Modeling Evo-Devo:
Broken Hierarchies and Multiple Scales of Organization and Complexity

Diego Rasskin Gutman, Borja Esteve Altava
Theoretical Biology Research Group
Institute Cavanilles for Biodiversity and Evolutionary Biology
University of Valencia
Spain

Abstract

Evo-Devo, the science that puts together in a common framework the dynamics of evolution with the processes of embryonic development is inherently multiscale. The hierarchical organization of life phenomena also contributes to the possibility of reducing its complexity to workable modules. However, the emphasis on a compositional, or blocks-within-blocks kind of hierarchy, implies a reductionistic perspective on multiscaling that ignores the irreducibility of some levels. The notion of generative hierarchies tackles this problem, introducing an organicist perspective that, while keeping levels of organization, acknowledges the existence of breaks in the hierarchy at the genomic, cellular, individual, and species levels. Whereas independent modeling in development or evolution has been done at each scale of organization, no multiscale approaches have so far been worked out that can account for the relationship between these two fundamental mechanisms that have shaped biodiversity throughout the history of life on Earth.

Organization and Complexity in Evo-devo

Departing from evolutionary theory and developmental theory, the field of evolutionary developmental biology (evo-devo) has flourished in recent years, fuelled by the discovery of the so called “developmental genetic toolkit,” a suit of genes used during development shared by most animals [1]. Before this renaissance, evo-devo had also a rich research tradition starting back in the 19th century with the French anatomist Etienne Geoffroy St Hilaire, and the German embryologists Karl von Baer and Ernst Haeckel (both famous for their law-like semiempirical developmental observations that ended up in Haeckel’s linkage of ontogeny with phylogeny), and later in the 20th century, outstanding figures such as Garstang, De Beer, Waddington, Gould, and Alberch [2,3 and references therein]. In essence, this “old” evo-devo, was focused on morphological issues in a comparative framework. The emphasis of this morphological evo-devo was on how anatomical parts differed in related species as a result of specific growth rates. This kind of research was subsumed under the all encompassing theme of heterochrony, which has still a very active role in the field. However, the morphological evo-devo tradition has reinvented itself, going beyond heterochrony as a main focus, to embrace other issues such as modularity, innovation and emergence of morphological traits, and phenotypic plasticity [4]. In the background of this tradition is the question of biological organization and biological complexity. Whereas organization has been “solved” by resorting to hierarchy and modularity, biological complexity is one of those concepts for which there are no universal metrics; hence, it has rarely been used as a proxy for evolution and never to infer systematic relationships [5,6,7]. An important contribution to the debate on complexity was Herbert Simon’s article “On
the architecture of complexity” [8] in which he argued that modularity design in living forms is the only way to achieve high degrees of complexity in an evolutionary context ([9] and references therein).

Levels and Hierarchy

Biological levels of organization span well over 10 orders of magnitude both in time and space. Wherever you are in your scale range, different physicochemical laws apply. For example, in fluid mechanics, the extremely small, subcellular environments have low Reynolds numbers. Thus, when modelling dynamics at the subcellular level, such as the dynamics of cilia movement in mouse embryonic cells [10], inertia can be rightly ignored. In addition to the large span of time and space scales, biological phenomena is also special insofar as it presents a typical hierarchical organization of “blocks within blocks” in which smaller components make up larger ones from atoms to individuals, populations, societies, species and up the taxonomic ranks. This organization has been apparent to many authors, a part of an old tradition in science: “The history of science has strongly encouraged the view that the existence of and generalizations about larger aggregations of matter, such as biological systems [...] can be explained by generalizations about the smaller aggregations of matter that compose them” [11].

A distinction can be made between compositional hierarchy and generative hierarchy, while component hierarchy is a reductionistic box-within-box portrayal of biological organization, generative hierarchy is a dynamic, modular view of biological processes that accommodates the processes by which organisms are formed (for other hierarchy distinctions, such as compositional versus procedural, see [12]). Both hierarchies need to be fully studied to reach a thorough understanding of biological phenomena (figure 1).

Under the reductionistic view, all what an organism is can be described in terms of its parts and interactions, including properties emerging at higher levels of organization. Thus, under this paradigm, biology can be subsumed to molecular biology and the latter to physics in a straightforward manner, making conceptual jumps that sometimes cannot be fully justified. In sharp contrast with this view, the integrative approach to nature (historically known as organicism) claims that dissociating organisms in their constituent parts prevents the understanding of emergence properties. From the point of view of biological organization, organicism clashes with reductionism in the interpretation of causality: while reductionists think in terms of bottom-up causality, organicists emphasize top-down and multilevel causality. A compositional view of biological organization asserts that each level is composed by smaller parts of the lower level combined under certain rules. These composition rules are level-specific and they usually are generalized as self-organization processes. For example, atoms compose molecules, molecules compose sub-cellular parts, sub-cellular parts compose cells, cells compose tissues and so on. As a result of the nested organization of compositional hierarchy, the ever inclusion of parts into larger part leads continuously from atoms to organisms. A direct consequence of compositional hierarchy of biological entities such as cells or multicellular organisms is their reduction to physical entities down to the level of molecules or even atoms. The analogy holds even in organisms and family groups, social groups, populations, or species.
Generative hierarchy is a description of processes. Rather than partitioning biological entities as a series of nesting boxes, generative hierarchy makes a statement about the dynamical biological mechanism that glues together the components at each level. For example, in compositional hierarchy, molecules are composed by atoms. In generative hierarchy, atoms generate molecules by means of electrical interactions. In compositional hierarchy, tissues are made of cells; in generative hierarchy, cell adhesion forms a tissue. This fundamental difference highlights an intriguing question: as we move up in the compositional hierarchy from atoms to species, we find breaking points where the generation of a higher-level component does not hold up anymore. Rather, in these levels in which the hierarchy is broken, the elements form not from lower-levels elements, but as a recursive, autonomous process, either at a developmental or at evolutionary scales. There are four levels of recursive generation: genome, cell, organism, and species. These four levels cannot be produced by assembling lower level parts, they need a pre-existing template provided by their own structure to produce a new element. Genomes are made out of preexisting genomes, cells are made out of preexisting cells, individuals arise from individuals, and species arise from species (see figure 2). So, generative hierarchy is exclusive of biological phenomena: it can only be portrayed once living phenomena occurred in the history of life. Once the first genome, the first cell, the first individual, or the first species arose, generative hierarchy could not be stopped.

The organization of new genomes, new cells, new whole organisms, and new species is not carried out by the dynamics of their lower constitutive levels nor by a combination of their components, but through a self-dividing dynamics, unavoidably attached to the evolutionary process. Other levels can be accessed in a compositional fashion, such as a sub-cellular part made by the self-organization of biomolecules. However, a cell cannot be built by assembling sub-cellular parts: a cell always originates from another cell. The relationship between lower and higher levels is fundamental in the understanding of natural organization and the generation of complexity. It also has a relevant impact in modeling multi-scale phenomena because reduction of larger scales to smaller ones only can be done before a hierarchical break occurs. Thus, organs and tissue behavior can be fully explained under a cellular perspective, but not by further lower levels. This claim does not neglect the importance of molecular mechanisms to understand the upward causation effect from lower levels to higher ones, which always occurs, but which cannot be used as a sufficient explanation.

In the next sections we will survey the types of models available in the literature at each of these levels for both development and evolution; it will be evident that the need for multiscale models in evo-devo is still there. One of the few exceptions is the landmark work by Salazar-Ciudad and Jernvall [13] on the development and evolution of mammal teeth, integrating genetic regulatory networks and morphology across generations.

Models in Developmental biology: a matter of scales

Two basic approaches have been used to model developmental mechanisms. One approach, uses the language, formalisms, and tools of physics, abstracting and describing the problem in terms of differential equations, i.e., by evaluating changes in the quantity of a variable such as concentration of a protein over time [14]. The second approach, Artificial Life (AL) draws more heavily from concepts and tools derived from computational sciences [15,16]. Efforts at
multiscale modeling are not abundant. Chatuverdi et al. [17] have implemented a computational multi-scale model for growing vertebrate limbs, integrating several models from each of the scales involved in morphogenesis: molecular, subcellular, cellular, tissues, and organs. Their implementation integrate different mathematical approaches to morphogenesis, with relative success in single-scale modelization, such as statistical mechanics, cellular automata and differential equations. These models are coupled by matching their temporal and spatial variables. Integration allows information to flow among any levels of the biological hierarchy involved in limb morphogenesis, enhancing the importance of multilevel causation approaches in multi-scale modelling.

Molecules

Most modeling in developmental biology has focused on the sub-cellular level. Much attention has been given to regulatory signaling networks and other molecular processes that govern cell differentiation processes, shaping our general notions of morphogenesis itself. Such is the case with positional information models provided by gradients [18,19,20]; the polar coordinates model [21]; or the limb growth progress zone model [22]. Pre-patterning models are a very influential model of this kind, such as Turing’s reaction-diffusion mechanisms that are able to generate very complex chemical pre-patterns given a set of boundary conditions [23,24]. The analysis of protein expression dynamics is a more classic approach using differential equations to capture the rate of change of the molecules involved in a signaling pathway. The number of models using this methodology is vast, both in kinds of signaling pathways and model organisms used, such as somite formation in vertebrates [25,26], or the segment polarity network in Drosophila [27]. Other models make use of Boolean networks [28], that is, on/off switches between involved genes in a signaling pathway, as in early specification of the sea urchin endomesoderm [29].

Cells

The cellular level is modeled by focusing on cell behavior types: migration, division, apoptosis, differentiation, and changes in cell adhesion. The latter is the single most important feature that drives these models. This kind of modeling has been especially successful in Dictyostelium discoideum, a mold which alternate between unicellular and multicellular states in response to nutrients during its life cycle. Work by Hogeweg and collaborators show in great detail the whole dynamical process of slime mold morphogenesis by using hybrid approaches of cellular automata and differential equations ([30, 31, 32]. The relation between gene networks and cell behavior has been explored in the evolution of artificial organisms [33,34,35].

In more complex organisms, a reaction-diffusion mechanism has been used to account for a set of patterning events leading to the array of prechondrogenic condensations and subsequent skeletal pattern in the vertebrate limb [36]. Main features of the reaction-diffusion model, such as positive autoregulations or induced pattern transitions have been confirmed by experimental studies [37,38,39]. In addition, a mechanochemical model was suggested [40], in which cell density, cell movement, forces between cells and the extra cellular matrix, and calcium concentration as a regulatory molecule determine three domains of morphogenetic stability: condensation, segmentation, and bifurcation.
Tissues

Another level of morphogenesis modeling is achieved by fully taking into account the mechanical properties of cells in a tissue. The focus is on the different types of cell adhesion and deformation given the dynamics of cell growth and embryo development [41]. The process of neurulation from ectodermal embryonic cells has been modeled by using finite element analysis, shedding some theoretical insights on the type of forces occurring during neurulation dynamics [42,43]. A similar approach was taken by Oster and collaborators to model the mechanisms involved in the primary invagination in sea urchin embryogenesis to form the archenteron [44], exploring the interaction among the cells and the extracellular matrix, and providing a suitable comparative framework to evaluate different theories on the mechanisms of primary invagination. Plant growth has been modeled by the computational approach based on L-sytems [45]. A “formal grammar” [46] indicates a cell what to do as a function of its position, allowing, among other things, the complex branching processes typical of plant growth. A related model of fractal growth has also been applied to marine sessile organisms, including corals, sponges and seaweed [47,48].

Models in Evolutionary Biology

Mathematical modeling has also a long tradition in evolutionary biology going back to work by Volterra and Lotka on the prey-predator relationship [49,50], as well as quantitative genetics in the work of Fischer, Haldane and Wright, see [51]. Of particular interest is Wright “adaptive landscape” model, in which genetic traits were mapped on a fitness function; this work alone gave rise to disparate modeling traditions, from Kauffman’s work on genetic networks in rugged fitness landscapes, and with it, the whole theory of complex adaptive systems [52,53] to the so called Theoretical Morphology in which models (morphospaces) of forms were assessed sometimes against some measure of adaptation whereas in other occasions these models were used to analyze the structure of the morphospace itself, with the aim of checking for morphological constraints [53,54,55,56].

Genes

The neutral theory for evolutionary change at the molecular level [57] transformed the predominant paradigm proposed by the Modern Synthesis, which only allowed for a mutation space divided into advantageous and deleterious genetic changes. Kimura’s neutral theory introduced mathematical models to analyze neutral changes, that where neither positive nor negative in terms of fitness, with the expectation that most changes were actually neutral. Nowadays, random and neutral genetic change is the main null hypothesis in genetic evolution. In addition, Kimura’s neutral models allowed the introduction of genetic models to establish phylogenetic relationships between genes [58]. Since then, molecular clocks have been studied for many genes from vertebrates [59] to viruses [60].

Another important aspect in genetic evolution concerns the evolution of the architecture of genes and genomes [61]. Quantitative genetic models analyze the evolution of the genetic variance-covariance matrix (G-matrix) which tries to establish relationships between genetic
and phenotypic variation in a directional selective framework [62]. Despite its relevance as a multiscale attempt to study together genes and phenotypes, quantitative genetics fails to construct useful models for genetic evolution itself, because of its simplifying assumptions that ignore the importance of drift and non-additive genetic interactions, such as epistasis and dominance [63]. Phenotype landscape Models introduce vectorial analysis with differential equations to study changes in the relation between genotype and phenotype linking the latter with fitness [64], which in fact is the opposite to what quantitative genetic models of the G-matrix do. In addition, network models are used in genomes and metabolisms [65,66]. In genetic programming, artificial gene networks can be simulated through an evolutionary process of mutation and recombination followed by selection of optimal configuration [67,68,69]. Using genetic algorithms to evolve the program coding allows for models of the evolution of complex systems, such as locomotion strategies [70] or developmental control [71].

**Populations**

Theoretical evolutionary ecology models the interaction among organisms using a wide array of mathematical tools, see [72]. Linear equations and logistic models model the predator-prey and competition dynamics, both between and within populations or species [73]. Other models using coupled differential equations have been used to model life tables of growing cohorts [74] or reproduction and survival trade-offs to maximise fitness functions [75]. Game theory applies to interaction analysis when different strategies are available in terms of choice between cost and benefit as in male:female ratios [76]. Using game theory, Maynard Smith and Price [77] introduced a key concept in the evolutionary theory of interacting populations: Evolutionary Stable Strategies (ESS), which are behavioural, structural, or physiological strategies adopted by interacting organisms during their evolution. Network theory also has been used in analysis of community structure and evolution from a topological and dynamical perspective [78]. The failure of deterministic models in evolutionary ecology pointed out the importance of introducing stochastic and chaotic models [79] to test theoretical models and predictions with real data.

**Species**

When evolution is viewed at higher scales, beyond population dynamics, we enter the sphere of macroevolution, which is defined as evolutionary change at species or above species level. Macroevolutionary models concern basically three main processes: species origination, species extinction, and species transformation. The raw data to model this scale is morphology, because it is shared by both extant and extinct species. The main speciation models in biology are the allopatric, sympatric, and peripatric depending on the way the new species is originated from the parental one. However, speciation by other mechanisms, such as hybridization, is also possible [80], and it is still unknown the relative importance of each model of speciation [81].

Extinction models at species level can be found in the work of Raup and Sepkoski [82] on marine fossil extinctions, although there is a long tradition in long-scale evolution and extinction models specific for very particular fossil groups or geologic ages [83]. The observation that evolutionary change tends to be concentrated within speciation events lead Stanley [84] to propose a model of evolution at species level, independent of microevolutionary
processes. In the same direction, Eldredge [85] formalised species sorting (also called species selection) as differential speciation and extinction rates originating macroevolutionary patterns. Models of phyletic change has been statistically formalized recently by Hunt [86]. Hunt models capture a long tradition in evolutionary studies about the tempo and mode of evolutionary change, and his work points out key results about the nature of long-scale trends (see [87] and references therein).

In Geometric Morphometrics, the analysis of form and form transformation is carried out using sophisticated multivariate statistical and geometric tools for comparative purposes both at intraspecific and interspecific levels. Most studies are based on Procrustes Geometry, although alternatives such as coordinate-free Euclidean analysis are also used. Allometry, heterochrony, or fluctuating symmetry models can be efficiently tested in Geometric Morphometrics, making of this method a good candidate to carry out multiscale modelling above molecular levels ([88,89] and references therein). Work by West et al. [90] has introduced elegant fractal modelling to account for empirical allometric relationships in a variety of taxonomic groups.

Conclusions

True evo-devo modeling is absent from the literature (but see [13]). Its multiple scales (within individual life and across generations) make it a daunting task. Models are usually suitable for a specific time and space scale. In development, models have been applied to describe, explore, and predict the behavior and dynamics of molecules, cells, and tissues during embryonic growth. In evolution, modeling has two separate domains: simulations of molecular dynamics and analysis of macroevolutionary dynamics. Whereas modeling evolution at the molecular level is a relatively new and very well defined area of systems biology, mostly what is called bioinformatics, macroevolutionary modeling is heterogeneous, including patterns and trends of morphological transformations (mapping of morphological transformation in consensued cladograms), patterns of extinctions and speciation, the genotype-phenotype mapping based on quantitative genetics models. Multiple scales and multiple causes make the task of modeling development and evolution (evo-devo) not trivial.

Undoubtedly, any integrative modeling approach must go beyond generating statistical correlations among the four levels of biological organization as described above. A host of hybrid approaches will have to be used and new kinds of sophisticated schemas worked out to offer a complete integration in a biologist user-friendly fashion. A global modeling strategy will have to be developed combining what is already available starting with network analysis of regulatory pathways, systems of differential equations to account for the dynamics of molecular reactions, cellular automata to simulate the behavior of biological cells along with biological “grammars” and finite element-like strategies to describe the mechanical properties of tissues, which, as we have outlined, are already been used by the modeling community. These developmental models will have to be made across generations, to also account for evolutionary dynamics. Morphological data, using geometric and topological approaches will also enter the equation. Modularity and plasticity both in the design of the modeling strategy as well as to account for these fundamental properties of evo-devo dynamics will be at the center of evo-devo modeling. Although the combination of these approaches both analytically and numerically will pose formidable challenges, the rewards on having complete simulation platforms will be well worth the effort.
FIGURE 1. Left: Traditional schema of biological organization from atoms to species, where levels are composed by aggregation of elements of the lower levels. Right: Proposed schema of biological organization, where breaks are introduced between levels that are not developmentally related directly from the lower ones. Biological organization as a generative hierarchy is concerned with ontological processes. Several levels are generated by auto-organization processes, from lower level components, whereas others can only be formed from a pre-existing template (the case of genomes, cells, individuals, and species).
FIGURE 2. Hierarchical nested scheme of biological organization. Development and evolution are multilevel processes. Cell division (including genome division), multicellular organisms reproduction and speciation, are level-specific generative processes not reducible to lower levels. The rules that govern genomes, cells, organisms and species generation are neither in laws of lower levels nor in certain combination of their constitutive components, but in their own particular dynamics, which has arisen evolutionarily.

FIGURE 3. Evo-devo strives to elucidate the discrete organization of morphospace, in which all theoretical morphologies can be represented. Development and evolution with their respective scales and dynamics coalesce to offer an explanation for forms in nature. Models will have to accommodate the multiple scales of within-individuals and across-generation that are ubiquitous to the phenomenon of life.
Acknowledgements

We would like to thank Klaus Stiefel and Robert Sinclair for their kind invitation to participate at the Multiscale workshop in Okinawa. This work has been supported by grant BFU2008-00643 from the Ministerio de Ciencia e Innovación, Spain.

References


