2014

Formalizing Theories Of Development: A Fugue On The Orderliness Of Change

Scott F. Gilbert
Swarthmore College, sgilber1@swarthmore.edu

J. Bard

Recommended Citation
Scott F. Gilbert and J. Bard. (2014). "Formalizing Theories Of Development: A Fugue On The Orderliness Of Change". *Towards A Theory Of Development.* 129-143. DOI: 10.1093/acprof:oso/9780199671427.003.0008
https://works.swarthmore.edu/fac-biology/453

CHAPTER 8

Formalizing theories of development: a fugue on the orderliness of change

Scott F. Gilbert and Jonathan Bard

A pluralism of developmental perspectives

This essay must be tempered throughout with humility. In the past 50 years, developmental biology has recapitulated in rapid order the Industrial Revolution's succession of creator from person (organism), to apparatus (cell, molecule), to algorithm (program). In 50 more years, we will be very lucky if our essays are considered 'prescient,' because it is doubtful that they will be considered 'science' by those standards (see Flüsser, 1987).

We also should not be confined by the 'progressive' flow of such displacement from organism to program. Indeed, the notion that these different levels of agency succeed and displace one another is a Modernist notion that should be avoided. Creation, as Paul Weiss (1967, 1977) noted in his early systems theories of development, is found at all levels—the molecular, the cellular, the tissue, the organismal, and the ecological[1]—through the integration and recombination of lower-level entities into higher-level orders, and through the selection of viable possibilities by the upper-level agents. In studying this re-ordering, it is important to remember that while lower-level orders are the components of higher-level orders, the higher-level orders provide the context/niche for selecting possible lower-level structures (see Auletta, 2011; El-Hani & Emmeche, 2000; Ellis, 2012; Longo et al., 2012; Soto et al., 2009).

We should also respect a plurality of explanatory perspectives (Pirsig, 1974; Winther, 2011). In his analysis of part-whole explanations, Winther (2011) catalogues three major modes of developmental explanation: (i) structuralist (top-down) explanation, in which emergent organization is what needs to be explained and mathematical-logical formalisms carry the weight of explanation; (ii) mechanistic (bottom-up) explanation, in which parts and their causal interactions can explain developmental phenomena; and (iii) historical explanation, where development (both of parts and wholes) is placed in a larger, evolutionary, narrative. This paper sees such perspectives as being 'in resonance' with one another. The metaphor is that of electrons in a benzene ring. No perspective, alone, provides a complete account of developmental phenomena.

Relations and downward causation

This chapter takes a systems approach to development in that it tries to step back from the normal minutiae of developmental phenomena and asks how should one start to unpack the complexity of development in a way that captures both the parts and the whole; and a first step is to look at the relationship between them. Higher-order structure provides the 'interpretation' of the lower-level parts and processes. Using a linguistic analogy, the statement

[1] Interestingly, the molecular level, which is 'lower' than the biological levels, may have become a biological level after its appropriating a particular context within the cell. Newman (2012; this book) hypothesizes that the morphological properties of cell division, migration, and ordering were originally physical ('generic') properties of semi-solid deformable materials that later became taken over and canalized by the genome ('genetic'). At this later stage, the molecular level could become part of the biological levels of organization.

Towards a Theory of Development. Edited by Alessandro Minelli and Thomas Pradeu

'The party leaders were split on the platform' demonstrates that words not only define the sentence but that the sentence also defines the meaning of each word. Similarly, the supposedly true headline 'Prostitutes appeal to the Pope' (Russell-Rose, 2011) shows that context determines the meaning of the sentence. Bone morphogenic protein 4 (BMP4) can be a signal for growth, differentiation, or apoptosis within the same organism. What it does depend on is the historical context of the cells receiving it. (This and the other cited developmental examples are discussed in Gilbert, 2010). As Leo Rosten (2003) remarked, the sentence 'I should buy two tickets for her concert?' has seven different meanings depending on which word is emphasized!

Development is all about the interpretation of relationships. The fertilized egg inherits DNA; it does not inherit 'genes'. Genes and gene products are constructed anew in each cell in the developing embryo by the relationships between DNA, transcription factors, and RNA-splicing factors. Only certain regions of the DNA are constructed into genes, and different regions of the genome can be genes in different cell types. Note that the 'gene' is a higher-order structure than DNA, and that the interpretation of 'what is a gene' is done by the cell, an even higher-order structure (Stotz et al., 2006). As John Stamatoyannopoulos (2012: 1603), one of the leaders of the ENCODE project, recently summarized, 'Although the gene has been conventionally viewed as the fundamental unit of genomic organization, on the basis of ENCODE data it is now compellingly argued that this unit is not the gene but rather the transcript . . . On this view, genes represent a higher-order framework . . . creating a polyfunctional entity that assumes different forms under different cell states'.

Oyama (1985) has famously called this 'the ontogeny of information'. The organism does not inherit a 'program' as much as it inherits DNA and a cytoplasmic interpretation device (Gilbert, 1991; Nijhout, 1999). The same programmed music score can be interpreted in numerous ways by different orchestras. Every performance is different, even from the same score and the same orchestra. Indeed, it must be. Compare, for instance the recording of Pachelbel's *Canon* played by the English Chamber Orchestra under the baton of Johannes Somary with the same piece played by Musica Antiqua Köln, directed by Richard Goebel. Moreover, a concert A of 440 Hz is heard very differently when played by a cello or a trumpet. Even the interpretation of concert A differs geographically: concert A is 440 Hz in the United States and Britain, while it is usually 442 Hz in continental Europe. The interpretation of the score differs even in the pitch of the notes.

So there must be interaction between score and instrument (and orchestra, more largely), and there must be interaction between DNA and transcription factors. That the performance of a phenotype depends on its wider context has been long known by embryologists (see Gilbert & Sarkar, 2000) and is manifest in four major categories:

(i) *Plasticity.* Temperature-dependent pigmentation in butterflies, nutrition-dependent caste determination in hymenopterans, and site-specific sex determination in certain invertebrates were all known to early embryologists (see Hertwig, 1894). More recently, it has been seen that almost all, if not all, organisms have some developmental plasticity, and the inherited DNA determines a repertoire of phenotypes, not a specific phenotype. The environment can instruct which of the possible phenotypes to form. Species have evolved such that their genomes are responsive to environmental agents (see Gilbert & Epel, 2009). It is worth noting that the model systems often used in developmental biology have been specifically selected for their canalization (i.e. a lack of environmental agency) so that the genetics of development can be elucidated (Bolker, 2012; Gilbert, 2009).

(ii) *Organicism.* The parts of the organism determine the development of the whole and the whole developing organism reciprocally determines the properties of its parts. Lenoir (1982) has argued that the founders of modern embryology, Döllinger, Pander, von Baer, and Rathke, subscribed to the organicism set forth in Kant's *Critique of Judgment* (quoted in Lenoir, 1982: 25). Said Kant: 'The first principle required for the notion of an object conceived as a natural purpose is that the parts, with respect to both form and being, are only possible through their relationship to the whole . . . Secondly, it is required

that the parts bind themselves mutually into the unity of a whole in such a way that they are mutually cause and effect of one another.' Oskar Hertwig (1892), one of the leaders of embryology, proposed organicism as the true middle ground between reductionism and vitalism. He wrote that the parts of the organism develop in relation to each other, that is, the development of the part is dependent on the development of the whole. This was reiterated by Hans Spemann, who wrote, 'We are standing and walking with parts of our body which could have been used for thinking had they developed in another part of the embryo' (Spemann, 1943: 158–159; transl. by Horder & Weindling, 1986: 219). This emerging order was also thought to be critical to any philosophy of development by Paul Weiss, who said, 'Wherever we study such emergent order, we recognize it to be of tripartite origin, involving (1) elements with an inner order, (2) their orderly interactions, and (3) an environment fit to sustain their ordered group behavior' (Weiss, 1955: 296).

(iii) *Phenotypic heterogeneity*. The same mutation can produce a different phenotype in different individuals (Nijhout & Paulsen, 1997; Wolf, 1997, 2002). Phenotypic heterogeneity comes about because genes are not autonomous agents. Rather, genes interact with other genes and gene products, becoming integrated into complex pathways and networks. Thus, in addition to developmental plasticity dependent upon an environment external to the cell, genes can function differently depending on other genetic parameters. Bellus et al. (1996) found that the effects of the same mutant FGFR3 gene on limb development differed from person to person, with the phenotypes ranging from relatively mild anomalies to potentially lethal malformations. Similarly, the effects of particular mutant genes causing holoprosencephaly differ in the different family members having the same mutant gene (Dubourg et al., 2004; Marini et al., 2003). The severity of a mutant gene's effect often depends on the *other* genes, whose products have become part of the environment of the gene, as well as on environmental factors, and it will take a systems approach to find out how.

(iv) *Co-development*. All this regulation occurs through normal physico-chemical interactions. No higher-level process occurs in any other way. However, selection into viable networks and functional circuits occurs at a higher level, permitting only a subset of possible networks to evolve. As Leibniz (1697), one of the philosophers who most influenced Darwin, realized, while all permutations may be possible, very few will be compossible. By this he meant that not all possibilities could be actualized, because not all parts can function together to make coherent wholes. Ecosystems are examples of compossible systems: a squirrel and a whale are both possible, but not compossible in the same habitat. The fourth example of such higher-level phenomena, then, is the 'holobiont' created by the interactions of the 'host' with its symbionts. The host and symbiont are united anatomically, physiologically, immunologically, developmentally, and even evolutionarily (Gilbert et al., 2012; Nyholm and McFall-Ngai, this volume; Pradeu, 2011). Metabolic pathways initiated in the microbial symbiont get completed in the host, and vice versa; developmental pathways initiated in the host become completed by the symbiont, while the symbiont's metabolism is altered by signals from the host. Indeed, the gut and immune system of mammals is often completed by chemical signals originating from bacterial cells. We are literally 'becoming together' with the outside environment. Microbes are part of our post-embryonic developmental patterning, and the microbiome is our eleventh organ system.

This tells us that downward causation can be brought about in several ways. First, entities at higher levels place constraints on which lower-level interactions are viable and maintainable. Second, the parts must be compossible to form a greater whole. The bacteria that constitute our gut microbiota are not selected for their species; rather, they are selected for their functions (Faust et al., 2012; HMPC, 2012). Third, the higher-level entities also interpret the lower-level agents: a signal for apoptosis in one cell is a signal for proliferation in another. Fibroblast growth factors promote growth in some

cases and prevent growth in others. The transcription activated by paracrine factors depends upon the receiving cell's developmental history.

And there is a fourth mechanism: the higher-level structures give the physical location in which the lower-level modules function. For literal 'top-down' causation, one can't beat the dorsal-ventral patterning of vertebrates or *Drosophila*. But for these processes to occur—for the top to become distinguished from the bottom—one needs the placement of mRNAs in particular places within the cell. The *gurken* mRNA has to be placed dorsally in the *Drosophila* oocyte; the *Vg1* message has to be placed ventrally in the amphibian oocyte. Chordin must be made by the dorsal cells of the vertebrate embryo; while the homologous protein must be made in the ventral cells of the fly embryo. And both arise by the interactions of numerous tissues (see El-Hani, manuscript submitted). Indeed, the ventral cells of *Drosophila* arise only because the Dorsal protein, a transcription factor, is placed into the ventral cells' nuclei by interactions between the oocyte and the ventral follicle cells. If the Dorsal protein enters all cells, the entire embryo is ventralized. And this is regulated by the positioning of the *gurken* mRNA in the future dorsal region of the oocyte. Thus, the higher-level cell structure can regulate the places of transcription factor-gene regulation and so generate patterning.

Relations and upward causation

This phenomenon of 'downward' causation meets with and interacts with the phenomena of 'upward' causation. Two principles must be recalled in every discussion of upward causation in embryology. First, there is Haraway's principle (2008: 25–26) that 'relationships are the smallest possible pattern for analysis'. Information is not about essence; it is about relations. This is germane to the above discussion and will be continued in the discussion below. Second, development acts almost exclusively through stereocomplementarity (Gilbert & Greenberg, 1984). Stereocomplementarity is the interaction between shapes, and it is one of the great unifying principles of biology. It is literally 'fitness', that is, things that fit together: enzymes/substrates; antibodies/antigens; DNA/transcription factors; paracrine factors/receptors; sperm/egg; the interlocking components of signal transduction pathways; the interlocking components of ribosomes. Keys must fit only into certain locks. It is all 'copulation', literally, the binding together. Thus, information in development is about the interaction of complementary shapes. Our 'information' is in-form-ation. That is, information takes shape; it is not abstract, although the rules for interactions may become so. Rather, even though we may represent information flow with arrows to indicate causation, direction, and temporality, we are really discussing the interactions of shaped objects.

There are ways other than stereocomplementarity through which nature transfers information. Mechanical transduction is used occasionally in development, especially in the production of the circulatory system and skeletal elements (Culver & Dickinson, 2010; Gilbert & Epel, 2009; Tang et al., 2004). Frequency, which is used in echolocation and insect mating systems, is rarely used in development, the major examples being the predator-induced hatching of red-eyed tree frog larvae and the settlement of coral larvae (Vermeij et al., 2010; Warkentin et al., 2006). Stereocomplementarity is the major way that information is embodied in developmental processes. And stereocomplementarity implies reciprocal relation. The stereocomplementary molecules mediating gamete recognition are the fastest diverging proteins known (Palumbi, 2009). And for each change on the protein of one gamete, there has to be a corresponding change on the other.

So one might ask: what is the stereocomplementary relation that defines 'reality' for the embryo? What type of interaction determines whether an entity is a real (i.e. functional or morphological) unit for development? Let us consider that the primary unit of reality for the embryo is the enhancer–transcription factor relation. If a gene for a marker protein (such as β-galactosidase or green fluorescent protein) is ligated to a promoter, enhancer traps can determine what the embryo considers 'real'. This reality might not correspond to adult 'reality'. Surely, enhancers will activate genes in the retina and the gut tube. But one enhancer will activate genes in the *medial* rib, while a different enhancer will activate genes in the *lateral* portion of the rib (Guenther et al., 2008). Apparently, 'lateral rib' is

an anatomical construction unit recognized by the embryo. Similarly, enhancer traps of the *Drosophila* embryo shows that the embryo comprises numerous compartments that are not apparent in the adult, but which are building units of the embryo (Buszczak et al., 2007).

But this is just the primary relationship and is in the nucleus. In order to understand or model cells and, indeed, embryos, one has to relate what happens in the nucleus to what happens in the cell cytoplasm and cell membrane.

Scoring development: we all live in recursive subroutines, recursive subroutines, recursive subroutines

When one starts to think about the principles that guide embryogenesis, one might be given the impression that all the decades of developmental biology research have shown is that the development of a particular simple tissue depends partly on its parent tissue (lineage) and partly on its neighbours (signalling). While its development can be understood with hindsight, there are still no formal predictors or rules as to what might happen. Development still lacks the sorts of underlying principles that make physics tractable and laws that gives quantitative predictions; one reason is that there are no elementary particles, and another is that any laws have proved elusive. Worse, it lacks a natural notation for writing the score, and that is one area that is touched on here.

The best that we have been able to do is to borrow the language of physics and try to constrain as much of development as possible into differential equations that describe change and predict performance, a tradition that started with the classic paper of Turing (1952) on molecular pattern formation. Today, there is a considerable amount of research in this area (see Barkai & Perrimon, 2011). Nevertheless, while there have been impressive successes in modelling a few phenomena such as signalling pathways (e.g. Witt et al., 2011), *Drosophila* segmentation (Ingolia, 2004) and somitogenesis (Goldbeter & Pourquié, 2008), general principles that help understand development have yet to emerge from this approach other than to affirm the integral importance of upwards and downwards causation.

The main reason for this is that development is very complicated and that what seems a simple development change is actually underpinned by the coordinated activity of hundred of proteins. What one soon notes is that, whatever the embryo, change is based on a relatively small set of protein networks whose outputs are the processes that drive patterning, signalling, proliferation, differentiation, and morphogenesis, and these themes are used over and over again (Bard, 2013). This simplicity stands in strong contrast to the complexity of the full developmental score with its swarms of genes, molecules, and tissues. Table 8.1 gives some idea of the numbers of these components for the mouse and human. The figures for protein-coding genes and proteins are well known; the number of developmental networks and output actions comes mainly from Gilbert (2010) and the numbers of tissues from Bard (2012). Figure 8.1 indicates how these events are integrated.

The number of processes is surprisingly small, and they fall into two groups (Table 8.1 and Table 8.2). First, there are the gene regulatory networks (Levine & Davidson, 2008), which comprise ~10 signal-activated networks and an unclear number of patterning and timing networks. These control the second group, which we can think of as process networks that actually lead to phenotypic change: here there are 5–10 pathways associated with proliferation, ~3 apoptosis networks, 5–10 morphogenetic networks, and a hierarchy of differentiation pathways (Bard, 2012). The number of high-level differentiation pathways is less clear because major cell types have subtypes, but one pointer here comes from the options available to neural crest cells. These include mesenchymal cells (bone, muscle, cartilage, fibroblasts), epithelia of various sorts, neurons and neuron-support cells, and melanocytes but not the other major lineage of blood cells and their many subtypes. There are perhaps ~10 main

Table 8.1 Levels and numbers.

Protein-coding genes	35 000
Proteins	~70 000
Developmental networks	~60
Output processes	~60
Simple tissues	~10 000

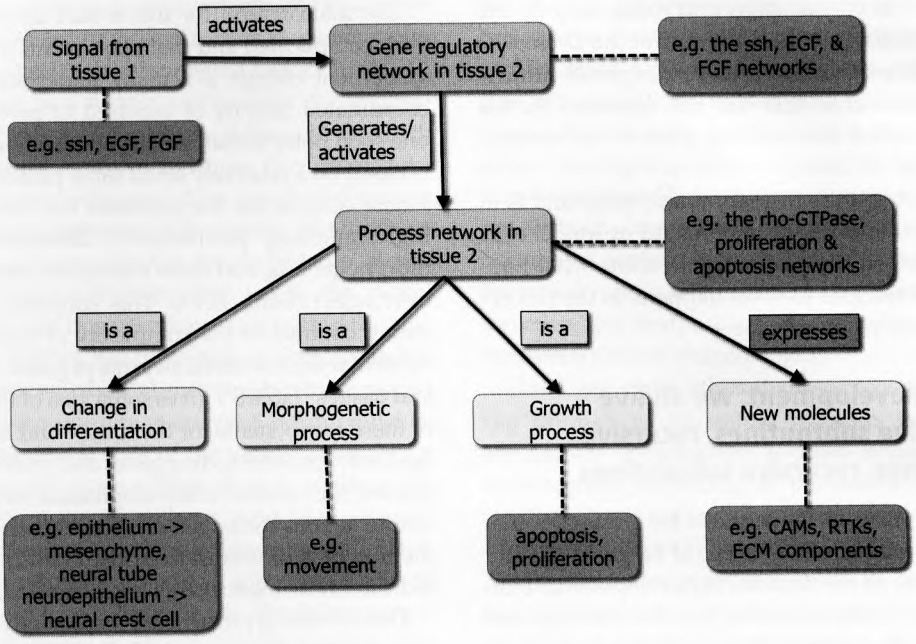

Figure 8.1 Graph showing the effects of signalling pathways. Examples are in darker grey boxes. The 'is a' link represents a typing or classification. Examples are in grey boxes. (From Bard, 2011b. Reproduced with permission from John Wiley & Sons.)

Table 8.2 Some major networks whose output are the processes that drive development.

Gene regulatory networks	Process networks	
Signalling	***Differentiation to***	***Morphogenesis***
ERK/MAPK	haematopoiesis lineage	boundary formation
FGF, JAK/STAT	erythroid lineage	(Eph-ephrin)
Notch–delta	lymphocyte lineage	epithelium
Shh, SMAD	myeloid lineage	branching
TGFβ, VEGF, Wnt	epithelium	folding
	mesenchyme	migration
Patterning	chondrocyte	rearrangement
Hox patterning	fibroblast	mesenchyme
RTK patterning	muscle	adhesion
Notch oscillator system	osteoblast	migration
signalling gradients (e.g. Shh)	neuron	
etc.	neuron-support cell	***Apoptosis***
	pigment-producing cell	caspase, fas
Timing		cellular apoptosis
Nothing is known of these		
		Proliferation
		cyclin+downstream events

cell differentiation routes. Taking a broad brush to the topic, there are ~50 major processes that underpin development (perhaps 60 if we allow for the possibility that a few more will be discovered), with some having several outputs (Figure 8.2).

These leitmotifs are used over and over again in each complex, multicellular animal as it develops. The fine details are not of course the same in each: evolutionary change means that the exact details of the networks and their outputs vary from organism to

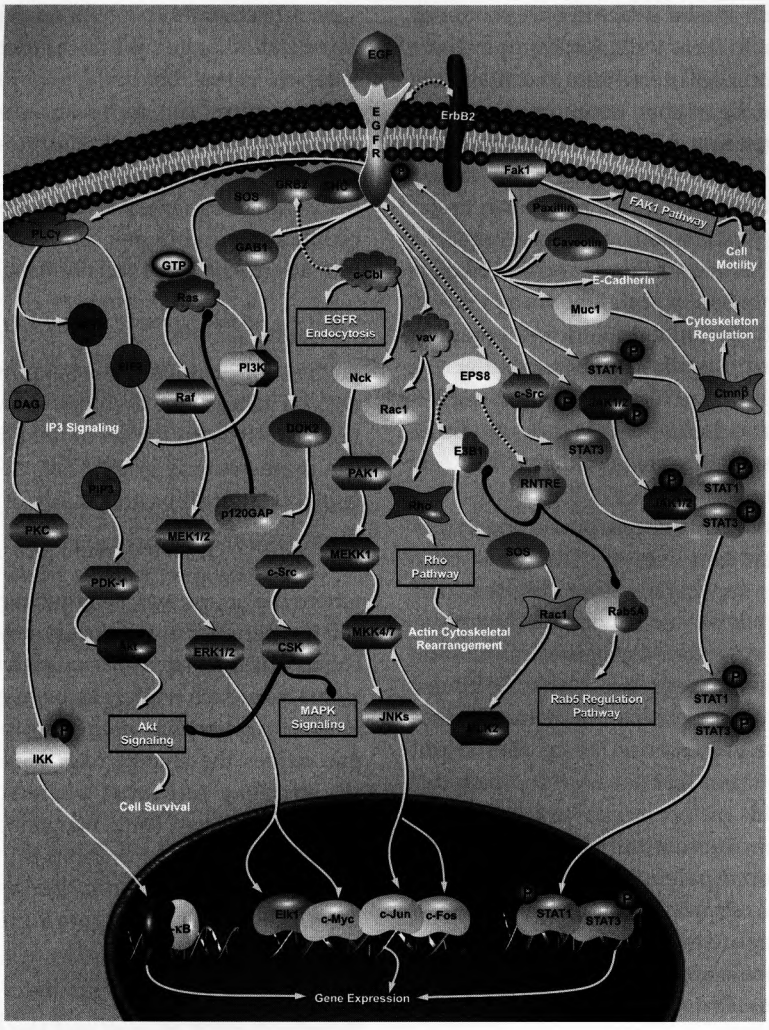

Figure 8.2 The EGF network (>60 proteins) activates the cell cycle. (modified from www.sabiosciences.com/pathwaycentral.php)

organism and from tissue to tissue. It should also be pointed out that the repertoire of developmental networks (Table 8.2) excludes the many more networks that 'run' the biochemical, physiological, and neurological systems. Nevertheless, if there is any underlying simplicity to be found in developmental biology, it centres around a basic set of molecular networks[2] whose outputs are the processes that drive embryogenesis (Figure 8.1). Not that these networks are simple: they contain ~10–50 interacting proteins (see http://www.sabiosciences.com/pathwaycentral.php and Figure 8.2). While elucidating the components and the organization of these networks has been a triumph of the last decade of research in molecular genetics, we still don't know how they work qualitatively, let alone quantitatively.

Processes are the subroutines of development

It is easiest to see how frequently the same processes are used by looking at the emerging anatomy of a

[2] What are called networks here are more commonly called pathways. The former is the preferable term because these assemblages of proteins often include alternate routes and end-points.

vertebrate, in which similar structures are produced across the embryo. In the mouse, for example, there are ~200 long bones, >50 vertebrae, and many examples of muscles, ligaments, neuronal nuclei and ganglia, and bifurcating tubes. Similar structures are produced over and over again with minor, locale-specific features that do no more than tinker with the numerical parameters of the process. The central difference between a femur and a phalange, for example, is only one of scale: the former is ~40 times the length of the latter. The development of classes of standard modules is ubiquitous across complex organisms and reflects the regular and frequent use of the processes that build these modules, albeit that their activities are modulated by local molecular constraints.

Modular development has an interesting implication within a systems context. There has been some discussion in the literature as to whether part at least of the genome should be viewed, metaphorically at least, as a database of genomic information available to developing cells (e.g. Noble, 2010). Indeed, it is hard not to visualize the networks that generate dynamic processes as being the output of genomic subroutines that are used in many different contexts. This metaphor can be taken a little further: as program subroutines have outputs that depend on their input parameters, so the output of process subroutines depend on the details of the cell types in which they are expressed. Each organism's development arises from an evolutionarily canalized set of compossible subroutines (Huang et al., 2009; Kauffman, 1987).

A formal language for development

The language of differential equations is sadly of limited applicability for development: we just don't know enough about the participants, their interactions, or their rate constants to be able to use the alphabet of mathematics to describe what is going on. The events of development do, however, give us some clues as to how to start describing things with some formality. Development involves events at levels from the genome through gene expression, signalling, networks, and processes, to tissues. In this list, it is clear that processes stand out as different: while genes, protein, networks, cells, and tissues reflect states, processes reflect activities: they drive state changes. While the former are nouns, the latter are verbs!

There turns out to be an area of mathematics known as graph theory that captures this difference. A mathematical graph is nothing like a data graph because it doesn't deal with numerical data. It turns out that many complex stories can be decomposed into a series of small facts of the general form

<state 1> <relationship 1> <state 2>[3]

Each is, for obvious reasons, known as a triplet and a given state can be involved in two or more triplet. For a given story, the set of linked triplets comprise the mathematical graph. For development, these triplet relationships are mainly of the form

<noun a> <verb x> <noun b>

where the nouns may be anything from a tissue to a cell to a network to a molecule, and verbs reflect processes (differentiates into, migrates, apoptoses, etc.). In practice, each triplet can be seen as a simple fact, with the relationship often being the *activity* or *process* that drives the change (e.g. *<SHH><activates><the shh signalling pathway>*, where shh stands for the signalling protein sonic hedgehog). The other core relationship is <is_a> and this is used as a classification tag (e.g. *<ectoderm> <is_a> <epithelium>*). Here, it is worth noting that Figure 8.1 is actually a formal graph.

Developmental change and the notation of graphs

The use of graphical notation to describe developmental change turns out to be useful in several ways (Bard, 2011b, 2013). First, the representation is visual; second, the format is web accessible and can be linked to other resources; third, the format lends itself to being updated as new information is discovered; fourth, making the graph highlights gaps in knowledge and so suggests experiments; fifth, it shows the centrality of the relation as the fundamental unit of development. Nevertheless, the format does have limitations. It is not easy, for example

[3] In graph theory, the standard terminology is *<node> <edge> <node>*.

to represent the internal structure of chemical reactions and biochemical pathways: they either need the insertion of dummy intermediates or they require a richer formulation than triplets (see www.sbgn.org/). A more difficult problem is including the full complexity of a developmental event: representing the networks underpinning the processes other than by their names would be unwieldy. In practice, this could only be done by listing the triplets and handling them computationally.

It turns out that many developmental phenomena can be represented as a graph where the nodes are biological entities scaling from proteins upwards and the edges are relationships (Bard, 2011b); and there are several advantages in doing this:

(i) They unpack the complexity of development by reducing it to a set of simple but integrated facts, albeit that the set may be quite large.
(ii) Extra triplets can be added as new parts of the story are discovered.
(iii) Where nodes have ontology IDs, links to associated data can be included.
(iv) IDs from PubMed can be used as citations of facts

There are, however, further advantages in representing the graphs as diagrams that show the general organization. For developmental biology, these include:

(v) They emphasize that control of development is widely distributed.
(vi) Gaps in the diagram highlight areas where further work is needed.
(vii) Colour can be used to reduce the complexity of the narrative.

Together, these advantages mean that the mathematical graph can be seen as a terse, updatable review of a developmental event. Further, because databases are continually being upgraded with new data, the ID links ensure that the associated data is also up to date. This is not to suggest that the graphical notation should be seen as a step towards a more general theory but rather that the formalism articulates the sort of clarity that makes theorization one step easier.

The information required to make a graph of how change takes place in an embryo comes from experimentation, and not only involves signals and the activation of processes but also a clear understanding of what these processes do. Some of this information is not yet available, and the resulting graphs will skate over some details (e.g. a network can be represented by a single node rather than by a intricate sub-graph whose details may not be germane to the problem being considered). Things are more complicated where morphogenesis takes place, as the final structure will not only depend on signals and networks but on such as physical activity by cells that is constrained by the geometry of the tissues (Figure 8.3). In a sense though, making the triplets and so producing the formal representation of the developmental event is relatively straightforward.

Integrating all this information for a real example in a single clear diagram is often hard to do, partly because so much is going on and partly because it can be difficult to maintain the sense of the dynamics. One trick that is helpful here is to embed molecular nodes within the blocks for the tissues. Another is to use different colours for different aspects of the diagram. Note that a classic graph, the London Underground map, where the relationship is *<connects with>*, uses colours to distinguish paths through the network.

An example clarifies this. All developing tissues, once they reach a critical size, need a blood supply, and this is achieved by the local mesenchyme secreting the signal protein vascular endothelial growth factor (VEGF). Research on mouse embryos has shown that this signal diffuses into the local environment, providing a concentration gradient that decreases with distance. Receptors on nearby blood vessels bind this signal, signal transduction activates the proliferation pathway locally, Notch–Delta activation ensures that the new cells form a single capillary, and this extends up the concentration gradient towards and into the original tissue (for reviews, see Chung & Ferrara, 2011; Suchting et al., 2007). There are some 30 small facts associated with this event, and each can be described as a triplet (with further triplets linking these facts with the original publications, as stored in PubMed, Gene Ontology and GXD, the mouse gene-expression database).

The key elements of the data on angiogenesis are shown in a graphical representation in Figure 8.4.

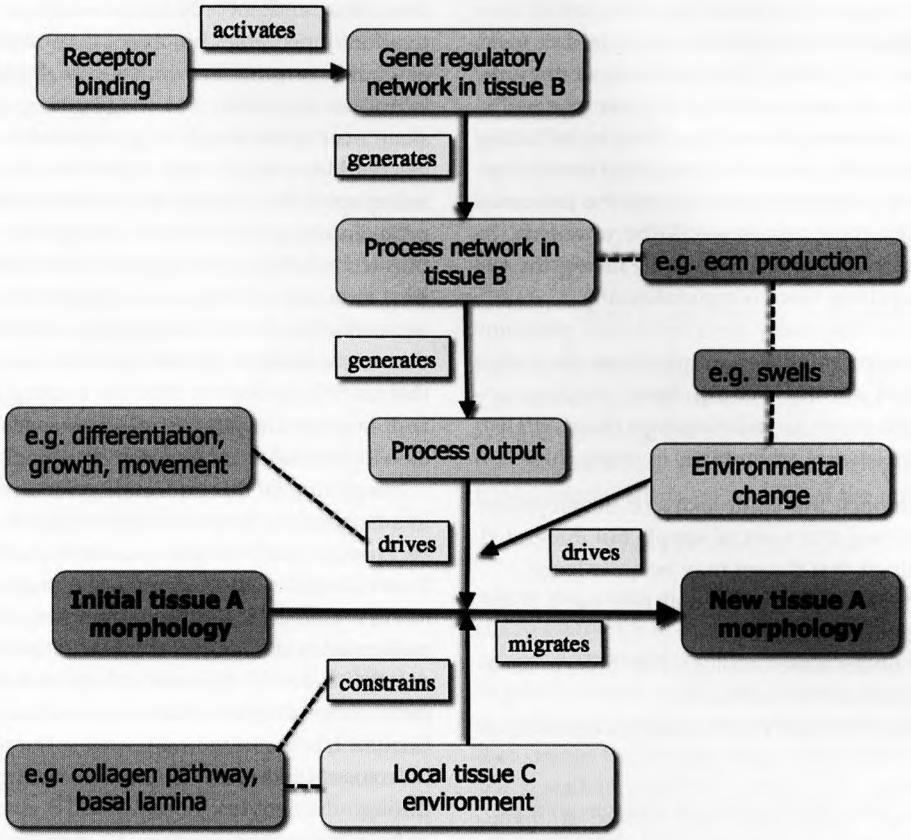

Figure 8.3 A graph showing the modelling of morphogenesis either by the downstream effects of gene activity (upper example is the effect of extracellular matrix production) or through existing boundary effects in the environment (lower example is a collagen track used for contact guidance). (From Bard, 2011b. Reproduced with permission from John Wiley & Sons.)

The use of shading and the embedding of molecular nodes within tissue ones enable the key features to be easily grasped. While links to gene-expression data and PubMed citations could be added, they would make the graph unwieldy, but could be included, with some trouble but little difficulty in a formal listing of the triplets. One advantage of the pictorial representation is that it becomes easy to see gaps in the story. Obvious questions that are yet to be answered are: how is blood circulation established, what is the range of the VEGF gradient, and how is Notch activated?

It should be emphasized that this graph is produced to demonstrate that the complexities of development can be represented in a compact visual format rather than a computational entity. While mathematicians will correctly assert that using a single triplet is inadequate to describe complex chemical interactions, the diagram does however provide an intuitive and clear understanding or what is going on. Equally important, this representation shifts the focus from the signal that activates angiogenesis to the actual process of angiogenesis. The different shades represent tissue states, processes, and networks, and it is worth pointing out that, in the graphical context, all nodes and all levels have equal status—there is no preferred level, a well-known property of systems biology analyses (Noble, 2010).

Discussion and conclusion

As was said in the opening section, this chapter takes a systems approach to development in that

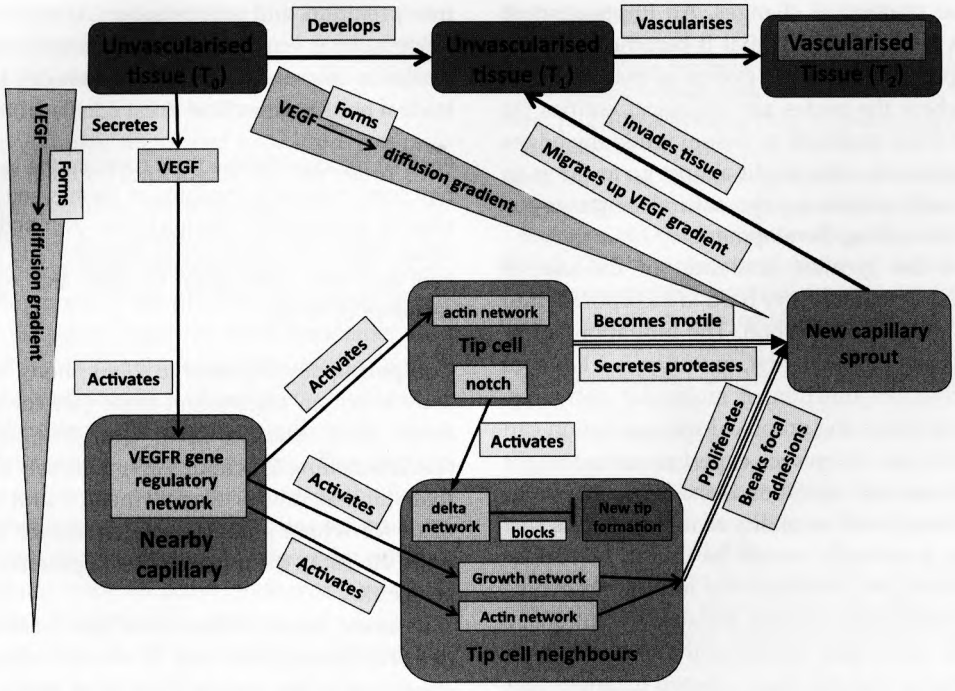

Figure 8.4 A graph describing some of the core events underpinning angiogenesis. Tissues are in darker boxes while molecular events are in lighter boxes, and processes are in pale boxes. (From Bard, 2011b. Reproduced with permission from John Wiley & Sons.)

it tries to step back from the normal minutiae of developmental phenomena. Perhaps the key point that it makes is that development always involves all levels from the genome to the environment, with causation working in both directions mainly to activate ~60 networks whose outputs are the processes that drive development (see Saetzler et al., 2011).

In the wider context, a theory of development cannot be a subset of a theory of genetics because much of development is not run by the genome. Genomic activity is neither cell- nor tissue-autonomous but acts as a resource to be activated by signals from other tissues. Any theory of the development of a tissue involves the prior history of that tissue, knowledge of the tissue's environment, and a description the geometry of that tissue's environment. The music is written in several parts. In a wider context, Waddington (1975) was less than impressed with a simple genomic description of development, noting that these were three perspectives on 'diachronic biology'. Gilbert et al. (1996) used the notion that development is the first derivative of gene expression, anatomy, and physiology, and that evolution is the first derivative of development. In this view, genetics is the means by which the same processes of development become inherited from one generation to the next, and evolution is seen as changes in the developmental processes, thereby giving new anatomical or physiological properties. The purpose of this discussion is to look at some of the implications of this approach.

Development as integrated processes

At first sight, this emphasis on processes might seem to contradict one of the few general principles of systems biology, that there is no preferred level of control. This focus on processes is not, however, a matter of control or levels, but accepts the reality that it is only processes that generate actions and that these can affect anything from a gene up to a tissue. Actions are verbs and everything else is a noun! Processes can thus affect change at any level, and in development the core processes are those

that cause anatomical changes. An implication of processes being actions is that it becomes possible to describe developmental events as mathematical graphs where the nodes are biological entities (at any level from molecule to tissue) while the edges are the processes. One might almost go as far as to say that such graphs are the natural language for formally describing development.

Perhaps the greatest limitation of the use of graphs here is that it is very hard to incorporate into them the effects of mutation. The limitation is important in two contexts: first, mutation is a key tool in exploring the function of molecular networks, and the most that the graphical representation can do is to indicate where such experimentation might be helpful; second, mutation is the driver of evolutionary change, and an ability to represent this sort of change graphically would be useful. Nevertheless, focussing on processes has interesting evolutionary implications, mainly because evolutionary change is essentially mutation-induced developmental change that has been selected (changes that are lost are normally viewed as congenital abnormalities), as has been apparent for almost a century (Goldschmidt, 1927).

Simple inspection shows that the key mutations that have driven anatomical change over the last few hundred million years are those that affect the dynamics of developmental processes. In vertebrates, these are for minor patterning and growth: there is good reason to suggest that, once vertebrates reached land, most further change was essentially quantitative. While mutations affect the function of individual proteins, the actual downstream phenotype depends on the role of that protein within the networks in which it plays some role. It is certainly sensible to suggest that mutations affect the dynamics of the networks and so modulate the output process (Bard, 2010). This focus on processes has a further advantage: mutations that affect the output of processes are integral to the network, so that those particular mutations are easily inherited (Bard, 2011a).

Development as performance

If there is any analogy for development, per se, it is performance. Performance is a mixture of score, interpretation, and improvisation. The notion of development as musical performance was mentioned earlier in this essay, and this conceit can be traced back at least as far as Karl Ernst von Baer (1864: 281):

> For that reason, I believe I can compare the various life-processes to musical thoughts or themes and call them creative ideas, which construct their own bodies themselves. What we call in music harmony and melody is here type (the combination of parts) and rhythm (the sequence of forms).

Comparing development to a symphony or a rhapsody is not an uncommon trope (see for example, Keim, 2012; Marino, 2004; Qiu, 2006). Schelling (1802; Schelling & Schott, 1989) famously remarked that 'architecture is frozen music', and to those of us for whom anatomy is architecture (as in the word *Bauplan*), the music of development is not frozen at all.

In music theory, a chord is a 'simultaneity', a series of different notes, each of which is played at the same time as the other pitches of its group. Thus, a chord progression is called a 'succession of simultaneities'. Chord progressions are the homologies of music. They are the underlying unity amidst the apparent diversity. The I-VI-IV-V progression (e.g. C-A^{m}-D^{m}/-G^{7}) originated in Western music in the 40s. It is the underlying progression of *Heart and Soul*, *The Way You Look Tonight*, and hundreds of others. The I-IV-I-V (C-F-C-G) theme is also characteristic of Western music, although it is a much earlier clade. It was very common in Elizabethan English music, and it is still extant, where this progression forms the basis for *Goodnight, Ladies* and *The Lion Sleeps Tonight*. There are only so many chords that work together. It's not what's possible. It's what's compossible.

Evolution occurs by changing development. Improvisation—playing something novel with other musicians—is not complete freedom. Rather, it is a mutual understanding of the chord progressions (Gorow, 2002). A good improviser has to know the cord changes, even if he or she decides to experiment with them. Each improvisation has to work within the musical context provided by the other performances. This is the mutually constructed niche that 'enables' the particular improvisation (Longo et al., 2012). So not everything is possible.

But within the rules and within the context, there is an infinite number of possibilities. Each animal has most of the same notes. But it is where you play the notes (in combination with what other notes), how long you play them, and how loud you play them, that matter. Homologies are the chord progressions of evolution. Each species is its own song. Each individual is a performance of that song, with its own idiosyncratic improvisation on the score. Graphs may provide the notation by which we can visualize the score. And we must remember that each score is not merely for a song, but for a choreographed performance of interacting shapes, a dance. Like dance, development is brought about by the interacting of pliable surfaces. The 'idea' of the dance is not the dance, the score, or the fleshy agents constructing it. The dance is its performance. Scoring such choreography has been very difficult and continues to be an active endeavor (Benesh & Benesh, 1983; Neagle & Ng, 2003).

Thus, development is an ongoing performative act. It involves a score (DNA), an orchestra for interpretation (to choose what DNA is a gene, what the function of BMP4 is in any particular cell, etc.), and improvisation (regulating gene expression such that most knockout mice have minimally altered phenotypes; altering anatomy by changing gene expression patterns). Like an ensemble group, no conductor is needed—just some ion transport as sperm meets egg is enough to start the show. The relationships between cell surfaces generate morphogenetic fields, tissues, and organs. The body builds itself as it develops, each whole becoming a part of something larger that it generates, and each whole defining the context of its parts. Development is a creative choreography of molecules, cells, tissues, organisms, and ecosystems. As each organism is a new developmental performance, we are left with Yeats' (1929) question, 'How can we know the dancer from the dance?'

Acknowledgements

SFG is funded by the Academy of Finland and a faculty research grant from Swarthmore College. He thanks S. Friedler and G. Levinson for their discussions, and A. Soto for her provocative critiques of the draft.

References

Auletta, G. (2011). *Cognitive Biology: Dealing with Information from Bacteria to Minds*. Oxford University Press, New York.

Bard, J. (2010). A systems view of evolutionary genetics. *Bioessays*, **32**, 559–63.

Bard, J. (2011a). The next evolutionary synthesis: from Lamarck and Darwin to genomic variation and systems biology. *Cell Communication and Signaling*, **9**, 30.

Bard, J. (2011b). A systems biology formulation of developmental anatomy. *Journal of Anatomy*, **218**, 591–9.

Bard, J. (2012). A new ontology (structured hierarchy) of human developmental anatomy for the first 7 weeks (Carnegie stages 1–20). *Journal of Anatomy*, **221**, 406–16.

Bard, J. (2013). Driving developmental and evolutionary change: a systems biology view. *Progress in Biophysics and Molecular Biology*, **111**, 83–91.

Barkai, N., and Perrimon, N. (2011). The era of systems developmental biology. *Current Opinion in Genetics and Development*, **21**, 681–811.

Bellus, G.A., Gaudenz, K., Zackai, E.H., et al. (1996). Identical mutations in three different fibroblast growth factor receptor genes in autosomal dominant craniosynostosis syndromes. *Nature Genetics*, **14**, 174–6.

Benesh, R., and Benesh, J. (1983). *Reading Dance: The Birth of Choreology*. McGraw-Hill, New York.

Bolker, J. (2012). There's more to life than rats and flies. *Nature*, **491**, 31–3.

Buszczak, M., Paterno, S., Lighthouse, D., et al. (2007). The Carnegie protein trap library: a versatile tool for *Drosophila* developmental studies. *Genetics*, **175**, 1505–31.

Chung, A.S., and Ferrara, N. (2011). Developmental and pathological angiogenesis. *Annual Review of Cell and Developmental Biology*, **27**, 563–84.

Culver, J.C., and Dickinson, M.E. (2010). The effects of hemodynamic force on embryonic development. *Microcirculation*, **17**, 164–78.

Davidson, E.H. and, Levine, M.S. (2008). Properties of developmental gene regulatory networks. *Proceedings of the National Academy of Sciences USA*, **105**, 20063–6.

Dubourg, C., Lazaro, L., Pasquier, L., et al. (2004). Molecular screening of SHH, ZIC2, SIX3, and TGIF genes in patients with features of holoprosencephaly spectrum: mutation review and genotype-phenotype correlations. *Human Mutation*, **24**, 43–51.

El-Hani, C.N. (2014). Downward determination as a propensity-changing non-causal relation. In press.

El-Hani, C.N., and Emmeche, C. (2000). On some theoretical grounds for an organism-centered biology: property emergence, supervenience, and downward causation. *Theory in Biosciences*, **119**, 234–75.

Ellis, G.F.R. (2012). Top-down causation and emergence: some comments on mechanisms. *Journal of the Royal Society Interface Focus*, **2**, 126–40.

Faust, K., Sathirapongsasuti, J.F., Izard, J., et al. (2012). Microbial co-occurrence relationships in the human microbiome. *PLoS Computational Biology*, **8**:e1002606.

Flüsser, V. (1987). *Vampyroteuthis infernalis*. Atropos Press, New York.

Gilbert, S.F. (1991). Commentary: cytoplasmic action in development. *Quarterly Review of Biology*, **66**, 309–16.

Gilbert, S.F. (2009). The adequacy of model systems for evo-devo. In A. Barberousse, M. Morange and T. Pradeu, eds., *Mapping the Future of Biology: Evolving Concepts and Theories*. Springer, Dordrecht, pp. 57–68.

Gilbert, S.F. (2010). *Developmental Biology*, 9th ed. Sinauer Press, Sunderland.

Gilbert, S.F., and Epel, D. (2009). *Ecological Developmental Biology*. Sinauer Associates, Sunderland.

Gilbert, S.F., and Greenberg, J.P. (1984). Intellectual traditions in the life sciences II. Stereocomplementarity. *Perspectives in Biology and Medicine*, **28**, 18–34.

Gilbert, S.F., Opitz, J.M., and Raff, R.A. (1996). Resynthesizing evolutionary and developmental biology. *Developmental Biology*, **73**, 357–72.

Gilbert, S.F, Sapp, J., and Tauber, A.I. (2012). A symbiotic view of life: we have never been individuals. *Quarterly Review of Biology*, **87**, 325–41.

Gilbert, S.F., and Sarkar, S. (2000). Embracing complexity: organicism for the 21st century. *Developmental Dynamics*, **219**, 1–9.

Goldbeter, A., and Pourquié, O. (2008). Modeling the segmentation clock as a network of coupled oscillations in the Notch, Wnt and FGF signaling pathways. *Journal of Theoretical Biology*, **252**, 574–85.

Goldschmidt, R. (1927). *Physiologische Theorie der Vererbung*. Springer, Berlin.

Gorow, R. (2002). *Hearing and Writing Music: Professional Training for Today's Musician*, 2nd ed. September Publishing, Gardena.

Guenther, C., Pantalena-Filho, L., and Kingsley, D.M. (2008). Shaping skeletal growth by modular regulatory elements in the *Bmp5* gene. *PLoS Genetics*, **4**:e1000308.

Haraway, D.J. (2008). *When Species Meet*. University of Minnesota Press, Minneapolis.

Hertwig, O. (1892). Urmund und Spina bifida. *Archiv für mikroskopische Anatomie*, **39**, 353–503.

Hertwig, O. (1894). *The Biological Problem of To-day: Preformation or Epigenesis?* Macmillan, New York.

HMPC (Human Microbiome Project Consortium) (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, **486**, 207–14.

Horder, T.J., and Weindling, P.J. (1986). Hans Spemann and the organiser. In T.J. Horder, J.A. Witkowski and C.C. Wylie, eds., *A History of Embryology*. Cambridge University Press, Cambridge, pp. 183–242.

Huang, S., Ernberg, I., and Kauffman S.A. (2009). Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. *Seminars in Cell and Developmental Biology*, **20**, 869–76.

Ingolia, N.T. (2004). Topology and robustness in the *Drosophila* segment polarity network. *PLoS Biology*, **2**: e123.

Kauffman, S.A. (1987). Developmental logic and its evolution. *BioEssays*, **6**, 82–7.

Keim, B. (2012). Cellular symphony of embryo development. http://www.wired.com/wiredscience/2012/06/embryo-development-video/

Leibniz, G.W. (1697). On the ultimate origination of things. Quoted in B.C. Look, *Leibniz's Modal Metaphysics*. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/leibniz-modal/

Lenoir, T. (1982). *The Strategy of Life: Teleology and Mechanics in Nineteenth Century German Biology*. D. Reidel, Dordrecht.

Longo, G., Montévil, M., and Kauffman, S. (2012). No entailing laws, but enablement in the evolution of the biosphere. *ACM Proceedings of the genetic and evolutionary computation conference, GECCO'12, Philadelphia, 7–11 July 2012*, Arxiv 1201.2069.

Marini, M., Cusano, R., De Biasio, P., et al. (2003). Previously undescribed nonsense mutation in SHH caused autosomal dominant holoprosencephaly with wide intrafamilial variability. *American Journal of Medical Genetics*, **117A**, 112–15.

Marino, M. (2004). Biography of Philip A. Beachy. *Proceedings of the National Academy of Sciences USA*, **101**, 17897–9.

Neagle, R.J., and Ng, K.C. (2003). *Machine-Representation and Visualisation of a Dance Notation*. Proceedings of Electronic Imaging and the Visual Arts—London July 2003.

Newman, S.A. (2012). Physico-genetic determinants in the evolution of development. *Science* **338**, 217–9.

Nijhout, H.F. (1999). Control mechanisms of polyphenic development in insects. *Bioscience*, **49**, 181–92.

Nijhout, H.F., and Paulsen, S.M. (1997). Developmental models and polygenic characters. *American Naturalist*, **149**, 394–405.

Noble, D. (2010). Biophysics and systems biology. *Philosophical Transactions of the Royal Society of London A*, **368**, 1125–39.

Oyama, S. (1985). *The Ontogeny of Information: Developmental Systems and Evolution*. Duke University Press, Durham.

Palumbi, S.R. (2009). Speciation and the evolution of gamete recognition genes: pattern and process. *Heredity*, **102**, 66–76.

Pirsig, R.M. (1974). *Zen and the Art of Motorcycle Maintenance: An Inquiry into Values*. William Morrow, New York.

Pradeu, T. (2011). A mixed self: the role of symbiosis in development. *Biological Theory*, **6**, 80–8.

Qiu, J. (2006). Epigenetics: the unfinished symphony. *Nature*, **441**, 143–5.

Rosten, L. (2003). *The New Joys of Yiddish*. Random House, NY. Excerpt at http://everything2.com/title/YiddishNew York.

Russell-Rose, T. (2011). Prostitutes appeal to Pope: text analytics applied to search. Accessed at http://isquared.wordpress.com/2011/07/28/prostitutes-appeal-to-pope-text-analytics-applied-to-search/

Saetzler. K., Sonnenschein, C., and Soto, AM. (2011). Systems biology beyond networks: generating order from disorder through self-organization. *Seminars in Cancer Biology*, **21**, 165–74.

Schelling, F.W.J. (1802). Philosophie der Kunst;quoted in Bright, M. (1984). *Cities Built to Music*. Ohio State University Press, p. 82.

Schelling, F.W.J. von, and Stott, D. W (1989). *The Philosophy of Art/Friedrich Wilhelm Joseph Schelling; edited, translated and introduced by Douglas W. Stott; foreword by David Simpson*. University of Minnesota Press, Minneapolis.

Soto, A.M., Rubin, B.S., and Sonnenschein, C. (2009). Interpreting endocrine disruption from an integrative biology perspective. *Molecular and Cellular Endocrinology*, **304**, 3–7.

Spemann, H. (1943). *Forschung und Leben*. Engelhorn, Stuttgart.

Stamatoyannopoulos, J.A. (2012). What does our genome encode? *Genome Research*, **22**, 1602–11.

Stotz, K.C., Bostanci, A., and Griffiths, P.E. (2006). Tracking the shift to 'postgenomics'. *Community Genetics*, **9**, 190–6.

Suchting, S., Freitas, C., le Noble, F., et al. (2007). The Notch ligand Delta-like 4 negatively regulates endothelial tip cell formation and vessel branching. *Proceedings of the National Academy of Sciences USA*, **104**, 3225–30.

Tang, G.H., Rabie, A.B., and Hagg, U. (2004). Indian hedgehog: a mechanotransduction mediator in condylar cartilage. *Journal of Dental Research*, **83**, 434–8.

Turing, A.M. (1952). The chemical theory of morphogenesis. *Philosophical Transactions of the Royal Society of London B*, **237**, 37–72.

Vermeij, M.J., Marhaver, K.L., Huijbers, C.M., et al. (2010). Coral larvae move toward reef sounds. *PLoS One*, **5**:e10660.

von Baer, K.E. (1864). Welche Auffassung der Natur ist die richtige? und wie ist diese Auffassung auf die Entomologie anzuwenden? In K.E. von Baer, *Reden gehalten in wissenschaftlichen Versammlungen und kleinere Aufsätze vermischten Inhalts. Erster Theil. Reden*. H. Schmitzdorff, St. Petersburg, pp. 237–84.

Waddington, C.H. (1975). *Evolution of an Evolutionist*. Edinburgh University Press, Edinburgh.

Warkentin, K.M., Caldwell, M.S., and McDaniel, J.G. (2006). Temporal pattern cues in vibrational risk assessment by embryos of the red-eyed treefrog. *Agalychnis callidryas*. *Journal of Experimental Biology*, **209**, 1376–84.

Weiss, P. (1955). Beauty and the beast: life and the rule of order. *Scientific Monthly*, **81**, 286–99.

Weiss, P. (1967). One plus one does not equal two. In G.C. Quarton, T. Melnechuk, and F.O. Schmitt, eds., *The Neurosciences: A Study Program*. Rockefeller University Press, New York, pp. 801–21.

Weiss, P. (1977). The system of nature and the nature of systems: empirical holism and practical reductionism harmonized. In K.E. Schaefer, H. Hensel and R. Brady, eds., *Toward a Man-Centered Medical Science*. Futura, Mt. Kisko, pp. 17–63.

Winther, R.G. (2011). Part-whole science. *Synthese*, **178**, 397–427.

Witt, J., Barisic, S., Sawodny, O., et al. (2011). Modeling time delay in the NFκB signaling pathway following low dose IL-1 stimulation. *EURASIP Journal on Bioinformatics and Systems Biology*, **3**, 1–6.

Wolf, U. (1997). Identical mutations and phenotypic variation. *Human Genetics*, **100**, 305–21.

Wolf, U. (2002). Genotype and phenotype: genetic and epigenetic aspects. In A. Grunwald, M. Gutmann, and E.M. Neumann-Held, eds., *On Human Nature: Anthropological, Biological, and Philosophical Foundations*. Springer, New York, pp. 111–20.

Yeats, W.B. (1929). Among school children. Accessed at http://poetry.about.com/od/poems/l/blyeatsamongchildren.htm