently attested to coherent feedforward motifs is that they ensure the delay of downstream gene expression. As the Boolean model in Fig. 1*C* shows, assuming AND logic between all regulatory inputs at the downstream node indeed results in the temporal delay of *sm50* expression in respect to its upstream activators (Fig. 1 *C1*). To test whether information for temporal delay is an intrinsic feature of feedforward circuitry, we compare its behavior to the linear pathway (*ets*>*alx1*>*dri*>*sm50*). Remarkably, the Boolean computation shows that the linear pathway results in a similar temporal delay of downstream gene activation (Fig. 1 *C2*). However, these two circuits behave differently once the initial input is turned off. In the feedforward constellation, if the input into the subcircuit disappears, the entire subcircuit is turned off within two time intervals (Fig. 1 *C3*). In the linear constellation, turning off the initial input only gradually inactivates the subcircuit, and expression of *sm50* occurrs for four additional time intervals (Fig. 1 *C4*). If instead of AND logic we assume that all regulatory interactions operate in OR logic, the subcircuit is turned on very rapidly, within two time intervals, but shows a slow turnoff rate similar to the linear pathway constellation (*SI Appendix*). OR logic functions may also contribute to the control of gene expression levels. Conversely, AND logic gates lead to a temporal delay from initial input to full activity similar to a linear pathway, but ensure a rapid inactivation of the subcircuit once the initial input or any of the upstream factors are no longer available.

The function of a subcircuit of this type is controlled by following components (Fig. 1*C*). Activation is mediated by interaction ①. Delay of downstream gene expression is encoded by the linear hierarchy of this subcircuit, given by interactions ②, ③, and ⑤, but only if these inputs are in AND logic with the upstream transcription factor (③ AND ④; ⑤ AND ⑦). This delay increases with the number of steps in the feedforward circuit. The dependence on all upstream inputs for expression of the downstream gene is controlled by interactions ⑤, ⑥, and ⑦, if all of them operate in AND logic, or by ③ AND ④ together with ⑤ AND ⑦. Both regulatory logics render expression of *sm50* conditional to all earlier regulatory events and, curiously enough, ensure the absence of stability in this subcircuit. Thus, contrary perhaps to intuition, the specific regulatory information contributed by the feedforward subcircuit is not delay of downstream gene expression, but rapid loss of downstream gene expression if any of the upstream GRN components are not available. Feedforward circuitry frequently controls expression of downstream differentiation genes that determine cellular function. This subcircuit represents a safety device guaranteeing that downstream cellular functions are installed only where and as long as all required inputs in the upstream specification GRN have been activated.

Incoherent Feedforward Subcircuit. In the incoherent feedforward subcircuit, the direct regulation of downstream genes by the upstream transcription factors and the indirect regulation through the intermediate transcription factor convey opposite regulatory information. Thus, one branch will lead to repression of the downstream gene, whereas the other will lead to activation. Surprisingly, the endomesoderm GRN model does not contain an incoherent feedforward subcircuit, although this constellation occurs frequently in other contexts (25). For purpose of completeness, we here consider an abstract subcircuit, as shown in Fig. 1*D*. In this

example, an upstream transcription factor A activates the expression of regulatory genes B and C, whereas transcription factor B, in turn, represses regulatory gene C.

The developmental function of this subcircuit becomes apparent when considered spatially, because this subcircuit can control the separation of two cell fate domains. As shown in the Boolean model in Fig. 1*D*, assuming that genes A and B are driven by different initial inputs, and that the input into A (Fig. 1 *D1*, input A) is expressed in a broader domain (domain 1 plus domain 2) than input B (domain 2), then this subcircuit installs two different regulatory states in domains 1 and 2. In the first domain, A activates C, whereas B is not expressed due to the absence of its initial input. In the second domain, A and B are expressed, and, in turn, C is expressed for a short time until repressor B becomes available. The result is that domain 1 expresses A and C, whereas domain 2 expresses A and B. The function of this subcircuit is particularly useful for the subdivision of a progenitor domain into two or more subdomains, where transcription factor A represents activators present earlier in development. When considering this subcircuit without the feedforward interaction, the linear path of A > B–|C results in transient activation of C wherever A is expressed, and no spatial separation of regulatory states occurs (Fig. 1 *D2*).

The determinant features of this subcircuit are interaction ①, controlling where A is expressed, and interactions ② and ③ operating in AND logic, such that B is expressed only where A and input B are present. In the case that interactions ② and ③ operate in OR logic, the result of the subcircuit is the same as the linear pathway, and B is expressed wherever A is present without spatial subdivision. In addition, the repression of gene C by B (interaction ④), has to be dominant over activation to install spatial subdivision. Similar to the coherent feedforward subcircuit, the incoherent feedforward instructs a connection to earlier regulatory events by using an upstream regulatory factor as activating input for downstream genes. This function is mediated by inputs ② and ⑤. This subcircuit is also comprehensive, in that gene expression is regulated in all cells of the initial field (interactions ②–⑤). The Boolean exclusion function ensuring that all cells must assume either one of two regulatory states is mediated by interactions ② and ④. Thus, for the purpose of precise spatial expression of two regulatory states, a transient sloppiness is the cost, in that gene C is transiently expressed before being terminally turned off.

Mutual-Repression Subcircuit. A subcircuit architecture frequently encountered in the literature is the mutual-repression circuit, consisting of two antagonizing repressors. Thus, a first regulatory gene encodes a repressor of a second regulatory gene, which, in turn, represses expression of the first gene. Simultaneous expression of both genes can generate a bistable state that shifts with the slightest difference in expression levels to exclusive expression of only one of the two regulatory genes. In dynamic models, when both regulatory genes are assumed to be coexpressed, stochastic differences in expression levels can trigger a random switch in one or the other way. Indeed, this subcircuit structure has been observed in GRNs responding to gradients of regulatory factors, where mutually repressing regulatory genes are offering a mechanistic explanation for the expression of distinct regulatory states in response to small differences in activating inputs (8, 9, 26, 27).

In developmental GRNs, the outcome of this subcircuit is far from random. In the only example within the endomesoderm GRN, shown in Fig. 1*E*, the two regulatory genes *alx1* and *gcm* are repressing each other in a mutual-repression subcircuit, although evidence for direct regulatory linkages is so far missing. In addition, both genes are regulated by different inputs, indicating that the decision between the two genes is predetermined. *Alx1*, encoding a transcription factor essential for skeletogenesis, is directly controlled by the regulatory mechanism installing the skeletogenic regulatory state (13). Conversely, *gcm* expression occurs exclusively in NSM cells, controlled by Delta/Notch signaling. During normal development, *gcm* and *alx1* are thus never coexpressed, and expression of either one or the other regulatory gene is carefully controlled by the upstream GRN circuitry.

A Boolean model of this subcircuit demonstrates this function (Fig. 1*E*). The double-negative gate circuit ("DN gate") activates expression of *alx1* in skeletogenic cells, thus promoting the skeletogenic cell fate (Fig. 1 *E1*). In NSM cells, Delta/Notch signaling ("D/N signal") activates expression of *gcm* (Fig. 1 *E2*). Thus, the upstream inputs essentially determine the spatial expression of each regulatory gene and the outcome of the switch. However, if the circuitry is altered in silico such that both regulatory genes respond to the same regulatory inputs, the result is quite different. If *gcm* and *alx1* are both activated downstream of Delta/Notch signaling, the model predicts oscillating expression of both regulatory genes and no stabilization of cell fates (Fig. 1 *E3*). That is, if two regulatory genes respond in the same way to the same regulatory input, this circuit cannot decide between the two possible outcomes. Although in this constellation this circuit produces bistable behaviors as predicted, this function is not useful for programs that determine the organization of the body plan. However, if a small bias is introduced in the control of gene expression, such that for example Delta/Notch signaling activates *gcm* before *alx1*, the subcircuit again produces deterministic outcomes such that Delta/Notch signaling always initiates expression of *gcm*, the product of which represses *alx1* expression (Fig. 1 *E4*). Given the combinatorial control of gene expression, differences in the timing and/or levels of regulatory gene expression even downstream of a shared regulatory input can be just as hardwired and deterministic as in the example shown here, as long as the *cis*-regulatory control systems of the two regulatory genes are not identical.

The important regulatory interactions in the mutual-repression circuit shown in Fig. 1*E* are interactions ② and ③, executing mutual repression. These interactions ensure that the two regulatory genes are never coexpressed. Because Alx1 and Gcm operate upstream as well as downstream in the GRN hierarchy, their exclusive expression is fundamental for the separation of the two regulatory states. Where the two cell fates are specified is determined by regulatory interactions ① and ④, the external inputs controlling expression of the two regulatory genes. Thus, the mutual-repression subcircuit does not make decisions, it only ensures that an earlier regulatory decision is executed in a binary manner, offering a deterministic response to alternative regulatory inputs.

Double-Negative Gate Subcircuit. A double-negative gate subcircuit again includes two regulatory genes with repressive functions, organized in tandem. In this circuit, an upstream regulatory gene encodes a repressor of a second regulatory gene, the product of which represses the expression of downstream genes. In consequence, expression of genes controlled by the double-negative gate is permitted wherever the first regulatory gene is expressed. In addition to the two repressors, this subcircuit receives broadly distributed activating inputs, whereas the spatial patterning function is determined by repression.

The example from the endomesoderm GRN controls the earliest patterning process in sea urchin development (1, 3, 4, 13). As shown in Fig. 1*F*, *pmar1* expression in skeletogenic cells is controlled by maternally localized inputs and leads to repression of *hesC* (13, 15). *HesC* expression itself is controlled not only by Pmar1, but also by maternal ubiquitous activators, leading to expression of *hesC* in all cells of the sea urchin embryo, except where *pmar1* is expressed. In turn, HesC controls the repression of *ets1/2*, *alx1*, and *tbr*, regulatory genes functioning upstream in the skeletogenic GRN (13). Thus, only where *pmar1* is expressed is activation of the skeletogenic GRN permitted, whereas everywhere else, HesC represses regulatory genes essential for this GRN.

A Boolean model of the subcircuit in Fig. 1*F* shows that expression of *hesC* and *ets1/2* downstream of a ubiquitous activator

initially occurs in all cells of the embryo. However, in skeletogenic cells, localized activators drive *pmar1* expression, leading to repression of *hesC*. In the absence of HesC, this subcircuit supports expression of the skeletogenic GRN, even after the transient expression of Pmar1 is turned off. Thus, this subcircuit constitutes a binary patterning device discriminating two exclusive regulatory states by expression of either one of two repressors.

In this subcircuit, the activation functions are mostly separate from the patterning functions. Thus, regulatory interactions ⑤, ⑥, and ⑧ are required for the activation of the skeletogenic GRN, whereas spatial expression is essentially determined by interactions ①, ②, ④, ⑦, and ⑨. This subcircuit is particularly useful in early development, where several maternally expressed ubiquitous transcription factors are available for control of gene expression. Because of the irreversible nature of repressive interactions, this subcircuit will install a stable expression pattern, even if the upstream interactions ① and ② are only transient. Thus, the transient availability of a localized activating input is sufficient to initiate the activity of this subcircuit, which then continues to control spatially restricted gene expression.

In summary, this analysis shows that, even though different types of subcircuit can be identified based on the architecture of regulatory interactions, circuit structure does not completely define circuit function. Both topology and number of regulatory interactions limit the range of subcircuit functions and contribute to its regulatory information. In addition, subcircuit function depends on the *cis*-regulatory logic by which these interactions are processed. In several instances, we predict *cis*-regulatory logic functions required for the operation of the subcircuit. Thus, evaluating the contribution of given regulatory features to developmental subcircuit function provides the means to assess regulatory information beyond the control of individual circuit nodes.

Regulatory Information in Circuitry Mediating Intercellular Signaling Interactions

Regulatory interactions across cell boundaries depend on signaling interactions, where regulatory information is transmitted from cells expressing signaling ligands to cells responding to the signal. Gene regulation in response to signaling interactions is executed by signal response transcription factors (SRTFs) that are regulated downstream of specific signaling pathways (see ref. 1 for discussion). In classical induction experiments, ectopically expressed signal ligands have been shown to induce ectopic expression of target genes. Recent insights, however, suggest that signaling interactions operate more complex regulatory functions. For instance, the control of context-specific target genes by given signaling interactions indicates that signaling interactions operate in combination with cell-fate–specific regulatory states. Furthermore, an increasing number of experimental observations show that SRTFs are required to prevent ectopic expression of target genes (19, 28, 29), opposite to what would be expected if signaling interactions control gene expression by inductive activation. Indeed, many SRTFs, such as, for example, Tcf/Lef (Wnt signaling), Su(H) (Delta/Notch signaling), and Gli (Hedgehog signaling), mediate a binary toggle switch function: They activate expression of target genes in response to the presence of the signal and repress target genes in absence of the signal (30).

The example in Fig. 2 demonstrates the regulatory information encompassed in such signaling systems. In sea urchins, *wnt1* is expressed in the anterior endoderm, signaling to adjacent posterior endoderm cells (24). In signal-receiving cells, Tcf associates with coactivator β-catenin and, together with a lineage-specific transcription factor, Eve, activates the expression of *hox11/13b* (12, 31). However, in ectodermal cells where no Wnt1 signal is received, *hox11/13b* is repressed by Tcf and its cofactor Groucho (31). If the regulatory input from Tcf is removed, conversely, expression of *hox11/13b* is activated by Eve alone. Thus, in the absence of Tcf, Eve is sufficient to activate *hox11/13b* expression. The corresponding Boolean model shows the gene-expression pattern operated by this circuit (Fig. 2).

The regulatory information in signaling interaction toggle switch circuits is contributed by activating inputs from a context-specific transcription factor (Eve) and from Tcf/β-catenin (Fig. 2, *Lower Left*, inputs ① and ②, and by repression through Tcf/Groucho (Fig. 2, *Lower Left*, input ③). In all systems where mutation of binding sites for an SRTF leads to ectopic expression of the target gene, the activating form of SRTF has to operate in OR logic to other activating inputs. These context-specific transcription factors are sufficient to activate gene expression, however, only in the absence of the repressive form of SRTF. Thus, context-specific transcription factors can be expressed at earlier developmental times or within a broader domain without inducing expression of signal-regulated genes. Only when and where SRTF-mediated repression is removed by signaling interaction, SRTF target genes start to be expressed. We here conclude that this form of inductive

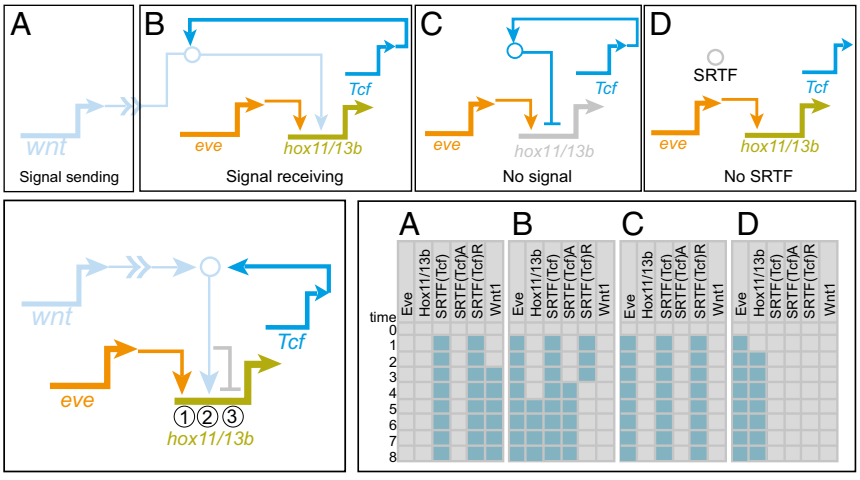

Fig. 2. Structure and function of signaling interactions mediated by toggle switch circuitry. Components of this subcircuit as expressed in different cellular domains are shown on top. Cells expressing *wnt* (*A*) signal to adjacent cells (*B*), where *hox11/13b* is expressed downstream of Tcf/β-catenin and Eve. In cells not receiving Wnt signaling (*C*), *hox11/13b* is repressed by Tcf/Groucho, whereas in the absence of the SRTF (*D*), *hox11/13b* expression is activated by Eve. The complete subcircuit is shown in *Lower Left*, and gene expression in each condition a–d as determined by Boolean modeling is shown in *Lower Right*. SRTF, signal response transcription factor. For model equations, see *SI Appendix*.

signaling is broadly used, where differential regulation of signaling target genes depends not on inductive activation, but on inductive derepression. In essence, gene regulation by signaling interactions is thus determined through interaction ③. The context specificity of target gene regulation downstream of the signal is mediated by interaction ①, which is necessary for target gene activation. As a special feature of this circuit, interactions ② and ③ are both mediated by the SRTF and operate through the same binding sites, thus ensuring the Boolean regulation of its target genes.

Information Encoded in the Hierarchical Organization of GRNs

Development is organized hierarchically in that earlier events will determine what follows. Developmental hierarchy is a result of the unidirectionality and irreversibility of regulatory interactions. In addition to development, network hierarchy is also of profound importance for evolutionary processes, because the consequence of alterations in regulatory linkages or regulatory nodes will depend on their position within a GRN (1, 32, 33). Thus, although change in a patterning circuit operating downstream in a GRN may affect the location of given cell types, change in upstream patterning circuits may alter the position or identity of entire body parts. Furthermore, the intrinsic structure of GRNs may bias the chance for evolutionary change, such that certain network architectures such as kernels show conservation over large evolutionary distances, whereas other network components, in particular signaling interactions, can be subject to more rapid changes (34–36). These observations suggest that the organization of subcircuits within GRN hierarchy is an important determinant of network information that impacts development as well as evolutionary process.

Fig. 3 shows the occurrence of each type of subcircuit within the sea urchin endomesoderm GRN. Thus, the endomesoderm GRN is composed of at least 23 subcircuits, including all types defined above, except the incoherent feedforward subcircuit. Almost every node in this model contributes to at least one, but often several different, subcircuits. The most frequently deployed subcircuits are the coherent feedforward and the positive feedback subcircuits. Much rarer encounters are the mutual repression, double-negative gate, and the community-effect subcircuit. Interestingly, the two frequently deployed subcircuits both operate alternative aspects in cell-fate–specification GRNs. Thus, positive feedback subcircuits provide stabilization of gene expression, whereas coherent feedforward circuits provide quality control, ensuring that downstream genes are only expressed where all upstream inputs are present. Other less-represented subcircuits all contribute to developmental patterning, a function that is also executed by signaling interactions, which occur frequently within the endomesoderm GRN (Fig. 3) and all other developmental GRNs.

The perhaps most unexpected but perfectly logical result of this analysis is the extent of overlap between individual subcircuits within this network (Fig. 3). Thus, for example, *alx1* in the skeletogenic GRN (Fig. 3, PMC) is simultaneously involved in a coherent feedforward circuit, a mutual-repression circuit, and the double-negative gate circuit. In the aboral mesoderm (Fig. 3, aboral NSM), *gcm* participates in a mutual repression with *alx1*, in a positive feedback with *gatae* and *six1/2*, in a coherent feedforward circuit with multiple regulatory genes controlling differentiation genes, and is itself controlled by Delta/Notch toggle switch signaling. However, not only network nodes are multifunctional. Individual regulatory interactions may also participate in several subcircuit functions, particularly those contributing to positive feedbacks and coherent feedforward subcircuits. As shown in

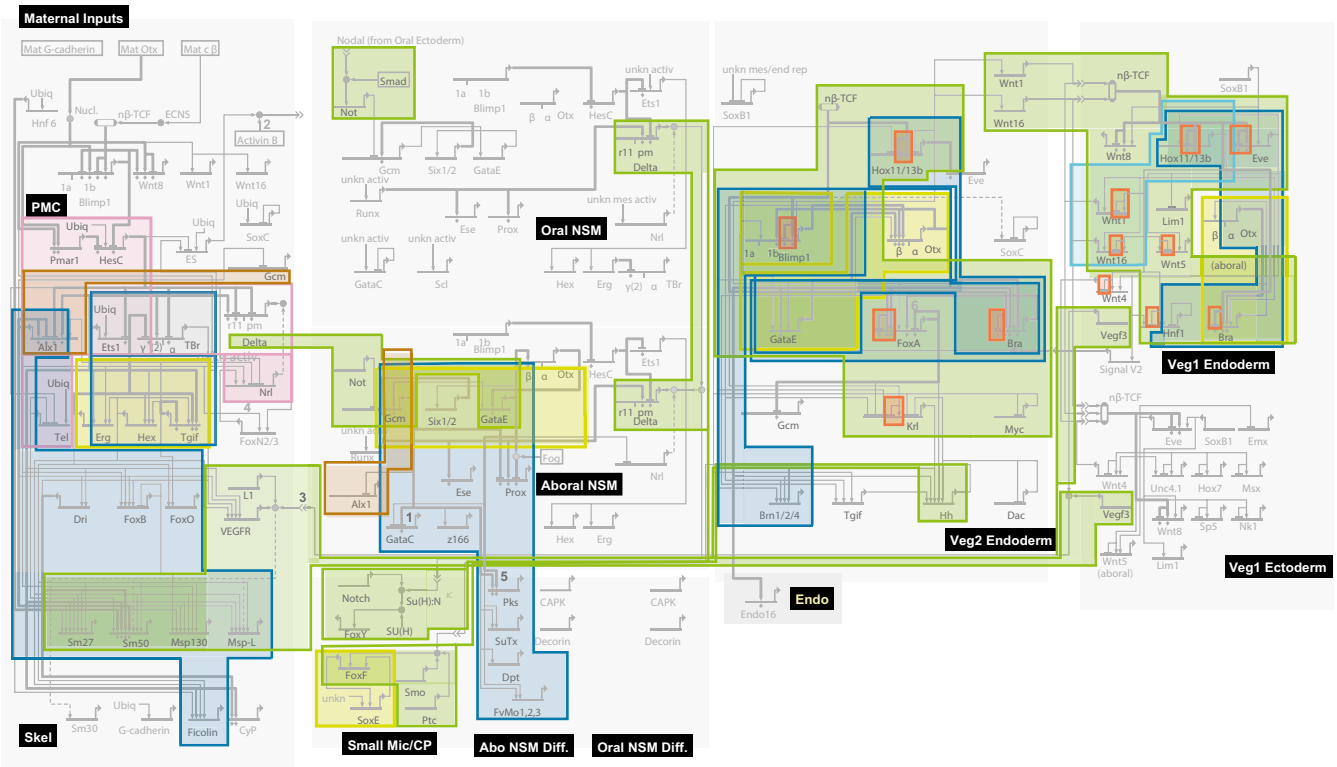

Fig. 3. Distribution of subcircuits in the endomesoderm GRN. Subcircuits of each type identified in the endomesoderm GRN model are color-coded as follows: pink, double-negative gate; dark blue, coherent feedforward subcircuit; light blue, community-effect subcircuit; yellow, positive feedback subcircuit; green, signaling interaction; red, toggle switch circuitry; and brown, mutual-repression subcircuit. For recent updates of the endomesoderm GRN model, see grns.biotapestry.org/SpEndomes/.

Fig. 3, activating coherent feedforward circuits often occur not in isolation, but are organized into clusters of multiple subcircuits of this type, thus connecting genes vertically throughout the GRN hierarchy.

Individual subcircuits within a GRN therefore do not operate in a linear hierarchy, but are strongly intertwined. In addition, network nodes at the intersection of subcircuits are often controlled by signaling interactions. Multiple intercellular signaling events connect individual cell-fate GRNs within the endomesoderm network. Thus, the specification of different mesodermal cell fates are connected through Delta/Notch signaling, the specification of endodermal cell fates through Wnt signaling, the specification of coelomic pouches to the specification of foregut endoderm through Hh signaling, the specification of oral mesoderm to the oral/aboral axis through Nodal signaling, and the differentiation of skeletogenic cells to the specification of posterior endoderm and perianal ectoderm through Vegf signaling (Fig. 3). Each signaling interaction may control a large number of genes within signal-receiving GRNs (24). Among the target genes responding to signaling interactions or also other patterning subcircuits are often precisely those regulatory genes that are strongly connected to other genes in the GRN, through multiple network subcircuits.

This structural organization of subcircuits has consequences for development as well as evolution. In development, strong interactions between individual subcircuits ensure the coordinated activation of developmental functions. For an assessment of regulatory information contributed by individual nodes and linkages, evaluating the number of subcircuits they participate in will be a relevant measure. Thus, nodes and linkages positioned at the intersection of subcircuits will control many functions of a GRN, whereas other nodes control specific subfunctions, such as particular patterning processes. Such analysis will enable the assessment of regulatory information of given regulatory features in respect to the overall function of GRNs. Furthermore, we can assume that the organization of subcircuits within GRNs will also affect the rate of evolutionary change at given network nodes. Overlapping subcircuit structures, particularly at upper levels of GRN hierarchy, constrain the downstream developmental process and its alteration in evolution. As a consequence, evolutionary change of nodes and interactions at the intersection of several subcircuits can have pleiotropic effects on GRN function and be less favorable than evolutionary changes of other GRN nodes. How general these patterns of GRN organization are and how they contribute to regulatory information in development and evolution will be revealed upon the solution of additional developmental GRNs.

Conclusion

All regulatory information in GRNs is ultimately encoded in individual *cis*-regulatory elements that compute upstream inputs into downstream outputs. They determine how individual network nodes are wired within a developmental GRN, how individual subcircuits are organized to process developmental functions, and how subcircuits are connected within the overall network hierarchy. As this analysis shows, every aspect of regulatory information encoded within regulatory sequence, including the identity of regulatory inputs as well as the regulatory logic by which inputs are computed, has potentially profound consequences on the functionality of developmental GRNs. Thus, although mechanistically the source of regulatory information has to be found at the *cis*-regulatory level, the significance of each *cis*-regulatory feature for the overall developmental process, as envisioned in the 1990s (37), becomes apparent only when considered in the context of all levels of network organization.

ACKNOWLEDGMENTS. We thank Deanna Thomas for her help in the preparation of the figures. This work was supported by National Institutes of Health Grant HD 037105.

1. Peter IS, Davidson EH (2015) *Genomic Control Process, Development and Evolution* (Academic/Elsevier, London).
2. Davidson EH, et al. (2002) A genomic regulatory network for development. *Science* 295(5560):1669–1678.
3. Davidson EH, Levine MS (2008) Properties of developmental gene regulatory networks. *Proc Natl Acad Sci USA* 105(51):20063–20066.
4. Peter IS, Davidson EH (2009) Modularity and design principles in the sea urchin embryo gene regulatory network. *FEBS Lett* 583(24):3948–3958.
5. Alon U (2007) Network motifs: Theory and experimental approaches. *Nat Rev Genet* 8(6):450–461.
6. Pimanda JE, et al. (2007) Gata2, Fli1, and Scl form a recursively wired gene-regulatory circuit during early hematopoietic development. *Proc Natl Acad Sci USA* 104(45):17692–17697.
7. Narula J, Smith AM, Gottgens B, Igoshin OA (2010) Modeling reveals bistability and low-pass filtering in the network module determining blood stem cell fate. *PLOS Comput Biol* 6(5):e1000771.
8. Balaskas N, et al. (2012) Gene regulatory logic for reading the Sonic Hedgehog signaling gradient in the vertebrate neural tube. *Cell* 148(1-2):273–284.
9. Cotterell J, Sharpe J (2010) An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients. *Mol Syst Biol* 6:425.
10. Peter IS, Faure E, Davidson EH (2012) Predictive computation of genomic logic processing functions in embryonic development. *Proc Natl Acad Sci USA* 109(41):16434–16442.
11. Peter IS, Davidson EH (2010) The endoderm gene regulatory network in sea urchin embryos up to mid-blastula stage. *Dev Biol* 340(2):188–199.
12. Peter IS, Davidson EH (2011) A gene regulatory network controlling the embryonic specification of endoderm. *Nature* 474(7353):635–639.
13. Oliveri P, Tu Q, Davidson EH (2008) Global regulatory logic for specification of an embryonic cell lineage. *Proc Natl Acad Sci USA* 105(16):5955–5962.
14. Sorrells TR, Johnson AD (2015) Making sense of transcription networks. *Cell* 161(4): 714–723.
15. Revilla-i-Domingo R, Oliveri P, Davidson EH (2007) A missing link in the sea urchin embryo gene regulatory network: HesC and the double-negative specification of micromeres. *Proc Natl Acad Sci USA* 104(30):12383–12388.
16. Rafiq K, Shashikant T, McManus CJ, Ettensohn CA (2014) Genome-wide analysis of the skeletogenic gene regulatory network of sea urchins. *Development* 141(4):950–961.
17. Saunders LR, McClay DR (2014) Sub-circuits of a gene regulatory network control a developmental epithelial-mesenchymal transition. *Development* 141(7):1503–1513.
18. Ransick A, Davidson EH (2012) Cis-regulatory logic driving glial cells missing: Self-sustaining circuitry in later embryogenesis. *Dev Biol* 364(2):259–267.
19. Ransick A, Davidson EH (2006) cis-regulatory processing of Notch signaling input to the sea urchin glial cells missing gene during mesoderm specification. *Dev Biol* 297(2):587–602.
20. Materna SC, Ransick A, Li E, Davidson EH (2013) Diversification of oral and aboral mesodermal regulatory states in pregastrular sea urchin embryos. *Dev Biol* 375(1):92–104.
21. Croce JC, McClay DR (2010) Dynamics of Delta/Notch signaling on endomesoderm segregation in the sea urchin embryo. *Development* 137(1):83–91.
22. Gurdon JB (1988) A community effect in animal development. *Nature* 336(6201): 772–774.
23. Bolouri H, Davidson EH (2010) The gene regulatory network basis of the "community effect," and analysis of a sea urchin embryo example. *Dev Biol* 340(2):170–178.
24. Cui M, Siriwon N, Li E, Davidson EH, Peter IS (2014) Specific functions of the Wnt signaling system in gene regulatory networks throughout the early sea urchin embryo. *Proc Natl Acad Sci USA* 111(47):E5029–E5038.
25. Johnston RJ, Jr, et al. (2011) Interlocked feedforward loops control cell-type-specific Rhodopsin expression in the Drosophila eye. *Cell* 145(6):956–968.
26. Briscoe J, Small S (2015) Morphogen rules: Design principles of gradient-mediated embryo patterning. *Development* 142(23):3996–4009.
27. Jaeger J (2011) The gap gene network. *Cell Mol Life Sci* 68(2):243–274.
28. de-Leon SB, Davidson EH (2010) Information processing at the foxa node of the sea urchin endomesoderm specification network. *Proc Natl Acad Sci USA* 107(22):10103–10108.
29. Ozdemir A, Ma L, White KP, Stathopoulos A (2014) Su(H)-mediated repression positions gene boundaries along the dorsal-ventral axis of Drosophila embryos. *Dev Cell* 31(1):100–113.
30. Barolo S, Posakony JW (2002) Three habits of highly effective signaling pathways: principles of transcriptional control by developmental cell signaling. *Genes Dev* 16(10):1167–1181.
31. 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(1):252–267.
32. Peter IS, Davidson EH (2011) Evolution of gene regulatory networks controlling body plan development. *Cell* 144(6):970–985.
33. Erwin DH, Davidson EH (2009) The evolution of hierarchical gene regulatory networks. *Nat Rev Genet* 10(2):141–148.
34. Davidson EH, Erwin DH (2006) Gene regulatory networks and the evolution of animal body plans. *Science* 311(5762):796–800.
35. Hinman VF, Davidson EH (2007) Evolutionary plasticity of developmental gene regulatory network architecture. *Proc Natl Acad Sci USA* 104(49):19404–19409.
36. 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(23):13356–13361.
37. Davidson EH (1990) How embryos work: A comparative view of diverse modes of cell fate specification. *Development* 108(3):365–389.