HHS Public Access

Author manuscript

Mar Genomics. Author manuscript; available in PMC 2020 April 01.

Published in final edited form as:

Mar Genomics. 2019 April ; 44: 3–12. doi:10.1016/j.margen.2019.01.003.

Encounters across networks: windows into principles of genomic regulation

Ellen V. Rothenberg

Division of Biology & Biological Engineering, California Institute of Technology, Pasadena, California USA

Abstract

Gene regulatory networks account for the ability of the genome to program development in complex multi-cellular organisms. Such networks are based on principles of gene regulation by combinations of transcription factors that bind to specific cis-regulatory DNA sites to activate transcription. These cis-regulatory regions mediate logic processing at each network node, enabling progressive increases in organismal complexity with development. Gene regulatory network explanations of development have been shown to account for patterning and cell type diversification in fly and sea urchin embryonic systems, where networks are characterized by fast coupling between transcriptional inputs and changes in target gene transcription rates, and crucial *cis*-regulatory elements are concentrated relatively close to the protein coding sequences of the target genes, thus facilitating their identification. Stem cell-based development in post-embryonic mammalian systems also depends on gene networks, but differs from the fly and sea urchin systems. First, the number of regulatory elements per gene and the distances between regulatory elements and the genes they control are considerably larger, forcing searches via genome-wide transcription factor binding surveys rather than functional assays. Second, the intrinsic timing of network state transitions can be slowed considerably by the need to undo stem-cell chromatin configurations, which presumably add stability to stem-cell states but retard responses to transcription factor changes during differentiation. The dispersed, partially redundant *cis*-regulatory systems controlling gene expression and the slow state transition kinetics in these systems already reveal new insights and opportunities to extend understanding of the repertoire of gene networks and regulatory system logic.

Keywords

Gene regulatory networks; cis-regulatory elements; transcription factor binding; network response timing; chromatin; combinatorial logic

Construction of benchmark gene regulatory networks for development

Gene regulatory networks as a decoder of development: outline of a legacy

This article examines the theory of gene regulatory networks for development that is a central legacy of Eric H. Davidson’s work. The conceptual framework for this theory has

evroth@its.caltech.edu.

proven to be prescient, and it has become an influential guide to thinking about development in the past 20 years in many more recently-studied systems, far beyond those studied directly by Davidson himself (Hobert 2006; Singh and Pongubala 2006; Sinner et al. 2006; Georgescu et al. 2008; Laslo et al. 2008; Ririe et al. 2008; Arda et al. 2010; Swiers et al. 2010; Bruex et al. 2012; Grocott et al. 2012; Kueh and Rothenberg 2012; Monteiro 2012; Rottinger et al. 2012; Arda et al. 2013; Pires et al. 2013; Clark et al. 2014; Hobert 2014; Wang et al. 2014; Simoes-Costa and Bronner 2015; Martik et al. 2016; Parker et al. 2016; Buckingham 2017; Charney et al. 2017; Laurent et al. 2017; Longabaugh et al. 2017; Ramirez et al. 2017). However, here the hidden assumptions of the theory are examined in light of developmental kinetics of systems with very different properties, and in engagement with new results emerging from genome-wide analyses of transcription factor deployment in specific cell types.

Eric H. Davidson was deeply influenced by Theodor Boveri's work [rev by (Satzinger 2008; Arnone et al. 2018; Caianiello 2018)], especially by Boveri's inference that embryonic development depends on the participation of a full complement of chromosomes (Laubichler and Davidson 2008). The question Davidson considered central was how the process details of embryogenesis could be encoded in the physical constituents of the genome. A century after Boveri's work, vastly more was known about chromosomes, genes and their products, i.e. the genomic "hardware", but for many, the "software" capable of coding for a complex, emergent process such as development still seemed difficult to fathom. Eric Davidson argued strongly that the genetic coding of development can be understood comprehensively because it operates through hierarchical transcriptional regulatory networks (Davidson et al. 2002a; Davidson et al. 2003; Davidson 2006; Davidson and Erwin 2006; Erwin and Davidson 2009; Davidson 2010; Peter et al. 2012; Peter and Davidson 2015). Thus, he pointed out that the way the genome encodes the program for development is through the matching of the genomic *cis*-regulatory sequences that control any gene to be activated with the DNA-binding specificities of transcription factors that will control it, which are themselves encoded in the genome and are themselves controlled by similar genomic *cis*-regulatory sites. The potential for such a network to increase the complexity of gene expression patterns progressively in time is an outcome of the distinct specificities of different transcription factors, which allow new DNA sequences to be used for regulation as these factors become expressed, and the rule that more than one transcription factor affects the expression of a given gene, so that the number of regulatory combinations increases steeply as progressively more factors become active.

Taking these "information science" principles together with basic embryological boundary conditions, Davidson showed how the forward drive of development and the generation of complexity from simplicity could be explained [e.g. in (Davidson 1990; Davidson 1993; Davidson 2006; Davidson 2010; Peter and Davidson 2015)]. The zero-time point of this process is fixed: it is the state of a fertilized egg, pre-furnished with a limited signaling transduction machinery capable of activating a small number of pre-existing transcription factors. Once this previously existing set of transcription factors was triggered to induce expression of new factors, then the quorum needed to bind and activate some new regulatory site in the genome could be filled, the gene regulated by that site should then be activated, and if the product was a transcription factor, it would then join the mix for the next round of

target gene regulation. During cleavage, any asymmetric signaling that differentially affected one of two daughter cells could trigger activation of additional specific transcription factors, via known signal transduction biochemistry. Because this response would initiate only in the daughter cell that received the signal, the two daughters would initiate different endogenous regulatory cascades. This would not only make them express different genes but also allow them to begin to signal asymmetrically, each to one another's descendants, triggering additional steps in the cascade of regulatory differentiation.

The genomic encoding of this process would reside (1) in the sequences of the genes encoding the transcription factors, determining their binding specificities and determining which would be signal-dependent rather than constitutively active, (2) in the sequences of each of the *cis*-regulatory elements that these factors should bind, and (3) in the genomic linkages of these binding sites to other coding sequences, determining what genes would be regulated. In principle, a mutation that changed the conditions in which a transcription factor was expressed, or a mutation that changed its binding site in a *cis*-regulatory element, should have a directly predictable effect, both on binding of the DNA by the affected transcription factor and on the expression patterns of its targets. The fundamentally progressive nature of transcriptional activation and specialization would be affected through all the descendants of the cell in which transcription factor activity was altered.

This comprehensive view that Davidson proposed equated cell type with the combination of transcription factors expressed in the cell (= "regulatory state") and explained such cellular identities in terms of cumulative past transcription factor activities in the cells' progenitors. In framing each time point of development this way, Davidson threw down the gauntlet to those who consider developing embryonic cells as signal-responsive but basically obscure "black boxes". His vision challenged the developmental biology field to reach for predictiveness in modeling the development of organisms via deterministic regulatory networks.

Historical context of developmental gene network models: Systems and logic

Gene regulatory networks were envisioned and discussed in theoretical terms (Britten and Davidson 1969) well before concrete examples of the molecular mechanisms were understood, and decades before adequate molecular biology tools were available to dissect development in multicellular organisms. Bacterial and especially phage genetics were pivotal for the validation of a gene network view of biological decisions, because they were genetically simple and accessible via some of the earliest tools available (Jacob and Monod 1961; Monod et al. 1963; Ptashne et al. 1980; Johnson et al. 1981). A foundational advance was the discovery that the phage lambda lysogeny control system was based on mutual repression, which in fact implemented a logic gate based on defined gene products and defined DNA regulatory sequences. This demonstrated the possibility that complex biological decision-making "software" could be defined in molecular terms (Ptashne 1992). In other words, it predicted that the logic of developmental genetics was implemented by gene expression biochemistry and encoded in the biophysics of transcription factor binding to DNA.

Logic has been integral to the late 20th century field of “Genetics” in both approach and analytical methods. To use genetics, powerful phenotypic screens are first cleverly devised to isolate extremely rare but informative mutants that affect some endpoint. Once mutants are identified, the pathway architecture through which the products of the mutant genes normally interact is inferred from the phenotypes of these mutations in combination. Thus, classical genetics always starts from a clearly observable phenotype and identifies the most potent mechanistic components responsible for it by working backwards through the impacts of different mutational perturbations, not by biochemical measurements on a “normal” system. It makes minimal assumptions about completeness and its definition of important pathway components does not even depend on the kinds of gene products that may be involved, a question that is deferred until after they are identified. These underlying principles are profoundly different from those of modern systems biology and functional genomics, a contrast pivotal for what follows. But for biological systems that can be screened in large numbers and with readouts that can be scored easily at high throughput, genetics is immensely powerful at finding numerous components that ***need to be included*** in a gene network, even if they do not comprise all the components that ***make it work***.

Thus, the first complex gene networks for multicellular organism development were inferred for the initial patterning of the early *Drosophila melanogaster* embryo through gap gene and pair-rule gene expression control, the components of which had been identified via genetic screens of embryos mutated into extreme, nonviable mutant genotypes (Nusslein-Volhard and Wieschaus 1980; Hülskamp et al. 1990; Hülskamp and Tautz 1991; Stanojevic et al. 1991; Small et al. 1992; Struhl et al. 1992; Schulz and Tautz 1994; Morisato and Anderson 1995; Ingham 2016; Wieschaus and Nusslein-Volhard 2016). The idea of a network responsible for such a process was a great conceptual improvement over the popular interpretation of single transcription factors acting as “master regulators” (Weintraub et al. 1991; Gehring 1996). The use of this embryonic system was important not only because it was handy for genetics, but also because it was especially favorable for identifying mutations affecting transcription factor genes. Fortuitously, because the *Drosophila* embryo remains a syncytium until initial axis patterning has occurred, the role of inter-cellular signaling is initially less salient than in some other developmental systems. Accordingly, among the first gene products inferred to control patterning, there was a preponderance of transcription factors (TF). Because roles of TFs in transcriptional regulatory function are known to depend on their DNA binding, and because their genomic actions reflect their distinctive, individual target sequence specificities, a straightforward prediction could be made as to how mutations in TF genes should cause their effects. Levine and collaborators, among others, capitalized on this prediction in a succession of elegant papers that anchored the links (nodes) in the embryo patterning network directly to TF-target DNA binding. The diffusional properties of locally synthesized TFs along the axes of the syncytial embryos allowed expression of a “patterned” target gene to be interpreted as a pure biophysical outcome of TF concentration vs. binding affinity for the target sites (Small et al. 1991; Stanojevic et al. 1991; Ip et al. 1992; Small et al. 1992; Jiang and Levine 1993; Schulz and Tautz 1994; Stathopoulos and Levine 2002a; Stathopoulos and Levine 2002b; Stathopoulos et al. 2002). Parallel work on later events in Drosophila development, for example by Sean Carroll, emphasized the power of transcriptional regulation networks for patterning (Kim et

al. 1996; Weatherbee and Carroll 1999). In many ways, the fly embryo patterning network model was the triumphant flowering of a conceptual framework started by ideas of Britten and Davidson, implemented by the mechanisms first demonstrated in the phage λ lysogeny control system.

In the 1990's and 2000's, the Davidson group used different research methodology but convergent philosophy to study the sea urchin *Strongylocentrotus purpuratus*, ultimately constructing an exceptionally complete model of the regulatory gene network circuitry in that formed all the cell types in the endomesodermal half of the embryo, well into gastrulation (Kirchhamer and Davidson 1996; Yuh et al. 1998; Davidson et al. 2002b; Oliveri et al. 2002; Oliveri et al. 2008; Peter and Davidson 2010; Peter and Davidson 2011; Peter et al. 2012). Interestingly, in this animal classical genetics were not used; instead, the basis for gene identification was transcriptome enrichment screening, and the perturbations were tested by gene cloning, gene injection, antagonist injection, and blastomere transplantation rather than by mutagenesis and organismal phenotypic screening. This meant that Davidson's group built the sea urchin gene regulatory network bottom-up, very concretely based on the regulatory DNA sequences that guided correct expression of individual injected genes, rather than on abstract epistasis relationships among genes of unknown function. However, the structure-function relationships assumed to operate at TF-target DNA interaction sites were essentially the same as in the *Drosophila* early embryo work, and the outcome of TF perturbation in organismal phenotype was similarly assumed to be clear. Numerous links to individual cis-regulatory elements for genes in the network were defined biochemically and shown to be capable of mediating the logic functions ascribed to them. At the end, the fully-developed model was tested by formal Boolean modeling and shown to be logically sufficient to account for the overwhelming majority of observed gene expression dynamics in the living embryo (Peter et al. 2012).

The networks patterning the fly embryo and specifying the sea urchin endomesoderm emerged as two great triumphs of reading the logic of molecular biology in development. These successes suggested that the rules for gene network operation in these systems could be universal for creating and maintaining tissue and cell type complexity in multicellular organisms. The ways that these two gene network systems could be solved, however, were significantly assisted by features of the biological systems themselves, which made them compatible with available molecular biology technologies and yielded readily interpretable phenotypes. One non-trivial feature, often underestimated, was the fact that early embryonic blastomeres (in many organisms) usually change fate without dying when subjected to regulatory perturbation. The ability to recover and track cells that were responding abnormally was vital to open a clear window on the inner working of the fate-decision networks. Another key feature was the way the anatomical patterning of gene expression in two or three dimensions in these systems could rapidly reveal combinatorial effects, for example, by showing that altering one of several inputs caused a distortion of the expression pattern in one dimension without effects in another. A third critical feature was the role of time in these systems, discussed below.

Gene network logic principles at the molecular level

Combinatoriality of gene control: network nodes, not pathways

Understanding of the transcriptional regulatory biochemistry of multicellular organisms is crucial for exploring the rules for gene network operation. Two major paradigms initially shed most light on how input-output logic was mediated by *cis*-regulatory elements: first, the developmentally regulated patterning genes in non-vertebrate systems just described, such as flies and sea urchins, and second, immune response genes in mammals. These were systems studied by different groups of scientists, many of them virtually ignorant of work in the alternative type of system, but who converged on discovery of the same essential principles. Both established the central point that the decision to express a target gene does not depend on a single transcription factor but rather on the combinatorial interaction of several transcription factors. This was codified in an important review (Arnone and Davidson 1997) which, as a rare synthesis, drew from both types of systems.

In the case of the invertebrate developmental patterning systems, timing of gene expression was known *a priori* to be rapidly changing, and the question of interest was how the spatial boundaries of a gene's expression were set in the embryo. In a beautiful demonstration of near-Boolean logic, expression was shown to appear in domains of overlap between TFs that provided positive inputs to "AND" logic gates controlling the gene of interest, and bounded by territories in which specific negative factors involved in "AND NOT" logic were expressed (Small et al. 1991; Stanojevic et al. 1991). An analogous process patterned the dorsal/ventral axis . The factors themselves were localized in the embryo by a combination of cell lineage (nuclear lineage, in the *Drosophila* syncytial blastoderm) and current signal exposure, both tightly associated with prior spatial elements of the developing embryo.

In the case of mammalian immune response genes, the genes of interest were usually cytokine genes encoding extremely powerful intercellular communication factors, the expression of which could have explosive consequences for the behavior of many other cells in the organism. For analysis of the *cis*-regulatory systems of such genes, the issue was to determine the correct combination of circumstances that could determine the time, not space *per se*, in which the gene should be allowed to be expressed. Here the emphasis was on factors that are activated by distinct signaling pathways in the same cell, so that the gene of interest could be activated only in the right combination of signaling circumstances (Thanos and Maniatis 1995; Rothenberg and Ward 1996; Ramirez-Carrozzi et al. 2009), or in the right lineage of activated cells under appropriate signaling (Avni and Rao 2000; Lee et al. 2000; Murphy and Reiner 2002; Nakayama and Yamashita 2008; Sekimata et al. 2009; Yang et al. 2011; Ciofani et al. 2012; Oestreich and Weinmann 2012).

In both embryological and immunological cases, the logic of *cis*-regulation was found to reflect the need to establish occupancy of the target gene's *cis*-regulatory system by a combination of TFs, not simply by one alone. A logical transformation of the constituent inputs had to be invoked to activate transcription. Thus, in both kinds of cases, the global expression pattern of the target gene is dependent on, but distinct from, the expression patterns of each of the input regulators.

Quorum binding at the *cis*-regulatory element as a microprocessor

As emphasized by Levine, Davidson, Maniatis, and others, the physical switch for transcriptional regulation in a gene regulatory network is the *cis*-regulatory element in the DNA. As a template for TF binding in a particular 3-D space, the unique sequence of the *cis*-regulatory element for a given gene determines how the conditions needed for gene expression have to be satisfied, and the engagement of this *cis*-regulatory element by transcription factors *in vivo* is the sensor. A gene can have more than one *cis*-regulatory element, allowing its expression to be induced under more than one set of conditions, but the contribution to gene activity from each element is determined by the output from the regulatory element ***as a unit***. This means that individual TF molecules provide contributions to a "quorum", which at each *cis*-element is determined by the combination of transcription factor binding sites it contains. Binding results in output activity only where the quorum is complete.

These principles were illustrated in the 1990's by laborious dissection of specific transcriptional regulatory elements, and they have been repeatedly reconfirmed, as genome-wide methods for mapping chromatin and transcription factor states have been applied to gene regulation. Functionally active regulatory elements usually also become "open" to nuclease or transposase activity in chromatin (measured by DNase hypersensitivity, micrococcal nuclease sensitivity, or ATAC-seq), as the collection of bound TFs protects the DNA less efficiently than the nucleosomes that they displace (Elgin 1988; Jenuwein et al. 1993; Boyes and Felsenfeld 1996; Takemoto et al. 2000; Rao et al. 2001; Follows et al. 2006; Hoogenkamp et al. 2009; Buenrostro et al. 2013). Biochemically, at least some cases show that binding of individual transcription factors at regulatory elements is only stabilized when others are present as well (discussed below). Even in cases where factors are able to bind stably at incompletely occupied elements, success at assembling a quorum is probably "certified" by recruitment of a coactivator such as Ep300 or Crebbp and Mediator complex [e.g., (Vahedi et al. 2012)]. This assembly then enables the enhancer to communicate with the promoter.

Davidson recognized the analogies between *cis*-regulatory element operation and computation in the control of single genes, beginning years before developing gene regulatory network models (Yuh et al. 1996; Yuh et al. 1998; Yuh et al. 2001; Istrail and Davidson 2005). In fact, this was an essential point. It was the need for multiple inputs working combinatorially to control each output that made gene networks -- rather than pathways -- conceptually essential to explain development. In turn, the ability to observe clear spatiotemporal effects in the fly and sea urchin embryo systems, as an immediate consequence of altering those inputs, was essential to reveal the logic operations for progressive cell type diversification within gene regulatory networks.

Inter-network comparison parameters: Role of time in gene network operation

In the gene regulatory systems that have been successfully explained, changes in expression of target genes follow directly upon changes in their inputs, i.e. changes in completeness of

their transcription factor quorums. The biochemistry of transcription at open gene loci makes this a fast response, a tight temporal coupling. Thus, in mammalian immune cells, activation-dependent transcription factors can be mobilized to the nucleus within minutes and new transcripts of target genes like *Il2* detected from formerly silent loci within a quarter or half an hour. In *Drosophila melanogaster*, important patterning genes begin being expressed within the 13th nuclear cycle which is less than 15 min long, and early patterning gene expression is extremely dynamic. In the Anterior/Posterior axis of *Drosophila*, the transcriptional activation speed even outpaces the logic as expression begins in broad initial domains that are then refined to thin, sharp stripes, displaying their best-known expression patterns as repression begins to affect expression, but only for a short time before fading out completely. “Step times” of only 15 min from factor activation to target activation are seen (Clark and Akam 2016). In the sea urchin embryo, living at temperatures colder than those of *Drosophila* or mammals by 10–20°C respectively, responses take longer to become detectable, as expected for known effects of temperature on transcription and translation rates (Ben Tabou de-Leon and Davidson 2009). Still, here too from expression of the RNA encoding the TFs to appearance of transcripts from the target genes of those TFs, the step time is not longer than 3 h (Bolouri and Davidson 2003; Peter et al. 2012).

The rapid and dynamic response kinetics of these genes implies (1) that the regulatory elements are open for occupancy by changing sets of factors, and (2) that these changing occupancy patterns encounter no obstacle to altering transcriptional initiation rates. Thus, TFs are free to read the genome and cause changes in transcription immediately; indeed, the TFs expressed in a cell in this kind of situation would completely define its regulatory state.

This makes sense for an early non-mammalian embryo, where the organism needs to differentiate multiple cell types quickly, or in an immune response where strong but accurately controlled reactions may be needed within minutes or hours. In reality, however, another level of regulation can be needed to earmark the subset of genes that will be susceptible to these immediate changes in input. In *Drosophila*, early embryonically expressed genes are pre-marked by binding of a common zinc finger TF, Zelda, before their specific regulators are fully assembled (Harrison et al. 2011). Another contributor to the fast response of zygotic genes in *Drosophila* may be the widespread pre-poising of RNA polymerase at the promoters of many developmentally relevant genes well before they become active. Thus, these polymerases require only release from a tethering complex in order to begin productive transcription (Zeitlinger et al. 2007). In inflammatory responses in mammalian cells, analogously, a first wave of target genes for inflammatory signals is also pre-sensitized for responsiveness, in this case by promoter DNA sequences that specifically disfavor nucleosome assembly, again lowering the threshold for RNA polymerase recruitment (Ramirez-Carrozzi et al. 2009).

If the step that couples transcription factor occupancy to RNA polymerase activation operates faster than the steps involved in TF binding to *cis*-regulatory elements, then it can be simplified to a non-regulatory constant when modeling the dynamics of these network linkages. Thus, assuming that the coupling is fast, target gene expression dynamics can be predicted in terms of TF binding dynamics alone. This prediction has been the basis for most of the successful gene regulatory network models for development. Peter, Faure and

Davidson in fact needed the assumption of a stereotypical, invariant step time when they generated their highly successful predictive model of the early embryonic gene regulatory network in the sea urchin *Strongylocentrotus purpuratus* (Peter et al. 2012). The model used an innovative variant of Boolean logic with an absolute time scale to predict the dynamics of gene activation and repression. The model calculated expression of ~50 individual genes encoding TFs and signaling components needed as regulatory inputs to other genes, hour by hour and cell by cell, over the whole sea urchin embryo endomesoderm up to 30 h after fertilization, based primarily on the known input requirements of those individual genes. The assumption of fast temporal coupling, within a step time of 3 h, not only made the model feasible to construct on this ambitious, inclusive scale, but also made it easy to validate against measured gene expression. The elegant models that have been developed for anterior-posterior axis patterning in insects (Manu et al. 2009; Jaeger et al. 2012; Clark and Akam 2016; Verd et al. 2017) similarly rely on matching TF activity patterns with immediate transcriptional responses, with step times of ~15 min.

It is important to emphasize the unstated assumption about tight temporal coupling of factor binding to transcriptional output in these systems, however, because it probably does not apply in all cases of real-life developmental gene expression. In particular, this assumption comes with corollaries about chromatin accessibility and about the sites in the genome that particular TFs may or may not be able to reach, in a given cell context. It also comes with further assumptions about the biophysical parameters that relate a given site's occupancy by the requisite transcription factor quorum – i.e. factor residency times and cooperativities – to the likelihood that this occupancy will cause an RNA polymerase II complex to start transcribing a gene in response. Do we understand these biophysical parameters well enough to make such assumptions in biological systems in general?

The challenge of a different vantage point

As described below, one issue is how DNA binding by transcription factors relates to transcriptional regulatory output. Another is the problem of how regulatory inputs are coupled to outputs in time. As a result of both, it is harder to explain the relative irreversibility of development, over long time scales, exclusively in terms of the activity of well-studied transcriptional control mechanisms.

Differences among perceived rules for different gene networks may be ontological ("ground truths"), or, at least in part, epistemological. Starting with the epistemology, one cannot do the same experiments that work so well in early *Drosophila* and *Strongylocentrotus* embryos in all developmental systems. Seeking to generalize the impressive modeling successes in fly and sea urchin embryos to regulatory networks operating in organisms with long lifespans and large genomes, researchers have run into challenges. For developmental events much later than the cleavage-stage to gastrulation events that were the focus in *Strongylocentrotus* and *Drosophila,* different methods to test gene network perturbations have been required than genetic screens and direct injection of genes or gene expression antagonists into fertilized eggs. The difficulties include generational timescales too long for genetic selection, multiple paralogous regulatory genes to consider for each effect, and difficulties of gene transfer into the cells of interest, all limiting the power of exhaustive genetic and

molecular-genetic perturbation experiments like those that established TF-target relationships in gene networks in early embryo model systems. Therefore, not only have many researchers in mammalian systems chosen to study later developmental decisions, but they have also attempted to study their gene networks using different techniques.

First, the advent of powerful genome-wide analysis technologies like ChIP-seq and ATAC-seq would appear to offer a shortcut to map direct TF-target interactions that operate in these systems. If the basis of gene network state switches is simply the “pure” presence or absence of a factor that can bind a given genomic site, this should identify the same input-target node interactions as genetic tests. However, the shortcut has turned out to be somewhat of a detour. It has revealed a high level of apparently non-functional albeit site-specific TF binding across the genome, discussed in the next section, which implies that another level of the regulatory code itself exists, determining which binding is functional. This difficulty was not evident in the *Drosophila* and *Strongylocentrotus* systems previously studied because of the strongly phenotype-based strategies that were used to identify the sites where TFs worked in their regulatory networks, and which were the focus of study.

ChIP-seq measurement of TF-DNA binding, therefore, does not provide the same information about TF—target interaction as does direct perturbation analysis, even though it probably reveals novel rules about how TFs really work on the genome. Yet the differences among networks are not exclusively epistemological. The kinetics of gene network responses during cell fate determination in mammals can be much slower with respect to inputs than in the sea urchin and fly embryonic systems, requiring days till the response is manifest. This suggests that the system properties themselves could depart from the rules operating in the early embryonic invertebrate systems, in scientifically interesting ways. The next sections delve further into these epistemological and ontological issues.

Transcription factor binding vs. transcription factor function: combinatoriality in action?

Transcription factors are seen to be distributed across the genome in notably cell type-specific patterns, and even in developmental stage-dependent patterns, which represent only a small fraction of the potential binding motifs in genomic sequence. Thus, clearly TF’s do not simply “read genomic DNA sequence” in an autonomous way. Detailed studies provide evidence that certain TF binding patterns are indeed strongly influenced by their collaboration with certain partner TFs (Garrity et al. 1994; Chlon et al. 2012; Ptasinska et al. 2012; Hosokawa et al. 2018). When the factor depends on another factor for its binding to certain sites, as has been shown for certain factor pairs or complexes (Ets1:Runx1, Pax5:Ets1, Tal1:E2A:Lmo2:Ldb1:Gata factor “pentameric complexes”)(Wadman et al. 1997; Wheat et al. 1999; Hollenhorst et al. 2009; El Omari et al. 2013; Hoang et al. 2016), then ChIP-seq binding peaks might convey more information than simply the TF being monitored, i.e., implying the presence of the collaborating factor – which may have its own developmental regulation – as well as the factor being monitored.

However, the impact of ChIP-seq on gene network analysis has been more limited due to the discovery that many TFs engage the genome in widespread “nonfunctional” binding. This

problematic binding is not random background; it is specific insofar as it often occurs at sites with some version of the TF's preferred motif. Most troublesome for distinguishing cause from correlation, such nonfunctional binding is often seen at active promoters and other sites associated with genes that are active in the cell at the time analyzed. Thus, a naïve analysis based only on static correlation of TF binding with target gene expression may conclude that these represent positive regulatory sites for the factor. However, the observation is that in most cases, many or most of the genes linked to TF binding sites may be completely ***insensitive*** to either loss or gain of function of the TF under study (e.g., (McManus et al. 2011; Ungerbäck et al. 2018)). From ChIP-seq data alone, there may be a 90% or greater chance that a given TF occupancy site is ***not*** associated with immediate regulatory function by that TF. Thus, at these sites the TF binding is either inadequate to mediate function, or adventitious, where other factors more important for function may be creating a preferential binding environment.

The lack of immediate response to a TF's binding can have different causes (Ghisletti et al. 2010; Smale 2010; Rothenberg 2013; Vahedi et al. 2013), of course. Thus, it is possible that some of the unassigned TF occupancies could play a role in the timing events discussed next. Also, sites where multiple TFs, independently measured, can all be seen to bind together within the same cell type are very likely to indicate the sites of active regulatory regions (Wilson et al. 2010), even if it is not clear *a priori* which of the bound factors is most important for their activity. Finally, the "true", sensitively regulated target genes are certainly included among the genes linked to TF binding sites by ChIP-seq, even though they are not the only genes nor the majority of the genes. Thus, the ChIP-seq data being collected can be a valuable resource for understanding the sites through which a TF regulates the expression of particular target genes, once they are identified functionally. However, the hope that mapping of TF binding would reveal gene network relationships without requiring actual perturbation tests has been frustrated by the large number of occupancy sites found to be non-functional (see below). Furthermore, it has become clear that much remains to be learned about how the same TF can be functional when it occupies one binding site and non-functional when it occupies another, and how these different outcomes can be evolutionarily selected.

Slow timing: competing network states, or epigenetic barriers?

Timing is a clue that there may be an ontological "ground truth" difference, not simply an epistemological difference, between gene networks working in early embryos of fast-developing organisms and in later mammalian cell type differentiation. In marked contrast to fly, sea urchin, and also zebrafish, frog, ascidian, and nematode embryos, developmental progression steps in later mammalian development are extraordinarily slow. In one of the best-studied systems, hematopoietic stem cell differentiation to diverse blood cell types in mammals, full differentiation of each cohort of progeny cells can take weeks, accompanied by extensive cell proliferation through most of the process (Upadhaya et al. 2018). Gene expression analysis of populations and individual cells along these pathways suggest that even when upstream regulators are expressed, some responses to their presence are intrinsically slower to occur in these later-differentiating cells (Nerlov and Graf 1998;

Dionne et al. 2005; Iwasaki et al. 2006; Laiosa et al. 2006; Taghon et al. 2007; Miyai et al. 2018).

It is tempting to think of later-differentiating cell programs to be slowed by metastability of the intermediate states that may not be available to embryonic states. Two general classes of mechanisms could contribute to this stability. One could be the establishment of a self-reinforcing gene network state with multiple positive feedbacks among its regulatory genes. Another could be a requirement for physical changes in the chromatin states across the genome that sharply raises the threshold for causing broad gene expression change. To maintain positive regulation, chromatin changes associated with transcriptional activation make DNA more accessible for the binding of many TFs, in some cases measurably reducing the affinity of binding required to establish occupancy [e.g. (Ungerbäck et al. 2018)], and thus setting up a potential positive feedback. To preserve gene silence, local chromatin states associated with resistance to gene activation include intranuclear compartmentalization (Goldmit et al. 2005; Reddy et al. 2008; Lin et al. 2012) and/or local DNA CpG methylation and histone H3K9 or H3K27 methylation (Bintu et al. 2016; Bogdanovic and Lister 2017). The development of specific late-arising cell types in late mammalian fetal or postnatal life, e.g. in blood, brain, or reproductive tissues, provides a test case for the roles of such mechanisms, since these cell types emerge from precursors which have descended from the same fertilized egg as embryonically-differentiating tissues in the same organisms, but which have had their own terminal differentiation postponed.

Chromatin barriers have been much discussed in the highly abnormal forced transformation of fully differentiated cells into induced pluripotent cells using the four TFs called “Yamanaka factors” (Ho et al. 2011; Doege et al. 2012; Watanabe et al. 2013; Takahashi and Yamanaka 2016; Bogdanovic and Lister 2017). However, the natural developmental timing of a transition can also take advantage of chromatin structure as a downshifting mechanism, especially in a case where absolute synchrony among cells is not important. In the case of one particularly slow developmental step, natural timing has been shown to be determined not only by combinatorial input availability but also by the need to overcome a strong chromatin state barrier. This is the lineage commitment step for developing T cells, which is tightly correlated with the onset of expression of the TF Bcl11b for the first time. The activation of the *Bcl11b* locus from a silent state depends on AND logic involving Notch signaling and three other developmentally regulated transcription factors (Kueh et al. 2016). The Notch signaling actually drives a feed-forward circuit to activate at least two of the other factors with which it then collaborates to turn on *Bcl11b*. However, the timing of *Bcl11b* activation is still delayed for days even after all these requirements are apparently met. In fact, two of the positive inputs have already exerted their major functions for enabling *Bcl11b* activation, and have become dispensable, at least two days before the *Bcl11b* gene actually turns on (Kueh et al. 2016). Clearly, additional transcriptional inputs might be needed, but alternatively there might be possible cis-acting chromatin constraints. These mechanisms could be distinguished by using a two-color allelic tagging system to test whether the two alleles of the gene in the same cell are activated synchronously or not (Ng et al. 2018). The results showed clearly that the two alleles were activated asynchronously, in random order throughout the population, but with variable time delays of up to 4 days between the times of activation of the two alleles in the same cells. Expression of even one

allele proved that a cell already had the full complement of trans-acting factors needed to activate the gene, as shown by one allele, thus, if the other equivalent allele was not expressed yet, there must be some local cis-acting constraint. The advantage for one allele relative to the other was often shared among multiple clonal progeny of single precursor cells, though varying randomly between different clones, implying that the mechanism delaying one allele's expression was heritable across several cell cycles. These results suggested that a slow *cis*-acting step was needed to undo local repressive chromatin states that otherwise constrained gene activation even after trans-acting TF input requirements were met. The time scale of this cis-opening mechanism could account for most if not all of the slow response to the inputs to this gene (Ng et al. 2018).

For gene network analysis, this mechanistic delay could have important consequences for the accuracy of network inference. If immediate response to perturbation were the only criterion to identify the regulators of this gene, then most or all of the inputs could have been missed. The time scale needed to observe the functional impact of the TFs that bind to the *Bcl11b cis*-regulatory sites was also long enough for other potential intermediate events to occur. Perturbations had to be carried out over a variety of time intervals, and binding as well as perturbation data had to be considered, and even so contributions from additional regulators could not be ruled out. The effort was only made because the sharp, irreversible off to on transition in the expression of the *Bcl11b* gene was previously shown to be crucial for the timing of developmental commitment. However, if similar mechanisms operate at multiple nodes throughout a gene network, the slowing of responses in biological contexts where chromatin-structure resistance is involved can make gene network solution much more difficult.

Logic implications of *in vivo* DNA occupancy patterns

TF genomic occupancy, in light of differential chromatin accessibility, is thus a result of prior TF action to cause selective, possibly slow opening of distinct genomic regions, a readout of history as much as a way to predict the next gene expression changes. There is a further implication of the observed pattern of transcription factor binding on the DNA, with its frequent pattern of strong occupancy at sites near genes that seem unaffected by changes in that transcription factor's availability. The assumption that a "true" target gene ought to respond immediately to perturbations in any of its functionally significant regulators is itself an assumption of "AND" logic. The validation of the connection comes from the brittleness of this regulatory relationship. But what if this is not what evolution selects for in all cases? And what if the ChIP-seq data, with their implied excess of non-functional sites, are actually showing a mechanism through which "OR" logic could be carried out?

Molecularly, there may be no clear dividing line between "OR" logic and "AND" logic. Davidson and colleagues frequently noted "additive OR" relationships, where removal of a positive regulator reduced the target's expression significantly enough to call, but did not eliminate it. This kind of result is the rule rather than the exception in many analyses of acute responses to developmental gene regulatory perturbation in mammals, i.e., in cases where the phenotype is scored before the affected cells may die or otherwise disappear. Scoring the connection in such cases depends on the strength of the effect relative to a

confident detection threshold. Success in matching “additive OR” effects to a Boolean model can thus emerge from an expert assessment of the right threshold to use for distinguishing expression from non-expression. However, TF assemblies in fact are likely to have a continuum of stabilities. Whether or not the loss of a given factor reduces complex stability, and thus reduces the resulting expression of a target gene, enough to have the effect perceived may not be an absolute. Recent evidence clarifies that much of the output from an “active” enhancer complex is probabilistic anyway, increasing the rate of “transcriptional bursting” from a promoter, but with quiet spells in between bursts, even when all regulators are (at least nominally) present. Depending on RNA processing and turnover times in a cell, and depending on the level of averaging of measurements over a population, a slightly reduced probability of transcriptional firing in a particular cell state may or may not be detected. The effect of time on the measurement window needed to see effects (previous section) is also important here.

Recall that, whereas an individual TF binding site need not have any detectable function, it appears that there is often regulatory significance for DNA sites where multiple TFs bind together in the same cell type (Wilson et al. 2010). Active enhancers frequently recruit multiple transcription factors more effectively than isolated sites, each through interactions with the other factors as well as with the DNA. Notably, a given factor may even be seen more commonly at an active enhancer with a mediocre binding site, where it can bind with partners, than at an inactive region with a superior binding site (Hosokawa et al. 2018; Ungerbäck et al. 2018). Thus, even leaving out effects of chromatin state (discussed above), in many cell types the TFs may not be reading the genome for cognate target sites in a fully independent way. When they bind at the same regulatory element, even if they are not all necessary for transcriptional output from the element, they may all make probabilistic contributions to the overall stability of the complex and to the robustness of the regulatory element’s activity.

Network perturbation insensitivity and evolutionary selection

Robustness is an urgent concern, more than previously guessed, because it has turned out that many enhancers are not evolutionarily conserved. Instead, over relatively short evolutionary times, they are often highly fluid in terms of site numbers, site spacings and organizations, and even the exact types of sites that are present. This was demonstrated in a series of *Drosophila* species for specific known enhancers (Hare et al. 2008) and later, more broadly, for many mammalian regulatory elements (Schmidt et al. 2010). The paradox had seemed to be that the regulation of the target gene by the same factors was conserved even if the regulatory element sequence in the DNA was not. However, the “stickiness” of active enhancers and promoters, and the superfluity of regulatory factor binding that results, could provide a buffer against *cis*-regulatory mutations that reduce the quality of particular TF sites (Heinz et al. 2013). This is not to say that all binding contributes to function. However, these considerations raise the possibility that many gene network connections could be based on biochemical “OR” logic, and that the mechanism for such “OR” logic could very well be embedded in aspects of TF occupancy patterns as shown by ChIP-seq studies.

An evolutionary need for robustness could make “OR” logic preferred in long-selected cell type gene expression programs. In mammals and flies alike, it also appears that genes with developmentally important expression are usually served by multiple cis-regulatory elements; single-element deletion experiments imply that OR logic between such active enhancers for the same gene often reinforce its expression in a given cell state (Barolo 2012; Kieffer-Kwon et al. 2013; Gonzalez et al. 2015; Cannavò et al. 2016). In the genomic evolution of jawed vertebrates, buffering through TF “OR” logic has been taken even further to the genome-wide level, by triplication, quadruplication, or higher-order copy number increases of TF coding loci themselves (Anderson and Rothenberg 2000; Panopoulou et al. 2003). The duplication of regulatory as well as coding sequences leave substantial overlap (despite drift toward specialization) in the expression patterns of similar factors from now-different loci. “AND” logic makes gene networks crisp to predict and easy to validate, whereas “OR” logic makes for difficulties in proving or disproving a network connection; but “OR” logic may offer deep advantages to the organisms themselves.

Concluding remarks

A legacy of both Boveri and Davidson is our understanding that the genome encodes development through the evolution of cis-regulatory systems as well as through the evolution of protein coding genes, and that we can ultimately understand development at a system level if we focus on the cis-regulatory systems that control expression of the TFs themselves. Early embryos of several fast-developing organisms, especially *Drosophila melanogaster* and *Strongylocentrotus purpuratus*, have provided elegant demonstrations of how well this vision works. In turn, a legacy of the results from these systems has been appreciation of universal rules for gene regulation molecular biology in complex multicellular organisms.

However, the attempt to push beyond the early embryo has also revealed aspects of the coding for gene regulation that are not as well understood as many imagine. We do not yet understand why so much of TF binding across the genome appears to be nonfunctional. This adds to the very incomplete state of knowledge we have about what makes the same TF act as an activator at one site and as a repressor at another. We also have profound ignorance about how response timing at any given network node is programmed. There are likely to be different levels of repression applied to genes in more and less reversible silenced states, but we do not know how to predict which ones are easier (faster) or harder (slower) to reverse. Finally, although it has been clear that multiple regulatory elements can control the same gene, we still have very incomplete understanding about the precedence rules and inter-element syntax that determines the outcome.

The impetus to solve the actual operation of real developmental systems to the point of understanding them predictively has been one of the greatest drivers of progress in genomic science. It should continue to be an inspiration to reveal these real-life mechanisms that remain to be understood.

Acknowledgments

This perspective is dedicated to the memory of Eric H. Davidson and the community of scientific inquiry that he inspired. An extensive literature and conversations with this community over decades contributed to the ideas

discussed here. References for many general points in this article could be given only as examples or reviews, and apologies are offered to many additional authors for further citations that are not included. Gene network studies in the author's group have been supported by grants from the USPHS (R33HL089123, R01HL119102, R01AI095943, and RC2CA148278), the L. A. Garfinkle Memorial Laboratory Fund, the Al Sherman Foundation, and the Albert Billings Ruddock Professorship of Biology.

REFERENCES

Anderson MK, Rothenberg EV. 2000 Transcription factor expression in lymphocyte development: clues to the evolutionary origins of lymphoid cell lineages? Curr Top Microbiol Immunol 248: 137–155. [PubMed: 10793477]

Arda HE, Benitez CM, Kim SK. 2013 Gene regulatory networks governing pancreas development. Dev Cell 25: 5–13. [PubMed: 23597482]

Arda HE, Taubert S, MacNeil LT, Conine CC, Tsuda B, Van Gilst M, Sequerra R, Doucette-Stamm L, Yamamoto KR, Walhout AJ. 2010 Functional modularity of nuclear hormone receptors in a Caenorhabditis elegans metabolic gene regulatory network. Mol Syst Biol 6: 367. [PubMed: 20461074]

Arnone MI, Davidson EH. 1997 The hardwiring of development: organization and function of genomic regulatory systems. Development 124: 1851–1864. [PubMed: 9169833]

Arnone MI, Oliveri P, Martinez P. 2018 A conceptual history of the "regulatory genome": From Theodor Boveri to Eric Davidson. Mar Genomics

Avni O, Rao A. 2000 T cell differentiation: a mechanistic view. Curr Opin Immunol 12: 654–659. [PubMed: 11102768]

Barolo S 2012 Shadow enhancers: frequently asked questions about distributed cis-regulatory information and enhancer redundancy. Bioessays 34: 135–141. [PubMed: 22083793]

Ben Tabou de-Leon S, Davidson EH. 2009 Modeling the dynamics of transcriptional gene regulatory networks for animal development. DevBiol 325: 317–328.

Bintu L, Yong J, Antebi YE, McCue K, Kazuki Y, Uno N, Oshimura M, Elowitz MB. 2016 Dynamics of epigenetic regulation at the single-cell level. Science 351: 720–724. [PubMed: 26912859]

Bogdanovic O, Lister R. 2017 DNA methylation and the preservation of cell identity. Curr Opin Genet Dev 46: 9–14. [PubMed: 28651214]

Bolouri H, Davidson EH. 2003 Transcriptional regulatory cascades in development: initial rates, not steady state, determine network kinetics. Proc Natl Acad Sci U S A 100: 9371–9376. [PubMed: 12883007]

Boyes J, Felsenfeld G. 1996 Tissue-specific factors additively increase the probability of the all-or-none formation of a hypersensitive site. EMBO J 15: 2496–2507. [PubMed: 8665857]

Britten RJ, Davidson EH. 1969 Gene regulation for higher cells: a theory. Science 165: 349–357. [PubMed: 5789433]

Bruex A, Kainkaryam RM, Wieckowski Y, Kang YH, Bernhardt C, Xia Y, Zheng X, Wang JY, Lee MM, Benfey P et al. 2012 A gene regulatory network for root epidermis cell differentiation in Arabidopsis. PLoS Genet 8: e1002446. [PubMed: 22253603]

Buckingham M 2017 Gene regulatory networks and cell lineages that underlie the formation of skeletal muscle. Proc Natl Acad Sci U S A 114: 5830–5837. [PubMed: 28584083]

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013 Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10: 1213–1218. [PubMed: 24097267]

Caianiello S 2018 Mechanistic philosophies of development: Theodor Boveri and Eric H. Davidson. Mar Genomics

Cannavò E, Khoueiry P, Garfield DA, Geeleher P, Zichner T, Gustafson EH, Ciglar L, Korbel JO, Furlong EE. 2016 Shadow enhancers are pervasive features of developmental regulatory networks. Curr Biol 26: 38–51. [PubMed: 26687625]

Charney RM, Paraiso KD, Blitz IL, Cho KWY. 2017 A gene regulatory program controlling early Xenopus mesendoderm formation: Network conservation and motifs. Semin Cell Dev Biol 66: 12–24. [PubMed: 28341363]

Chlon TM, Dore LC, Crispino JD. 2012 Cofactor-mediated restriction of GATA-1 chromatin occupancy coordinates lineage-specific gene expression. Mol Cell 47: 608–621. [PubMed: 22771118]

Ciofani M, Madar A, Galan C, Sellars M, Mace K, Pauli F, Agarwal A, Huang W, Parkurst CN, Muratet M et al. 2012 A validated regulatory network for Th17 cell specification. Cell 151: 289–303. [PubMed: 23021777]

Clark E, Akam M. 2016 Odd-paired controls frequency doubling in Drosophila segmentation by altering the pair-rule gene regulatory network. Elife 5: e18215. [PubMed: 27525481]

Clark MR, Mandal M, Ochiai K, Singh H. 2014 Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. Nat Rev Immunol 14: 69–80. [PubMed: 24378843]

Davidson EH. 1990 How embryos work: a comparative view of diverse modes of cell fate specification. Development 108: 365–389. [PubMed: 2187672]

Davidson EH. 1993 Later embryogenesis: regulatory circuitry in morphogenetic fields. Development 118: 665–690. [PubMed: 7915668]

Davidson EH. 2006 The Regulatory Genome: Gene Regulatory Networks in Development and Evolution Academic Press, San Diego.

Davidson EH. 2010 Emerging properties of animal gene regulatory networks. Nature 468: 911–920. [PubMed: 21164479]

Davidson EH, Erwin DH. 2006 Gene regulatory networks and the evolution of animal body plans. Science 311: 796–800. [PubMed: 16469913]

Davidson EH, McClay DR, Hood L. 2003 Regulatory gene networks and the properties of the developmental process. Proc Natl Acad Sci U S A 100: 1475–1480. [PubMed: 12578984]

Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, Yuh CH, Minokawa T, Amore G, Hinman V, Arenas-Mena C et al. 2002a A genomic regulatory network for development. Science 295: 1669–1678. [PubMed: 11872831]

Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, Yuh CH, Minokawa T, Amore G, Hinman V, Arenas-Mena C et al. 2002b A provisional regulatory gene network for specification of endomesoderm in the sea urchin embryo. Dev Biol 246: 162–190. [PubMed: 12027441]

Dionne CJ, Tse KY, Weiss AH, Franco CB, Wiest DL, Anderson MK, Rothenberg EV. 2005 Subversion of T lineage commitment by PU.1 in a clonal cell line system. Dev Biol 280: 448–466. [PubMed: 15882585]

Doege CA, Inoue K, Yamashita T, Rhee DB, Travis S, Fujita R, Guarnieri P, Bhagat G, Vanti WB, Shih A et al. 2012 Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. Nature 488: 652–655. [PubMed: 22902501]

El Omari K, Hoosdally SJ, Tuladhar K, Karia D, Hall-Ponsele E, Platonova O, Vyas P, Patient R, Porcher C, Mancini EJ. 2013 Structural basis for LMO2-driven recruitment of the SCL:E47bHLH heterodimer to hematopoietic-specific transcriptional targets. Cell Rep 4: 135–147. [PubMed: 23831025]

Elgin SCR. 1988 The formation and function of DNase-I hypersensitive sites in the process of gene activation. J Biol Chem 263: 19259–19262. [PubMed: 3198625]

Erwin DH, Davidson EH. 2009 The evolution of hierarchical gene regulatory networks. Nat Rev Genet 10: 141–148. [PubMed: 19139764]

Follows GA, Dhami P, Gottgens B, Bruce AW, Campbell PJ, Dillon SC, Smith AM, Koch C, Donaldson IJ, Scott MA et al. 2006 Identifying gene regulatory elements by genomic microarray mapping of DNaseI hypersensitive sites. Genome Res 16: 1310–1319. [PubMed: 16963707]

Garrity PA, Chen D, Rothenberg EV, Wold BJ. 1994 IL-2 transcription is regulated in vivo at the level of coordinated binding of both constitutive and regulated factors. Mol Cell Biol 14: 2159–2169. [PubMed: 8114746]

Gehring WJ. 1996 The master control gene for morphogenesis and evolution of the eye. Genes Cells 1: 11–15. [PubMed: 9078363]

Georgescu C, Longabaugh WJ, Scripture-Adams DD, David-Fung ES, Yui MA, Zarnegar MA, Bolouri H, Rothenberg EV. 2008 A gene regulatory network armature for T lymphocyte specification. ProcNatlAcadSciUSA 105: 20100–20105.

Ghisletti S, Barozzi I, Mietton F, Polletti S, De Santa F, Venturini E, Gregory L, Lonie L, Chew A, Wei CL et al. 2010 Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. Immunity 32: 317–328. [PubMed: 20206554]

Goldmit M, Ji Y, Skok J, Roldan E, Jung S, Cedar H, Bergman Y. 2005 Epigenetic ontogeny of the Igk locus during B cell development. NatImmunol 6: 198–203.

Gonzalez AJ, Setty M, Leslie CS. 2015 Early enhancer establishment and regulatory locus complexity shape transcriptional programs in hematopoietic differentiation. Nat Genet 47: 1249–1259. [PubMed: 26390058]

Grocott T, Tambalo M, Streit A. 2012 The peripheral sensory nervous system in the vertebrate head: a gene regulatory perspective. Dev Biol 370: 3–23. [PubMed: 22790010]

Hare EE, Peterson BK, Iyer VN, Meier R, Eisen MB. 2008 Sepsid even-skipped enhancers are functionally conserved in Drosophila despite lack of sequence conservation. PLoS Genet 4: e1000106. [PubMed: 18584029]

Harrison MM, Li XY, Kaplan T, Botchan MR, Eisen MB. 2011 Zelda binding in the early Drosophila melanogaster embryo marks regions subsequently activated at the maternal-to-zygotic transition. PLoS Genet 7: e1002266. [PubMed: 22028662]

Heinz S, Romanoski CE, Benner C, Allison KA, Kaikkonen MU, Orozco LD, Glass CK. 2013 Effect of natural genetic variation on enhancer selection and function. Nature 503: 487–492. [PubMed: 24121437]

Ho R, Chronis C, Plath K. 2011 Mechanistic insights into reprogramming to induced pluripotency. J Cell Physiol 226: 868–878. [PubMed: 20945378]

Hoang T, Lambert JA, Martin R. 2016 SCL/TAL1 in hematopoiesis and cellular reprogramming. Curr Top Dev Biol 118: 163–204. [PubMed: 27137657]

Hobert O 2006 Architecture of a microRNA-controlled gene regulatory network that diversifies neuronal cell fates. Cold Spring Harb Symp Quant Biol 71: 181–188. [PubMed: 17381295]

Hobert O. 2014 Development of left/right asymmetry in the Caenorhabditis elegans nervous system: from zygote to postmitotic neuron. Genesis 52: 528–543. [PubMed: 24510690]

Hollenhorst PC, Chandler KJ, Poulsen RL, Johnson WE, Speck NA, Graves BJ. 2009 DNA specificity determinants associate with distinct transcription factor functions. PLoS Genet 5: e1000778. [PubMed: 20019798]

Hoogenkamp M, Lichtinger M, Krysinska H, Lancrin C, Clarke D, Williamson A, Mazzarella L, Ingram R, Jorgensen H, Fisher A et al. 2009 Early chromatin unfolding by RUNX1: a molecular explanation for differential requirements during specification versus maintenance of the hematopoietic gene expression program. Blood 114: 299–309. [PubMed: 19339695]

Hosokawa H, Ungerbäck J, Wang X, Matsumoto M, Nakayama KI, Cohen SM, Tanaka T, Rothenberg EV. 2018 Transcription factor PU.1 represses and activates gene expression in early T cells by redirecting partner transcription factor binding. Immunity 48: 1119–1134.e1117. [PubMed: 29924977]

Hülskamp M, Pfeifle C, Tautz D. 1990 A morphogenetic gradient of hunchback protein organizes the expression of the gap genes Kruppel and knirps in the early Drosophila embryo. Nature 346: 577–580. [PubMed: 2377231]

Hülskamp M, Tautz D. 1991 Gap genes and gradients--the logic behind the gaps. Bioessays 13: 261–268. [PubMed: 1679987]

Ingham PW. 2016 Drosophila segment polarity mutants and the rediscovery of the hedgehog pathway genes. Curr Top Dev Biol 116: 477–488. [PubMed: 26970635]

Ip YT, Park RE, Kosman D, Bier E, Levine M. 1992 The dorsal gradient morphogen regulates stripes of rhomboid expression in the presumptive neuroectoderm of the Drosophila embryo. Genes Dev 6: 1728–1739. [PubMed: 1325394]

Istrail S, Davidson EH. 2005 Logic functions of the genomic cis-regulatory code. Proc Natl Acad Sci U S A 102: 4954–4959. [PubMed: 15788531]

Iwasaki H, Mizuno S, Arinobu Y, Ozawa H, Mori Y, Shigematsu H, Takatsu K, Tenen DG, Akashi K. 2006 The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages. Genes Dev 20: 3010–3021. [PubMed: 17079688]

Jacob F, Monod J. 1961 Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol 3: 318–356. [PubMed: 13718526]

Jaeger J, Manu, Reinitz J. 2012 Drosophila blastoderm patterning. Curr Opin Genet Dev 22: 533–541. [PubMed: 23290311]

Jenuwein T, Forrester WC, Qiu RG, Grosschedl R. 1993 The immunoglobulin μ enhancer core establishes local factor access in nuclear chromatin independent of transcriptional stimulation. Genes Dev 7: 2016–2032. [PubMed: 8406005]

Jiang J, Levine M. 1993 Binding affinities and cooperative interactions with bHLH activators delimit threshold responses to the dorsal gradient morphogen. Cell 72: 741–752. [PubMed: 8453668]

Johnson AD, Poteete AR, Lauer G, Sauer RT, Ackers GK, Ptashne M. 1981 λ Repressor and cro--components of an efficient molecular switch. Nature 294: 217–223. [PubMed: 6457992]

Kieffer-Kwon KR, Tang Z, Mathe E, Qian J, Sung MH, Li G, Resch W, Baek S, Pruett N, Grontved L et al. 2013 Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. Cell 155: 1507–1520. [PubMed: 24360274]

Kim J, Sebring A, Esch JJ, Kraus ME, Vorwerk K, Magee J, Carroll SB. 1996 Integration of positional signals and regulation of wing formation and identity by Drosophila vestigial gene. Nature 382: 133–138. [PubMed: 8700202]

Kirchhamer CV, Davidson EH. 1996 Spatial and temporal information processing in the sea urchin embryo: modular and intramodular organization of the CyIIIa gene cis-regulatory system. Development 122: 333–348. [PubMed: 8565846]

Kueh HY, Rothenberg EV. 2012 Regulatory gene network circuits underlying T cell development from multipotent progenitors. Wiley Interdiscip Rev Syst Biol Med 4: 79–102. [PubMed: 21976153]

Kueh HY, Yui MA, Ng KKH, Pease SS, Zhang JA, Damle SS, Freedman G, Siu S, Bernstein ID, Elowitz MB et al. 2016 Asynchronous combinatorial action of four regulatory factors activates Bcl11b for T cell commitment. Nat Immunol 17: 956–965. [PubMed: 27376470]

Laiosa CV, Stadtfeld M, Xie H, de Andres-Aguayo L, Graf T. 2006 Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBPα and PU.1 transcription factors. Immunity 25: 731–744. [PubMed: 17088084]

Laslo P, Pongubala JM, Lancki DW, Singh H. 2008 Gene regulatory networks directing myeloid and lymphoid cell fates within the immune system. SeminImmunol 20: 228–235.

Laubichler MD, Davidson EH. 2008 Boveri's long experiment: sea urchin merogones and the establishment of the role of nuclear chromosomes in development. Dev Biol 314: 1–11. [PubMed: 18163986]

Laurent F, Girdziusaite A, Gamart J, Barozzi I, Osterwalder M, Akiyama JA, Lincoln J, Lopez-Rios J, Visel A, Zuniga A et al. 2017 HAND2 Target Gene Regulatory Networks Control Atrioventricular Canal and Cardiac Valve Development. Cell Rep 19: 1602–1613. [PubMed: 28538179]

Lee HJ, Takemoto N, Kurata H, Kamogawa Y, Miyatake S, O'Garra A, Arai N. 2000 GATA-3 induces T helper cell type 2 (Th2) cytokine expression and chromatin remodeling in committed Th1 cells. J Exp Med 192: 105–115. [PubMed: 10880531]

Lin YC, Benner C, Mansson R, Heinz S, Miyazaki K, Miyazaki M, Chandra V, Bossen C, Glass CK, Murre C. 2012 Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. Nat Immunol 13: 1196–1204. [PubMed: 23064439]

Longabaugh WJR, Zeng W, Zhang JA, Hosokawa H, Jansen CS, Li L, Romero-Wolf M, Liu P, Kueh HY, Mortazavi A et al. 2017 Bcl11b and combinatorial resolution of cell fate in the T-cell gene regulatory network. Proc Natl Acad Sci U S A 114: 5800–5807. [PubMed: 28584128]

Manu, Surkova S, Spirov AV, Gursky VV, Janssens H, Kim AR, Radulescu O, Vanario-Alonso CE, Sharp DH, Samsonova M et al. 2009 Canalization of gene expression in the Drosophila blastoderm by gap gene cross regulation. PLoS Biol 7: e1000049. [PubMed: 19750121]

Martik ML, Lyons DC, McClay DR. 2016 Developmental gene regulatory networks in sea urchins and what we can learn from them. F1000Res 5: e7381.

McManus S, Ebert A, Salvagiotto G, Medvedovic J, Sun Q, Tamir I, Jaritz M, Tagoh H, Busslinger M. 2011 The transcription factor Pax5 regulates its target genes by recruiting chromatin-modifying proteins in committed B cells. EMBO J 30: 2388–2404. [PubMed: 21552207]

Miyai T, Takano J, Endo TA, Kawakami E, Agata Y, Motomura Y, Kubo M, Kashima Y, Suzuki Y, Kawamoto H et al. 2018 Three-step transcriptional priming that drives the commitment of multipotent progenitors toward B cells. Genes Dev 32: 112–126. [PubMed: 29440259]

Monod J, Changeux JP, Jacob F. 1963 Allosteric proteins and cellular control systems. J Mol Biol 6: 306–329. [PubMed: 13936070]

Monteiro A 2012 Gene regulatory networks reused to build novel traits: co-option of an eye-related gene regulatory network in eye-like organs and red wing patches on insect wings is suggested by optix expression. Bioessays 34: 181–186. [PubMed: 22223407]

Morisato D, Anderson KV. 1995 Signaling pathways that establish the dorsal-ventral pattern of the Drosophila embryo. Annu Rev Genet 29: 371–399. [PubMed: 8825480]

Murphy KM, Reiner SL. 2002 The lineage decisions of helper T cells. Nat Rev Immunol 2: 933–944. [PubMed: 12461566]

Nakayama T, Yamashita M. 2008 Initiation and maintenance of Th2 cell identity. Curr Opin Immunol 20: 265–271. [PubMed: 18502111]

Nerlov C, Graf T. 1998 PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors. Genes Dev 12: 2403–2412. [PubMed: 9694804]

Ng KKH, Yui MA, Mehta A, Siu S, Irwin B, Pease S, Hirose S, Elowitz MB, Rothenberg EV, Kueh HY. 2018 A stochastic epigenetic switch controls the dynamics of T-cell lineage commitment. Elife 7: e37851. [PubMed: 30457103]

Nusslein-Volhard C, Wieschaus E. 1980 Mutations affecting segment number and polarity in Drosophila. Nature 287: 795–801. [PubMed: 6776413]

Oestreich KJ, Weinmann AS. 2012 Master regulators or lineage-specifying? Changing views on $CD4^+$ T cell transcription factors. Nat Rev Immunol 12: 799–804. [PubMed: 23059426]

Oliveri P, Carrick DM, Davidson EH. 2002 A regulatory gene network that directs micromere specification in the sea urchin embryo. Dev Biol 246: 209–228. [PubMed: 12027443]

Oliveri P, Tu Q, Davidson EH. 2008 Global regulatory logic for specification of an embryonic cell lineage. Proc Natl Acad Sci U S A 105: 5955–5962. [PubMed: 18413610]

Panopoulou G, Hennig S, Groth D, Krause A, Poustka AJ, Herwig R, Vingron M, Lehrach H. 2003 New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes. Genome Res 13: 1056–1066. [PubMed: 12799346]

Parker HJ, Bronner ME, Krumlauf R. 2016 The vertebrate Hox gene regulatory network for hindbrain segmentation: Evolution and diversification: Coupling of a Hox gene regulatory network to hindbrain segmentation is an ancient trait originating at the base of vertebrates. Bioessays 38: 526–538. [PubMed: 27027928]

Peter IS, Davidson EH. 2010 The endoderm gene regulatory network in sea urchin embryos up to mid-blastula stage. Dev Biol 340: 188–199. [PubMed: 19895806]

Peter IS, Davidson EH. 2011 A gene regulatory network controlling the embryonic specification of endoderm. Nature 474: 635–639. [PubMed: 21623371]

Peter IS, Davidson EH. 2015 Genomic Control Process: Development and Evolution Academic Press, Elsevier, San Diego, CA.

Peter IS, Faure E, Davidson EH. 2012 Predictive computation of genomic logic processing functions in embryonic development. Proc Natl Acad Sci U S A 109: 16434–16442. [PubMed: 22927416]

Pires ND, Yi K, Breuninger H, Catarino B, Menand B, Dolan L. 2013 Recruitment and remodeling of an ancient gene regulatory network during land plant evolution. Proc Natl Acad Sci U S A 110: 9571–9576. [PubMed: 23690618]

Ptashne M 1992 A Genetic Switch: Phage λ and Higher Organisms Cell Press, Cambridge, Ma.

Ptashne M, Jeffrey A, Johnson AD, Maurer R, Meyer BJ, Pabo CO, Roberts TM, Sauer RT. 1980 How the λ repressor and cro work. Cell 19: 1–11. [PubMed: 6444544]

Ptasinska A, Assi SA, Mannari D, James SR, Williamson D, Dunne J, Hoogenkamp M, Wu M, Care M, McNeill H et al. 2012 Depletion of RUNX1/ETO in t(8;21) AML cells leads to genome-wide changes in chromatin structure and transcription factor binding. Leukemia 26: 1829–1841. [PubMed: 22343733]

Ramirez-Carrozzi VR, Braas D, Bhatt DM, Cheng CS, Hong C, Doty KR, Black JC, Hoffmann A, Carey M, Smale ST. 2009 A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. Cell 138: 114–128. [PubMed: 19596239]

Ramirez RN, El-Ali NC, Mager MA, Wyman D, Conesa A, Mortazavi A. 2017 Dynamic Gene Regulatory Networks of Human Myeloid Differentiation. Cell Syst 4: 416–429 e413. [PubMed: 28365152]

Rao S, Procko E, Shannon MF. 2001 Chromatin remodeling, measured by a novel real-time polymerase chain reaction assay, across the proximal promoter region of the IL-2 gene. J Immunol 167: 4494–4503. [PubMed: 11591776]

Reddy KL, Zullo JM, Bertolino E, Singh H. 2008 Transcriptional repression mediated by repositioning of genes to the nuclear lamina. Nature 452: 243–247. [PubMed: 18272965]

Ririe TO, Fernandes JS, Sternberg PW. 2008 The Caenorhabditis elegans vulva: a post-embryonic gene regulatory network controlling organogenesis. Proc Natl Acad Sci U S A 105: 20095–20099. [PubMed: 19104047]

Rothenberg EV. 2013 Epigenetic mechanisms and developmental choice hierarchies in T-lymphocyte development. Brief Funct Genomics 12: 512–524. [PubMed: 23922132]

Rothenberg EV, Ward SB. 1996 A dynamic assembly of diverse transcription factors integrates activation and cell-type information for interleukin-2 gene regulation. Proc Natl Acad Sci USA 93: 9358–9365. [PubMed: 8790334]

Rottinger E, Dahlin P, Martindale MQ. 2012 A framework for the establishment of a cnidarian gene regulatory network for “endomesoderm” specification: the inputs of ss-catenin/TCF signaling. PLoS Genet 8: e1003164. [PubMed: 23300467]

Satzinger H 2008 Theodor and Marcella Boveri: chromosomes and cytoplasm in heredity and development. Nat Rev Genet 9: 231–238. [PubMed: 18268510]

Schmidt D, Wilson MD, Ballester B, Schwalie PC, Brown GD, Marshall A, Kutter C, Watt S, Martinez-Jimenez CP, Mackay S et al. 2010 Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. Science 328: 1036–1040. [PubMed: 20378774]

Schulz C, Tautz D. 1994 Autonomous concentration-dependent activation and repression of Kruppel by hunchback in the Drosophila embryo. Development 120: 3043–3049. [PubMed: 7607091]

Sekimata M, Perez-Melgosa M, Miller SA, Weinmann AS, Sabo PJ, Sandstrom R, Dorschner MO, Stamatoyannopoulos JA, Wilson CB. 2009 CCCTC-binding factor and the transcription factor T-bet orchestrate T helper 1 cell-specific structure and function at the interferon-γ locus. Immunity 31: 551–564. [PubMed: 19818655]

Simoes-Costa M, Bronner ME. 2015 Establishing neural crest identity: a gene regulatory recipe. Development 142: 242–257. [PubMed: 25564621]

Singh H, Pongubala JM. 2006 Gene regulatory networks and the determination of lymphoid cell fates. CurrOpinImmunol 18: 116–120.

Sinner D, Kirilenko P, Rankin S, Wei E, Howard L, Kofron M, Heasman J, Woodland HR, Zorn AM. 2006 Global analysis of the transcriptional network controlling Xenopus endoderm formation. Development 133: 1955–1966. [PubMed: 16651540]

Smale ST. 2010 Pioneer factors in embryonic stem cells and differentiation. CurrOpinGenetDev 20: 519–526.

Small S, Blair A, Levine M. 1992 Regulation of even-skipped stripe 2 in the Drosophila embryo. EMBO J 11: 4047–4057. [PubMed: 1327756]

Small S, Kraut R, Hoey T, Warrior R, Levine M. 1991 Transcriptional regulation of a pair-rule stripe in Drosophila. Genes Dev 5: 827–839. [PubMed: 2026328]

Stanojevic D, Small S, Levine M. 1991 Regulation of a segmentation stripe by overlapping activators and repressors in the Drosophila embryo. Science 254: 1385–1387. [PubMed: 1683715]

Stathopoulos A, Levine M. 2002a Dorsal gradient networks in the Drosophila embryo. Dev Biol 246: 57–67. [PubMed: 12027434]

Stathopoulos A, Levine M. 2002b Linear signaling in the Toll-Dorsal pathway of Drosophila: activated Pelle kinase specifies all threshold outputs of gene expression while the bHLH protein Twist specifies a subset. Development 129: 3411–3419. [PubMed: 12091311]

Stathopoulos A, Van Drenth M, Erives A, Markstein M, Levine M. 2002 Whole-genome analysis of dorsal-ventral patterning in the Drosophila embryo. Cell 111: 687–701. [PubMed: 12464180]

Struhl G, Johnston P, Lawrence PA. 1992 Control of Drosophila body pattern by the hunchback morphogen gradient. Cell 69: 237–249. [PubMed: 1568245]

Swiers G, Chen YH, Johnson AD, Loose M. 2010 A conserved mechanism for vertebrate mesoderm specification in urodele amphibians and mammals. Dev Biol 343: 138–152. [PubMed: 20394741]

Taghon T, Yui MA, Rothenberg EV. 2007 Mast cell lineage diversion of T lineage precursors by the essential T cell transcription factor GATA-3. Nat Immunol 8: 845–855. [PubMed: 17603486]

Takahashi K, Yamanaka S. 2016 A decade of transcription factor-mediated reprogramming to pluripotency. Nat Rev Mol Cell Biol 17: 183–193. [PubMed: 26883003]

Takemoto N, Kamogawa Y, Lee HJ, Kurata H, Arai KI, O'Garra A, Arai N, Miyatake S. 2000 Cutting edge: chromatin remodeling at the IL-4/IL-13 intergenic regulatory region for Th2-specific cytokine gene cluster. J Immunol 165: 6687–6691. [PubMed: 11120785]

Thanos D, Maniatis T. 1995 Virus induction of human IFNβ gene expression requires the assembly of an enhanceosome. Cell 83: 1091–1100. [PubMed: 8548797]

Ungerbäck J, Hosokawa H, Wang X, Strid T, Williams BA, Sigvardsson M, Rothenberg EV. 2018 Pioneering, chromatin remodeling, and epigenetic constraint in early T-cell gene regulation by SPI1 (PU.1). Genome Res 28: 1508–1519. [PubMed: 30171019]

Upadhaya S, Sawai CM, Papalexi E, Rashidfarrokhi A, Jang G, Chattopadhyay P, Satija R, Reizis B. 2018 Kinetics of adult hematopoietic stem cell differentiation in vivo. J Exp Med 215: 2815–2832. [PubMed: 30291161]

Vahedi G, Kanno Y, Sartorelli V, O'Shea JJ. 2013 Transcription factors and CD4 T cells seeking identity: masters, minions, setters and spikers. Immunology 139: 294–298. [PubMed: 23586907]

Vahedi G, Takahashi H, Nakayamada S, Sun HW, Sartorelli V, Kanno Y, O'Shea JJ. 2012 STATs shape the active enhancer landscape of T cell populations. Cell 151: 981–993. [PubMed: 23178119]

Verd B, Crombach A, Jaeger J. 2017 Dynamic maternal gradients control timing and shift-rates for Drosophila gap gene expression. PLoS Comput Biol 13: e1005285. [PubMed: 28158178]

Wadman IA, Osada H, Grutz GG, Agulnick AD, Westphal H, Forster A, Rabbitts TH. 1997 The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. EMBO J 16: 3145–3157. [PubMed: 9214632]

Wang S, Sengel C, Emerson MM, Cepko CL. 2014 A gene regulatory network controls the binary fate decision of rod and bipolar cells in the vertebrate retina. Dev Cell 30: 513–527. [PubMed: 25155555]

Watanabe A, Yamada Y, Yamanaka S. 2013 Epigenetic regulation in pluripotent stem cells: a key to breaking the epigenetic barrier. Philos Trans R Soc Lond B Biol Sci 368: 20120292. [PubMed: 23166402]

Weatherbee SD, Carroll SB. 1999 Selector genes and limb identity in arthropods and vertebrates. Cell 97: 283–286. [PubMed: 10319808]

Weintraub H, Davis R, Tapscott S, Thayer M, Krause M, Benezra R, Blackwell TK, Turner D, Rupp R, Hollenberg S. 1991 The myoD gene family: nodal point during specification of the muscle cell lineage. Science 251: 761–766. [PubMed: 1846704]

Wheat W, Fitzsimmons D, Lennox H, Krautkramer SR, Gentile LN, McIntosh LP, Hagman J. 1999 The highly conserved beta-hairpin of the paired DNA-binding domain is required for assembly of Pax-Ets ternary complexes. Mol Cell Biol 19: 2231–2241. [PubMed: 10022910]

Wieschaus E, Nusslein-Volhard C. 2016 The Heidelberg screen for pattern mutants of Drosophila: A personal account. Annu Rev Cell Dev Biol 32: 1–46. [PubMed: 27501451]

Wilson NK, Foster SD, Wang X, Knezevic K, Schutte J, Kaimakis P, Chilarska PM, Kinston S, Ouwehand WH, Dzierzak E et al. 2010 Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. Cell Stem Cell 7: 532–544. [PubMed: 20887958]

Yang XP, Ghoreschi K, Steward-Tharp SM, Rodriguez-Canales J, Zhu J, Grainger JR, Hirahara K, Sun HW, Wei L, Vahedi G et al. 2011 Opposing regulation of the locus encoding IL-17 through direct, reciprocal actions of STAT3 and STAT5. Nat Immunol 12: 247–254. [PubMed: 21278738]

Yuh CH, Bolouri H, Davidson EH. 1998 Genomic cis-regulatory logic: experimental and computational analysis of a sea urchin gene. Science 279: 1896–1902. [PubMed: 9506933]

Yuh CH, Bolouri H, Davidson EH. 2001 Cis-regulatory logic in the endo16 gene: switching from a specification to a differentiation mode of control. Development 128: 617–629. [PubMed: 11171388]

Yuh CH, Moore JG, Davidson EH. 1996 Quantitative functional interrelations within the cis-regulatory system of the S. purpuratus Endo16 gene. Development 122: 4045–4056. [PubMed: 9012524]

Zeitlinger J, Stark A, Kellis M, Hong JW, Nechaev S, Adelman K, Levine M, Young RA. 2007 RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. Nat Genet 39: 1512–1516. [PubMed: 17994019]