CrossMark

Sameness, novelty, and nominal kinds

David Haig

Abstract Organisms and their genomes are mosaics of features of different evolutionary age. Older features are maintained by 'negative' selection and comprise part of the selective environment that has shaped the evolution of newer features by 'positive' selection. Body plans and body parts are among the most conservative elements of the environment in which genetic differences are selected. By this process, well-trodden paths of development constrain and direct paths of evolutionary change. Structuralism and adaptationism are both vindicated. Form plays a selective role in the molding of form.

Keywords Adaptation · Developmental constraint · Evolvability · Formal cause · Homology · Novelty · Strategic gene · Transposable elements

> One simply cannot escape the conclusion that the brain of a rat and a human are actually the 'same' in spite of their obvious differences. (Wagner 1989)

Interface

Homology, Genes, and Evolutionary Innovation (Wagner 2014; henceforth *HGEI*) diagnoses a conflict between structuralist and functionalist (or adaptationist) styles of thinking. Functionalists explain organismal traits by their adaptive value whereas structuralists explain why things are the way they are by appeal to structural constraints and capacities. Wagner proposes "to overcome this conflict by addressing a specific biological phenomenon for which the conflict often crystallizes: the question of homology … At its core, the question is whether

D. Haig (✉)
Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA
e-mail: dhaig@oeb.harvard.edu

homologs exist—that is whether they are natural members of the 'furniture of the world' or whether they are only transient traces of the phylogenetic past. In the latter case, they would have no biological, conceptual, or causal significance. In the former case, homologs would have to play a central role among the concepts of evolutionary theory" (*ibid.* p. 8). Wagner opts for their central importance.

For Wagner (2014), "the realization that complex organisms/systems have unique and historically contingent variational constraints and biases paves the way for a seamless unification of functionalist and structuralist agenda" (*ibid.* p. 19). In this synthesis, conserved structural properties have a causal role in determining how structures vary, and fail to vary, over evolutionary time. Although an olive branch of unification is offered, *HGEI* is written to correct the myopia and astigmatism of adaptationism. Adaptationists, it is suggested, have belittled, misrepresented and misunderstood structuralists and rapprochement should occur on structuralist terms (as befits the injured party). An adaptationist who describes similar phenomena in functionalist language is likely to feel misrepresented and misunderstood and to insist that rapprochement occur on functionalist terms. And thus, an underlying consensus may be obscured by semantic arguments because human nature is quicker to recognize when we are misunderstood than when we have misunderstood.

My adaptationist commentary on *HGEI* is an attempt to seek areas of consensus with structuralists and identify where different perspectives of the same scene seem to tell different stories. Deeply conserved structures are indeed significant parts of the living world and have had profound influences on the course of evolution. This is something an adaptationist should be willing to concede without having to renounce natural selection as a cause of conserved structures.

Positive and negative selection

Adaptation by natural selection replaces 'old' less-adapted gene sequences by 'new' more-adapted sequences, but natural selection does not cease once the 'old' is replaced by the 'new' because these 'newly old' sequences will inevitably deteriorate unless mutations that impair their function continue to be eliminated by selective deaths. Natural selection is now associated with the failure to replace 'old' more-adapted sequences with 'new' less-adapted sequences. These two faces of natural selection are sometimes labeled positive selection (associated with *origin* of novel function) and negative selection (associated with *maintenance* of existing function). The selective reasons for the origin of a character may differ from the selective reasons for its maintenance.

Wagner (2014) makes extensive use of 'loss-of-function' mutations to infer 'normal functions' of genes but does not explicitly consider the role of negative selection in the maintenance of conserved aspects of development, rather the role of negative selection is implicit. The genes and regulatory networks that are responsible for developmental constraints are conserved, not because functional changes do not occur, but because functional changes are not tolerated.

Adaptationists and structuralists agree that conserved morphological and genomic features are maintained by negative selection, but often disagree on the

interpretation. Many adaptationists would interpret the maintenance of structure in the face of mutation as an expression of the power of natural selection whereas many structuralists would interpret the absence of positive selection as evidence of a constraint on what can evolve. This has been a recipe for mutual misunderstanding.

Morphological and genetic homology

Owen's (1848) definition of a homolog as "the same organ in different animals under every variety of form and function" is often cited as the first clear statement of the concept of homology although Owen (1846) himself wrote that in "illustrating the term *homology*, I have always felt and stated that I was merely making known the meaning of a term introduced into comparative anatomy long ago, and habitually used in the writings of the philosophical anatomists of Germany and France." He saw his own conceptual contribution as distinguishing among general, special, and serial homology: general homology was the relation in which a part stood to "the ideal or fundamental type"; serial homology the relation between repeated parts within the same body; and special homology the "essential correspondence" of the same parts in different species.

Berkeley's (1857) *Introduction to cryptogamic botany* illustrates how analogy and homology were used shortly before publication of *On the origin of species* (Darwin 1859): "Analogy … is always liable to seduce an inattentive or ignorant observer into wrong notions as to the relation of beings between which it exists" but "Homology is of far more value; for when true it is founded on a deep knowledge of structure, and is indicative of either close or remote relation." Analogy was "resemblance of function". Homology was "correspondence of structure or origin". Homologous structures were "identical in essence and origin" (Berkeley 1857, pp. 39–41). The use of 'origin' in these definitions probably referred to ontogeny rather than phylogeny.

Darwin (1859) proposed that special homology was explained by descent from a common ancestor: "If we suppose that the ancient progenitor, the archetype as it may be called, had its limbs constructed on the existing general pattern, for whatever purpose they served, we can at once perceive the plain signification of the homologous construction of the limbs throughout the whole class" (*ibid*., p. 435). Moreover, with archetypes replaced by ancestors, general homology could be understood as special homology between present and past.

After the acceptance that descent with modification was one possible cause of structural resemblance, Lankester (1870) recommended that 'homology' be discarded and replaced by two new terms: homogeny and homoplasy. He wrote "Structures which are genetically related, in so far as they have a single representative in a common ancestor may be called *homogenous*." By contrast, "When identical or nearly similar forces, or environments, act on two or more parts of an organism which are exactly or nearly alike, the resulting modifications of the various parts will be exactly or nearly alike. … I propose to call this type of agreement *homóplasis* or *homóplasy*." Homoplasy included "all cases of close resemblance of form which are not traceable to homogeny, all *details* of

resemblance not homogenous, in structures which are broadly homogenous, as well as in structures having no genetic affinity." Homogeny depended "simply on the inheritance of a common part" whereas homoplasy depended "on a common action of evoking causes or moulding environment on such homogenous parts, or on parts which for other reasons offer a likeness of material to begin with."

The difference between Darwin's and Lankester's explanation of serial homology is interesting. For Darwin (1859), serial homology was jointly explained by correlations of growth and descent. "The several parts of the body which are homologous, and which, at an early embryonic period, are alike, seem liable to vary in an allied manner: we see this in the right and left sides of the body varying in the same manner; in the front and hind legs, and even in the jaws and limbs, varying together, for the lower jaw is believed to be homologous to the limbs" (*ibid.*, p. 143). Repeated parts were primordially similar but had diverged in form by natural selection to perform diverse functions. Nevertheless, "we need not wonder in discovering in such parts or organs, a certain degree of fundamental resemblance, retained by the strong principle of inheritance" (*ibid.*, p. 438).

For Lankester (1870), serial homology was homoplasy. Forelimbs and hindlimbs could not descend from the same part. Therefore, their concerted evolution must be explained by similar external forces acting on similar, but independently malleable, parts. Lankester probably believed in the direct inheritance of form, with inheritance akin to memory. A change to the forelimb of a progenitor could be 'remembered' as a corresponding change in the forelimbs of descendants, but how could a change to an ancestral forelimb be 'remembered' by subsequent hindlimbs? What is missing in Lankester's account is Darwin's appeal to 'internal' correlations of growth.

Weismann (1890) recognized that contemporary enthusiasm for the inheritance of acquired characteristics was based on a model of inheritance in which changes to a parental part were directly communicated to the filial part ('direct inheritance of form'). He championed an alternative model in which offspring inherited determinants of form. The determinants were inherited in nuclear chromatin but were expressed in cytoplasm. Serial homology and 'correlations of growth' could now be explained by the expression of the same determinants in different locations within the body. These hereditary factors were conceptualized as genes in the years after rediscovery of Mendel's experiments.

A concept of genetic homology arose from the theory that genes were physical structures with precise locations on chromosomes. Chromosomes of the same form paired at meiosis and were recognized as homologs. Genes at the same location on homologous chromosomes were alleles. If alleles could trace their lineage back to a single material gene then genetic homology could be rooted in common ancestry. Greater precision was given to the concept of genetic homology with the discovery of the double helix. Two DNA sequences were homologous if both were derived from a common ancestral template via unbroken chains of replication.

The ancestry of a set of nucleic acid sequences coalesces in their most recent common template. Once genetic homology was defined as descent from the same template, homologous sequences need no longer exist at the same chromosomal locus nor need different parts of a sequence have the same ancestry. An ancestral sequence could leave descendants at multiple chromosomal locations by gene duplication and

genome rearrangements (Fitch 1970). Moreover, cutting and splicing could create novel sequences with parts that coalesce in different ancestral templates.

At first sight, genetic homology provides a means of identifying morphological homologs: two forms are homologous if their development is determined by homologous genes. However, epigenetic processes mean there can be no simple mapping of genotypic determinants to phenotypic forms. Body parts are generated anew in each generation by interaction of many genes in environmental context (Waddington 1957). Although the question whether two DNA sequences are derived from a common template is relatively well-defined, the question whether two parts have evolved from the same ancestral part is problematic. "Characters are not literally derived from other characters. Organs are not descended from other organs, nor are they inherited from ancestors" (Cartmill 1994).

One resolution of this problem would be to reject talk of morphological homology as conceptually incoherent and restrict attributions of homology to DNA sequences. Such an approach would investigate the genetic architecture that underlies a character's development and ascertain how these networks and associated characters have been transformed in multiple lineages over evolutionary time. From this eliminationist perspective, the morphologist's question whether the phyllode of an *Acacia* is 'really' a leaf or a petiole (Boke 1940) would be seen as wrong-headed. Rather, the developmental mechanisms that produce 'phyllodes' would be dissected and the similarities and dissimilarities of this machinery with the mechanisms that produce 'leaves' and 'petioles' in related taxa would be described. The ontological question whether two parts belonged to the same *or* different kinds would lose meaning as more was learnt about the mechanistic reasons for their similarities *and* dissimilarities. One could still talk of 'leaves,' 'petioles,' and 'phyllodes' as terms of convenience without having an ontological commitment to leaves and petioles as natural kinds.

The problem with defining homologs as 'the same organ under every variety of form' is that there are no unambiguous criteria for deciding when organs of different form are the same and when they are different. This problem was not resolved by replacing archetypes by ancestors. Nor was it resolved by replacing belief in the direct inheritance of form with belief in the inheritance of determinants of form, especially when determinants were identified by their effects on form.

By contrast to problems with attributions of morphological homology, templated replication of nucleic acids does provide a simple interpretation of claims that two sequences are derived from a common ancestral sequence. For many purposes, coalescence of gene sequences provides a perfectly adequate concept of homology. Nevertheless, practical situations will remain in which information about genetic mechanisms is unavailable but one would like to make statements about common 'ancestry' of morphological features. One could then employ concepts of morphological homology as heuristic tools without a metaphysical commitment to homologs as natural kinds.

Characters and states

Wagner (2014) advocates investigation of how the gene networks underlying character development have evolved, but his purpose is not to supplant the classical

concept of morphological homology with one based on genetic homology. Rather, he seeks to identify the aspects of genetic control that form the mechanistic basis of character identity. Homology, he writes, "reflects the developmental organization of organisms" (*ibid*. p. 72). The division of bodies into homologs carves an organism at its joints, sometimes literally. Body parts exhibit evolutionary continuity that transcends changes in most aspects of the genetic control of their development.

Wagner's model of character development distinguishes character identities from character states and recognizes three tiers of genetic control. The first tier is responsible for positional cues and the third for the realization of character state. The crucial middle tier is the Character Identity Network (ChIN). ChINs are mutually-exclusive functional units that are more strongly conserved than the other tiers, thus creating continuity of character identity and quasi-autonomy from other characters despite changes of character state and location. Character identity in this model corresponds to the "same organ" and character state to "every variety of form and function" in Owen's definition of homologs. As a corollary, *HGEI* distinguishes between evolutionary modification of character states (adaptation) and origin of new character identities (novelty). Wagner proposes that the genetic changes that create novelty often differ in kind from those that produce adaptation.

The vertebrate lens exemplifies a body part that can be homologized among species. Wagner (2014) uses the lens to illustrate how homologous organs need not have the same developmental origin nor follow the same ontogenetic pathway. When the lens of a newt is removed, a new lens develops from marginal cells of the iris even though a lens usually develops from ectodermal cells of the embryonic eye cup (*ibid*. p. 84). One might say that the old and regenerated lenses are serially homologous in time.

Mature lens cells lack nuclei and consist of 30–50 % protein in aqueous solution. The most abundant lens proteins are called crystallins and function in the refraction of light (Graw 2009). Different enzymes have been recruited as crystallins in different taxa (Wistow 1993; Piatigorsky 2007). For example, lactate dehydrogenase A (LDH-A) has been recruited as the υ-crystallin of the platypus (van Rheede et al. 2003) whereas lactate dehydrogenase B (LDH-B) has been recruited as the ε-crystallin of crocodiles and birds (Wistow et al. 1987; Brunekreef et al. 1996). The genes that encode LDH-A and LDH-B diverged from each other after a whole-genome duplication that occurred early in vertebrate evolution (Stock et al. 1997). The principal requirement to serve as a crystallin is that a protein remain in solution at high concentration to maintain lens transparency without aggregation to form cataracts. This property must be exceptionally stable because lenses must function for the life of an animal without protein turnover. Conserved properties of lactate dehydrogenases, such as solubility and compact structure, probably predisposed LDH-A and LDH-B for 'independent' cooption as crystallins in archosaurs and monotremes.

Can a principled distinction be maintained between characters and states? The presence of ε- and υ-crystallins are character states of the lens, but, does an enzyme's cooption as a crystallin also create a new character? Is this a minor regulatory change, a mere tweaking of expression levels in the lens, or a novelty, an enzyme doing double-duty as a structural protein? There are many features of lenses

about which one could make meaningful statements of shared or non-shared ancestry. Wagner (2014) sees the failure of the historical concept of homology to clearly distinguish between characters and their states as a weakness because it replaces a notion of sameness with one of residual similarity (*ibid*. p. 73), but a rigid state/identity distinction risks restricting attributions of homology to individualized parts with chiseled ChINs.

Natural and nominal kinds

Wagner (2014) recognizes boundaries are blurred between character identities and states but denies that the existence of a continuum undermines his advocacy of a state/identity distinction because things toward one end of the continuum are clearly character states whereas things toward the other end are clearly character identities (*ibid*. p. 198). More generally, he recognizes a continuum of conservation from the relatively invariant to the evolutionarily ephemeral, but champions a philosophical position that highly-conserved attributes are sufficiently stable to be defining properties of natural kinds.

Wagner's (2014) major reason for considering biological kinds to be classes rather than evolutionary individuals is that homologs and taxa can possess invariant properties. Thus, "the 'essence' of being a eukaryote resides in the manner in which their cells are organized and how their genetic material is packaged" (*ibid*. p. 236). But this essence excludes most dinoflagellates from the ranks of eukaryotes. Dinoflagellate nuclei contain abundant DNA uncomplexed with histones and organized as permanently condensed, liquid crystalline chromosomes. Other morphological and molecular characters however clearly indicate that dinoflagellates evolved from within the alveolate clade of eukaryotes (Gornik et al. 2012). One could undoubtedly modify the 'defining characters' of eukaryotes to include dinoflagellates but this new definition would be vulnerable to the next exception that probed the rule. It seems simpler to define eukaryotes as an 'individual' with a unique evolutionary history than as a class defined by particular characters.

Consider vertebrate kidneys. Kidneys develop from nephrogenic cords of mesoderm. The embryonic pronephros differentiates from the cranial end of a nephrogenic cord. The mesonephros subsequently develops from more caudal regions of the cord and replaces the pronephros as it degenerates. The adult kidney of fish and amphibians is considered to be a mesonephros, but the mesonephros is replaced as the functional adult kidney of reptiles (including birds) and mammals by a metanephros that develops from the far caudal cord, although parts of the mesonephros persist as ducts of the male epididymis. Neither fish nor amphibians possess a metanephros (Hamilton et al. 1947).

Pronephros, mesonephros, and metanephros are serial homologs. Mesonephric kidneys of amphibians and epididymal ducts of mammals are special homologs. How can one define kidneys as a class of which all members are mutual homologs but no non-members are homologs? Is the epididymis a kidney in disguise or do epididymides possess their own character identity? Do the mesonephric kidneys of amphibians and metanephric kidneys of mammals belong to the same or different

classes? Answers could be found for each of these questions that would rescue 'kidney' or 'metanephros' or 'epididymis' as natural kinds, at least temporarily, but such answers must be sought anew for each and every homologous part and such ad hoc solutions must be revised each time something is found that does not fit neatly an existing definition.

Wagner (2014) prefers precise definitions to fuzzy concepts and rejects the idea that homologs are nominal kinds, but his depiction of nominal kinds as "simple arbitrary summaries" and "human-made distinctions for our own convenience … otherwise meaningless" (*ibid*. pp. 229–230) stacks the deck against nominalism. A nominalist could counter that all categories are human constructs but some are less arbitrary, and more useful, than others. Indeed, a consistent nominalist would not object to useful categories being labelled natural kinds because she would accept that 'natural kind' is itself a nominal kind that represents whatever a linguistic community chooses it to mean.

Wagner (2014) disarmingly acknowledges that nothing in *HGEI* should be construed as a formal definition. Rather he sees his book as presenting models that will evolve into precise definitions as more is learnt about developmental mechanisms (*ibid*. pp. 242–4). The opposite may be true. As more is learnt, more forms may sit uncomfortably with respect to any definition one may care to propose. The nominalist would concede that homology concepts are useful tools but would affirm that different tools are suited to different tasks. In her pluralist view, different disciplines will inevitably adopt different homology concepts, just as different disciplines employ different species concepts.

The identification of homologous parts in disparate organisms is often presented as conclusive evidence for their evolutionary descent from a common ancestor but this commonplace argument hides an ontological tension. Evolutionary thought is inimical to rigorous definitions and well-defined categories because it concerns processes by which things of one kind become things of other kinds, whereas attributions of homology attempt to capture that which remains unchanged despite transformations of form and function. The more things change, the more they remain the same.

Novelty and adaptation

Wagner (2014) argues that the "origin of homologs" is distinct from "their modification by natural selection" (*ibid*. p. 43). He writes "It is conceptually … necessary to distinguish between the evolution of adaptations and the origin of novelties" because "there are a number of features characteristic of novelties that make it unlikely that the adaptationist program will give us satisfactory answers" (*ibid*. pp. 121–3). In particular, the genetic rewiring of regulatory networks that creates novelties differs in kind from the mere tweaking of existing pathways that produces adaptation. As a consequence, "innovation is a different kind of process than is adaptation, which is usually studied within populations at the micro-evolutionary level" (*ibid*. p. 209). Novelties are both rare and pregnant with possibilities.

There is more than one way to carve evolutionary biology at the joints. Wagner (2014, pp. 10–12) equates adaptationism with population genetics, but many population geneticists would deny the adaptationist label and many adaptationists would agree with Wagner that the genetic variation currently segregating in populations is not typical of the genetic changes that were responsible for major evolutionary innovations. As Eshel (1996) has argued, short-term and long-term evolution proceed by "radically different rules" because the short-term models of population genetics employ different equilibrium concepts from the long-term models of adaptationists. The postulate that "the behaviour of the long-term process can be fully understood by extrapolation of the analytically well-defined short-term process … is mathematically wrong."

Wagner (2014) consistently downplays the role of adaptation in the genesis of novelty and it is here that he parts company with adaptationists: "the specific new potential of a novelty can hardly be 'seen' by natural selection that originally selected the new trait"; and "it is unlikely that natural selection can provide a satisfactory account of the fact that feathers turned out to be able to support flight, whereas hair did not" (*ibid.* p. 123). But no adaptationist would ascribe prescience to natural selection. The evolutionary potential of an innovation is always recognized in hindsight, with some changes judged retrospectively more significant than others.

Wagner (2014) ascribes significant responsibility for determining the direction of evolutionary change to sources of variation rather than assigning the sole directive role to the sifting of this variation by natural selection. In particular, he proposes that rewiring of genetic networks does not occur by natural selection of point mutations but by co-option of ready-made promoters from transposable elements that are only episodically active and lineage specific. "Evidence that transposable elements play a major role in the evolution of gene regulatory networks affects various uniformitarian ideas that are broadly accepted in evolutionary biology. … It is not far-fetched then to think that the evolutionary fate of a lineage is strongly influenced and different from that of other lineages, in part, because of the nature of genomic parasites that infect its genome at any point in time" (*ibid.* p. 207).

Fisher (1934) articulated the standard adaptationist critique of theories that ascribe the "effective guidance of the evolutionary process to the agencies which cause mutation." He acknowledged mutation "as a condition which renders evolution possible" but assigned the creative agency to the selective processes that eliminate all but a small minority of mutations. Most adaptationists would now construe 'mutation' to include insertions of transposable elements and would emphasize that most insertions, like most point mutations, are deleterious or selectively neutral. Deleterious insertions are eliminated by negative selection whereas neutral insertions either drift to extinction or are eventually degraded by mutation. Only a small minority of insertions are preserved by positive selection, those that serendipitously enhance adaptation (Haig 2012a).

The analogy of circuit rewiring may be helpful. Transistors and integrated circuits have revolutionized what is technologically possible, but not all circuits can be readily reconfigured as new devices. Wagner suggests adaptationists overemphasize the role of unconstrained natural selection and pay insufficient attention to

the properties of components and existing circuits. Adaptationists argue that components do not self-assemble into novel gadgets without an electrical engineer. They emphasize the role of the engineer (natural selection), both in assembly of circuits and origin of components.

Consider the expression of prolactin (*PRL*) in the endometrium of elephants, rodents, and primates but not in the endometrium of rabbits, pigs, dogs, armadillos, or opossums (Emera et al. 2012). Three independent origins of endometrial *PRL* expression are associated with insertions from four families of transposable elements (TEs) that occurred at different times during mammalian evolution (Lynch et al. 2008; Emera et al. 2012). Some insertions were not associated with endometrial expression for many millions of years (Emera and Wagner 2012a). All these insertions survived the sieve of natural selection, but there were undoubtedly many more insertions in *PRL* genes that have not left descendants or detectable traces in extant genomes. Are we to understand that endometrial expression would be adaptive in all mammals but an endometrial promoter never arose in the lineages of rabbits, pigs, dogs, or armadillos because of the absence of the right kind of TE? Or should we conclude that endometrial expression has arisen many times but has only been retained by natural selection in three lineages? Are TEs responsible for the origin of novelty or is the creative agency the winnowing of insertions by natural selection? Do TEs rewire regulatory networks or does natural selection rewire networks using their promoters as handy components?

TEs (including retroviruses) are suggested to have rewired endometrial and placental regulatory networks (Cohen et al. 2009; Lynch et al. 2011; Emera and Wagner 2012b; Chuong 2013; Chuong et al. 2013). Although TE insertions can occur almost anywhere in the genome, including in crystallin genes (Nag et al. 2007), their promoters have never been reported to confer lens-specific expression, or rewire regulatory networks of the lens, despite the recruitment of diverse genes as crystallins during vertebrate evolution (Wistow 1993). Retroviruses are probably a source of placenta-specific and endometrium-specific promoters because retroviral expression in these tissues facilitates infectious transmission between mother and child (Haig 2012a, 2013), but retroviruses are not expected to possess lens-specific promoters because replication in lenses does not facilitate retroviral transmission.

The endometrium and placenta appear to be more rapidly evolving than the lens. Wagner would assign much of the credit for accelerated evolution at the maternal-fetal interface to TEs as a source of saltatory genetic variation. An adaptationist would explain rapid evolution of tissues that separate mother and fetus as an outcome of antagonistic selection between genes expressed in the maternal endometrium and genes expressed in the fetal placenta; between genes of maternal and paternal origin in the placenta; and between retroviral adaptations and the defenses of maternal and filial hosts (Haig 1993, 2008a, 2012a).

Modularity and evolvability

A genome's nucleic acid sequence can be likened to the software of a robotic control system. Among the important tasks that it must control is the assembly of its

own hardware. Such metaphors, of organisms as robots and genomes as software, are often dismissed because the metaphors devalue organismal autonomy and privilege genes over environment. But the objections seem overstated. Useful robots regularly make autonomous decisions and adjust their behavior in response to environmental inputs (Haig 2008b). Comparisons between genetic and robotic control systems are useful, both for the similarities and dissimilarities revealed.

Software engineering encompasses synchronic goals—writing software that is useful now—and diachronic goals—writing software that will be easy to modify in the future. With respect to future modification, software should be robust so that changes do not break what already works and open-ended so new functions can be integrated with minimal change (Calcott 2014). Software development is usually decomposed into manageable parts that can be independently programmed with a separate interface that calls upon modules as needed. Modular design has synchronic benefits—division of labor (separate teams can work simultaneously on distinct tasks without needing to constantly communicate) and comprehensibility (system function can be studied one module at a time)—overlapping with diachronic benefits—reduced pleiotropy (changes within a module do not ramify to other parts of the program) and re-use (self-contained modules can be adapted for novel functions) (Parnas 1972; Calcott 2014).

The diachronic benefits of modular software probably translate to genetic evolution but the synchronic benefits may not. In software engineering, each module is tested and debugged before interactions among modules and global functioning are tested. When the assembled system does not perform as intended, modularity facilitates the isolation and correction of problems (trouble-shooting). Natural selection does not divide labor among teams, cannot comprehend the code that it generates, and does not isolate problems before correcting them. Modules (if these exist) are not tested individually but as an ensemble. Negative selection to maintain existing genetic code is costly. An organism must die without progeny to eliminate a bug in any part of the system. These costs of negative selection create an advantage for using the same code for multiple functions (pleiotropy) because a lethal failure to perform one function purifies shared code for all functions.[1]

Modularity and evolvability can be designed features of software but can they be evolved properties of genetic systems? Programmers can anticipate future needs, but natural selection lacks foresight. For Lynch (2007), "there is no compelling empirical or theoretical evidence that complexity, redundancy or other features of genetic pathways are promoted by natural selection" whereas for Calcott (2014), "complex integrated systems, whether evolved or engineered, share structural properties that affect how easily they can be modified to change what they do." Lynch sees most proposals about the evolution of evolvability as adaptationist overreach whereas Calcott sees evolvability as distinct from adaptation. Evolutionary arguments about evolvability are a semantic morass (Sniegowski and Murphy 2006).

---

[1] Costs of selective death are reduced if there is postzygotic provisioning of offspring because death can occur before full commitment of effort (Hamilton 1966; Hastings 2000). Some copying errors are detected and corrected by proof-reading mechanisms during DNA replication. These mechanisms are analogous to error-correcting properties of computer code.

When algorithms evolve to solve a task rather than are designed to solve the task, the evolved algorithms provide mixed support for the idea that natural selection favors modular architectures. Nonmodular algorithms tend to outperform modular algorithms when given a unitary task. Modularity, if initially present, tends to break down because there are many possible connections that break modularity and increase fitness. On the other hand, algorithms spontaneously evolve modularity if given multiple tasks that alternate in a regular fashion (Kashtan and Alon 2005).

If the source code of a successful software package were examined, some parts of the code would be highly conserved and others would have undergone extensive change since the earliest versions. Some modules might have entirely changed their code while maintaining conserved function and conserved links to other modules. By contrast, if the evolution of a once successful software package were examined, the source code would be found to have initially undergone updating and addition of new functions, then updates slowed and eventually ceased, although the package continued to be used by an ever dwindling number of users until its eventual extinction. There are many reasons why packages become 'extinct' but one factor could be structural features of the source code that were not conducive to efficient updating and modification for new uses. By such a process, one would observe preferential survival of more evolvable software.

Every genome encodes features that are anciently conserved because changes are not compatible with a viable organism. Some of these developmental constraints may be more or less conducive to changes in other features that allow adaptation to changing environments. Thus differential extinction will result in the preferential survival of lineages with more evolvable genomes. This selection among lineages can be considered to choose among developmental constraints. Over evolutionary time, 'good' constraints that promote long-term survival in a changing world, or prevent a lineage falling into short-term evolutionary traps, will tend to be preserved whereas 'bad' constraints that impede adaptive change will be eliminated.

I believe consensus exists between structuralists and adaptationists about what might be called 'clade selection' of evolvability or evolutionary constraints where the latter are conserved features that influence evolutionary outcomes. However, clade selection is not sufficient. Conserved features must be maintained by negative selection within populations. One can ask whether loss of evolvability is one of the reasons for a feature's maintenance by negative selection or whether effects on evolvability are incidental byproducts of negative selection for individual benefits. In general, mechanisms maintained by immediate benefits to organisms will be more robust than mechanisms maintained by effects on evolvability. One could also ask about the reasons for the origin of constraints by positive selection and whether there are reasons why positive selection for immediate benefits should favor mechanisms that have enhanced evolvability as a byproduct.

Hypoxic cells of vertebrates release signals that trigger nearby blood vessels to grow toward the hypoxic region, alleviating hypoxia. This mechanism enhances evolvability because body parts are automatically supplied with blood vessels as they evolve new shapes (Gerhart and Kirschner 2009), but the mechanism is maintained by individual-level selection because organisms in which it malfunctions experience immediate costs. The benefit of enhanced evolvability for the

lineage is an incidental byproduct of the individual-level benefit. Perhaps plastic mechanisms, as exemplified by the vascular response to hypoxia, are superior to rigid mechanisms because plastic mechanisms allow adaptive responses to environmental or within-population variation. Thus, facultative responses favored by individual benefits may also enhance evolvability.

One can retrospectively conclude that morphologically diverse clades have had more evolvable genomes than extinct clades or 'living fossils.' Many will be tempted to extrapolate that clades that have been evolvable in the past will be evolvable in the future, but past extinctions of previously dominant clades should temper enthusiasm for this prediction. One can more confidently predict that some lineages that are now judged to have been evolvable will become extinct.

Formal causes

Selection chooses from a set of alternatives. In the formalism of the strategic gene (Haig 2012b), natural selection of 'allelic' differences requires three components of choice: a genetic difference; a phenotypic difference; and a selective environment. In this formalism, phenotype contains all things that differ, and environment all things that are the same, for the items of choice. Thus conserved features of bodies and genomes are part of the selective environment that chooses among genetic alternatives based on their phenotypes. The genetic variant chosen constitutes a record of the choice (Haig 2014). When the choices of the environment are consistent and repeated, then one of the alternatives can become a fixed part of the selective environment for other choices. By this means, natural selection converts the variable (that which is selected) into the invariant (that which selects).

Processes by which genetic differences cause phenotypic differences can be considered efficient causes and those by which phenotypic differences cause differential genetic replication can be considered final causes. Such processes take place in evolved structures (let us call them bodies) that consist of physical stuff (material cause) and inherited information about past choices (formal cause) of which the genome is the 'textual' record. For each particular choice, the selective environment includes all aspects of bodies and genomes shared by the alternatives. Efficient and final causes are differences that make a difference (they involve implicit comparisons between things). Material and formal causes constitute what a thing *is* without comparison (Haig 2014).

Genomic features are highly conserved if most genetic differences (mutations) have phenotypic effects that cause elimination of the difference (negative selection). These 'invariant' features of the genome, and the bodily forms they determine, are part of the selective environment that chooses among differences in less conserved parts of the genome. Conserved features of the body and genome are often more conservative than aspects of the 'external' environment and form the mechanistic basis of unity of type. These features confirm "the structuralist intuition that complex systems play a causal role in determining their evolutionary fates" (Wagner 2014, p. 18).

Vertebrates with two pairs of lateral appendages have moved from water, onto land, into air, back to water many times, and share many of their habitats with six-

legged insects and eight-legged spiders. The original reasons why an ancestral vertebrate evolved two pairs of fins, whereas an ancestral insect evolved six pairs of legs and an ancestral spider eight pairs of legs, are probably lost in the depths of time. And these 'original' reasons are distinct from the reasons why these numbers have been maintained ever since by negative selection. Paired pectoral and pelvic appendages have been among the most highly-conserved elements of the selective environment in which the adaptive radiation of vertebrates has taken place. The body plans (*Baupläne*) of vertebrates, insects, and spiders can be judged retrospectively to have been stable platforms for evolutionary discovery.

Evolution is a recursive process that dissolves traditional distinctions between cause and effect when considering causal relations between biological kinds rather than between instances of those kinds (Haig 2014). Genes have a causal role in the production of bodies and bodies a causal role in determining which genes survive the filter of natural selection. Form is shaped by genetic networks but form may persist while those networks are radically refashioned under the selective constraints of form. Where does continuity of form reside in the flux of efficient causes? Wagner (2014) would locate that continuity in Character Identity Networks (ChINs). These "are the most conserved part of the gene regulatory network that underlies character development and, thus, are most consistently associated with manifest character identity" (*ibid*. p. 186). But, given enough time, could not a ChIN be changed beyond recognition and yet a character remain 'the same'?

Capsid proteins of diverse viruses exhibit structural similarities that are unlikely to be explained by convergence despite an absence of detectable similarity in amino acid sequence (Bamford 2003). These proteins are encoded by DNA sequences that descend from an ancestral sequence that encoded an ancestral capsid protein more than a billion years ago. Although structural similarities suggest that capsid proteins are genetically homologous by the criterion of descent of their genes from a common template, no similarity can be detected in either the nucleic acid or amino acid sequence. We have returned full-circle to comparative morphology. Common ancestry is suggested by shared possession of 'double-barrel trimers' and 'jelly rolls' recognized by three-dimensional gestalt rather than linear sequence (Benson et al. 2004; Bamford et al. 2008). Form itself has dictated what has been conserved.[2]

References

Bamford DH (2003) Do viruses form lineages across different domains of life? Res Microbiol 154:231–236
Bamford DH, Grimes JM, Stuart DI (2008) What does structure tell us about virus evolution? Curr Opin Struct Biol 15:655–663
Benson SD, Bamford JKH, Bamford DH, Burnett RM (2004) Does common architecture reveal a viral lineage spanning all three domains of life? Mol Cell 16:673–685
Berkeley MJ (1857) Introduction to cryptogamic botany. Bailliere, London

---

[2] The 'story' has been retold so many times with embellishments, abbreviations, and subtle rewording that evidence of common ancestry is no longer detectable in the text but confined to structural features of the plot.

Boke NH (1940) Histology and morphogenesis of the phyllode in certain species of *Acacia*. Am J Bot 27:73–90

Brunekreef GA, Kraft HJ, Schoenmakers JGG, Lubsen NH (1996) Mechanism of recruitment of the lactate dehydrogenase-B/ε-crystallin gene by the duck lens. J Mol Biol 262:629–639

Calcott B (2014) Engineering and evolvability. Biol Philos 29:293–313

Cartmill M (1994) A critique of homology as a morphological concept. Am J Phys Anthropol 94:115–123

Chuong EV (2013) Retroviruses facilitate the rapid evolution of the mammalian placenta. BioEssays 35:853–861

Chuong EB, Rumi MAK, Soares MJ, Baker JC (2013) Endogenous retroviruses function as species-specific enhancer elements in the placenta. Nat Genet 45:325–329

Cohen CJ, Lock WM, Mager DL (2009) Endogenous retroviral LTRs as promoters for human genes: a critical assessment. Gene 448:105–114

Darwin C (1859) On the origin of species by means of natural selection or the preservation of favoured races in the struggle for life. John Murray, London

Emera D, Wagner GP (2012a) Transformation of a transposon into a derived prolactin promoter with function during human pregnancy. Proc Natl Acad Sci USA 109:11246–11251

Emera D, Wagner GP (2012b) Transposable element recruitments in the mammalian placenta: impacts and mechanisms. Brief Funct Genomics 11:267–276

Emera D, Casola C, Lynch VJ, Wildman DE, Agnew D, Wagner GP (2012) Convergent evolution of endometrial prolactin expression in primates, mice, and elephants through the independent recruitment of transposable elements. Mol Biol Evol 29:239–247

Eshel I (1996) On the changing concept of evolutionary population stability as a reflection of a changing point of view in the quantitative theory of evolution. J Math Biol 34:485–510

Fisher RA (1934) Indeterminism and natural selection. Philos Sci 1:99–117

Fitch WM (1970) Distinguishing homologous from analogous proteins. Syst Zool 19:99–113

Gerhart J, Kirschner M (2009) The theory of facilitated variation. Proc Natl Acad Sci USA 104:8582–8589

Gornik SG, Ford KL, Mulhern TD, Bacic A, McFadden GI, Waller RF (2012) Loss of nucleosomal DNA condensation coincides with appearance of a novel nuclear protein in dinoflagellates. Curr Biol 22:2303–2312

Graw J (2009) Genetics of crystallins: cataract and beyond. Exp Eye Res 88:173–189

Haig D (1993) Genetic conflicts in human pregnancy. Q Rev Biol 68:495–532

Haig D (2008a) Placental growth hormone-related proteins and prolactin-related proteins. Placenta 29(suppl A):S36–S41

Haig D (2008b) Conflicting messages: genomic imprinting and internal communication. In: D'Ettorre P, Hughes DP (eds) Sociobiology of communication. Oxford University Press, Oxford, pp 209–223

Haig D (2012a) Retroviruses and the placenta. Curr Biol 22:R609–R613

Haig D (2012b) The strategic gene. Biol Philos 27:461–479

Haig D (2013) Genomic vagabonds: endogenous retroviruses and placental evolution. BioEssays 35:845–846

Haig D (2014) Fighting the good cause: meaning, purpose, difference, and choice. Biol Philos. doi:10.1007/s10539-014-9432-4

Hamilton WD (1966) The moulding of senescence by natural selection. J Theor Biol 12:12–45

Hamilton WJ, Boyd DJ, Mossman HW (1947) Human embryology. Williams & Wilkins, Baltimore

Hastings IM (2000) Models of human genetic disease: how biased are the standard formulae? Genet Res 75:107–114

Kashtan N, Alon U (2005) Spontaneous evolution of modularity and network motifs. Proc Natl Acad Sci USA 102:13733–13778

Lankester ER (1870) On the use of the term homology in modern zoology and the distinction between homogenetic and homoplastic agreements. Ann Mag Nat Hist 6:34–43

Lynch M (2007) The evolution of genetic networks by non-adaptive processes. Nat Rev Genet 8:803–813

Lynch VJ, Tanzer A, Wang Y, Leung FC, Gellersen B, Emera D, Wagner GP (2008) Adaptive changes in the transcription factor HoxA-11 are essential for the evolution of pregnancy in mammals. Proc Natl Acad Sci USA 105:14928–14933

Lynch VJ, LeClerc RD, May G, Wagner GP (2011) Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. Nat Genet 43:1154–1159

Nag N, Peterson K, Wyatt K, Hess S, Ray S et al (2007) Endogenous retroviral insertion in *Cryge* in the mouse *No3* cataract mutant. Genomics 89:512–520

Owen R (1846) Observations on Mr. Strickland's article on the structural relations of organized beings. Lond Edinb Dublin Philos Mag J Sci (third series) 28:525–527

Owen R (1848) On the archetype and homologies of the vertebrate skeleton. John van Voorst, London

Parnas DL (1972) On the criteria to be used in decomposing systems into modules. Commun ACM 15:1053–1058

Piatigorsky J (2007) Gene sharing and evolution. Harvard University Press, Cambridge

Sniegowski PD, Murphy HA (2006) Evolvability. Curr Biol 16:R831–R834

Stock DW, Quattro JM, Whitt GS, Powers DA (1997) Lactate dehydrogenase (LDH) gene duplication during chordate evolution: the cDNA sequence of the LDH of the tunicate *Styela plicata*. Mol Biol Evol 14:1273–1284

van Rheede T, Amons R, Stewart N, de Jong WW (2003) Lactate dehydrogenase A as a highly abundant eye lens protein in platypus (*Ornithorhynchus anatinus*): upsilon (υ)-crystallin. Mol Biol Evol 20:994–998

Waddington CH (1957) The strategy of the genes. George Allen & Unwin, London

Wagner GP (1989) The biological homology concept. Annu Rev Ecol Syst 20:51–69

Wagner GP (2014) Homology, genes, and evolutionary innovation. Princeton University Press, Princeton

Weismann A (1890) Prof. Weismann's theory of heredity. Nature 41:317–323

Wistow G (1993) Lens crystallins: gene recruitment and evolutionary dynamism. Trends Biochem Sci 18:301–306

Wistow GJ, Mulders JWM, de Jong WW (1987) The enzyme lactate dehydrogenase as a structural protein in avian and crocodilian lenses. Nature 326:622–624