Integrative and Comparative Biology

Integrative and Comparative Biology, volume 58, number 4, pp. 640–653
doi:10.1093/icb/icy061

Society for Integrative and Comparative Biology

SYMPOSIUM

Inference of Developmental Gene Regulatory Networks Beyond Classical Model Systems: New Approaches in the Post-genomic Era

Selene L. Fernandez-Valverde,[1,*] Felipe Aguilera[2,†] and René Alexander Ramos-Díaz[*]

*CONACYT, Unidad de Genómica Avanzada, Laboratorio Nacional de Genómica para la Biodiversidad (Langebio), Centro de Investigación y de Estudios Avanzados del IPN, Irapuato, Guanajuato, Mexico and †Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias Biológicas, Universidad de Concepción, Chile

From the symposium "From Small and Squishy to Big and Armored: Genomic, Ecological and Paleontological Insights into the Early Evolution of Animals" presented at the annual meeting of the Society for Integrative and Comparative Biology, January 3–7, 2018 at San Francisco, California.

[1]E-mail: selene.fernandez@cinvestav.mx

[2]E-mail: faguilera@udec.cl

Synopsis The advent of high-throughput sequencing (HTS) technologies has revolutionized the way we understand the transformation of genetic information into morphological traits. Elucidating the network of interactions between genes that govern cell differentiation through development is one of the core challenges in genome research. These networks are known as developmental gene regulatory networks (dGRNs) and consist largely of the functional linkage between developmental control genes, *cis*-regulatory modules, and differentiation genes, which generate spatially and temporally refined patterns of gene expression. Over the last 20 years, great advances have been made in determining these gene interactions mainly in classical model systems, including human, mouse, sea urchin, fruit fly, and worm. This has brought about a radical transformation in the fields of developmental biology and evolutionary biology, allowing the generation of high-resolution gene regulatory maps to analyze cell differentiation during animal development. Such maps have enabled the identification of gene regulatory circuits and have led to the development of network inference methods that can recapitulate the differentiation of specific cell-types or developmental stages. In contrast, dGRN research in non-classical model systems has been limited to the identification of developmental control genes via the candidate gene approach and the characterization of their spatiotemporal expression patterns, as well as to the discovery of *cis*-regulatory modules via patterns of sequence conservation and/or predicted transcription-factor binding sites. However, thanks to the continuous advances in HTS technologies, this scenario is rapidly changing. Here, we give a historical overview on the architecture and elucidation of the dGRNs. Subsequently, we summarize the approaches available to unravel these regulatory networks, highlighting the vast range of possibilities of integrating multiple technical advances and theoretical approaches to expand our understanding on the global gene regulation during animal development in non-classical model systems. Such new knowledge will not only lead to greater insights into the evolution of molecular mechanisms underlying cell identity and animal body plans, but also into the evolution of morphological key innovations in animals.

Introduction

The emergence of multicellularity was one of the most remarkable events in animal evolution. Perhaps even more remarkable has been the subsequent evolution of cell types and animal body plans, which reflects the capacity of multicellular animals to give rise to a wide range of different cell types and complex organs in a highly ordered and reproducible manner (Arendt et al. 2016; Sebé-Pedrós et al. 2017). In most—if not all—animals, the multicellular state is established in each generation through serial divisions of the zygote, where daughter cells produced by these divisions become an independent and fully specialized cell type. This functional specialization occurs largely during development and involves the tight coordination of cell proliferation, cell

Advance Access publication June 18, 2018

differentiation, tissue growth, and developmental genetic programs (Peter and Davidson 2011a; Brunet and King 2017).

In the past few years, many studies have revealed that genes encoding transcription factors and signaling molecules are critical controllers of pattern formation and cell fate specification during development (Zeitlinger and Stark 2010). Notably, most of these genes are highly conserved across animals (i.e., metazoans) and even their closest unicellular relatives (Larroux et al. 2008; de Mendoza et al. 2013; Sebé-Pedrós et al. 2017). This striking level of conservation suggests that cell types and animal body plans are, at least partially, controlled by the regulatory capacities of these highly conserved genes. Yet, we cannot help but be intrigued by how such a conserved set of genes with few examples of gene expansions and little changes in their functionality (Carroll 2008; Sebé-Pedrós and de Mendoza 2015) can lead to the vast diversity of cell types and body plan forms found in animals.

Developmental control genes have been characterized for decades. This characterization has been focused on comparisons of the spatiotemporal expression patterns of these genes among divergent species, suggesting that transcription factors and signaling molecules participate in multiple, independent developmental processes (Arthur 2000; Martindale 2005; Carroll 2008; Martín-Durán et al. 2018). Most importantly, however, it appears that changes in the spatial regulation of developmental control genes are related to morphological divergence, suggesting that the changes in morphology are the result of nucleotide substitutions in *cis*-regulatory elements and amino-acid substitutions in transcription factors affecting the regulation of gene expression (Lynch and Wagner 2008; Lynch et al. 2011).

This notion is not new and was hypothesized more than four decades ago by Britten and Davidson (1969, 1971), who postulated that intergenic genomic regions and mutational changes in protein-coding sequences have an important role in determining differences in gene regulatory patterns, and consequently, in animal body plan diversity. Therefore, animal body plan evolution, along with cell type evolution, is controlled by the precise regulation of gene expression in time and space, which in turn is driven by developmental gene regulatory networks (dGRNs) (Davidson and Erwin 2006; Peter and Davidson 2011a). Such networks comprise a constellation of elements including regulatory genes (transcription factors, signaling molecules, noncoding RNAs), regulatory sequences (*cis*-regulatory modules, enhancers, promoters, insulators), and target genes (differentiation and structural genes), which interplay tightly to trigger induction or repression of gene expression (Levine 2010; Peter and Davidson 2011a; Rebeiz et al. 2015). Therefore, the right execution of this molecular choreography, repeated anew in every generation, is fundamental to the life of every animal on Earth.

Recent advances in evo-devo have focused on determining the regulatory gene interactions that underpin the dGRNs and how these interactions control the process of regulation of gene expression during animal development. In this sense, many molecular mechanisms that underlie dGRNs and influence cell type and animal body plan evolution have been characterized in vertebrates (Loose and Patient 2004; Parfitt and Shen 2014; Goode et al. 2016; Gouti et al. 2017; Hamey et al. 2017), sea urchin (Peter and Davidson 2010, 2011b; Rafiq et al. 2014; Erkenbrack 2016), *Drosophila melanogaster* (Nasiadka et al. 2002; Schroeder et al. 2004; Sandmann et al. 2007), and *Caenorhabditis elegans* (Maduro and Rothman 2002; Inoue et al. 2005; Ririe et al. 2008). Results obtained from these classical model systems have shown that changes in animal body plan forms are the result of evolutionary changes in the architecture of these dGRNs (Davidson 2010a; Peter and Davidson 2011a; Hinman and Cheatle Jarvela 2014). Thus, the evolution of dGRNs plays a key role in the emergence of animal diversity.

Beyond these classical model systems, little information exists concerning the transcriptional developmental dynamics in other animals (Leininger et al. 2014; Israel et al. 2016; Levin et al. 2016; Dylus et al. 2018), or the changes in *cis*-regulatory modules controlling embryonic development (Schwaiger et al. 2014; Gaiti et al. 2017). However, this has started to change due to rapid advances in high-throughput sequencing (HTS) technologies. Here, we review how dGRNs have been deciphered, discuss how recent advances in HTS technologies have helped to refine the inference of dGRNs, and how these new technologies have been (and can be) used to unravel dGRNs in non-classical model systems. We expect that findings obtained from classical model systems will be coupled with data from other phylogenetically informative non-classical model systems in order to achieve a better understanding of how cell types and animal body plans might have evolved (Achim et al. 2015).

dGRN architecture

dGRNs can be represented as complex logic maps that state in detail the interactions between

developmental control genes (transcription factors and components of cell signaling pathways) and *cis*-regulatory modules (promoter, enhancers, and insulators) in order to visualize how differentiation and structural genes (target genes) are turned off or on at a given time and location during development (Levine and Davidson 2005). In addition, dGRNs have a modular architecture, consisting of multiple sub-circuits—each in charge of individual regulatory tasks defined by a set of specific developmental control genes and their *cis*-regulatory modules (Erwin 2009; Davidson 2010a). Building on this modular architecture, dGRNs are hierarchical as they are divided into different components. For example, the components controlling the initial stages of development are at the top of the hierarchy, while the portions governing intermediate processes, such as spatial subdivision and morphological patterns are in the middle, and the components controlling more specific functions, including cell differentiation and organogenesis/morphogenesis are at the periphery (Erwin 2009).

The hierarchical organization of dGRNs leads to diverse evolutionary rates in components that occupy different parts of the network. For instance, the components known as kernels, which consist of conserved interactions among transcription factors (Davidson 2010b), are highly conserved regulatory interactions showing slow changes during evolution; thus, it is believed that they are responsible for the progenitor states of a developing structure (Davidson and Erwin 2006). In contrast, other components of the network, known as intermediate and peripheral, are more evolvable and often co-opted to different functions during development or subjected to gene loss or duplication events over evolution; therefore, changes in these components occur at high evolutionary rates and have great impacts on the phenotype (Wittkopp et al. 2002; Aguilera et al. 2017; Peter and Davidson 2017). Although there have been reviews which develop hypotheses on how dGRNs might evolve (Carroll 2008; Peter and Davidson 2011a), there are currently few comparisons of dGRN topologies (Hinman and Cheatle Jarvela 2014; Andrikou and Arnone 2015; Kittelmann et al. 2018). Therefore, understanding how the components of dGRNs have evolved among animal species is a central problem in evo-devo, which can only be resolved through comprehensive comparisons of dGRN architecture in classical model systems with those of non-classical model systems with important phylogenetic positions in the animal tree of life. Despite the paucity of comparative dGRN analyses, some studies suggest that the evolution of cell types and animal body plans depends upon changes in the architecture of dGRNs (Carroll 2008; Peter and Davidson 2011a; Hinman and Cheatle Jarvela 2014), while others suggest that molecular regulatory networks can be diverse and the high evolvability of dGRNs does not generate drastic changes in the animal body plan (Sommer 2012).

Initial strategies for unraveling dGRNs

The first theoretical model describing the mechanisms controlling gene regulation in higher eukaryotes was postulated by Eric H. Davidson and his long-time colleague Roy J. Britten (Britten and Davidson 1969; 1971). These seminal papers described how trans-acting factors might regulate diverse batteries of genes, and how gene regulation could have evolved and shaped the evolution of animal body plans (Britten and Davidson 1969, 1971). Since then, dGRNs have been primarily studied via the careful and minute characterization of regulatory interactions in specific tissues and developmental stages in various organisms (Levine and Davidson 2005; Carroll 2008; Peter and Davidson 2011a). Eric H. Davidson was also the founder and pioneer of studying the molecular interactions happening during development as a whole system. For nearly 30 years, Davidson's lab was dedicated to untangling all specific interactions occurring during sea urchin (*Strongylocentrotus purpuratus*) development through targeted experiments. This journey was largely complemented with concurrent advances in DNA technologies.

For instance, with the development of recombinant DNA in the early 1980s and through a series of experiments, Davidson and his colleagues characterized the genomic *cis*-regulatory sequence of the *Endo16* gene (an endoderm-specific gene that is expressed at the onset of sea urchin gastrulation) (Yuh et al. 1994; Yuh and Davidson 1996). These studies showed that protein-coding genes are controlled by nearby DNA regulatory sequences that serve as binding-sites for transcription factors, showing modularity of these binding-sites and demonstrating that different modules governed different temporal aspects of gene expression (Yuh et al. 1994; Yuh and Davidson 1996). Over the years, many elegant experiments were conducted in order to understand how transcription of individual genes is controlled in terms of time, space, and abundance (Calzone et al. 1988; Hough-Evans et al. 1990; Wang et al. 1995; Zeller et al. 1995; Kirchhamer and Davidson 1996; Coffman et al. 1997; Arnone et al. 1998). These and other experiments involved the

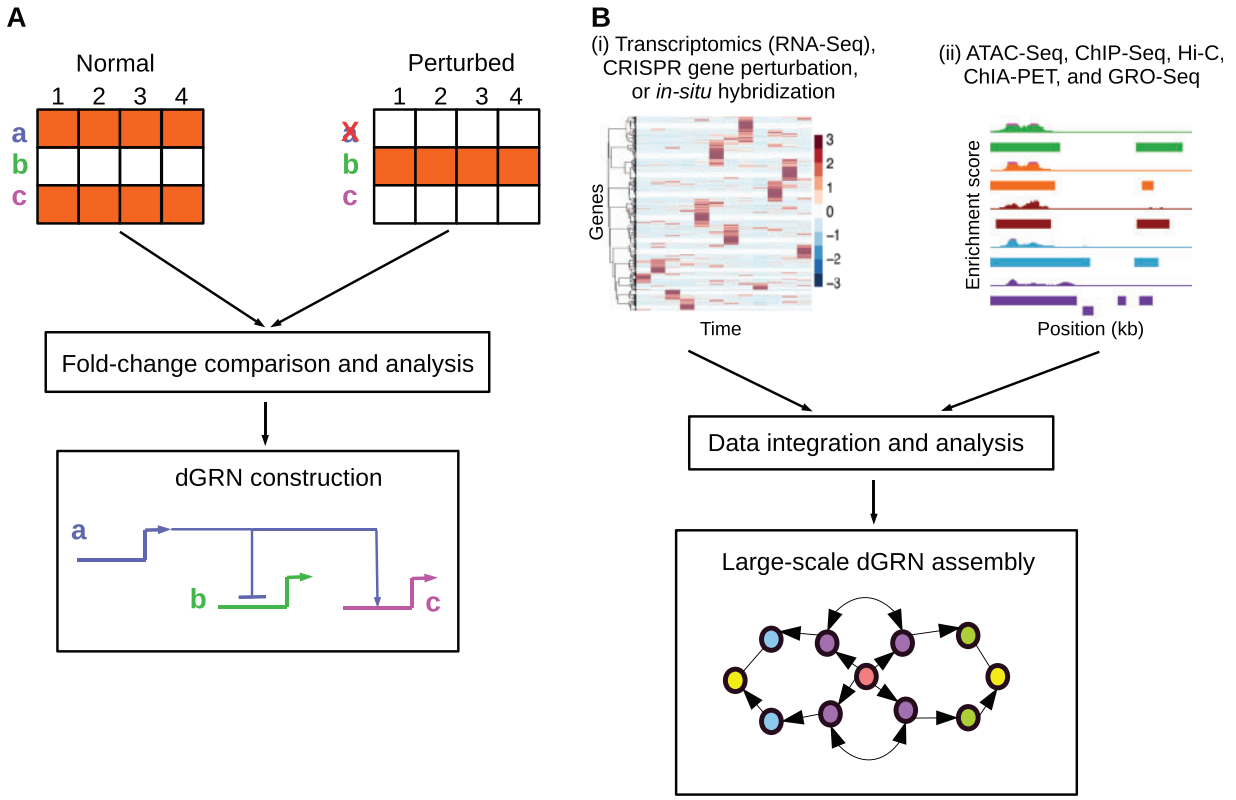

Fig. 1 *Two avenues for dGRN inference*. **(A)** Classical approach to identify regulatory interactions during animal development, originally developed by Davidson, consisting of knockout experiments used to infer candidate regulators by comparing a perturbed gene expression profile with a normal gene expression profile during a specific stage of the developmental process. For example, in an experiment performed in four embryos at the same developmental stage where *gene a* is knocked-out, it is observed that the expression level of *gene b* significantly increases while the expression of *gene c* decreases compared with a normal (without perturbation) gene expression profile of the three genes at the same developmental stage. It can be concluded that *a* activates *c* but inhibits the expression of *b*. Perturbation experiments provide a robust reproducible method to infer regulatory interactions. However, they are difficult to apply to large sets of genes. **(B)** A computational network assembly can be an alternative approach to explore dGRNs in classical- and non-classical model systems. In general, at least two types of data sets are required in this approach: (i) spatiotemporal gene expression such as developmental transcriptomics (RNA-Seq), large scale CRISPR perturbation or *in-situ* hybridization data to infer spatial co-expression; (ii) datasets that can provide information about candidate active regulatory processes such as ATAC-Seq, ChIP-Seq, Hi-C, ChIA-PET, GROSeq. This approach poses two computational challenges: (1) heterogeneous data standardization and integration, for example, combining time-course developmental transcriptomics with ATACSeq to infer which regulatory interactions could be active whether regulators are transcribed and their target *cis*-regulatory regions are accessible during a developmental stage; (2) dGRN assembly that implies identifying which of the inferred interactions should be kept and which ones should be dropped using quantitative criteria. If these challenges are resolved, it would be possible to improve comparative analyses between dGRNs in different animal taxa and facilitate the exploration of dGRNs in non-classical model systems providing a large-scale dGRN containing candidate regulatory interactions for experimental validation.

systematic analysis of spatiotemporal patterns of gene expression, coupled with targeted mutational and/or perturbation analyses, during sea urchin development to determine the regulatory interactions (i.e., positive or negative feedbacks) between developmental control genes, *cis*-regulatory modules, and differentiation genes (Fig. 1A).

After years of these painstaking analyses regarding the function of developmental control genes and their *cis*-regulatory modules during developmental progression in sea urchin, the description of the dGRN controlling the specification of the endomesoderm was published (Davidson et al. 2002). This study gave support to Britten and Davidson's theoretical model envisioned in 1969 (Britten and Davidson 1969). After the inference of the dGRN that governs endomesoderm specification, the same group unraveled a dGRN controlling the endoderm specification (Peter and Davidson 2010, 2011b). Altogether, these studies (Davidson et al. 2002; Peter and Davidson 2010, 2011b) give strong evidence that embryonic development is directed by the dynamic progression of regulatory states, which means a series of sub-circuits, each of which endows particular developmental functions

(Davidson and Erwin 2006; Davidson 2010b). The approach performed by Davidson was employed by other researchers to unravel dGRNs in other species including *Drosophila* (Levine and Davidson 2005) and *Xenopus* (Loose and Patient 2004).

But this was not the end of Davidson's ambitions. He was aware that mathematical models are necessary to predict the behavior of a dGRN. In this context, a computational Boolean model that uses gene expression and interaction data was generated to predict how the dGRN would respond to perturbations (Peter et al. 2012; Peter and Davidson 2017). Although this computational model works well in sea urchin, its inherent requirement for detailed gene expression measurements linked to temporal and spatial information in the developing embryo means that such precise dGRN models cannot be readily built for most other taxa, making comparisons at this level challenging.

Despite great advances in the inference of dGRNs by Davidson and his colleagues, many of his hypotheses could not be extensively tested or challenged in systems other than the sea urchin, mainly because relevant data on gene regulation could not be collected on a high-resolution scale due to the difficulties in identifying regulatory elements in the genome. However, these limitations are now being overcome, and we foresee that many dGRNs will be deciphered in the near future (Hockman et al. 2018), along with more mechanistic studies about how changes in the genome cause changes in development.

Accurate deciphering dGRNs requires two interconnected efforts. The first is collecting relevant and high-throughput data on gene expression and regulation mechanisms across species, while the second effort is developing computational methods that use these expression and regulatory data to infer gene regulatory networks (GRNs), compare networks across species, and infer evolutionary histories (Thompson et al. 2015). Fortunately, tremendous technological and computational advances have been made not only to accurately measure gene expression and regulation at all levels, but also to systematically infer dGRNs from comparative data, and by doing so, to study the evolution of these complex networks. We here describe some of these technological and computational advances and how they have been applied to provide new insights into the inference and evolution of dGRNs.

High-throughput temporal and spatial nucleic acid surveys

The last decade has seen a radical change in the way we perform developmental and evolutionary biology experiments, largely propelled by the advent of HTS, which has radically fostered our ability to survey most molecules present in specific cell types, in temporal and spatial contexts, in a high-throughput manner. For instance, transcriptome sequencing (RNA-Seq) has been successfully leveraged to identify genes that may be relevant to embryonic development in diverse taxa. In this regard, transcriptional profiling of developmental transitions has been successfully carried out in several animals (Conaco et al. 2012; Akbari et al. 2013; Yang et al. 2013; Geib et al. 2014; Li et al. 2014; Nakanishi et al. 2015; Basu et al. 2016; Niu et al. 2016; Qi et al. 2016; Gaitán-Espitia and Hofmann 2017; Simon et al. 2017; Xu et al. 2017; Torres-Oliva et al. 2018). Additionally, large scale single embryo RNA-Seq studies leveraging HTS technology have been done in at least 10 different animal taxa throughout their developmental time course (Levin et al. 2016). These studies highlight how these technological advancements have facilitated the identification of gene cohorts expressed throughout the development of different species. However, it is clear that the analytical methods still have room for optimization, including how to more clearly defining orthology between different genes as well as taking into account the phylogenetic positioning of the sampled species (Altenhoff et al. 2016; Dunn et al. 2018). Finally, most of these studies have thus far been carried out in whole organisms, making it difficult to extrapolate the information acquired to direct inference of dGRNs, as the spatial positioning of the detected transcripts within the embryo is lost. To overcome this limitation, computational techniques are being developed to integrate high-quality expression data (i.e., *in situ* hybridization atlases) with transcriptomic approaches (Achim et al. 2015; Satija et al. 2015).

Beyond the identification of expressed transcripts, developmental transcriptomics has also allowed us to query the conservation of co-expressed gene modules that can be compared across different taxa (Israel et al. 2016). The identification of co-expression cues in transcriptomic experiments can provide a point of convergence between co-regulation and GRNs through the assumption that genes which participate in similar biological processes will share regulatory programs and, as a consequence, be co-expressed (Ruprecht et al. 2017). However, such approaches may be unable to capture expression dynamics that are largely constrained in time. For example, if a group of genes is only transiently co-regulated but this co-regulation is lost in subsequent stages, the co-regulation signal may not be identified

as significant due to its transient nature in comparison to signals present across the whole developmental time course.

Another crucial advancement in transcriptomics is the recent development of techniques to sequence both DNA and RNA extracted from single cells (Tang et al. 2009; Navin et al. 2011) as well as the development of methods to analyze such data. This is an area of rapid growth and development in the scientific community, as single-cell data have a number of particular caveats such as high levels of systematic bias due to several rounds of amplification, low sequencing coverage, and high-dropout rate (Bacher and Kendziorski 2016). Nevertheless, recent studies have successfully characterized the embryonic single cell transcriptomes of classical model systems such as the mouse (Scialdone et al. 2016; Mohammed et al. 2017), *Xenopus tropicalis* (Briggs et al. 2018), zebrafish (Farrell et al. 2018; Wagner et al. 2018), the fruit fly (Karaiskos et al. 2017), and the nematode (Cao et al. 2017), as well as non-classical model species such as bovines (Lavagi et al. 2018). Remarkably, these studies have demonstrated that this technique can be successfully applied to a variety of systems. It will not be long before the cost is reduced enough to enable such studies to not be limited to a few developmental stages but for a comprehensive developmental time course. This feat has thus far only been achieved in human and mouse preimplantation embryos (Guo et al. 2017; Stirparo et al. 2018).

Several recently developed HTS techniques have already become instrumental to mapping the interactors with genomic regulatory elements. These include techniques such as 3-C based methods (e.g., 3-C, 4-C, Hi-C, and ChIA-PET; reviewed in de Wit and de Laat (2012)), which allow us to directly query proximity of genomic regions via crosslinking, thus unveiling the 3D interactions of seemingly distant DNA stretches, for example, the interaction between distal enhancers with core-developmental control genes, facilitating the establishment of the regulatory hierarchy. Other techniques such as ATAC-Seq help us identify open regions in the chromatin which are available for molecular interactions (Buenrostro et al. 2013), while ChIP-Seq allows us identify which genomic sites are bound by a specific regulatory protein (e.g., transcription factor, RNA-polymerase, modified histones) (Barski et al. 2007; Johnson et al. 2007), and GRO-Seq enables the identification of RNAs that are being actively elongated in a particular condition (Core et al. 2008). In spite of their recent development, these techniques have already revealed general chromatin organization patterns, including that evolutionary divergent chromatin organization loci tend to harbor developmental genes (Chambers et al. 2013) and that *cis*-regulatory units are reorganized during embryonic development (Freire-Pritchett et al. 2017; Hockman et al. 2018). Furthermore, many of these techniques have begun to be successfully applied to single cells. For example, ATAC-Seq has been recently used to assess chromatin accessibility during *C. elegans* and *D. melanogaster* development, enabling the identification of developmental enhancers (Daugherty et al. 2017; Cusanovich et al. 2018), while ChiP-Seq has been applied to identify histone marks in mixed populations of embryonic stem (ES) cells, permitting identification of cell identity solely through these epigenetic marks (Rotem et al. 2015). The caveat that, generally, a single-cell can only be surveyed for a single type of measurement (either DNA, RNA-Seq, ChiP-Seq, etc.), meaning the information acquired at different levels may not be from the exact same cell, is being quickly overcome with recent techniques allowing simultaneously profile chromatin accessibility, methylation, and transcriptome (Macaulay et al. 2017; Clark et al. 2018).

Complementary to high-throughput quantification techniques for nucleic acids (DNA and RNA), automated technologies have greatly enhanced our capacity to generate detailed spatial maps of the intra-organismal and intracellular localization of nucleic acids. This includes well established techniques, such as robotic whole-mount *in situ* hybridization, which automates and parallelizes the spatial surveying of mRNAs (Quiring et al. 2004). In fact, public *in situ* data was instrumental to spatially organize all single cells sequenced in *Drosophila* embryos via the correlation of high expression of gene markers that had previously been spatially mapped as part of the Berkeley *Drosophila* Transcription Network Project (Karaiskos et al. 2017). This approach was so successful that virtual *in situ* results generated from the transcriptomic data were later corroborated experimentally (Karaiskos et al. 2017). Other recent computational advancements that have led to the development of image identification software through the use of machine learning have facilitated the automated and accurate identification of subcellular compartments in high-throughput *in situ* and single-cell immune fluorescence experiments (Lin et al. 2015). Additionally, techniques such as laser capture microdissection and high throughput Fluorescence-activated cell sorting now enable the precise selection of minute pieces of tissue and/or specific cell populations that can be used for sequencing (Datta et al. 2015; Handley et al. 2015).

One of the most recent technological advances is the possibility of doing perturbations in dGRNs using genome editing techniques such as CRISPR and its derivatives. In model systems, CRISPR/Cas9 has already been used for interrogating specific components of GRNs. For example, in mice it has been used to corroborate key genetic elements of a GRN responsible for retina differentiation (Wang et al. 2014) as well as the involvement of *Wnt* in the activation of *Hox* genes (Neijts et al. 2016). In insects, this system has also enabled uncoupling neurogenic GRNs in *Drosophila* (Rogers et al. 2017) and the identification of core genes involved in gene patterning in a butterfly (Prakash and Monteiro 2018). Although these studies have shown the applicability of CRISPR/Cas9 as a genetic tool to infer gene function, the perturbation assay is only made in single-locus. Recently, a new technique known as Perturb-Seq, which combines sc-RNA-Seq and CRISPR-based perturbations in a high-throughput manner, has emerged as a genetic screen tool able to perform multi-locus gene perturbation assays that allow the identification of individual gene targets, gene signatures, and cell states affected by individual perturbations and their genetic interactions (Dixit et al. 2016). However, we believe both single- and multi-locus gene perturbation assays through CRISPR/Cas9 are extremely important for dGRN studies as most perturbation assays in non-classical model species rely on antisense morpholinos to knockdown genes, a technique known to have unpredictable off-target effects, variations in efficiency, and potential toxicity (Lin and Su 2016). On the other hand, CRISPR/Cas9 has now been implemented in a diverse array of animals, including *Nematostella* (Ikmi et al. 2014), *Acropora* (Cleves et al. 2018), *Crepidula* (Perry and Henry 2015), *Parhyale* (Martin et al. 2016), *Daphnia* (Mohamad Ishak et al. 2017), sea urchin (Lin and Su 2016), and *Ciona* (Stolfi et al. 2014), opening the possibility of conducting perturbation experiments to specific GRNs and expediting the experimental characterization of dGRNs.

Major strategies for modeling GRNs

Ever since the metaphorical proposition of differentiation through canalization was proposed (Waddington 1957), several efforts have been undertaken to try to predict how the delimitations of cell differentiation occur from a molecular perspective. Network inference from transcriptomic data has been popular since the inception of microarrays (Chasman et al. 2016). Network inference approaches have been recently expanded with the use of transcriptomic data generated by HTS techniques. Researchers have made several efforts to try to use quantitative information and infer both the structure and the regulatory interactions from biological data. Several computational and modeling approaches can be used to study GRNs and other types of biological networks, and can be classified into two major categories: (1) network inference and reconstruction (Chasman et al. 2016) and (2) evolutionary network analysis (Thompson et al. 2015).

Regardless of the type of approach, the most difficult challenge is the adequate integration of heterogeneous data in the analysis. For example, gene co-expression data can provide a general idea of which genes are participating in a specific biological process, but they do not provide solid evidence for physical interactions or causality, and usually algorithms based only on time-course gene expression data do not perform better than random assignment of regulatory interactions (Lopes and Bontempi 2013). Hence, prior biological knowledge is required for more robust analyses (Banf and Rhee 2017a). Several types of biological data can be used as prior knowledge for network inference and reconstruction, for example, DNA-binding motifs for known specific transcription factors (Chasman et al. 2016), transcription factor orthology, and *cis*-regulatory module conservation (Thompson et al. 2015).

GRN reconstruction applied to embryonic development has the additional challenge, which is both experimental and computational, of integrating both spatial and temporal gene expression data as well as gene perturbation data in order to infer the actual regulators and targets in the dGRN (Li and Davidson 2009; Lopes and Bontempi 2013). Despite the complexity of dGRNs and the current challenges mentioned above, there are various examples of successful inference and reconstruction of GRNs in animals. For example, Ocone et al. (2015) proposed a method for learning the GRN dynamics during the differentiation of mouse hematopoietic stem cells from single-cell snapshot data. This method first reconstructs a dGRN and then estimates the gene expression dynamics based on the inferred network to compare it with the observed gene expression data (Ocone et al. 2015). Recently, Sanchez-Castillo et al. (2018) implemented a Bayesian method for dGRN inference and tested it with two temporal single-cell datasets: one spanning the zygote to blastocyst transition in mouse with measurements by qPCR; and the other corresponding to a differentiation time course of zebrafish hematopoietic stem cells with RNA-Seq. This method has reported a better performance than

previous inference methods that rely only on temporal gene expression measurements (Sanchez-Castillo et al. 2018). Additional methods have now been developed to enable dGRNs reconstruction from single cell data (Babtie et al. 2017); however, none of these methods integrates any other type of regulatory or functional data beyond gene expression yet.

Recently developed dGRN reconstruction methods in plants do have this capacity to integrate diverse high-throughput datasets to improve the network inference. For example, GRACE integrates gene co-expression with ChIP-Seq data to improve the inference of candidate regulatory interactions and successfully recovered most of the interactions involved in cell cycle control in *Arabidopsis thaliana* (Banf and Rhee 2017b). A second example includes the integration of detailed spatiotemporal transcriptomic data to infer a dGRN involved in *A. thaliana* root stem (de Luis Balaguer et al. 2017). Although these examples of dGRN reconstruction are focused on plants, they illustrate how the integration of different datasets, such as ChIP-Seq, chromatin accessibility, or spatiotemporal expression datasets, could improve the network inference in comparison to approaches that use gene expression data exclusively (de Luis Balaguer et al. 2017; Banf and Rhee 2017b).

Models showing that the state of regulatory genes is not random and can converge to predictable or stable states have been used to explain why differentiated cells do not dedifferentiate spontaneously and how their identities are maintained through time by reaching stable gene inactivation or activation patterns at the cellular level (Kauffman 1969, 1974). Using this theoretical framework and a mathematical definition for epigenetic landscape as the state space of all possible cell states borrowed from dynamical systems theory, Huang et al. (2005) proposed a model based on the mathematical representation of the cell states to model the transition from undifferentiated to differentiated cells. This kind of model has been also used to model differentiation in cancer (Huang et al. 2009), induced pluripotent stem cells (Huang 2010), and stem cells (Huang 2011). For a general discussion and details about landscape modeling, see Huang (2010) and Huang (2012). A modified version of landscape modeling was proposed more recently based on evidence from single-cell experiments: Moris et al. (2016) propose that cell fate transitions are both stochastic and discontinuous and discuss how this idea differs from the classical view of continuous cell fate transition in the Waddington's epigenetic landscape (Waddington 1957). There are also more recent approaches to the problem of cell state transition during differentiation. Herberg and Roeder (2015) review some of the models that had been developed to model pluripotency control in ES cells and discuss how a successful GRN modeling should help to explain how the control of pluripotency drives cell fate during normal development and why induced pluripotency is experimentally hard to achieve. Other models try to uncover the general mechanisms that drive cell differentiation at the single-cell level, for example, Stumpf et al. (2017) created a method to model stem cell differentiation in mouse neuronal lineage to explain transcriptional heterogeneity, this is, the observed differences in gene expression observed between cells derived from the same progenitor cell. This study concludes that a non-Markov stochastic model could explain cell differentiation in animal stem cells and account for cell heterogeneity whether the stochasticity of gene expression produced by the underlying pluripotency regulatory network is considered.

As these models remain to be applied into an organismal development context, the question that remains is whether we will ever be able to predict how cells will be changing identity through differentiation space by measuring the concentration of molecules within this cell during the transition.

The integration challenge of different high-throughput data and GRNs

As discussed earlier, we have now all the data necessary to do a thorough characterization of dGRNs. Although we can now generate high-throughput information on molecules composition, abundance, and spatial and temporal organization, we still lack precise ways to detect interactions accurately in a one to one manner. Furthermore, the sheer scale of HTS measurements inherently entails a number of errors (false positives and false negatives) that has to be taken into account when using this data for dGRN reconstruction. To add complexity, data integration itself is challenging and relies on what question we want to answer with our data in order to choose an unbiased or supervised integration strategy (Hawkins et al. 2010). These caveats highlight the great challenge in integrating results from diverse HTS experiments to understand the logic of complex systems such as dGRNs (Hawkins et al. 2010). We also now have experimental techniques to specifically perturb gene expression in selected tissues and developmental time-points as well as techniques to infer regulatory relationships using quantitative data with relative success. Therefore, a new question arises: How do we integrate this information within

a coherent and potentially predictive framework where regulatory genes, *cis*-regulatory modules, and differentiation genes and their interactions affected by perturbations are modeled to predict cell type differentiation or specific animal body plans?

As previously mentioned, there are already successfully integrated sets of data allowing the dissection of well-defined dGRNs. Some prime examples are the endomesoderm and endoderm specification dGRNs in the sea urchin *S. purpuratus*. Given its comprehensive characterization at molecular and spatiotemporal resolution, attempts have been made at modeling its interactions using Boolean models to predict stable states that may correlate with differentiated cell types (Peter et al. 2012; Peter and Davidson 2017). These networks have also been used in combination with transcriptomic information to do thorough comparative analysis of GRNs between sea urchins and sea stars (Hinman and Cheatle Jarvela 2014; Cary and Hinman 2017). These comparisons have unveiled remarkable conservation in the regulatory architecture of such networks, improving our comprehension of morphogenetic processes (e.g., echinoderm skeletogenesis) and how regulatory interactions drive modifications to developmental outcomes (e.g., morphological novelties), which ultimately, has provided a better understanding of the evolution of animal body plans. However, getting such detailed interaction networks took over 30 years of devoted work by the Davidson lab and others in the scientific community. Given that it remains highly impractical to carry out similar experimental dGRN characterization efforts in broad taxon sampling, can we capitalize on HTS technologies to accelerate our rate of understanding of dGRNs in non-classical model systems?

We believe the answer to this question lies in leveraging theoretical developments aimed at the inference of GRNs directly from quantitative measurements of the concentration of mRNA and other molecules combined with the recently developed technical approaches of surveying all mRNAs in shortly timed snapshots in single cells from full embryos (Fig. 1B). Multidimensional mapping of the transcription content of such sequencing has revealed that it is possible to differentiate groups of cells in the embryo (Cao et al. 2017; Karaiskos et al. 2017; Achim et al. 2018; Briggs et al. 2018; Farrell et al. 2018; Wagner et al. 2018); however, their localization cannot be solely inferred via their transcriptional content, as previously illustrated in the *Drosophila* and *Platynereis* embryos (Achim et al. 2015; Karaiskos et al. 2017). In this case, single-cell transcriptomic data had to be informed by *in situ* hybridization experiments to be able to accurately spatially place all sequenced cells in the embryo (Achim et al. 2015; Karaiskos et al. 2017). This highlights two facts: (1) there is no smooth spatial progression between the transcriptional information of different cell types, and (2) with current single-cell HTS technology, we are still unable to infer spatial localization solely from gene expression data.

This is not surprising given that discrete domains of expression and cell identity are well known to exist within developing embryos. Building on this idea, recent work has highlighted the possibility that cells are not smoothly transitioning between differentiation states, but that they undertake alternative highly heterogeneous expression patterns comprising a transition state, which is necessary for them to arrive to another differentiated or "attractor" state (reviewed in Moris et al. (2016)). A recent article has provided experimental evidence for this hypothesis, showing that single cells show higher transcriptional variation while they are transitioning between two differentiated states (Stumpf et al. 2017). These findings could help explain some of the high expression heterogeneity commonly observed in single-cell HTS data. However, this does not preclude us from using this integrated information to begin to computationally infer dGRNs at least at the level of identifiable cell populations through a species development.

Recent studies in basal metazoans have demonstrated HTS techniques can be used in non-classical model systems to acquire information about their underlying regulatory mechanisms. Two key studies have surveyed for histone modification marks associated with gene regulatory elements in a sponge (Gaiti et al. 2017) and a cnidarian (Schwaiger et al. 2014). Interestingly, both studies found evidence for typically metazoan looking enhancers, while lacking identifiable insulator elements. Indeed, orthologues of CTCF, a major effector protein which binds to insulators, seems to be found with high homology solely within Bilateria (Heger et al. 2012); however, the weaker association between TAD boundaries and CTCF in species such as *Drosophila* raises the possibility that other proteins could be carrying out similar functions to CTCF in non-bilaterians (Ramírez et al. 2018). How distal regulatory elements and nuclear topology influence gene expression in animal taxa with diverse chromosome number and genome size is one of the future frontiers of network inference approaches.

Final remarks and future perspectives

We are at a crucial moment in the development of HTS technologies and their integration. Unless we begin to use them to survey and model a larger part of the planet's biodiversity (Lewin et al. 2018) and untangle the GRNs in a number of organisms with a variety of life histories embryonic forms, informative phylogenetic position in the animal tree of life, we will be unable to acquire the necessary knowledge to understand the mechanisms of cell type and morphological evolution in animals. However, we believe this goal is now within reach through the integration of experimental and theoretical approaches to improve our predictive modeling of dGRNs in non-classical model systems that cannot be readily achieved using their genomic sequence alone.

In this review, we have argued for the use of mathematical modeling techniques as complementary tools for hypothesis generation in evolutionary developmental biology, specifically, testable hypothesis about dGRN architecture in non-classical model systems. We believe these approaches can be complementary to quantitative developmental biology techniques as these integrated approaches have already been proven useful to dissect and frame biologically the large amounts of quantitative data and spatially restricted data that we are now able to acquire using HTS technologies. We foresee that this experimental framework will enhance our ability to detect *cis*-regulatory sequences, gene interactions, and to predict dGRNs in different biological contexts such as cell type diversity and embryonic development. In addition, we believe that non-classical model systems are poised to most benefit from such approaches due to the rapid advancement of HTS technologies.

We expect in the near future that information derived from classical model system will be coupled with dGRNs from non-classical model systems to identify key evolutionary modifications that led to the fascinating diversity of animal forms seen in nature. Such integrated approaches will deliver new knowledge in dGRN evolution that will allow us to better understand the molecular mechanisms underpinning cell identity and development in animals and beyond.

Acknowledgments

The authors would like to thank Dr. Kevin Kocot and Dr. Erik Sperling for organizing the symposium "From Small and Squishy to Big and Armored: Genomic, Ecological and Paleontological Insights into the Early Evolution of Animals" as part of the 2018 SICB Annual Meeting.

Funding

S.L.F.-V. is supported by a Cátedras CONACYT Fellowship – 2015, F.A. is supported by a CONICYT PAI/Concurso Nacional Insercion en la Academia, Convocatoria 2017 (79170033) and R.A.R.-D. is supported by a CONACYT Master's Research Scholarship.

References

Achim K, Eling N, Vergara HM, Bertucci PY, Musser J, Vopalensky P, Brunet T, Collier P, Benes V, Marioni JC, et al. 2018. Whole-body single-cell sequencing reveals transcriptional domains in the annelid larval body. Mol Biol Evol 35:1047–62.

Achim K, Pettit J-B, Saraiva LR, Gavriouchkina D, Larsson T, Arendt D, Marioni JC. 2015. High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. Nat Biotechnol 33:503–9.

Aguilera F, McDougall C, Degnan BM. 2017. Co-option and de novo gene evolution underlie molluscan shell diversity. Mol Biol Evol 34:779–92.

Akbari OS, Antoshechkin I, Amrhein H, Williams B, Diloreto R, Sandler J, Hay BA. 2013. The developmental transcriptome of the mosquito *Aedes aegypti*, an invasive species and major arbovirus vector. G3; Genes|Genomes|Genetics 3:1493–509.

Altenhoff AM, Boeckmann B, Capella-Gutiérrez S, Dalquen DA, DeLuca T, Forslund K, Huerta-Cepas J, Linard B, Pereira C, Pryszcz LP, et al. 2016. Standardized benchmarking in the quest for orthologs. Nat Methods 13:425–30.

Andrikou C, Arnone MI. 2015. Too many ways to make a muscle: evolution of GRNs governing myogenesis. Zool Anz 256:2–15.

Arendt D, Musser JM, Baker CVH, Bergman A, Cepko C, Erwin DH, Pavlicev M, Schlosser G, Widder S, Laubichler MD, et al. 2016. The origin and evolution of cell types. Nat Rev Genet 17:744–57.

Arnone MI, Martin EL, Davidson EH. 1998. Cis-regulation downstream of cell type specification: a single compact element controls the complex expression of the CyIIa gene in sea urchin embryos. Development 125:1381–95.

Arthur W. 2000. The concept of developmental reprogramming and the quest for an inclusive theory of evolutionary mechanisms. Evol Dev 2:49–57.

Babtie AC, Chan TE, Stumpf MPH. 2017. Learning regulatory models for cell development from single cell transcriptomic data. Curr Opin Syst Biol 5:72–81.

Bacher R, Kendziorski C. 2016. Design and computational analysis of single-cell RNA-sequencing experiments. Genome Biol 17:63.

Banf M, Rhee SY. 2017a. Computational inference of gene regulatory networks: approaches, limitations and opportunities. Biochim Biophys Acta 1860:41–52.

Banf M, Rhee SY. 2017b. Enhancing gene regulatory network inference through data integration with Markov random fields. Sci Rep 7:41174.

Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. Cell 129:823–37.

Basu S, Hadzhiev Y, Petrosino G, Nepal C, Gehrig J, Armant O, Ferg M, Strähle U, Sanges R, Müller F. 2016. The *Tetraodon nigroviridis* reference transcriptome: developmental transition, length retention and microsynteny of long non-coding RNAs in a compact vertebrate genome. Sci Rep 6:33210.

Briggs JA, Weinreb C, Wagner DE, Megason S, Peshkin L, Kirschner MW, Klein AM. 2018. The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. Science 360:eaar5780.

Britten RJ, Davidson EH. 1969. Gene regulation for higher cells: a theory. Science 165:349–57.

Britten RJ, Davidson EH. 1971. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. Q Rev Biol 46:111–38.

Brunet T, King N. 2017. The origin of animal multicellularity and cell differentiation. Dev Cell 43:124–40.

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 10:1213–18.

Calzone FJ, Thézé N, Thiebaud P, Hill RL, Britten RJ, Davidson EH. 1988. Developmental appearance of factors that bind specifically to cis-regulatory sequences of a gene expressed in the sea urchin embryo. Genes Dev 2:1074–88.

Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, Daza R, Qiu X, Lee C, Furlan SN, Steemers FJ, et al. 2017. Comprehensive single-cell transcriptional profiling of a multicellular organism. Science 357:661–7.

Carroll SB. 2008. Evo-Devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell 134:25–36.

Cary GA, Hinman VF. 2017. Echinoderm development and evolution in the post-genomic era. Dev Biol 427:203–11.

Chambers EV, Bickmore WA, Semple CA. 2013. Divergence of mammalian higher order chromatin structure is associated with developmental loci. PLoS Comp Biol 9:e1003017.

Chasman D, Fotuhi Siahpirani A, Roy S. 2016. Network-based approaches for analysis of complex biological systems. Curr Opin Biotechnol 39:157–66.

Clark SJ, Argelaguet R, Kapourani C-A, Stubbs TM, Lee HJ, Alda-Catalinas C, Krueger F, Sanguinetti G, Kelsey G, Marioni JC, et al. 2018. scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. Nat Commun 9:781.

Cleves PA, Strader ME, Bay LK, Pringle JR, Matz MV. 2018. CRISPR/Cas9-mediated genome editing in a reef-building coral. Proc Natl Acad Sci U S A 115:5235–40.

Coffman JA, Kirchhamer CV, Harrington MG, Davidson EH. 1997. SpMyb functions as an intramodular repressor to regulate spatial expression of CyIIIa in sea urchin embryos. Development 124:4717–27.

Conaco C, Neveu P, Zhou H, Arcila ML, Degnan SM, Degnan BM, Kosik KS. 2012. Transcriptome profiling of the demosponge *Amphimedon queenslandica* reveals genome-wide events that accompany major life cycle transitions. BMC Genomics13:209.

Core LJ, Waterfall JJ, Lis JT. 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science 322:1845–8.

Cusanovich DA, Reddington JP, Garfield DA, Daza RM, Aghamirzaie D, Marco-Ferreres R, Pliner HA, Christiansen L, Qiu X, Steemers FJ, et al. 2018. The *cis*-regulatory dynamics of embryonic development at single-cell resolution. Nature 555:538.

Datta S, Malhotra L, Dickerson R, Chaffee S, Sen CK, Roy S. 2015. Laser capture microdissection: big data from small samples. Histol Histopathol 30:1255–69.

Daugherty AC, Yeo RW, Buenrostro JD, Greenleaf WJ, Kundaje A, Brunet A. 2017. Chromatin accessibility dynamics reveal novel functional enhancers in *C. elegans*. Genome Res 27:2096–107.

Davidson EH. 2010a. Emerging properties of animal gene regulatory networks. Nature 468:911–20.

Davidson EH. 2010b. The regulatory genome. Burlington (MA): Academic Press.

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

Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, Yuh C-H, Minokawa T, Amore G, Hinman V, Arenas MC, et al. 2002. A genomic regulatory network for development. Science 295:1669–78.

de Luis Balaguer MA, Fisher AP, Clark NM, Fernandez-Espinosa MG, Möller BK, Weijers D, Lohmann JU, Williams C, Lorenzo O, Sozzani R. 2017. Predicting gene regulatory networks by combining spatial and temporal gene expression data in *Arabidopsis* root stem cells. Proc Natl Acad Sci U S A 114:E7632–40.

de Mendoza A, Sebé-Pedrós A, Šestak MS, Matejčić M, Torruella G, Domazet-Lošo T, Ruiz-Trillo I. 2013. Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. Proc Natl Acad Sci U S A 110:E4858–66.

de Wit E, de Laat W. 2012. A decade of 3C technologies: insights into nuclear organization. Genes Dev 26:11–24.

Dixit A, Parnas O, Li B, Chen J, Fulco CP, Jerby-Arnon L, Marjanovic ND, Dionne D, Burks T, Raychowdhury R, et al. 2016. Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. Cell 167:1853–66.e17.

Dunn CW, Zapata F, Munro C, Siebert S, Hejnol A. 2018. Pairwise comparisons across species are problematic when analyzing functional genomic data. Proc Natl Acad Sci U S A 115:E409–17.

Dylus DV, Czarkwiani A, Blowes LM, Elphick MR, Oliveri P. 2018. Developmental transcriptomics of the brittle star *Amphiura filiformis* reveals gene regulatory network rewiring in echinoderm larval skeleton evolution. Genome Biol 19: 26.

Erkenbrack EM. 2016. Divergence of ectodermal and mesodermal gene regulatory network linkages in early development of sea urchins. Proc Natl Acad Sci U S A 113:E7202–11.

Erwin DH. 2009. Early origin of the bilaterian developmental toolkit. Philos Trans R Soc Lond B Biol Sci 364:2253–61.

Farrell JA, Wang Y, Riesenfeld SJ, Shekhar K, Regev A, Schier AF. 2018. Single-cell reconstruction of developmental

trajectories during zebrafish embryogenesis. Science 360:eaar3131.
Freire-Pritchett P, Schoenfelder S, Várnai C, Wingett SW, Cairns J, Collier AJ, García-Vílchez R, Furlan-Magaril M, Osborne CS, Fraser P, et al. 2017. Global reorganisation ofcis-regulatory units upon lineage commitment of human embryonic stem cells. eLife 6:e21926.
Gaitán-Espitia JD, Hofmann GE. 2017. Gene expression profiling during the embryo-to-larva transition in the giant red sea urchin *Mesocentrotus franciscanus*. Ecol Evol 7:2798–811.
Gaiti F, Jindrich K, Fernandez-Valverde SL, Roper KE, Degnan BM, Tanurdzić M. 2017. Landscape of histone modifications in a sponge reveals the origin of animal cis-regulatory complexity. eLife 6:e21926.
Geib SM, Calla B, Hall B, Hou S, Manoukis NC. 2014. Characterizing the developmental transcriptome of the oriental fruit fly, *Bactrocera dorsalis* (Diptera: tephritidae) through comparative genomic analysis with *Drosophila melanogaster* utilizing modENCODE datasets. BMC Genomics 15:942.
Goode DK, Obier N, Vijayabaskar MS, Lie-A-Ling M, Lilly AJ, Hannah R, Lichtinger M, Batta K, Florkowska M, Patel R, et al. 2016. Dynamic gene regulatory networks drive hematopoietic specification and differentiation. Dev Cell 36:572–87.
Gouti M, Delile J, Stamataki D, Wymeersch FJ, Huang Y, Kleinjung J, Wilson V, Briscoe J. 2017. A gene regulatory network balances neural and mesoderm specification during vertebrate trunk development. Dev Cell 41:243–7.
Guo F, Li L, Li J, Wu X, Hu B, Zhu P, Wen L, Tang F. 2017. Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. Cell Res 27:967–88.
Hamey FK, Nestorowa S, Kinston SJ, Kent DG, Wilson NK, Göttgens B. 2017. Reconstructing blood stem cell regulatory network models from single-cell molecular profiles. Proc Natl Acad Sci U S A 114:5822–9.
Handley A, Schauer T, Ladurner AG, Margulies CE. 2015. Designing cell-type-specific genome-wide experiments. Mol Cell 58:621–31.
Hawkins RD, Hon GC, Ren B. 2010. Next-generation genomics: an integrative approach. Nat Rev Genet 11:476–86.
Heger P, Marin B, Bartkuhn M, Schierenberg E, Wiehe T. 2012. The chromatin insulator CTCF and the emergence of metazoan diversity. Proc Natl Acad Sci U S A 109:17507–12.
Herberg M, Roeder I. 2015. Computational modelling of embryonic stem-cell fate control. Development 142:2250–60.
Hinman VF, Cheatle Jarvela AM. 2014. Developmental gene regulatory network evolution: insights from comparative studies in echinoderms. Genesis 52:193–207.
Hockman D, Chong-Morrison V, Gavriouchkina D, Green S, Amemiya CT, Smith JJ, Bronner ME, Sauka-Spengler T. 2018. A genome-wide assessment of the ancestral neural crest gene regulatory network (https://www.biorxiv.org/content/early/2018/03/12/275875.full.pdf+html).
Hough-Evans BR, Franks RR, Zeller RW, Britten RJ, Davidson EH. 1990. Negative spatial regulation of the lineage specific CyIIIa actin gene in the sea urchin embryo. Development 110:41–50.
Huang S. 2010. Cell lineage determination in state space: a systems view brings flexibility to dogmatic canonical rules. PLoS Biol 8:e1000380.
Huang S. 2011. Systems biology of stem cells: three useful perspectives to help overcome the paradigm of linear pathways. Philos Trans R Soc Lond B Biol Sci 366:2247–59.
Huang S. 2012. The molecular and mathematical basis of Waddington's epigenetic landscape: a framework for post-Darwinian biology? Bioessays 34:149–57.
Huang S, Eichler G, Bar-Yam Y, Ingber DE. 2005. Cell fates as high-dimensional attractor states of a complex gene regulatory network. Phys Rev Lett 94:128701.
Huang S, Ernberg I, Kauffman S. 2009. Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. Semin Cell Dev Biol 20:869–76.
Macaulay IC, Ponting CP, Voet T. 2017. Single-cell multiomics: multiple measurements from single cells. Trends Genet 33:155–68.
Ikmi A, McKinney SA, Delventhal KM, Gibson MC. 2014. TALEN and CRISPR/Cas9-mediated genome editing in the early-branching metazoan *Nematostella vectensis*. Nat Commun 5:5486.
Inoue T, Wang M, Ririe TO, Fernandes JS, Sternberg PW. 2005. Transcriptional network underlying *Caenorhabditis elegans* vulval development. Proc Natl Acad Sci U S A 102:4972–7.
Israel JW, Martik ML, Byrne M, Raff EC, Raff RA, McClay DR, Wray GA. 2016. Comparative developmental transcriptomics reveals rewiring of a highly conserved gene regulatory network during a major life history switch in the sea urchin genus *Heliocidaris*. PLoS Biol 14:e1002391.
Johnson DS, Mortazavi A, Myers RM, Wold B. 2007. Genome-wide mapping of in vivo protein-DNA interactions. Science 316:1497–502.
Karaiskos N, Wahle P, Alles J, Boltengagen A, Ayoub S, Kipar C, Kocks C, Rajewsky N, Zinzen RP. 2017. The *Drosophila* embryo at single-cell transcriptome resolution. Science358:194–9.
Kauffman S. 1974. The large scale structure and dynamics of gene control circuits: an ensemble approach. J Theor Biol 44:167–90.
Kauffman SA. 1969. Metabolic stability and epigenesis in randomly constructed genetic nets. J Theor Biol 22:437–67.
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 122:333–48.
Kittelmann S, Buffry AD, Franke FA, Almudi I, Yoth M, Sabaris G, Couso JP, Nunes MDS, Frankel N, Gómez-Skarmeta JL, et al. 2018. Gene regulatory network architecture in different developmental contexts influences the genetic basis of morphological evolution. PLoS Genet 14:e1007375.
Larroux C, Luke GN, Koopman P, Rokhsar DS, Shimeld SM, Degnan BM. 2008. Genesis and expansion of metazoan transcription factor gene classes. Mol Biol Evol 25:980–96.
Lavagi I, Krebs S, Simmet K, Beck A, Zakhartchenko V, Wolf E, Blum H. 2018. Single-cell RNA sequencing reveals developmental heterogeneity of blastomeres during major genome activation in bovine embryos. Sci Rep 8:4071.
Leininger S, Adamski M, Bergum B, Guder C, Liu J, Laplante M, Bråte J, Hoffmann F, Fortunato S, Jordal S, et al. 2014. Developmental gene expression provides clues to relationships between sponge and eumetazoan body plans. Nat Commun 5:3905.

Levin M, Anavy L, Cole AG, Winter E, Mostov N, Khair S, Senderovich N, Kovalev E, Silver DH, Feder M, et al. 2016. The mid-developmental transition and the evolution of animal body plans. Nature 531:637–41.

Levine M. 2010. Transcriptional enhancers in animal development and evolution. Curr Biol 20:R754–63.

Levine M, Davidson EH. 2005. Gene regulatory networks for development. Proc Natl Acad Sci U S A 102:4936–42.

Lewin HA, Robinson GE, Kress WJ, Baker WJ, Coddington J, Crandall KA, Durbin R, Edwards SV, Forest F, Gilbert MTP, et al. 2018. Earth BioGenome Project: sequencing life for the future of life. Proc Natl Acad Sci U S A 115:4325–33.

Li E, Davidson EH. 2009. Building developmental gene regulatory networks. Birth Defects Res C Embryo Today 87:123–30.

Li JJ, Huang H, Bickel PJ, Brenner SE. 2014. Comparison of *D. melanogaster* and *C. elegans* developmental stages, tissues, and cells by modENCODE RNA-seq data. Genome Res 24:1086–101.

Lin C-Y, Su Y-H. 2016. Genome editing in sea urchin embryos by using a CRISPR/Cas9 system. Dev Biol 409:420–8.

Lin J-R, Fallahi-Sichani M, Sorger PK. 2015. Highly multiplexed imaging of single cells using a high-throughput cyclic immunofluorescence method. Nat Commun 6:8390.

Loose M, Patient R. 2004. A genetic regulatory network for *Xenopus* mesendoderm formation. Dev Biol 271:467–78.

Lopes M, Bontempi G. 2013. Experimental assessment of static and dynamic algorithms for gene regulation inference from time series expression data. Front Genet 4:303.

Lynch VJ, May G, Wagner GP. 2011. Regulatory evolution through divergence of a phosphoswitch in the transcription factor CEBPB. Nature 480:383–6.

Lynch VJ, Wagner GP. 2008. Resurrecting the role of transcription factor change in developmental evolution. Evolution 62:2131–54.

Maduro MF, Rothman JH. 2002. Making worm guts: the gene regulatory network of the *Caenorhabditis elegans* endoderm. Dev Biol 246:68–85.

Martin A, Serano JM, Jarvis E, Bruce HS, Wang J, Ray S, Barker CA, O'Connell LC, Patel NH. 2016. CRISPR/Cas9 mutagenesis reveals versatile roles of *Hox* genes in crustacean limb specification and evolution. Curr Biol 26:14–26.

Martindale MQ. 2005. The evolution of metazoan axial properties. Nat Rev Genet 6:917–27.

Martín-Durán JM, Pang K, Børve A, Lê HS, Furu A, Cannon JT, Jondelius U, Hejnol A. 2018. Convergent evolution of bilaterian nerve cords. Nature 553:45–50.

Mohamad Ishak NS, Nong QD, Matsuura T, Kato Y, Watanabe H. 2017. Co-option of the bZIP transcription factor Vrille as the activator of Doublesex1 in environmental sex determination of the crustacean *Daphnia magna*. PLoS Genet 13:e1006953.

Mohammed H, Hernando-Herraez I, Savino A, Scialdone A, Macaulay I, Mulas C, Chandra T, Voet T, Dean W, Nichols J, et al. 2017. Single-cell landscape of transcriptional heterogeneity and cell fate decisions during mouse early gastrulation. Cell Rep 20:1215–28.

Moris N, Pina C, Arias AM. 2016. Transition states and cell fate decisions in epigenetic landscapes. Nat Rev Genet 17:693–703.

Nakanishi N, Camara AC, Yuan DC, Gold DA, Jacobs DK. 2015. Gene expression data from the moon jelly, *Aurelia*, provide insights into the evolution of the combinatorial code controlling animal sense organ development. PLoS One 10:e0132544.

Nasiadka A, Dietrich BH, Krause HM. 2002. Anterior-posterior patterning in the *Drosophila* embryo. Adv Dev Biol Biochem 12:155–204.

Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D, Esposito D, et al. 2011. Tumour evolution inferred by single-cell sequencing. Nature 472:90–4.

Neijts R, Amin S, van Rooijen C, Tan S, Creyghton MP, de Laat W, Deschamps J. 2016. Polarized regulatory landscape and Wnt responsiveness underlie Hox activation in embryos. Genes Dev 30:1937–42.

Niu D, Wang F, Xie S, Sun F, Wang Z, Peng M, Li J. 2016. Developmental transcriptome analysis and identification of genes involved in larval metamorphosis of the razor clam, *Sinonovacula constricta*. Mar Biotechnol 18:168–75.

Ocone A, Haghverdi L, Mueller NS, Theis FJ. 2015. Reconstructing gene regulatory dynamics from high-dimensional single-cell snapshot data. Bioinformatics 31:i89–96.

Parfitt D-E, Shen MM. 2014. From blastocyst to gastrula: gene regulatory networks of embryonic stem cells and early mouse embryogenesis. Philos Trans R Soc Lond B Biol Sci 369:20130542.

Perry KJ, Henry JQ. 2015. CRISPR/Cas9-mediated genome modification in the mollusc, *Crepidula fornicata*. Genesis 53:237–44.

Peter IS, Davidson EH. 2010. The endoderm gene regulatory network in sea urchin embryos up to mid-blastula stage. Dev Biol 340:188–99.

Peter IS, Davidson EH. 2011a. Evolution of gene regulatory networks controlling body plan development. Cell 144:970–85.

Peter IS, Davidson EH. 2011b. A gene regulatory network controlling the embryonic specification of endoderm. Nature 474:635–9.

Peter IS, Davidson EH. 2017. Assessing regulatory information in developmental gene regulatory networks. Proc Natl Acad Sci U S A 114:5862–9.

Peter IS, Faure E, Davidson EH. 2012. Predictive computation of genomic logic processing functions in embryonic development. Proc Natl Acad Sci U S A 109:16434–42.

Prakash A, Monteiro A. 2018. apterous A specifies dorsal wing patterns and sexual traits in butterflies. Proc R Soc Lond B Biol Sci 285:20172685 (doi: 10.1098/rspb.2017.2685).

Qi L, Fang Q, Zhao L, Xia H, Zhou Y, Xiao J, Li K, Ye G. 2016. De novo assembly and developmental transcriptome analysis of the small white butterfly *Pieris rapae*. PLoS One 11:e0159258.

Quiring R, Wittbrodt B, Henrich T, Ramialison M, Burgtorf C, Lehrach H, Wittbrodt J. 2004. Large-scale expression screening by automated whole-mount in situ hybridization. Mech Dev 121:971–6.

Rafiq K, Shashikant T, McManus CJ, Ettensohn CA. 2014. Genome-wide analysis of the skeletogenic gene regulatory network of sea urchins. Development 141:950–61.

Ramírez F, Bhardwaj V, Arrigoni L, Lam KC, Grüning BA, Villaveces J, Habermann B, Akhtar A, Manke T. 2018.

High-resolution TADs reveal DNA sequences underlying genome organization in flies. Nat Commun 9:189.
Rebeiz M, Patel NH, Hinman VF. 2015. Unraveling the tangled skein: the evolution of transcriptional regulatory networks in development. Annu Rev Genomics Hum Genet 16:103–31.
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 105:20095–9.
Rogers WA, Goyal Y, Yamaya K, Shvartsman SY, Levine MS. 2017. Uncoupling neurogenic gene networks in the *Drosophila* embryo. Genes Dev31:634–8.
Rotem A, Ram O, Shoresh N, Sperling RA, Goren A, Weitz DA, Bernstein BE. 2015. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nat Biotechnol 33:1165–72.
Ruprecht C, Vaid N, Proost S, Persson S, Mutwil M. 2017. Beyond genomics: studying evolution with gene coexpression networks. Trends Plant Sci 22:298–307.
Sanchez-Castillo M, Blanco D, Tienda-Luna IM, Carrion MC, Huang Y. 2018. A Bayesian framework for the inference of gene regulatory networks from time and pseudo-time series data. Bioinformatics 34:964–70.
Sandmann T, Girardot C, Brehme M, Tongprasit W, Stolc V, Furlong EEM. 2007. A core transcriptional network for early mesoderm development in *Drosophila melanogaster*. Genes Dev 21:436–49.
Satija R, Farrell JA, Gennert D, Schier AF, Regev A. 2015. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 33:495–502.
Schroeder MD, Pearce M, Fak J, Fan H, Unnerstall U, Emberly E, Rajewsky N, Siggia ED, Gaul U. 2004. Transcriptional control in the segmentation gene network of *Drosophila*. PLoS Biol 2:E271.
Schwaiger M, Schönauer A, Rendeiro AF, Pribitzer C, Schauer A, Gilles AF, Schinko JB, Renfer E, Fredman D, Technau U. 2014. Evolutionary conservation of the eumetazoan gene regulatory landscape. Genome Res 24:639–50.
Scialdone A, Tanaka Y, Jawaid W, Moignard V, Wilson NK, Macaulay IC, Marioni JC, Göttgens B. 2016. Resolving early mesoderm diversification through single-cell expression profiling. Nature 535:289–93.
Sebé-Pedrós A, de Mendoza A. 2015. Transcription Factors and the Origin of Animal Multicellularity. In: Ruiz-Trillo I, Nedelcu A, editors. Evolutionary transitions to multicellular life. advances in marine genomics. Vol. 2. Dordrecht: Springer.
Sebé-Pedrós A, Degnan BM, Ruiz-Trillo I. 2017. The origin of Metazoa: a unicellular perspective. Nat Rev Genet 18:498–512.
Simon S, Sagasser S, Saccenti E, Brugler MR, Schranz ME, Hadrys H, Amato G, DeSalle R. 2017. Comparative transcriptomics reveal developmental turning points during embryogenesis of a hemimetabolous insect, the damselfly *Ischnura elegans*. Sci Rep 7:13547.
Sommer RJ. 2012. Evolution of regulatory networks: nematode vulva induction as an example of developmental systems drift. Adv Exp Med Biol 751:79–91.
Stirparo GG, Boroviak T, Guo G, Nichols J, Smith A, Bertone P. 2018. Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. Development 145.
Stolfi A, Gandhi S, Salek F, Christiaen L. 2014. Tissue-specific genome editing in *Ciona* embryos by CRISPR/Cas9. Development 141:4115–20.
Stumpf PS, Smith RCG, Lenz M, Schuppert A, Müller F-J, Babtie A, Chan TE, Stumpf MPH, Please CP, Howison SD, et al. 2017. Stem cell differentiation as a non-Markov stochastic process. Cell Syst 5:268–82.e7.
Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A, et al. 2009. mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods 6:377–82.
Thompson D, Regev A, Roy S. 2015. Comparative analysis of gene regulatory networks: from network reconstruction to evolution. Annu Rev Cell Dev Biol 31:399n428.
Torres-Oliva M, Schneider J, Wiegleb G, Kaufholz F, Posnien N. 2018. Dynamic genome wide expression profiling of *Drosophila* head development reveals a novel role of Hunchback in retinal glia cell development and blood-brain barrier integrity. PLoS Genet 14:e1007180.
Waddington CH. 1957. The strategy of the genes. London: Allen and Unwin. Reprinted 2014.
Wagner DE, Weinreb C, Collins ZM, Briggs JA, Megason SG, Klein AM. 2018. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. Science 108 (doi:10.1126/science.aar4362).
Wang DG, Kirchhamer CV, Britten RJ, Davidson EH. 1995. SpZ12-1, a negative regulator required for spatial control of the territory-specific CyIIIa gene in the sea urchin embryo. Development 121:1111–22.
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 30:513–27.
Wittkopp PJ, Vaccaro K, Carroll SB. 2002. Evolution of yellow gene regulation and pigmentation in *Drosophila*. Curr Biol12:1547–56.
Xu EG, Mager EM, Grosell M, Stieglitz JD, Hazard ES, Hardiman G, Schlenk D. 2017. Developmental transcriptomic analyses for mechanistic insights into critical pathways involved in embryogenesis of pelagic mahi-mahi (*Coryphaena hippurus*). PLoS One 12:e0180454.
Yang H, Zhou Y, Gu J, Xie S, Xu Y, Zhu G, Wang L, Huang J, Ma H, Yao J. 2013. Deep mRNA sequencing analysis to capture the transcriptome landscape of zebrafish embryos and larvae. PLoS One 8:e64058.
Yuh CH, Davidson EH. 1996. Modular cis-regulatory organization of Endo16, a gut-specific gene of the sea urchin embryo. Development 122:1069–82.
Yuh CH, Ransick A, Martinez P, Britten RJ, Davidson EH. 1994. Complexity and organization of DNA-protein interactions in the 5'-regulatory region of an endoderm-specific marker gene in the sea urchin embryo. Mech Dev 47:165–86.
Zeitlinger J, Stark A. 2010. Developmental gene regulation in the era of genomics. Dev Biol 339:230–9.
Zeller RW, Griffith JD, Moore JG, Kirchhamer CV, Britten RJ, Davidson EH. 1995. A multimerizing transcription factor of sea urchin embryos capable of looping DNA. Proc Natl Acad Sci U S A 92:2989–93.