# THE PHENOGENETIC LOGIC OF LIFE

*Kenneth M. Weiss*

Abstract | For nearly a century we have understood that life works through genes, and so have had an elegant theory for general evolution. But this did not explain the kinds of traits that characterize organisms, nor how genes produce them. Advances in recent decades have opened the way for an understanding of the phenogenetic logic or relational principles of life, by which a few basic characteristics of genomes produce biological phenotypes through some basic developmental processes. This logic provides a general explanation of the nature and source of organismal design, and a powerful programme for research.

We owe it to Charles Darwin and Alfred Wallace for providing us with a powerful theory to explain how life reflects a history of divergence from common ancestry, through descent with modification, driven by change in gene structure, reproductive isolation, functional adaptation and chance[1]. Darwin had inadequate information to explain how phenotypic information was inherited, and it was Mendel who suggested the answer to that question. In the 1930s, through the 'modern evolutionary synthesis', population genetics provided a formal theoretical foundation for Darwinian evolution, defined as changes in the frequency of inherited variation[2,3]. Population genetics was developed as a field before the nature of inherited variation was known, but the discovery of DNA as a sequence-based repository of evolutionary memory and a code for amino-acid sequences was consistent with the theory of evolution as a process[4].

However, this theory of history tells us little about what traits will evolve and how genes produce them. Although development was important to the formulation of Darwin's evolutionary theory, it became severed from evolutionary biology through much of the twentieth century[5]. During that time, exquisite experimentation showed empirically how complex morphology emerges even from a single cell (a fertilized egg), and there were also comparable advances in the study of 'virtual' (physiological, non-physical) traits (for examples, see REFS 6,7). However, a more complete evolutionary synthesis, often classified under the catch-phrase the 'evolution of development' (EvoDevo), has been emerging, facilitated by advances in molecular genetics (BOX 1) that have revealed elements of a unifying phenogenetic logic of life — the phenomena that connect biological phenotypes with their underlying genetic bases[8–13]. 'Logic' is the operative concept, because unlike the stereotype according to which genes are independent, bead-like functional units that are linearly arranged along a chromosome, phenogenetic phenomena are the higher-order, 'emergent' results of structure and interaction.

Phenogenetic logic can be encapsulated by the general principle of duplication with variation, driven by change in gene expression, component sequestration, functional divergence and chance. The symmetry between phenogenetic logic and Darwin's principles is not accidental (TABLE 1). If evolution is the history of species, phenogenetics is history within organisms. Species evolution and embryological development are both nested phenomena of differentiation between related units — organisms in species evolution, and cells or gene products in development. A phenogenetic synthesis is possible because the nature of the DNA sequence explains both the memory and diversity of species evolution, and the memory and diversity of biological functions. But if there are logical similarities between individuals in species and between cells within organisms, there are also important differences.

How can the great diversity of biological traits on Earth — so complex in space and time — be produced by a linear sequence of nucleotides that constitutes the genomes of living organisms (FIG. 1)? If evolution is a contingent historical process, can there be generalizations that describe these processes beyond enumerating

*Department of Anthropology, 409 Carpenter Building, University Park, Penn State University, Pennsylvania 16802, USA.*
*e-mail: kenweiss@psu.edu*
doi:10.1038/nrg1502

Box 1 | **Recent technical advances for phenogenetic study**

**Technical advances have made it possible to identify many aspects of phenogenetics. In essence, these methods have revealed the nature, use and mechanisms of differential gene expression in cells, complex organisms and systems. A few of the principal tools now in use include:**

- **Whole-genome sequences for several model species and their online availability.**
- ***In situ* tissue hybridization with gene-specific antisense RNA probes to document cell or tissue-specific gene expression.**
- **Isolation and characterization of mRNA from specific cells by diverse methods both in series (ESTs,** SERIAL ANALYSIS OF GENE EXPRESSION **(SAGE),** DIFFERENTIAL DISPLAY**) and by using parallel assays (array technology).**
- **Use of genome-scale expression technology to identify correlated gene expression under experimentally modifiable or naturally occurring variable conditions, as well as to determine quantitative expression levels.**
- **Genome-scale surveys of protein–protein interactions.**
- **Transgenic reporters with context-specific expression drivers to document *in vivo* gene expression.**
- **Modification of gene expression in experimental transgenic animals and plants by homologous recombination or cell transfection (for example, using viral expression regulators and RNA inhibitors).**
- **Molecular assays for protein–DNA binding to identify regulatory proteins and their target sequences (for example,** GEL-SHIFT **and** YEAST ONE-HYBRID ASSAYS**).**
- **Experimental documentation and confirmation of *cis*-regulation of gene expression.**
- **Informatics as a means of identifying genome organization, gene families, regulatory elements, transposable elements and other modular or repetitive aspects of genomes.**

them case by case? Are there fundamental principles that apply similarly to plants, animals and even unicellular organisms?

The objective of this article is to show that we can identify such principles and generalizations. I first itemize a few basic ways in which genomes are organized that are fundamental to the way genetic information is used. Then, I identify a similarly modest list of developmental principles of the use of that genetic information in the construction of the physical and physiological traits of organisms. From principles such as these (BOX 2), various facts that span the diversity of species fall into place, adding to our ability to predict the nature of things still to be found, accounting for their evolution, and providing powerful tools for research. Phenogenetic logic helps to supplement classical evolutionary principles to add to a more comprehensive theoretical understanding of life.

## General characteristics of genomes

There are probably many ways to describe the main aspects of genomes that are relevant to phenogenetic logic. The following categorization is an attempt to do that.

*Modularity.* Life is modular. This is true at all levels of organization, ranging from genes to cells, tissues, organs, individuals, species and even ecosystems. Modular structure in genomes in turn facilitates the modular nature of traits at higher levels. Chromosomes are packed with modular functional units. We think immediately of the order of codons (nucleotide triplets), which specifies the corresponding order of amino acids in proteins. But further modularity in DNA is just as important and includes the exon segments of eukaryotic genes (which might themselves contain internal-repeat motifs), sequence duplicates in centromeres and telomeres, dispersed

SERIAL ANALYSIS OF GENE EXPRESSION
A method for quantitative and simultaneous analysis of a large number of transcripts; short sequence tags from the mRNAs that are produced in a cell are isolated, concentrated and cloned; their sequencing reveals a gene-expression pattern that is characteristic of the tissue or cell type from which the tags were isolated.

DIFFERENTIAL DISPLAY
A technique for detecting those genes that are expressed only under specific conditions; it involves isolation and comparison of mRNA from two or more populations of cells.

GEL-SHIFT ASSAY
An electrophoretic gel-based assay in which proteins that bind to a DNA fragment are detected by virtue of the reduced migration of the DNA. The assay is often used to detect transcription-factor binding.

YEAST ONE-HYBRID ASSAY
An assay that identifies DNA-binding proteins from cDNA libraries or known gene sequences.

Table 1 | **Comparative logic and symmetries between evolution and phenogenetics**

| Characteristic | Evolution | Phenogenetics |
|---|---|---|
| Overall principle | Descent with modification | Duplication with variation |
| Explains | Life through the history of organisms | Life through form and function, a somatic history within organisms |
| Source of variation | Gametic mutation | Differential gene expression |
| Means of sequestration | Mating barriers and speciation through mechanisms of genomic and ecological isolation | Partial sequestration with information transfer through signalling using arbitrary codes |
| Means of divergence | Adaptation by competitive response to external conditions | Functional specialization through programmed response to signalling and other internal conditions |
| Role of chance | Genetic drift | Molecular stochasticity in timing and concentration of components, somatic mutation and epigenetic modification (for example, methylation) |


Figure 1 | **A phenogenetic metaphor.** Linear DNA sequences produce the profusion of complex traits of living organisms through a plethora of branching and duplication processes, the spatial and temporal natures of which can be captured metaphorically by the harmonies of music. **a** | A representative DNA sequence. **b** | A profusion of the results of duplication with variation. Image courtesy of K.M.W.

transposable elements, splicing and other RNA-processing signals, microsatellites and other short tandem repeats, and binding sites for histones that package chromosomes, as well as those for transcription factors that regulate gene expression[10,11,14,15].

*Duplication.* If genomes can be characterized in a single phrase, it is as the product of billions of years of duplication events, and from early in its history, life has depended on these events[16] (FIG. 2a) — an insight gained early on by Susumu Ohno[17]. For example, gene families arise through replication errors that produce tandem arrays of related genes that might subsequently be transposed around the genome — individually or as clusters that can be expanded by further tandem duplication or contracted by gene deletion. Duplication generates functionally redundant elements that are the working material for the evolution of new functions[18–20]. Today, we automatically assume that a 'novel' (previously unknown) gene is not novel at all, but belongs to a family, which we expend great effort — usually successfully — to find. There is no spontaneous generation: even an orphan gene must have parents.

*Arrangement.* Duplication produces modular genomic units, the functions of which depend on their chromosomal arrangement; another basic aspect of phenogenetic logic. The correspondence between the arrangement of nucleotides as codons and amino acids is well known, but arrangement is vital in many other ways. Complex life has evolved only because cells in multicellular organisms — and even in cellular aggregates such as bacterial BIOFILMS and slime-mould colonies — have different functions. Differentiated cells, organs and tissues occur because a given cell typically expresses only a subset of the genes in its genome. Selective gene expression is largely controlled by *cis*-regulation, the binding of regulatory proteins to regulatory sequence elements (for example, enhancers, promoters, repressors and insulators) in the DNA near the regulated gene[9,10]. Expression might be affected by the local state of DNA, such as by unwinding or epigenetic modification by histone acetylation or CpG methylation, that affects the ability of transcription factors to reach their binding sites, or even by permanent, local sequence modification, as in the case of somatic recombination in mammalian antibody genes[15].

Genome-wide expression profiling in diverse species has revealed the high level of correlated expression of chromosomally clustered genes, which might consist of related or unrelated genes. Clustered genes are often under the coordinated control of one or more local regulatory element that directs the expression of the genes jointly or sequentially during development, or in response to changed circumstances. Arrangement-related expression of clustered genes is found in traits as diverse as anatomical development (for example, Hox genes in axial patterning, cadherins in neuronal migration)[9,21]; physiology (for example, globin genes in the case of changing oxygen binding requirements and genes that encode calcium-binding proteins in vertebrate mineralized tissue)[15,22,23]; immunity (such as genes that

BIOFILMS
Microbial biofilms are populations of microorganisms that are concentrated at an interface (usually solid–liquid) and are typically surrounded by an extracellular polymeric substance (EPS) matrix. Aggregates of cells that are not attached to a surface are sometimes termed 'flocs' and have many of the same characteristics as biofilms.

Box 2 | **Phenogenetic principles**

**The term 'phenogenetics' refers to the relationship between genes and the traits that they cooperate to produce. The following is a list of basic genomic characteristics related to phenogenetic phenomena, followed by a few basic processes that result from or depend on those genomic characteristics. The list is provided as a general guide, and reflects the specific perspective of the author.**

**Genome characteristics**

- **Genome organization is modular.**
- **Genomes are the products of a history of billions of years of diverse duplication events.**
- **The arrangement of genomic elements facilitates differentiation by selective gene expression.**
- **Cells acquire differentiated states as a result of information feedback to the genome.**
- **Complex systems are characterized by the partial sequestration of components.**
- **Function arises largely as the result of arbitrary coding.**
- **Functional differentiation is based on combinatorial gene expression.**

**Phenogenetic processes**

- **The generation of spatiotemporal asymmetries and polarity (such as body axes and precursors of fate maps).**
- **Branching and pouching (invagination and evagination).**
- **Segmentation, with partially sequestered units.**
- **Regional differentiation by dynamic inductive signalling.**
- **Repetitive patterning by quantitative interactions between antagonistic factors.**
- **Traits that emerge largely through nested epigenetic processes, rather than being specified as 'entities'.**

encode antibodies in animals and plant R-GENES)[8,11,24]; and sensory systems (such as genes required for the expression of red or green-sensitive photoreceptors, and olfactory receptors)[6,10,11,25–27]. However, co-expressed genes can be scattered across the genome so long as each has appropriately located regulatory elements. Similarly, although gene expression is mainly controlled by *cis*-regulation, the regulatory elements can flank their regulated gene on either side, or even be located within the gene. This shows again that what is fundamental is not a specific modular-arrangement rule, but the logical fact of arrangement.

***Phenogenetic information feeds back onto genomes.*** Selective gene expression reflects a profound difference between evolution and development. The central dogma of evolutionary theory is that DNA is a one-way, feedforward repository of information that evolves by differential reproduction (fitness) among individuals that function as competing free agents, whose variation arises randomly with respect to function — and not by Lamarckian inheritance of characteristics that are acquired from life experience. However, developmental change within an organism depends entirely on feedback from a cell's experience to its genome[10]. Cells are not free agents, but are actively induced by the coordinated nature of organisms to differentiate cooperatively to form tissues or respond to changing circumstances. This is what developmental 'fate-maps' are; lineages of cells differentiate in response to external signals or conditions, making commitments that are inherited by their mitotic descendant cells until some new contextual information induces a change.

Such somatic commitment involves gene-expression differences that can include direct modification of DNA; for example, by methylation of promoter regions in selected genes that affects their expression. Differentiation is the (somatic) inheritance of states that are acquired by a cell's experience. It is important to stress that, although this does not imply the 'striving' component that is usually associated with the idea of Lamarckian inheritance, somatic variation is fundamentally different from variation between individuals and arises through natural selection that acts on randomly occurring variation.

This might be seen as stretching the point, because somatic changes do not usually change the genome permanently in the same the way that germ-line mutations do. But that argument holds only if we restrict the term 'inheritance' to apply to DNA-nucleotide sequences, which is the traditional application, but is biologically unwarranted. In fact, DNA-changing somatic mutation of all types does occur (including random changes, as well as enzymatically driven rearrangements in immunoglobulin genes and even the self enucleation of red blood cells). Even germline mutations are only 'permanent' until some future mutation occurs in the gene or the same site is 'hit' again. This might be slower than somatic change (although many somatic changes are never reversed throughout the life-time of an organism, which can involve many cell divisions), but it is not qualitatively different. The key distinction is between variation that is acquired by experience, as occurs in the soma, and variation that is acquired by random change, which drives the evolution of organisms.

***Sequestration.*** Another characteristic of the organization of the genome that is related to differentiation and that also applies at higher levels, is that modular organization indicates a degree of autonomy or sequestration

R-GENES
Genes that are used by plants for immunological response to pathogens.

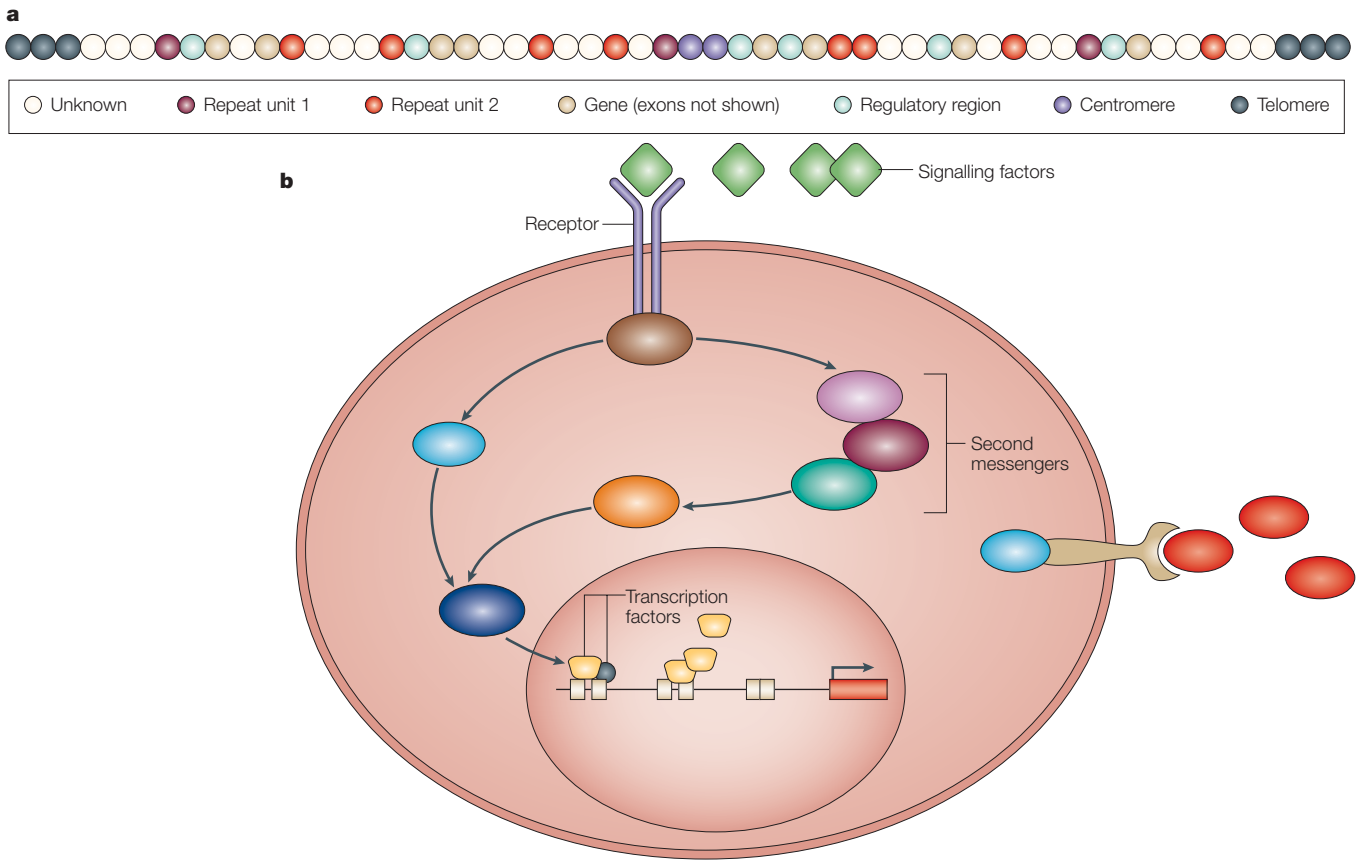


Figure 2 | **Basic principles of phenogenetic logic. a** | Modularity is ubiquitous in the genome. A schematic representation of a region of a chromosome illustrates the kinds of modular units that are typically found. **b** | Most gene expression involves arbitrary combinatorial coding. Labels identify proteins that are used in signalling. The signalling code specifies expression of the final gene (red rectangle within the nucleus).

among units. Differentiation is possible because the modular elements of genomes are functionally and physically sequestered and do not blend, as Darwin would have been happy to learn. But modular organization extends beyond genomes, because physical traits and physiological systems are also typically built of sequestered functional components. In a sense that is both trivial and profound, once cells evolved, 'all' that remained to make complex organisms was to construct them as aggregates of differentiated cells. The primaeval enabling step for organisms to arise was the evolution of the lipid membrane to make cells, and even components within cells, partly autonomous. This is because the environment within the cell was sequestered from the external environment and could modify its own internal conditions. Differentiation into tissues, organs, individuals and even species followed as a direct result.

However, sequestration is usually not complete. The parts must interact or there would be no organism. Partial sequestration is actively regulated. Cell membranes are littered with GAP JUNCTIONS, receptors and ion channels that are used to communicate with other cells, and to prevent complete isolation from them. Regulated partial sequestration is pervasive and occurs at all stages of life, including communication between cells (through growth factors and hormones), individuals (by pheromone signalling and territory marking) and even between species (for example, predator–prey relationships and pollination attractants). Circulatory and other dispersion systems aid or facilitate communication. A substantial fraction of genes (see the Gene Ontology Consortium web site) have functions of this sort that are required many steps before the ultimate physical and behavioural states are achieved — such as an organism's running speed or feathers — which have been the main concern of evolutionary biology since Darwin's time. Life has been designed around communication between partially sequestered units.

***Codes.*** The construction of biological traits is largely determined by the use of codes. A code is logically necessary for a developmental process to occur, but is functionally arbitrary in the sense that the physical nature of the elements of a code are not related to the physical nature of the result[11]. AGA is a nucleotide-sequence code that specifies the amino acid arginine, but AGA is unrelated to the properties of arginine in the millions of proteins in which it is found.

The protein code is only one aspect of biological coding. Similarly arbitrary is the use of enhancer, promoter, insulator and repressor sequences as codes that determine transcription-factor binding for the

GAP JUNCTION
A junction between two adjacent cells that consists of pores that allow the passage of molecules (up to 1 kD).

control of gene expression. The elements of signalling networks that comprise so much of the genome provide another example (FIG. 2b). So, a set of genes that includes the transcription factor PAX6 interact to code for different aspects of photoreception in most, if not all animals; however, the resulting types of eye are exceedingly diverse. In fact, even the protein-coding system itself requires an additional code to ensure that tRNA molecules with anticodons that match the codons in mRNA carry the correct amino acid[14].

*Combinatorial gene expression.* Modules at one level generally do not correspond directly with those at another in the way that codons do for amino acids[28]. There is not a one-to-one correspondence between individual units, such as leaves or toes, and the underlying genes. Instead, phenotypic modularity is achieved indirectly, through another manifestation of phenogenetic codes that is based on the combinatorial use of genes. Organisms possess a limited tool-kit that contains a modest number of regulatory factors, which are combined in different ways in different contexts. For example, most transcription and signalling factors are used several times in context-specific combinations within the same organism, or even at different times in the same structure. Physiological signalling systems and other regulatory networks are typically hierarchical, with nested cascades from initial to downstream points[29]. Evolutionary flexibility is achieved through the ability to apply and reapply existing code elements to new problems or to modify existing traits[30–32]. Combinatorial expression is another feature of phenogenetic logic: it is the combination — not the individual components — that specifies the function.

## Phenogenetic processes

The distinction between the characteristics of genomes that are of phenogenetic importance and the developmental processes in which they are used is arbitrary. The following attempt to categorize the principal generalizations that have wide applicability in the organismal world might contain omissions. But those included seem relevant and belong in such a list.

*Asymmetry.* The transmission of a cell's gene-expression state to its mitotic daughter cells until that state is subsequently modified produces the nested, hierarchical, cellular fate maps of which tissue and organ systems are built. This is achieved by a modest number of developmental processes, the genetic bases of which are becoming known. Early in embryonic development, such expression changes establish primary-axial

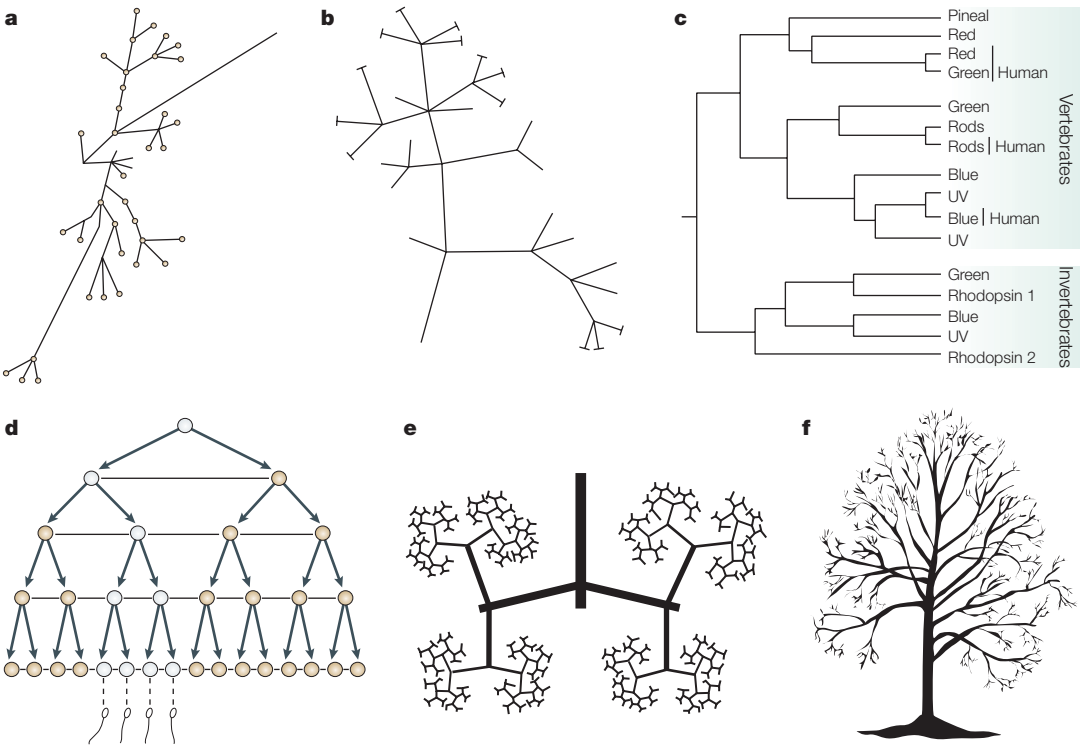


Figure 3 | **Ubiquity of branching structures in living organisms.** Multiple applications of branching logic show the logical symmetry of evolution among organisms and development within them. **a** | Charles Darwin's attempt to reconstruct Ernst von Baer's idea that embryos of contemporary similar species, such as vertebrates, have diverged from a common early-embryonic form. **b** | Darwin's sketch of his idea of divergence of species from a common ancestor. **c** | Divergence of the sequence of a single gene, or members of a gene family, from a common ancestor (the example of photoreceptors and olfactory receptors is shown). **d** | The divergence of tissues from a single cell within an organism. **e** | Schematic representation of the fractal-like mammalian bronchial (lung) tree. **f** | The source of the metaphor — real branches. Part **d** is modified from REF. 53 ; **a**–**c** and **e** are reproduced from REF. 11 © (2004) Wiley-Liss.

asymmetry, or polarity — usually in several dimensions[8–10,33] along which localized asymmetries subsequently develop. For example, the anterior–posterior and proximo-distal axes that form within limbs and teeth[6,34,35], or the production of leaves and flowers from shoot MERISTEMS[36].

*Branching.* One way to generate asymmetry is through branching — probably the most powerful single metaphor of life that captures the central ideas of both evolution and development — from phylogenetic trees to real ones[5,8,37] (FIG. 3). Branching is found both on the exterior of organisms (for example, animal digits, plant roots and branches) and inside them (for example, vessels, nerves and lungs). Multicellular organisms begin their lives organized as balls or tubes that become more complex through branching to produce hierarchically differentiated organs and structures. Branching is achieved through combinations of differential growth with accompanying changes in internal cytoarchitecture, loss of cohesion between cells and/or apoptosis. Variations include pouching into or out of an overall tubular plan, such as in arthropod TRACHEOLES, animal limbs, mammalian tooth cusps, milk ducts (which also involve internal branching) and tentacles. Branching might simply produce repeated units, as in lungs, blood vessels or plant roots, but the units might then be hierarchically differentiated, as in digestive organs that branch from the developing gut or functionally different regions in brain development.

*Segmentation.* A developmental characteristic that is largely produced by partial sequestration is segmentation. For centuries, biologists have pointed out the importance of segmentation in specific traits, but the pervasiveness of modular structures charaterized by segmentation into subunits — which are often themselves further and hierarchically differentiated — has not been fully appreciated or incorporated as a standard part of the literature on the general principles of organismal organization. The operational power of this fact has been seen in the discovery, predicted on the basis of the above developmental genetic considerations, that vertebrate brains are segmented in ways that had long been argued not to be present, and — in other instances where morphology did not clearly reveal functional boundaries — that are discovered by boundaries of gene-expression patterns, a kind of genetic, rather than physical septum. Segmentation is a characteristic of most physical and virtual traits in complex organisms[11].

*Dynamic inductive patterning.* Segmentation is often achieved through arbitrary coding by combinatorial gene expression in dynamic inductive patterning[38,39]. Inductive patterning works by information passing from one set of cells that induces gene-expression changes to other cells. This kind of communication works in various ways, including: gap junctions in adjacent cells; autoregulation (whereby cells receive their own signal); cell-surface proteins and receptors in neighbouring cells; locally diffusing signalling factors; internal, often long-distance, circulation of hormones and growth factors within organisms; and at longer distances, by pheromone signalling between organisms[11].

Signalling factors that are released into neighbouring cells or the extracellular space diffuse across a bed of similar cells, sometimes establishing a concentration gradient that decreases with distance from a local source. Cells detect the presence of these factors by expressing appropriate binding proteins or membrane-bound receptors, as well as second-messenger proteins to initiate response cascades that ultimately activate signal-specific transcription factors. The information is internalized when cells differentiate by changing gene expression along a 'morphogenetic field'; for example, at points where the signal concentration exceeds thresholds that the recipient cells can detect. They do this by various means, such as by the binding of signal molecules by enough copies of the cell's receptors. The process is dynamic because the result depends on the relative quantitative interactions between the factors, involving the production and diffusion rates of signalling factors, relative quantitative concentrations of activating and opposing inhibiting factors, the concentration of receptors on the surfaces of recipient cells, and the kinetics of receptor binding, second-messenger cascades and transcription activation.

*Repetitive patterning.* A particularly common kind of inductive patterning is repetitive patterning that produces traits with several similar units. Serial homology is the presence of multiple similar structures that seem to be developmental copies of the same, presumably ancestral, unit. This type of homology is pervasive throughout living organisms, as shown by traits as diverse as kidney nephrons, intestinal villi, papillae on tongues, fin rays, linear body segments in vertebrates, arthropods and annelids, and radial axes in echinoderms, leaves, flowers and hair. As long ago as 1892, William Bateson likened repetitive organ systems to dynamic interference patterns in physical wave-generating phenomena[40]. He used the analogy of Chladni figures by which violin-makers tune top and bottom plates by covering them with fine powder, applying specified tones through tuning forks to jostle the granules away from vibrating parts to quiescent nodes in the wood, and selectively shaving the wood until the desired resonance pattern is achieved (FIG. 4A).

Chladni figures are a good conceptual metaphor for periodic patterning in life, where the analogous interference patterns are due to the relative concentrations of diffusing activator and inhibitor substances: structural elements are induced to develop in areas that are dominated by the activator, and are surrounded by inhibition zones, as for example, in the wave-like molar-cusp pattern that is shown in FIG. 4B. Repetitive patterning occurs throughout embryogenesis, a silent symphony of life that simultaneously and harmoniously generates complex structures such as OMMATIDIA in insect

MERISTEM
In plants, this is a zone (for example, the apex of the shoot) that contains undifferentiated cells that continue to divide, providing cells for further growth and differentiation.

TRACHEOLES
Fine terminal branches of respiratory tubes.

OMMATIDIA
The elements of the compound eye of insects (in *Drosophila melanogaster*, the eye is formed from 800 ommatidia), each of which is an independent visual unit that contains eight photoreceptor cells, surrounded by four cone cells that secrete the lens, and seven pigment cells.

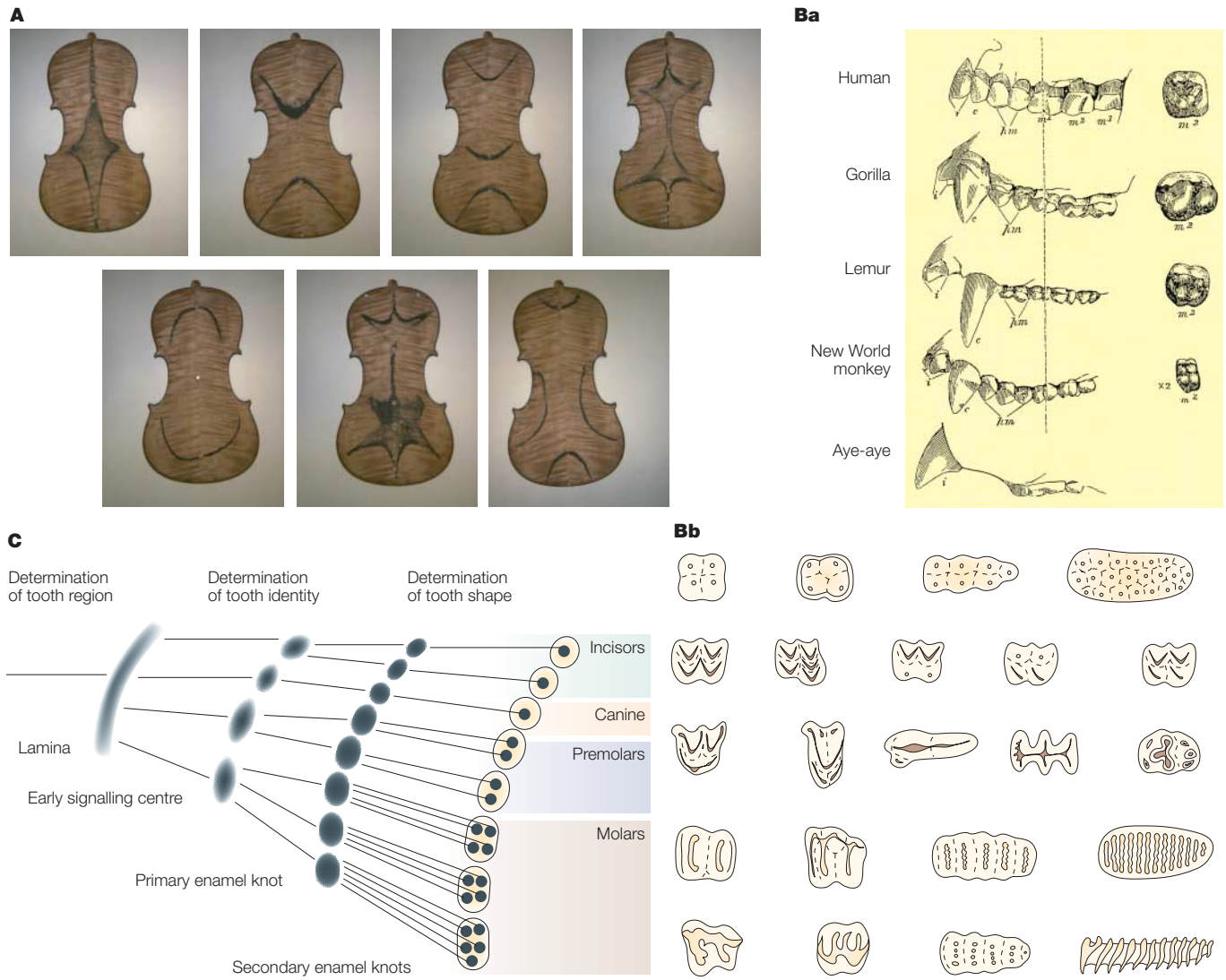


Figure 4 | **Dynamic patterning: complexity made simply. A** | Chladni figures as a metaphor for the harmony of patterned life. If different tone frequencies are applied to the same violin plate, each generates a different interference pattern (fine powder settles into quiescent nodes in which vibrations cancel out). From upper left to lower right, response patterns to 138, 231, 291, 306, 312, 392 and 1196 Hz. Reprinted, with permission, from E. Bossy and R. Carpenter, University of New South Wales. **B** | Many phenogenetic principles are illustrated by the development of the mammalian dental pattern. Mammals have rows of teeth that are symmetrically differentiated in left, right, upper and lower jaw quadrants. **Ba** | Primate dentitions. Reproduced from REF. 54. **Bb** | Mammalian molar-cusp surfaces provide compelling evidence of generation by a quantitative 'wave-generating' patterning process. Reproduced, with permission, from REF. 55 © (1989) Walter de Gruyter. **C** | Patterning processes can be nested. Mammalian teeth are generated by a hierarchical repetitive-patterning process that involves episodes of similar combinatorial gene expression, sequentially invoked (in areas schematically represented by dark shading) at different developmental stages. Modified, with permission, from REF. 34 © (2000) Elsevier Science Ltd.

eyes, hair, scales, coloration patterns, teeth, internal units within individual organs and, possibly, butterfly colour patterns (for examples, see REFS 8,33,40–44).

Repetitive patterning can occur in many ways. Similar structures can be continually generated from a source that is similar to a Roman-candle (for example, continual reptile tooth replacement, intestinal villus replacement or the annual growth of leaf-bearing plant meristems), is nested (feather or leaf structure), or from sources that are branched (for example, nephrons, pancreatic islets and milk glands), multidimensional (tetrapod limbs or pelage patterns in mammals), more-or-less simultaneously patterned (stripes of expression of the transcription factor Even-skipped in insect eggs or axial segments in Hox-regulated structures), scattered (gene-family arrangements in genomes) or cycled by intracellular signal-oscillators (vertebrate somites), and can also be produced in other ways[5,8,9,11]. Repetitive patterning can also occur along a line of cells that then divide, moving a row of descendant cells that remember their pattern-state away from the source, much as one unrolls a window shade, while the row of source cells

continue to generate and shed rows of newly patterned cells (examples include PROGRESS ZONES in limbs and perhaps teeth, and the colour patterning of seashells or of the bodies of fish)[6,34,35,45,46]. Gradient-threshold and repetitive patterning can work together, as for example, with concentrations of the short-range signalling factor Shh (Sonic hedgehog) that initially affect axial polarity in the early buds of tetrapod limbs, and then later affect the digits that form at the ends of the new limbs[6].

Similar dynamic interactions also affect virtual systems, in which circulating reagents trigger differential concentration-dependent responses. Mammalian fertility is one example; it is controlled in part by concentration-dependent signalling by various circulating hormones that are produced in and secreted from distant organs. The relative distribution of circulating lipids — such as cholesterol in LIPOPROTEINS of different density (for example, high-density and low-density lipoproteins), as the lipids are targeted for deposition in ADIPOCYTES, for metabolic use or for clearance — is part of another virtual system that is affected by concentration-dependence. Signal-based regulation of calcium for bone growth, calcium storage or physiological use might be another example.

Dynamic patterning is often layered, nested and regionally differentiated, as seen in the repetition of vertebrae that are similar within, but different between, the cervical, thoracic and lumbar regions; in the head, thorax and abdominal segments of insects; and in sepals and petals in flowers. Nesting is almost inevitable for at least two fundamental reasons. An early embryo must establish its primary axes before structures can be located regionally. One system induces another. Therefore, teeth cannot form before there is a jaw, which requires a head end to be established, and only then is jaw tissue produced that is responsive to the appropriate signals (FIG. 4C). Moreover, diffusing signals have no effect if the cells that they come into contact with have not already been induced to express the appropriate combination of receptors or binding proteins needed to detect the signal.

***Development as an epigenetic phenomenon.*** Dynamic patterning is a form of arbitrary coding by combinatorial gene-use that clearly illustrates what genes are not for. There is no gene specific to each hair, tooth or cusp. Instead, the same or similar gene-expression cascades generate each element. Much of the action is extra-cellular, and in a sense is an epiphenomenon relative to individual participating genes. Because the same genes are used in several ways in the same organism (in entirely different traits, or even in the same trait at different stages in its development), the genes participate in the logic — the process — but they do not individually correspond to specific elements or body parts.

Evolution works to a great extent by changing the logic, or by the use of higher-level mechanisms, to generate new or modified structures. This has important implications. Similar to shaving a tiny bit of wood from a violin plate, a simple change in the parameters of a quantitative patterning system, or a change in the elements of a Boolean (on–off) code, can have notable effects on form (for example, simple versus compound leaves, important differences between wild and domesticated maize, changes in the number of vertebrae or digits, or different cusp patterns in teeth of closely related species or in the upper and lower jaws of the same individual). The repetitive use of the same genes also has direct implications for our understanding of the fundamental concept of evolutionary homology[47–50]. We can count or identify structures (such as teeth, digits, ommatidia or bristles) but what is homologous is the process that makes them.

However, there are enumerative one gene–one function systems even for complex traits, as in R-genes for plant immunity or early vertebrate immunoglobulins, and the thousands of odorant-receptor genes in vertebrates and invertebrates. Some such systems have evolved into more logically compact systems; for example, in the use of somatic recombination to generate mammalian antibody diversity and the currently unknown means by which each olfactory-neuron cell expresses only one of its thousands of odorant-receptor genes[25].

## Conclusion

A simple logic with wide generality supplements the classical evolutionary principles to account for the diversity of complex traits that evolution has produced. I have tried to synthesize at least part of what is currently known into a cogent working framework. The overall principle, 'duplication with variation', is about as widely applicable as a general guide as Darwin's evolutionary principle, 'descent with modification', because development and evolution are different faces of the same phenomena, relating to what happens among germlines across generations of organisms and within the germline across generations of cells, respectively.

There is hardly an area of inferential or experimental biology in which these few simple principles are not in use, at least informally. But neither phenogenetic nor evolutionary principles are 'laws of nature' in the classical sense, and both have exceptions and specifics that are not usually predictable, except in general terms. However, the exception often proves the rule. To take one example, similar enhancer sequences are 'duplicated' and used across the genome to control genes that are regulated by a given transcription factor. But unlike whole genes, these sequences are short enough — only a few base pairs — to be generated by point mutations, rather than requiring duplication by replication slippage followed by transposition to become — by luck — juxtaposed to a new gene to change that gene's expression contexts and give rise to a new trait[32]. If this kind of duplication-by-mutation was not possible, complex life as we know it would not have evolved.

An important issue has been neglected in this discussion, and that is the role of chance. Chance is an essential part of both evolution and development[11,51,52]. The importance of chance in these processes is probably greatly under-appreciated. If there is a difference between evolution and development, it might be that

PROGRESS ZONE
The progress-zone model proposes that positional information in a growing system can be specified by the time that the cells spend in a zone where growth and differential signalling occur; in a vertebrate limb, the progress zone is specified by the apical ectodermal ridge.

LIPOPROTEINS
Complexes of cholesterol, triglycerides and proteins that transport lipids in the aqueous blood stream environment.

ADIPOCYTES
Fat cells that are found in the adipose tissue.

regardless of its role in adaptation[11,51], chance is essential to at least the initial generation of new evolutionary variation by mutation, and chance is built into segregation and reproduction. Chance is also omnipresent in development, but as a rule is less essential. But, as always, there are exceptions. Without chance in the somatic recombination that affects antibody expression, we would not be here to think about any of these problems. But to do justice to chance is beyond the scope of this article.

With elegant simplicity and symmetry, history and function — evolution and development — are one. They are facilitated by genes through the same modular, nested branching logic of life that is based on interaction between partially sequestered units that are coordinated by arbitrary coding. Genomes contain the sequence information and memory that facilitates Darwinian evolution, and the somatic memory that makes phenotypes possible. The result is a more unified view of the use of genes and the generation of diversity at all levels of life.


1. Darwin, C. *The Origin of Species* (Murray, London, 1859).
2. Mayr, E. *The Growth of Biological Thought* (Harvard Univ. Press, Cambridge, Massachusetts, 1982).
3. Provine, W. *The Origins of Theoretical Population Genetics* (Univ. Chicago Press, Chicago, 1971).
4. Moore, J. A. *Science as a Way of Knowing: the Foundations of Modern Biology* (Harvard Univ. Press, Cambridge, Massachusetts, 1993).
5. Gilbert, S., Opitz, J. & Raff, R. Resynthesizing evolutionary and developmental biology. *Dev. Biol.* **173**, 357–372 (1996).
6. Gilbert, S. *Developmental Biology* (Sinauer, Sunderland, Massachusetts, 2003).
   **This classic is an excellent, up-to-date treatment of animal development, with insightful consideration of the historic origins of the field and the recent advances in the genetics, as well as the diversity of experiments that demonstrate key processes in the generation of complex form.**
7. Alberts, B. *Molecular Biology of the Cell* (Garland Science, New York, 2002).
8. Gerhart, J. & Kirschner, M. *Cells, Embryos, and Evolution: Toward a Cellular and Developmental Understanding of Phenotypic Variation and Evolutionary Adaptability* (Blackwell Science, Malden, Massachusetts, 1997).
9. Carroll, S., Grenier, J. & Weatherbee, S. *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design* (Blackwell Science, Malden, Massachusetts, 2004).
   **A visually compelling and conceptually clear statement of the role of *cis*-regulation in the nature and evolution of complex traits. This book reflects Carroll's exemplary experimental abilities to show the power of regulatory mechanisms.**
10. Davidson, E. H. *Genomic Regulatory Systems: Development and Evolution* (Academic Press, San Diego, California, 2001).
    **A detailed presentation of the complex regulatory networks that are involved in the production of complex traits, demonstrating in detail how regulatory elements work and evolve. A new edition is forthcoming.**
11. Weiss, K. M. & Buchanan, A. V. *Genetics and the Logic of Evolution* (Wiley-Liss, New York, 2004).
    **This is a systematic book-length treatment, with supporting details and references, of the synthesis that is attempted in the current article. The authors relate these principles to the underlying genes and their evolution, attempting to create an integrated view of evolution and development that includes a deeper consideration of the role of chance in evolution.**
12. Wilkins, A. S. *The Evolution of Developmental Pathways* (Sinauer, Sunderland, Massachusetts, 2002).
13. Raff, R. A. *The Shape of Life: Genes, Development, and the Evolution of Animal Form* (Univ. Chicago Press, Chicago, 1996).
14. Weiss, K. M. "The" genetic code. *Evol. Anthropol.* (in the press).
15. Lewin, B. *Genes VIII* (Prentice Hall, Upper Saddle River, New Jersey, 2004).
16. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
17. Ohno, S. *Evolution by Gene Duplication* (Springer, Berlin; New York, 1970).
18. Cooke, J., Nowak, M. A., Boerlijst, M. & Maynard-Smith, J. Evolutionary origins and maintenance of redundant gene expression during metazoan development. *Trends Genet.* **13**, 360–364 (1997).
19. Ohta, T. Mechanisms of molecular evolution. *Phil. Trans. R. Soc. Lond. B* **355**, 1623–1626 (2000).
20. Ohta, T. Evolution by gene duplication revisited: differentiation of regulatory elements versus proteins. *Genetica* **118**, 209–216 (2003).
21. Redies, C. & Puelles, L. Modularity in vertebrate brain development and evolution. *Bioessays* **23**, 1100–1111 (2001).
22. Hardison, R. Hemoglobins from bacteria to man: evolution of different patterns of gene expression. *J. Exp. Biol.* **201**, 1099–1117 (1998).
23. Kawasaki, K., Suzuki, T. & Weiss, K. M. Genetic basis for the evolution of vertebrate mineralized tissue. *Proc. Natl Acad. Sci. USA* **101**, 11356–11361 (2004).
24. Roitt, I. M., Brostoff, J. & Male, D. K. *Immunology* 5th edn (Mosby, London; Philadelphia, 1998).
25. Mombaerts, P. Odorant receptor gene choice in olfactory sensory neurons: the one receptor–one neuron hypothesis revisited. *Curr. Opin. Neurobiol.* **14**, 31–36 (2004).
26. Kratz, E., Dugas, J. C. & Ngai, J. Odorant receptor gene regulation: implications from genomic organization. *Trends Genet.* **18**, 29–34 (2002).
27. Firestein, S. How the olfactory system makes sense of scents. *Nature* **413**, 211–218 (2001).
28. Wagner, G. P. & Schlosser, G. (eds) *Modularity in Development and Evolution* (Univ. Chicago Press, Chicago, 2003).
29. Barabasi, A. L. & Oltvai, Z. N. Network biology: understanding the cell's functional organization. *Nature Rev. Genet.* **5**, 101–113 (2004).
30. Rockman, M. V. & Wray, G. A. Abundant raw material for *cis*-regulatory evolution in humans. *Mol. Biol. Evol.* **19**, 1991–2004 (2002).
31. Wray, G. A. Transcriptional regulation and the evolution of development. *Int. J. Dev. Biol.* **47**, 675–684 (2003).
32. Wray, G. A. *et al.* The evolution of transcriptional regulation in eukaryotes. *Mol. Biol. Evol.* **20**, 1377–1419 (2003).
    **A rigorous survey of the evolution of short regulatory sequences, tying together classical molecular and phenotypic evolution, that notes the similarities and differences in the way that Darwinian evolution affects regulatory and protein-coding regions.**
33. Meinhardt, H. Organizer and axes formation as a self-organizing process. *Int. J. Dev. Biol.* **45**, 177–188 (2001).
34. Jernvall, J. & Thesleff, I. Reiterative signaling and patterning during mammalian tooth morphogenesis. *Mech. Dev.* **92**, 19–29 (2000).
35. Jernvall, J. & Jung, H. S. Genotype, phenotype, and developmental biology of molar tooth characters. *Am. J. Phys. Anthropol.* **31**, S171–S190 (2000).
36. Ng, M. & Yanofsky, M. F. Three ways to learn the ABCs. *Curr. Opin. Plant Biol.* **3**, 47–52 (2000).
37. Hogan, B. L. Morphogenesis. *Cell* **96**, 225–233 (1999).
38. Wolpert, L. One hundred years of positional information. *Trends Genet.* **12**, 359–364 (1996).
39. Wolpert, L. *et al. Principles of Development* (Oxford Univ. Press, London, 1998).
40. Weiss, K. M. Good vibrations: the silent symphony of life. *Evol. Anthropol.* **11**, 176–182 (2002).
41. Meinhardt, H. Models of biological pattern formation: common mechanism in plant and animal development. *Int. J. Dev. Biol.* **40**, 123–134 (1996).
42. Weiss, K. M., Stock, D. W. & Zhao, Z. Dynamic interactions and the evolutionary genetics of dental patterning. *Crit. Rev. Oral Biol. Med.* **9**, 369–398 (1998).
43. Murray, J. *Mathematical Biology* (Springer, Berlin, 2002).
    **This is an excellent formal treatment of the mathematical structure of repetitive-pattering (and other) mechanisms, many of which have subsequently been confirmed in broad outline in various organisms.**
44. Nijhout, H. F. *The Development and Evolution of Butterfly Wing Patterns* (Smithsonian Institution Press, Washington, 1991).
45. Meinhardt, H. *The Algorithmic Beauty of Sea Shells* (Springer, Heidelberg, 2003).
    **A beautiful, metaphorical, but also formalized demonstration of the phenotypic flexibility that can be achieved by simple patterning processes. The book includes a working programme so that readers can generate these kinds of patterns themselves.**
46. Kondo, S. & Asai, R. A reaction-diffusion wave on the marine angelfish *Pomocanthus*. *Nature* **376**, 765–768 (1995).
47. Larsson, H. C. & Wagner, G. P. Pentadactyl ground state of the avian wing. *J. Exp. Zool.* **294**, 146–151 (2002).
48. Wagner, G. P. A research programme for testing the biological homology concept. *Novartis Found. Symp.* **222**, 125–134; discussion 134–140 (1999).
49. Wagner, G. P. & Gauthier, J. A. 1,2,3 = 2,3,4: a solution to the problem of the homology of the digits in the avian hand. *Proc. Natl Acad. Sci. USA* **96**, 5111–5116 (1999).
50. Wray, G. A. & Abouheif, E. When is homology not homology? *Curr. Opin. Genet. Dev.* **8**, 675–680 (1998).
51. Weiss, K. M. The smallest grain in the balance. *Evol. Anthropol.* **13**, 122–126 (2004).
    **This and the other *Evolutionary Anthropology* citations are installments of a regular column that examines conceptual aspects of evolution and genetics. This specific reference deals with Darwinian determinism and alternative views that are based on population-genetic considerations.**
52. Carroll, S. Chance and necessity: the evolution of morphological complexity and diversity. *Nature* **409**, 1102–1109 (2001).
53. Wilder, H. H. *History of the Human Body* (Henry Holt, New York, 1909).
54. Huxley, T. H. *Evidence as to Man's Place in Nature* (Williams and Norgate, London, 1863).
55. Thenius, E. in *Handbook of Zoology* (ed. Kristensen, N. P.) (Walter de Gruyter Publishing, Berlin, 1989).



## Acknowledgements

This paper is an attempt to synthesize an expansive literature, rather than a literature review or report of new results. My citations have concentrated on reviews or overviews where detailed primary references can be found. Details on the covered topics can also easily be found by searching the internet. I thank Anne Buchanan and Sam Sholtis, and the reviewers for help in the development of these ideas and this manuscript.


## Competing interests statement

The author declares no competing financial interests.

## Online links

**FURTHER INFORMATION**

**Chladni figures for violin plates:** http://www.phys.unsw.edu.au/music/violin

**Gene Ontology Consortium:** http://www.geneontology.org/

**Kenneth Weiss' laboratory:** http://146.186.95.23/weiss_lab/index.html

**Access to this interactive links box is free online.**

## Author Biography

Ken Weiss was trained in mathematics at Oberlin College, in meteorology at St. Louis University, and in biological anthropology and genetics at the University of Michigan. His interests are in the amount, source and distribution of human variation; the practical and epistemological problem of inferring the contribution of that variation to disease; the evolutionary and developmental biology of patterned traits (specifically, vertebrate mineralization and dentition); and the history and philosophy of biological science. He is Evan Pugh Professor of Anthropology and Genetics at Penn State University, with a joint appointment in Biology.

## Online Summary

- There are parallels between Darwinian evolution — a history of individuals — and development — a history of cells within individuals.
- Darwin's theory explained how traits occur but not which traits will occur or their phenogenetics; that is, how genes produce biological traits.
- A key to these principles is their logic: they are relational principles that depend on the interaction of components, rather than the specific physical attributes of the components themselves.
- Phenogenetic logic comprises a small number of basic and simple characteristics of genomes, which can help to account for the diversity of biological traits through a few basic developmental processes.
- Fundamental to phenogenetics are duplication, modularity, the hierarchical organization and partial sequestration of components, inductive patterning — including dynamic repetitive patterning — the use of diverse types of arbitrary codes, and various kinds of budding and branching phenomena.
- These principles are fundamental to the nature of life, and have operational value for understanding life and for experimental as well as evolutionary biology.
- This review is an attempt to identify the elements of phenogenetic logic and to synthesize their role in the generating biological traits.
- Together, the symmetries of evolutionary processes and phenogenetic logic provide an elegant, simple and comprehensive view of the organization of life.

## Online Links

**Chladni figures for violin plates**
http://www.phys.unsw.edu.au/music/violin

**Gene Ontology Consortium**
http://www.geneontology.org/

**Kenneth Weiss' laboratory**
http://146.186.95.23/weiss_lab/index.html