Modeling Evolution of Developmental Gene Regulatory Networks

Renske M. A. Vroomans and Kirsten H. W. J. ten Tusscher

Contents

Abstract

The field of evo-devo studies what, how, and why developmental patterning processes have evolved. While comparative approaches based in experimental data are essential for answering the first two types of questions, evo-devo simulations studies are critical to answer why questions. By simulating evo-devo processes, the evolutionary tape can be replayed both under the same and different conditions, enabling us to answer questions on contingency, convergence, and constraints and their roles in determining evolutionary outcomes.

In this chapter, we describe the basic ingredients of computational models simulating evo-devo processes: gene expression regulation; cell and tissue behavior; and mutation-selection driven evolution. We describe for each of these model ingredients the choices that need to be made, e.g., whether the model simulates a one, two, or three-dimensional tissue, and how these affect computational efficiency as well as modeling outcomes. We focus on the importance of

R. M. A. Vroomans (✉) · K. H. W. J. ten Tusscher
Theoretical Biology Group, Department of Biology, Utrecht University, Utrecht, The Netherlands
e-mail: renske.vroomans@gmail.com

L. Nuño de la Rosa, G. B. Müller (eds.), *Evolutionary Developmental Biology*,
https://doi.org/10.1007/978-3-319-32979-6_118

incorporating a realistic, nonlinear, and evolvable genotype-phenotype map in evo-devo simulation models.

We end with an illustration of how evo-devo models have helped answer why questions in the field of animal body plan segmentation.

Keywords

Computational modeling · Evolutionary simulations · Gene regulatory networks · Genotype-phenotype mapping · Robustness and evolvability

Introduction

The field of evo-devo studies the reciprocal impact of evolution on development and development on evolution. The ultimate aim is to determine what, how, and why particular developmental patterning processes have evolved. Comparative approaches, such as anatomical comparisons, spatio-temporal gene expression mapping, or bioinformatic analysis of genome composition, may uncover which changes underlie the phenotypic differences we observe between extant species. They are instead less suitable for addressing *why* developmental programs have evolved and diverged along particular trajectories; for instance, it may be the case that a particular patterning process is discovered through mutations more easily than other patterning mechanisms, or instead, it may confer a selective advantage. The observation that evolution has followed a certain path in a certain species is insufficient to discern whether this path is just one of many possible evolutionary outcomes (all leading to different developmental programs), or rather that given a second or third chance, a similar trajectory would have been followed. In the latter case, even when unrelated species appear to have evolved similar developmental traits, it is hard to determine whether this convergence arose because this developmental mechanism confers the highest fitness advantage or because constraints arising from a limited toolkit of developmental genes or prior evolutionary processes reduced the evolutionary accessibility of alternative mechanisms Chipman (2010).

In-silico modeling provides us with a means to address these types of *why* questions. Depending on the particular approach, models can be used to investigate the evolutionary accessibility of different theoretically inferred developmental mechanisms that are capable of generating the same phenotype or to study their robustness to a variety of perturbations (Cotterell and Sharpe (2010); Jiménez et al. (2015); Salazar-Ciudad et al. (2001)). Furthermore, computer simulations allow us to "replay the evolutionary tape": letting the same developmental character evolve multiple times in silico to assess the likelihood of finding various alternative mechanisms to generate this character. With such simulations we can also compare the evolutionary consequences of a variety of different conditions, for instance, the presence or absence of gene expression noise or a morphogen gradient (Vroomans et al. 2016) – or what happens if two developmental

traits evolve at the same time (ten Tusscher and Hogeweg 2011). Finally, such simulations generate a perfect "*in-silico* fossil record" of both the genotype and phenotype of all ancestors leading up to an evolutionary outcome. This enables one to reconstruct the precise mutational trajectory leading up to this outcome and investigate whether convergent evolutionary outcomes arise from similar or different mutational routes (see chapter ▶ "Convergence").

In this chapter, we will discuss how *in-silico* models of evo-devo are built up and the different ways they have been used to tackle the *why* questions of developmental processes and their evolution.

Simulating Development

Development occurs on multiple levels: it involves processes that range from the subcellular polarization patterns within cells, via division, movement, cell fate and shape changes of individual cells to overall tissue-scale growth, patterning, and morphogenesis (see chapter ▶ "Levels of Organization in Evo-Devo"). Evo-devo models are necessarily simplified to keep them manageable in terms of required computational time and the ease with which results can be analyzed and understood. As a consequence, these models typically incorporate two organizational levels of development – cells and tissues – while subcellular patterning is usually ignored. Despite these simplifications, many choices still need to be made: how gene expression regulation and dynamics are modeled, what types of cell-cell communication are considered, whether cell division and growth are modeled explicitly, and more.

These modeling choices may have consequences for the types of developmental mechanisms that can be captured in the model, as well as the evolutionary questions that can be answered. In this section, we first discuss the building blocks necessary to simulate developmental processes. We discuss some of the different modeling choices that can be made, their advantages and disadvantages, and their consequences for the evolutionary process.

Within Cells

Gene Expression Levels There are two main ways in which gene expression levels can be modeled: Boolean or continuous (Fig. 1A). In models using boolean gene expression, only two levels of gene expression are distinguished: no expression (0) or full expression (1). Boolean models are computationally much more efficient and therefore attractive if one aims to investigate large networks containing many genes. Furthermore, Boolean models contain few parameters and therefore enable a qualitative analysis of network behavior when there is little information available on kinetic constants (Spirov and Holloway 2013). However, with the Boolean modeling formalism, the gradual activation or inhibition of a gene, or the graded expression of a gene across a tissue, cannot be simulated. To overcome this limitation and yet maintain the computational efficiency advantages, some modelers have extended the

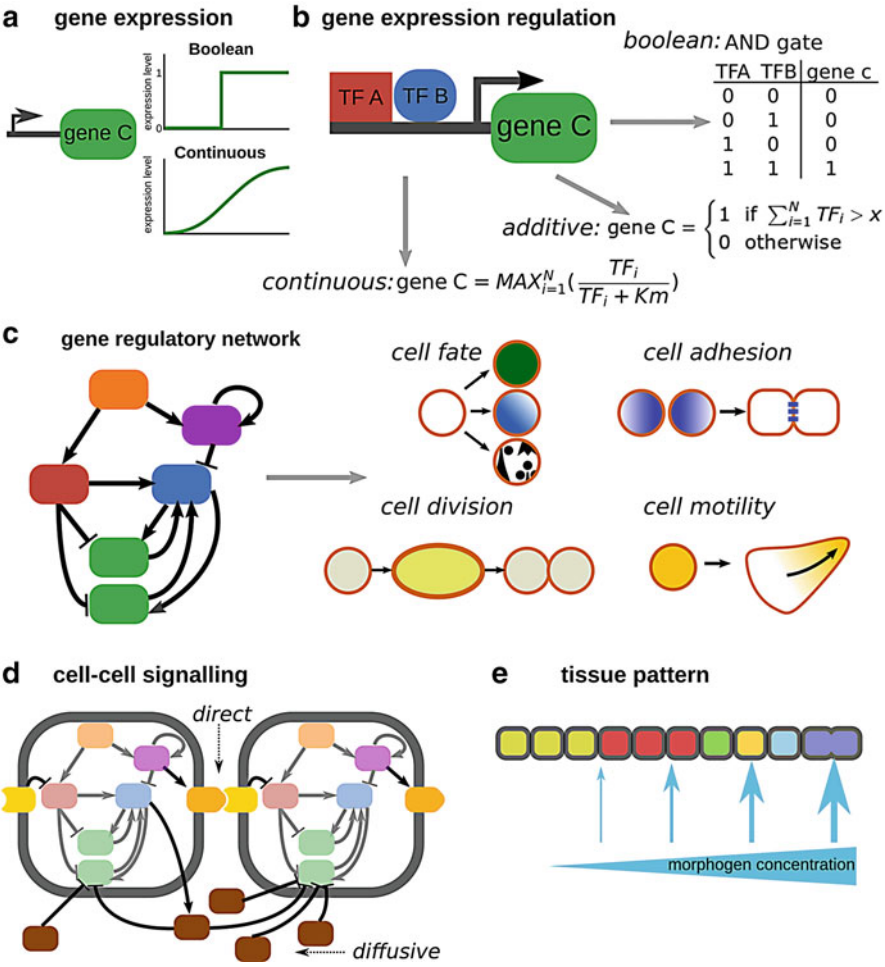

Fig. 1 Overview of the building blocks of evo-devo models

Boolean approach to include multiple discrete expression levels, for examples 0, 1, 2, and 3.

In models applying continuous gene expression, gene expression levels can take on any arbitrary expression value between zero and a superimposed or dynamically evolving maximum. While computationally less efficient, this approach is necessary if more gradual changes in expression are expected to be important for the developmental process under study, for example, a long-time-lag between the switching on and reaching full expression of a gene or the gradual amplification of initially small differences between cells to break symmetry.

Regulation of Gene Expression In multicellular eukaryotes, gene expression is regulated by a complex array of processes. Nuclear localization of the gene and its epigenetic state influence how easily the DNA can be accessed. Next, transcription factors control gene transcription in a complex, combinatorial manner via the promoter near the gene and via potentially multiple, modular enhancers that may even be present on a different chromosome. Finally, post-transcriptional processes like alternative splicing, protein modifications, and regulation of protein degradation are also involved. Current evo-devo models typically consider only the regulation of gene expression via transcription factors binding to the promoter, making use of one of three approaches: Boolean, additive, and continuous nonlinear integration of transcription factor input (Fig. 1B) (although one study did consider alternative control regions (Fujimoto et al. 2008)).

The Boolean approach uses so-called logical functions, or gates to integrate inputs, is typically combined with Boolean modelling of gene expression levels and usually assumes a constant number of transcription factors influencing each gene. For example, an AND gate requires that transcription factor A (TFA) and transcription factor B (TFB) are both expressed for the downstream gene to become expressed, while for an OR gate, the downstream gene becomes expressed if either TFA or TFB or both are expressed (Fig. 1B). To integrate larger numbers of inputs, more complex logical functions and combining of multiple logical functions are necessary. The disadvantage of this approach is that often, only a few regulatory inputs are relevant for the gene output, with others inconsequential due to the switch-like nature of Boolean gene expression. As a consequence, these models frequently overestimate the actual robustness of a developmental mechanisms (reviewed in Spirov and Holloway (2013)).

An alternative approach that is often used in combination with Boolean gene expression levels is additive integration. In this approach, TFs have an assigned weight – a positive value for activating and a negative value for repressive TFs – and gene expression is computed as a weighted sum of the expression levels of the TFs (ten Tusscher and Hogeweg 2009; Wagner 1996). This approach more easily allows for a variable number of TFs influencing each gene. The trade-off is that the additive approach is linear and therefore lacks some of the complex, nonlinear character typical of real gene expression regulation. When additive approaches are combined with continuous gene expression levels, the weighted sum of transcription factor inputs is often mapped to transcription rate via an additional, nonlinear function, thereby overcoming this linearity drawback (Salazar-Ciudad et al. 2001; Vroomans et al. 2016).

A final approach is continuous, nonlinear integration of gene inputs. Although existing in several varieties, they have in common that the input of an individual transcription factor on a downstream gene is modeled via a saturating Hill function (Fig. 1B). This mimics the saturation effect that occurs in vivo, where depending on TF concentration and binding site affinity, beyond a certain threshold all available regulatory sites will be occupied, so that an increase in TF concentration cannot further increase transcription of the downstream gene.

In Silico Gene Expression Dynamics During development, once a cell is formed it starts with an initial gene expression state that subsequently changes. At the start of development, this state is often maternally determined, while cells arising in the course of development typically inherit their state from the parent cell. To simulate this in evo-devo models, new *in-silico* cells are endowed with a particular initial gene expression state, where the extent to which different gene expression levels are discerned depends on the gene level model formalism chosen (Boolean or continuous). For cells already present from the start of *in-silico* development, a predefined, imposed gene expression state is used. Upon division, cells inherit their parental state. The subsequent unfolding of gene expression is dictated by the combination of activating and inhibiting signals from transcription factor genes upstream of each gene. Again, the chosen formalism of the model impacts exactly how transcription factor input is translated into gene expression levels at the next time instance.

Due to the small number of molecules of transcription factors and DNA polymerases, gene expression is an inherently noisy process. Thus, if one, for example, wishes to investigate whether noisy gene expression impacts the type of evolutionary outcomes by imposing selection for *developmental* robustness one needs to incorporate noise. In models with Boolean gene expression levels, noisy expression can be incorporated by using probabilistic update rules. For example, for an AND gate, if TF1 and TF2 are both expressed, the gene will become expressed with a probability of 90%, yet with a probability of 10% it remains not expressed. In case of continuous gene expression, a noise term can be added to the differential equations governing expression dynamics that modulates the average gene expression level (ten Tusscher and Hogeweg 2011).

Cell Behavior The differentiation of cells into distinct cell types is marked by the convergence of different cells on different subsets of stably expressed cell type defining genes. Thus, in evo-devo models, gene expression is required to converge to a stable pattern for successful differentiation. Apart from influencing the particular cell type, gene expression also influences cell behavior: adhesion to neighboring cells and extracellular matrix, growth, division, shape, and motility (Fig. 1C). These processes are crucial for understanding the interplay between tissue growth, morphogenesis, and patterning. Thus far, only a limited number of evo-devo models have incorporated genes affecting cell behavior beyond cell fate determination (Hogeweg 2000; Vroomans et al. 2016). However, recently new model formalisms have been developed to this end (Marin-Riera et al. 2016). Cell division, for example, may be implemented by incorporating a designated division gene whose levels need to exceed a threshold for division to occur (Vroomans et al. 2016). In the case of cell adhesion, the expression of a number of "adhesion" genes may generate a complex, cell-type-dependent adhesion profile (Hogeweg 2000).

Between Cells

Direct Cell-Cell Signalling In many developmental processes, extensive signaling takes place between directly neighboring cells to coordinate their gene expression dynamics. This can be used to minimize differences, as with Delta-Notch-mediated synchronization of the segmentation-clock in vertebrate somitogenesis. Cells can also use signaling to coordinate their polar orientation, as, for example, during *Drosophila* trichome patterning. Conversely, signaling may be used to amplify small initial differences, thereby enabling symmetry breaking. This process is often referred to as lateral inhibition, and, for example, patterns the hair cells of the chick inner ear.

Direct cell-cell signaling, emulating Delta-Notch-type signaling, has been incorporated into a few evo-devo models. To do so, a subset of the modeled genes is designated as signaling – rather than transcription factor genes. To simplify matters, expression of a signaling gene is assumed to directly regulate expression of downstream genes in the neighboring cells but not the cell in which it is expressed (Fig. 1D). Thus, one basically represents an entire signal transduction pathway as a single unit that evolution can use. Modeling separate ligands, receptors, kinases, nuclear receptors, etc., would make it highly unlikely for the *in-silico* evolutionary process to discover a functional cell-cell signaling system. Furthermore, all major signal transduction pathways were present in the evolutionarily most ancient, simple multicellular organisms, and multicellular complexity has mostly increased through the frequent reusage of these modules rather than inventing new pathways from scratch (Chipman 2010). Thus, implementing signaling genes in this simplified manner is deemed a reasonable approach.

A special type of direct cell-cell signaling is cell adhesion, which has been implemented in several developmental models but only a single evo-devo study (Hogeweg 2000). Differential cell adhesion, with cells either preferring to adhere to similar or to different cell types, has been shown to be a major driver of morphogenetic processes such as cell mixing, cell sorting, tissue engulfment (Graner and Glazier 1992), and convergent extension (Vroomans et al. 2015).

Long-Range Cell-Cell Signaling In addition to the short-range cell-cell signaling mediated by membrane bound receptor ligand pairs, long-range signaling mediated by diffusion of signaling molecules plays an important role in development. Well-known examples are the antagonistic FGF and RA gradients involved in vertebrate somitogenesis and the Bicoid gradient in early *Drosophila* development. Long-range signaling can be easily incorporated in evo-devo models by allowing diffusion of some gene products between cells (Cotterell and Sharpe 2010; Fujimoto et al. 2008) (Fig. 1D). Alternatively, morphogen gradients can be

superimposed (Fig. 1E). If one only wishes to investigate evolutionary processes arising after the prior evolution of the morphogen gradient, this latter approach is more computationally efficient (François et al. 2007; ten Tusscher and Hogeweg 2011). If in contrast, the question is how signaling centers and morphogen gradients may evolve, one needs to incorporate that certain genes may evolve the potential to be excreted and diffuse.

Tissue Level

Tissue Structure Developmental processes occur inside the complex, three-dimensional bodies of organisms. However, many developmental processes are restricted to a limited body region (e.g., eye development), occur on a largely flat surface (e.g., patterning wing veins in insects), or along a particular dimension (e.g., patterning along the anterior-posterior axis). This often allows one to focus modeling efforts to particular regions of the body or restrict simulations to two or even one dimension, reducing computational requirements and model complexity. Indeed, in many *in-silico* evo-devo studies of axial patterning, only a 1D tissue is considered, where cells form a single row (Fig. 1E) (Cotterell and Sharpe 2010; François et al. 2007; Fujimoto et al. 2008; Salazar-Ciudad et al. 2001; ten Tusscher and Hogeweg 2011; Vroomans et al. 2016). While such an approach is sufficient to study the basics of how gene regulatory networks underlying axial patterning may evolve, it also has clear limitations. For example, to investigate how patterning mechanisms evolved that ensure coherent boundaries, models should incorporate at least a two-dimensional tissue; Similarly, to take into account how cell movement contributes to patterning, considering higher dimensional tissues is essential.

Tissue Dynamics Depending on the developmental process under study, tissue patterning into different cell types may occur prior to or after processes such as cell division and motion that change overall tissue architecture (coined morphostatic patterning) or co-occur with tissue shape changes (morphodynamic) (Salazar-ciudad and Jernvall 2004). Axial patterning coincides with tissue growth and extension, and depending on the animal under study may also coincide with convergent extension (Vroomans et al. 2015). Still in many evo-devo studies this growth process is ignored and a fixed-size, one-dimensional tissue architecture is used (Cotterell and Sharpe 2010; François et al. 2007; Fujimoto et al. 2008; Salazar-Ciudad et al. 2001; ten Tusscher and Hogeweg 2011). In many cases, this is a reasonable approximation when no major reorganization of tissue occurs. However, cell division and tissue growth need to be explicitly incorporated in a model if one wishes, e.g., to investigate how the process of posterior elongation itself evolved. Depending on the exact research question, this incorporated growth process can be either imposed or regulated by the GRN (Hogeweg 2000; Vroomans et al. 2016) with levels of a designated gene deciding whether a cell is ready for division.

Tissue Pattern Development Given that all cells in a multicellular organism share the same genome, and hence the same regulatory networks governing gene expression dynamics, an initial symmetry breaking event is essential to enable different cells to obtain different fates. A famous example is the maternally deposited Bicoid mRNA in *Drosophila* that gives rise to a protein morphogen gradient via diffusion and enables different cells to start expressing different sets of genes. Somewhat similar to this, in sequentially segmenting animals such as vertebrates, but also in the beetle *Tribolium*, segmentation is controlled by gradients arising from the localized production of a stable mRNA or protein combined with localized growth. Alternatively, as is the case in, for example, *C. elegans,* development, fertilization may trigger a polarization process leading to the asymmetric division of the zygote into two cells with distinct fates.

In many evo-devo studies, the research question concerns developmental patterning downstream of the initial symmetry-breaking event. In this situation, simply superimposing the symmetry-breaking signal, such as a morphogen wavefront (ten Tusscher and Hogeweg 2011), differential gene expression (Salazar-Ciudad et al. 2001), or gradient (François et al. 2007; Fujimoto et al. 2008; Vroomans et al. 2016) is a valid approach. However, if the research question is concerned with the evolution of this symmetry-breaking event, either noisy gene expression or initial but non persistent differences between cells should be implemented, to investigate how these can be exploited by the *in-silico* evolutionary process as a trigger for symmetry breaking (Vroomans et al. 2016).

Evo-Devo Models

The field of evo-devo aims to answer how and why particular developmental mechanisms evolved. To illustrate how models have been used for this purpose, we will focus on a well studied developmental process: the subdivision of the animal anterior-posterior (A-P) axis into regular, repeating segments. The property of a segmented major body axis is shared among the distantly related vertebrate, arthropod and annelid clades. Furthermore, a number of animals in other clades seem to have a repeated A-P pattern (metamers) in some embryonic tissues. Most segmented animals generate their repeating units in a regular, sequential, anterior-to-posterior fashion from a posterior growth zone. Within the arthropods however, certain unrelated species develop their body segments simultaneously, the most famous example being the fruitfly *Drosophila*.

Together, these observations lead to many evolutionary questions. For example, it is still debated whether sequential segmentation evolved at least three times in parallel, or evolved once in the ancestor of bilateral animals and was subsequently fully or partially lost in many clades. For the first case, an obvious followup question is why this particular developmental mode would have evolved multiple times. Another major open question is why *Drosophila* uses such a complex, hierarchic regulatory cascade, where each segment is patterned by a unique combination of

genes. These questions make animal axial segmentation an excellent evo-devo study case, and computational modeling has been widely applied to it.

When answering *why* questions in evo-devo, we need to distinguish between why a particular developmental pattern – such as a segmented body axis – arose, and why a particular mechanism generating that pattern arose (see chapter ▶ "Proximate Versus Ultimate Causation and Evo-Devo"). The first question is hard to answer because it ultimately requires us to answer what purpose the developmental pattern may have originally served. In case of segments, perhaps there was selection for a larger body size, and segments were a simple, modular way to achieve that goal. Alternatively, there may have been selection for improved locomotive control of a large body, with segmental modules allowing independent control of different body regions.

Evo-devo models are particularly well suited for answering the second type of why questions. Central to answering these questions is an understanding of the nature of the genotype-phenotype map, and how it is molded by evolutionary processes. In biological organisms, the mapping of the genome into a phenotype via regulatory network architecture, gene expression dynamics, cell behaviour and developmental process is highly complex and non-linear. Since developmental models explicitly incorporate this genotype-phenotype mapping, they enable us to investigate which mutations are being buffered by the overall network dynamics and hence have no phenotypic effect, and which mutations cause a full collapse of the phenotype because they affect a regulatory hub impacting a large part of the network.

These models also allow us to determine -within the given boundary conditions- how many different types of developmental mechanism exist to generate a particular pattern and how often these different mechanisms occur. This may indicate that certain developmental mechanisms are more likely to occur than others. We can compare these different developmental mechanisms in terms of robustness to determine fitness advantages of one mechanism over the other. Alternatively, we can investigate their evolutionary nearness in terms of number of mutations and fitness of intermediate genotypes to assess the likelihood of evolutionary drift between equivalent mechanisms.

Finally, in models explicitly simulating the evolution of developmental processes we can investigate how mechanisms differ in evolvability, the ease in which evolution discovers and subsequently extends them. In these models we can trace how evolution shapes the genotype-phenotype mapping, tuning robustness and evolvability, and how this impacts the potential for incremental evolution of complex patterning.

Different Approaches

There are three main approaches to studying evo-devo questions with computational models. First, one can simulate the developmental process of interest, focusing on the robustness of the mechanism to noise in gene expression or mutations. With

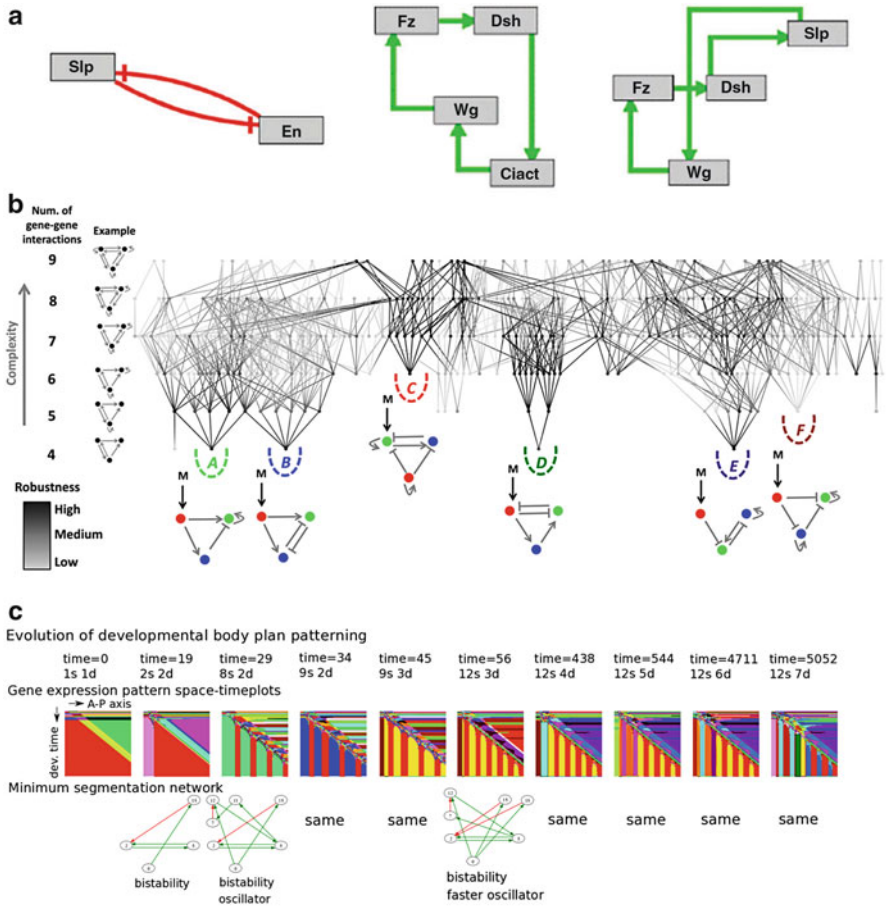

Fig. 2 *Three different approaches*. (**a**) The functional intracellular gene regulatory motifs identified for *Drosophila* segmentation gene network (Image from Sánchez et al. 2008). (**b**) The meta-network for segment-producing mechanisms. The letters indicate groups of networks that differ in developmental mechanism (Image from Cotterell and Sharpe 2010). (**c**) The *in-silico* fossil record of the evolution of a segmentation mechanism. Top row: the space-time plots of the development of several individuals in the evolutionary simulation (horizontal: space, vertical: developmental time). The colours indicate the different cell types. Bottom row: The corresponding minimal evolved gene regulatory networks that generate the cellular dynamics (Image adapted from ten Tusscher and Hogeweg 2011)

regards to body axis segmentation, such studies have been performed for both the pair rule and segment polarity networks (Sánchez and Thieffry 2003; Sánchez et al. 2008) (Fig. 2A). For both networks, the presence of mutual repression between genes was identified to play a major role in generating robust network dynamics. These results could be taken to suggest that these particular patterning mechanisms were selected for their high robustness. However, in absence of a comparison with

alternative patterning mechanisms resulting in similar downstream phenotypes, no strong claims of larger robustness than expected can be made.

A second approach is the so-called ensemble approach. In this approach, one investigates either all possible topologies of small size networks, or a large collection of randomly generated networks of a particular size (Cotterell and Sharpe 2010; Jiménez et al. 2015; Solé et al. 2002). Typically, the networks are sorted based on both the phenotype and the underlying developmental mechanism they encode. This method is efficient at finding many, if not all possible mechanisms for generating a certain phenotype, making it easier to compare them. For the small networks for which all possible topologies can be investigated, a meta-network can be created that connects similar gene regulatory networks (separated by a single difference) generating the same phenotype (Fig. 2B).

This meta-network has been used to study the mutational robustness of segmentation mechanisms, as this is determined by the number of interconnected networks generating the same mechanism (Cotterell and Sharpe 2010). Based on this approach, an alternative Turing-type mechanism for vertebrate segmentation was proposed which was found to be more robust than the classical clock-and-wavefront mechanism generally assumed to govern somitogenesis (Cotterell et al. 2015). The ensemble approach has also been used to study the evolvability from one segment-generating mechanism to the next, either for different mechanisms producing the same (Cotterell et al. 2015) or different phenotypes (Jiménez et al. 2015). A drawback of the ensemble approach is that it is thus far only feasible for small networks, that can perhaps best be interpreted as motifs of realistic, more complex developmental networks.

A final approach is to explicitly simulate the evolution of a developmental process (Fig. 2C). Darwinian evolution arises from the combination of reproduction with inheritance of parental properties, mutation to produce variety in offspring relative to parents, and selection which biases reproduction and survival to better adapted individuals. Simulating these processes requires the simulation of a population of individuals over many generations, imposing significant constraints on the complexity of the developmental process within a single individual that can be modeled.

In Silico Evolution

To build models that simulate Darwinian evolution, critical choices are the nature of the genome, the mutations operating on it, and the applied fitness criterion. In most evo-devo studies, the gene regulatory network is also considered the genome, and mutations operate directly on this network. Some studies, however, explicitly model a genome with genes and transcription factor binding sites, which encodes a gene regulatory network. Mutations then occur on the genome rather than on the regulatory network. Although this seems a minor difference, it may have important consequences for the evolutionary dynamics by impacting the mapping from genotype to phenotype.

In terms of mutations, most evo-devo models consider mutations that change which TFs influence a target gene, whether this influence is activating or repressive, the strength of this influence, deletion of a regulatory interaction, and insertion of a new regulatory interaction. In addition, some models incorporate mutations changing the maximum expression and degradation rate of a gene (Vroomans et al. 2016) and the diffusion constant of a gene product (Fujimoto et al. 2008). Finally, the models with an explicit genome incorporate duplication and deletion of genes, thereby allowing for variations in genome size. This substantially increases the degrees of freedom for the evolutionary process, and may hence impact the findability and evolvability of more complex developmental mechanisms. By implementing gene duplications such that the regulatory regions are duplicated together with the genes, evolution can tinker with one regulatory module, while another functional copy can be maintained.

These higher-level mutations have been suggested to increase evolvability (Spirov and Holloway 2013). However, large and complex genomes and networks may arise as a side effect of these extra degrees of freedom, with a high level of redundancy and many genes and interactions that have little effect. Unraveling how these genomes and networks translate into the observed developmental dynamics and final tissue pattern in these cases often requires pruning of the genomes and networks to identify the core mechanism.

In evo-devo simulations, a fitness criterion is typically used to ensure that the developmental pattern of interest evolves. The fitness score of an individual determines the reproduction rate of that individual, while leaving its death rate constant. Thus, in case of evo-devo studies which focus on body axis segmentation, fitness criteria evaluate the segmental pattern generated at the end of the development of an *in-silico* individual. However, different criteria may be applied, which vary in specificity. For example, one may simply let fitness increase with the number of generated segments, select for a particular number of segments, or even select for a particular spatial pattern of segments. The stricter the target, the more difficult it will be to evolve the desired phenotype because fewer evolutionary routes with intermediate fitness steps will be available (ten Tusscher 2013). Still, such a strict target may be important if one wishes to investigate how the regulatory mechanism changed due to evolutionary systems drift, while the developmental outcome remained constant (see chapter ► "Developmental System Drift").

A number of studies applied evolutionary simulations to evo-devo questions on segmentation. Collectively, these studies show that only a few distinct classes of mechanisms evolve for generating segments, and that which class emerges strongly depends on the applied morphogen dynamics and fitness criterion (reviewed in ten Tusscher (2013)). When the fitness criterion is very strict and/or the morphogen consists of a non-moving peak or gradient, segments are typically generated all at the same time. The mechanism used entails either a hierarchical cascade of gene expression involving many regulatory genes that mostly interact unidirectionally, or a self-organised emergent mechanism involving a limited number of mutually

interacting genes (François et al. 2007; Fujimoto et al. 2008; Kohsokabe and Kaneko 2016; Salazar-Ciudad et al. 2001).

The models show that while the emergent mechanisms more easily generate a larger number of segments the hierarchic mechanism is more robust to mutations as different segments depend on different genes. As a consequence, hierarchic mechanisms tend to replace emergent mechanisms over longer evolutionary time. The *in-silico* hierarchical mechanisms do to some extent resemble the *Drosophila* segmentation cascade. As such, the fact that they evolve under strict fitness criteria supports the idea that *Drosophila's* mechanism is secondarily evolved – hence, segment positions had to be strictly maintained relative to those generated by the ancestral mechanism (ten Tusscher 2013).

When, instead, simulations applied more general fitness criteria (supporting de-novo evolution of stripe patterning), and the morphogen was simulated to retract from anterior to posterior (emulating posterior growth), the most common evolutionary outcome is a sequential segmentation mechanism. This mechanism involved a continuous A-P transition from gene expression oscillations to a fixed segment pattern (François et al. 2007; ten Tusscher and Hogeweg 2011; Vroomans et al. 2016). By using the *in silico* fossil record generated in these simulations, it could be shown that this complex developmental mechanism evolves through the incremental evolution of network motifs, first generating bistability, then an oscillator and subsequently a sped up oscillator increasing the number of segments generated (François et al. 2007; ten Tusscher and Hogeweg 2011).

Thus, these studies provide powerful counterargument against the argument of irreducible complexity that is often made for complex novel phenotypes. Furthermore, by suggesting that evolution of sequential segmentation is relatively straightforward, they support the possibility of parallel evolution of this segmentation mode in the vertebrate, annelid and arthropod clades. These simulations also demonstrated that, compared to alternative mechanisms that occasionally evolved *in silico*, the sequential mode evolved more rapidly, was more evolvable and was more robust to noise in gene expression and division timing, and to mutations (François et al. 2007; Fujimoto et al. 2008; ten Tusscher and Hogeweg 2011; Vroomans et al. 2016).

Together, these studies thus suggest that, when growth occurs through posterior elongation, sequential segmentation is the expected evolutionary outcome. However, they leave open the question whether sequential segmentation is still the most likely evolutionary outcome if posterior elongation has not yet evolved (and therefore has to co-evolve). To address this, Vroomans and ten Tusscher (Vroomans et al. 2016) performed evo-devo simulations in which they selected both for axial growth and segmentation. In these simulations two mechanisms evolved, one in which growth and patterning occurred at the same time and across the entire tissue, and one in which both occurred sequentially from a posterior growth zone.

The simultaneous mechanism evolved tissue growth and segmentation concurrently, with new segments evolving as the tissue evolved to become larger. The sequential mechanism instead first evolved a large tissue and then evolved more and more segments. The simultaneous mechanism was the dominant outcome in simulations where only a transient morphogen signal was provided, while the sequential mechanism dominated when the morphogen was assumed to be maintained at a high

level in the posterior-most cell. Given the predominance of posterior elongation and sequential segmentation in extant organisms, these results suggest that these growth and segmentation modes arose after the earlier evolution of a posterior signalling center.

Discussion and Concluding Remarks

Most evo-devo research addresses *what* and *how* questions – what specific developmental mechanisms have evolved and how these have come about through mutations and selection. In addition to these types of questions, evo-devo modeling studies aim to also address *why* questions. We have shown here how computational models of evolution of development are constructed, and how they may provide deeper insights in why extant organisms use particular developmental mechanisms and not a theoretically possible alternative mechanism.

Answers that evo-devo models may provide could be that certain mechanisms are more easy to find for an evolutionary process given the nature of biological mutations or the prior evolutionary history and hence are statistically more likely to appear as an evolutionary outcome. Model outcomes may also demonstrate that certain mechanisms are more robust against mutations or developmental noise and therefore confer secondary fitness advantages, enabling them to evolutionary outcompete alternative patterning mechanisms. Finally, simulations may show that the need for coordination with simultaneously occurring other patterning mechanism may affect evolutionary outcome.

Constructing evo-devo models ultimately entails defining a genotype-phenotype mapping and how this mapping can be changed through evolution. Like all models, evo-devo models are by necessity simplifications. The choices made in terms of simulating gene expression, cell behaviour and tissue dynamics may affect the number and type of mechanisms that can generate a certain pattern in silico. As an example, in absence of diffusing gene products, no Turing-type patterning can arise. Similarly, choices for initial conditions, genome structure and mutational operators may affect evolvability and the potential for evolution to shape the genotype-phenotype mapping. Additionally, the choice for Boolean versus continuous gene expression modeling may substantially affect cellular differentiation dynamics and robustness of patterns to perturbations, again affecting what types of evolutionary outcomes are most likely to arise and persist. Thus, ideally, conclusions obtained in evo-devo modeling studies should be tested for their dependence on the modeling choices made.

As an example, if a study suggests the predominant evolution of a particular patterning mechanism, it is important to determine whether this depends on modeling assumptions or truly is a general outcome. If it is a general outcome, we can safely conclude that this particular mechanism is the expected evolutionary outcome. If instead the mechanism only dominates if particular assumptions are made, for example, the presence of a certain signaling center, this may reveal the critical dependence of the evolution of a trait on prior evolutionary events or certain aspects of the developmental genetic toolkit.

In this light, it is important to consider that while current evo-devo models are already quite complex, they do not yet incorporate major properties of metazoan genetic regulation. Incorporating cooperative activation and repression by nearby bound transcription factors, regulation by multiple modular enhancers and epigenetic regulation is likely to further increase the complexity of the genotype-phenotype mapping and the potential for evolution to fine-tune this mapping. It will be interesting to see how this may effect earlier modeling conclusions, and how it will increase our ability to explain how evolution converged to the developmental patterning mechanisms observed in extant organisms. Additionally, current evo-devo models focus on regulatory mutations affecting developmental patterning. Except for gene duplications, they ignore coding region mutations that could expand and modify the genetic toolkit available for development. Incorporating these types of mutations in evo-devo models is necessary to contribute to the debate on the relative importance of mutations in coding versus regulatory regions in evo-devo.

Cross-References

- ▶ A Macroevolutionary Perspective on Developmental Constraints in Animals
- ▶ Convergence
- ▶ Computational Modeling at the Cell and Tissue Level in Evo-Devo
- ▶ Evolvability
- ▶ Levels of Organization in Evo-Devo
- ▶ Modeling and Simulation in Evo-Devo
- ▶ Proximate Versus Ultimate Causation and Evo-Devo
- ▶ The Evolution and Development of Segmented Body Plans

References

Chipman AD (2010) Parallel evolution of segmentation by co-option of ancestral gene regulatory networks. BioEssays 32(1):60–70

Cotterell J, Sharpe J (2010) An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients. Mol Syst Biol 6(1)

Cotterell J, Robert-Moreno A, Sharpe J (2015) A local, self-organizing reaction-diffusion model can explain somite patterning in embryos. Cell Syst 1(4):257–269

François P, Hakim V, Siggia ED (2007) Deriving structure from evolution: metazoan segmentation. Mol Syst Biol 3(1)

Fujimoto K, Ishihara S, Kaneko K (2008) Network evolution of body plans. PLoS One 3(7):e2772

Graner F m c, Glazier JA (1992) Simulation of biological cell sorting using a two-dimensional extended potts model. Phys Rev Lett 69:2013–2016

Hogeweg P (2000) Evolving mechanisms of morphogenesis: on the interplay between differential adhesion and cell differentiation. J Theor Biol 203(4):317–333

Jiménez A, Munteanu A, Sharpe J (2015) Dynamics of gene circuits shapes evolvability. PNAS 112(7):2103–2108

Kohsokabe T, Kaneko K (2016) Evolution-development congruence in pattern formation dynamics: bifurcations in gene expression and regulation of networks structures. J Exp Zool B Mol Dev Evol 326(1):61–84

Marin-Riera M, Brun-Usan M, Zimm R, Välikangas T, Salazar-Ciudad I (2016) Computational modeling of development by epithelia, mesenchyme and their interactions: a unified model. Bioinformatics 32(2):219–225

Salazar-ciudad I, Jernvall J (2004) How different types of pattern formation mechanisms affect the evolution of form and development. Evol Dev 6(1):6–16

Salazar-Ciudad I, Newman SA, Solé RV (2001) Phenotypic and dynamical transitions in model genetic networks I. Emergence of patterns and genotype-phenotype relationships. Evol Dev 3(2):84–94

Sánchez L, Thieffry D (2003) Segmenting the fly embryo: a logical analysis of the pair-rule cross-regulatory module. J Theor Biol 224(4):517–537

Sánchez L, Chaouiya C, Thieffry D (2008) Segmenting the fly embryo: logical analysis of the role of the segment polarity cross-regulatory module. Int J Dev Biol 52:1059–1075

Solé RV, Salazar-Ciudad I, Garcia-Fernández J (2002) Common pattern formation, modularity and phase transitions in a gene network model of morphogenesis. Physica A 305(34):640–654

Spirov A, Holloway D (2013) Using evolutionary computations to understand the design and evolution of gene and cell regulatory networks. Methods 62(1):39–55

ten Tusscher K (2013) Mechanisms and constraints shaping the evolution of body plan segmentation. Eur Phys J E 36(5):1–12

ten Tusscher KH, Hogeweg P (2009) The role of genome and gene regulatory network canalization in the evolution of multi-trait polymorphisms and sympatric speciation. BMC Evol Biol 9(1):159

ten Tusscher KH, Hogeweg P (2011) Evolution of networks for body plan patterning; interplay of modularity, robustness and evolvability. PLoS Comput Biol 7(10):e1002208

Vroomans RMA, Hogeweg P, ten Tusscher KHWJ (2015) Segment-specific adhesion as a driver of convergent extension. PLoS Comput Biol 11(2):1–24

Vroomans RMA, Hogeweg P, ten Tusscher KHWJ (2016) In silico evo-devo: reconstructing stages in the evolution of animal segmentation. EvoDevo 7(1):14

Wagner A (1996) Does evolutionary plasticity evolve? Evolution 50(3):1008–1023