ONTOTENY AND SYSTEMATICS

Arnold G. Kluge and Richard E. Strauss

Museum of Zoology and Division of Biological Sciences, University of Michigan, Ann Arbor, Michigan 48109

INTRODUCTION

Literally hundreds of authors have attempted to interrelate ontogeny and systematics and to find the meaning in their connections (43). Few would doubt the inherent complexity of the subjects but clearly their study has been made overly difficult by imprecise and contradictory communication (34, 63). Even von Baer’s laws (100; see below), the very bases from which many theses have been developed, are so vaguely stated as to allow more than one interpretation. Further, the tendency has increased to try to interpret the older literature as if it were a coherent discourse that always led to a better understanding, when in all likelihood at least some of the early workers were simply confused and matters were, more often than not, left unresolved. It is also apparent, especially recently, that history is being rewritten in order to support a particular point of view (10, p. 38). A final complication relates to the fact that systematics is not always placed in a theoretical framework of evolution and speciation. Of course this is to be expected in the pre-Darwinian literature; however, that position has unexpectedly been readopted by some of our contemporaries (54, p. 17). The difficulty lies not so much in the new version of the old position, but in the ambiguous and inconsistent communication justifying the change.

Our review does not cover all relevant aspects of ontogeny and systematics, nor do we attempt to treat exhaustively any particular issue. Our effort is to survey some of the more important generalities. We also hope to clarify a few longstanding problems, as well as to highlight areas for further research. We focus largely on the descriptive issues of animal ontogeny. Processes and plants are well beyond our expertise. We assume that the genome and epigenetic phenomena produce a morphological superstructure, within which are embedded the traits we observe and describe. We do not consider the mechanisms
by which these traits develop, are inherited, and change across generations. In the discussion to follow, hierarchical implies a time-ordered, multiserial pattern, such as that denoted by a cladogram or phylogenetic tree. An ontogeny, a series of developmental events in an organism, can be viewed as hierarchical only under special circumstances, as discussed below. For the purposes of the present paper, we use cladogram in the sense of a hierarchical summary of character data, or of the taxa represented by those characters; it contains no necessary assumptions about, nor makes predictions of, evolution and speciation (70, p. 14).

HISTORICAL ACCOUNT

The terms evolution and development (the ontogeny of the individual) were synonymous until the 1840s, and increased specialization is still generally acknowledged to result from processes associated with both phenomena (19, 78). A strictly linear concept of nature, a single chain of being—a scala naturae—was generally accepted through the early part of the nineteenth century [e.g. as presented by Lorenz Oken in his Lehrbuch der Naturphilosophie (72)]. That view is effectively summarized by the Meckel-Serres law(s) in terms of ontogeny and systematics—comparative anatomy is repeated and reproduced by ontogeny, and (in a later formulation) higher forms repeat in their ontogeny the anatomical organization of adult organisms lower in that scale of life. This description of nature found its causal explanation in the pre-Darwinian concept of archetype and was recapitulatory in a nonphylogenetic sense.

Louis Agassiz (1, 2) also accepted the archetype concept; however, he recognized more than one chain of being. He claimed that, among the representatives of the same linear sequence, one could identify parallelisms between the structural gradation of their adult anatomies, the stages of their development, and the succession of representative fossils. Agassiz also observed that each such sequence had the appearance of progress, from the general to the special. The parallelism between the geological succession of animals and the embryonic growth of their representatives is recognized as Agassiz’s lasting contribution to systematics, even though his interpretation of the pattern was strictly typological—“the oldest representatives of every class may then be considered as embryonic types of their respective orders or families among the living” (2, p. 174). Furthermore, Agassiz seems to have inconsistently interpreted the meaning of the parallelisms that he identified. He was a creationist, explicitly denying common descent (78, p. 8). Still, he used fossils to interrelate the several chains of being that he recognized, and he referred to such groupings as genealogies. Agassiz’s denial contradicted the long held meaning of genealogy as well as the suggestions of the diagrams that he used to illustrate his ideas (1, Vol. 1, figure facing p. 170).
Karl Ernst von Baer was an Aristotelian essentialist throughout most, if not all, of his career as a biologist. Like Agassiz, he viewed nature as multiserial and divergent. While von Baer accepted the linear concept of grades of development (Ausbildung), he also firmly held to the idea of separate and divergent lineages on the basis of the type of organization. For example, he recognized four types of animal organization—radiate, articulate, molluscous, and vertebrate. Each type had its own archetype, the ideal form, of which all the members of the group were modifications. His research (100) clearly refuted the strictly linear concept of diversity embodied in the Meckel-Serres proposition (see above). He denied that there could be exact correspondences between the embryos of higher animals and the adults of lower animals. These resemblances, according to von Baer, are not between an early lizard embryo and an adult fish, but between the lizard embryo and a fish embryo up to a certain stage of development. In his later life, von Baer argued against miraculous creations, seeming to vacillate between explanations of descent with modification and speciation, and instantaneous formation of species consistent with natural laws. Von Baer’s four laws or rules are nonevolutionary, as can be inferred from the following translations: 1. The more general features of one of the large groups of animals appear earlier in their embryos than the more special features. 2. The less general structures develop from the more general, and so on, until finally the most specialized appears. 3. Every embryo of a given animal form [archetype], instead of passing through the other forms, rather becomes separated from them. 4. Fundamentally, therefore, the embryo of a higher form [archetype] resembles only the embryo of the lower form, never its adult (1828, Transl. 57). These summaries were meant to be about eternal, immutable patterns; inconsistencies between ontogeny and phylogeny were of no consequence. As will become evident below, confusion arises when one tries to make claims about both timeless patterns and phylogeny, which is clearly not timeless.

Phylogeny and character evolution were conceptually wedded to ontogeny by the mid-nineteenth century. For example, Darwin (29) believed that embryology was second to none in providing support for his doctrine of descent with modification. He not only had a firm grasp of ultimate causation, but he explained “the phenomena in terms of heredity and selection” (52, p. 323). Also, Spencer’s theorem [91, pp. 452–53; reformulated by Lovtrup (61, p. 352; 62)] provides an early example of how observations of ontogeny might be used to infer a hierarchy of ancestor-descendant groups, i.e., a phylogeny: “[I]n the course of their ontogeny the members of a set of twin taxa follow the same course of recapitulation up to the stage of their divergence into separate taxa.”

Additional generalizations, like Spencer’s theorem, have been proposed down through the years. For example, there is: 1. Darwin’s assertion (30, p. 466; see also proposition 6, below) that “Community in embryonic structure reveals community of descent; but dissimilarity in embryonic development
does not prove discommunity of descent, for in one of two groups the
developmental stages may have been suppressed, or may have been so greatly
modified through adaptation to new habits of life as to be no longer recogniz-
able”; 2. Haeckel’s (48) biogenetic law, “ontogeny recapitulates phylogeny;”
3. Crowson’s (28) restatement of the biogenetic law, “In general, . . . when the
immature stages of one kind of animal show resemblances to adults of some
other kind, this should be taken as prima facie evidence that the first kind of
animal has descended from ancestors resembling the second”; 4. Lovtrup’s
statement (61, p. 350) “If a feature x, present in members of the taxon T₁, is
epigentially dependent upon the feature y, present in members of the taxon T₂
at any stage of their ontogeny, then the individual members of T₁ must pass
through an ontogenetic stage distinguished by y”; 5. Naef’s statement [67;
translation from (52, p. 326); see also (51, p. 95)] “Morphogenetic processes of
an ontogenetic stage through whose modification a following stage arises are to
be looked upon as phylogenetically older than those which grew out of them;”
6. Von Baer’s law [100; as deduced by Lovtrup (61, p. 351)], “Ontogeny
recapitulates the taxonomic course and the mechanism of phylogeny”; and 7.
Nelson’s proposition (68, p. 327) that, “given an ontogenetic character
transformation from a character observed to be more general to a character
observed to be less general, the more general character is primitive and the less
general advanced.”

Ontogeny continues to figure prominently in current studies of both pattern
and process. Its present popularity is largely associated with recent de-
velopments in phylogenetic inference and heterochrony. These two areas form
the subject of most of the remainder of our review.

A SPECIAL KIND OF SYSTEMATICS

Recently, a few systematists have returned to typology. Nelson & Platnick (70,
p. 328) have in fact asserted that “systematists always have been, are, will be,
and should be typologists.” In addition, we are urged to discover pattern in
nature in a theoretically neutral context (20a). The particular theory to be
avoided is evolution, as was made clear by Patterson [(76, p. 57); see also
Brady (20a; p. 114) and Nelson & Platnick (70, 71)], who stated that “belief in,
or knowledge of, evolution is superfluous.” Further, in keeping with Huxley’s
proposal (55, p. 176) that development serve as the “sole basis of zoological
classification,” von Baer’s first and second laws are cited as fundamental, and
necessary for this program to reach its objectives (70, 76, 85). The justification
for replacing evolution with ontogeny is provided by Nelson’s (68, p. 336)
claim that “the concept of evolution is an extrapolation, or interpretation, of the
orderliness of ontogeny.” Thus, the existence of a natural hierarchy is assumed
(79), and it is further assumed that ontogeny is congruent with it. Ontogeny is
seen as orderly and always hierarchical; therefore, one classifies according to the generality of the observed ontogenetic transformations. The pattern observed, a cladogram, is a nested set of characters, or the taxa defined by those observations. Von Baer’s “scheme of the progress of development” in vertebrates (100, the figure facing p. 225) is considered by some to be the first cladogram based solely on ontogenetic observations (78). While we agree that von Baer’s laws may convince us of hierarchy in nature, it would be a revision of history to infer that von Baer interpreted such a pattern as phylogeny (78, p. 12).

The motive behind promoting a non-Darwinian view of systematics, more than 125 years after *On the Origin of Species*, is not entirely clear (11, 20, 20a, 70, 77, 80); however, it has been claimed that identifying patterns strictly empirically allows the underlying explanatory hypotheses to be evaluated independently (20; 70, p. 35; 79, p. 539; 85). Such independence is a laudable goal in science, but in the present case there is often a failure to distinguish between the general theory of evolution (descent with modification) and particular hypotheses, such as the mechanism of natural selection and its product adaptation (12, 13, 40). Moreover, this argument does not address the likely possibility that ontogeny is the product of those very same processes that one may wish to evaluate eventually, and it leaves unanswered several questions, such as what to do if alternative patterns are discovered. The phylogenetic systematist (see below) chooses among competing propositions by accepting the one that best fits the data, where hypotheses about evolution are the essential ingredient in making that choice, that is, *assumptions* of homoplasy are minimized. Parsimonious genealogies are those best supported by character distributions, and the likelihood of a genealogical hypothesis obtains from the probability it confers on the evolutionary transformation of characters (89, 90). The probability model proposed by Patterson (76, p. 41), while ostensibly nonevolutionary, does not provide a well-founded basis for choosing among alternative cladograms whose similarities and differences are not likely to be due to chance.

This special kind of systematics places severe limitations on what can be investigated. According to Papentin (75, p. 425), as paraphrased by Brooks & Wiley (23, p. 5), “Elucidation of a pattern based on analysis of natural complexity, without any information other than the pattern itself, is only a hypothesis concerning the reality of the pattern itself. It says nothing about the natural laws [processes] underlying its existence.” Moreover, the coherence and generality of such a theory-neutral program of research [referred to as “transformed” cladistics by Platnick (79); now known variously as natural order systematics, and pattern, modern, or methodological cladistics] are currently being vigorously debated. Philosophical considerations have been identified, and Beatty (11) has concluded that transformed cladistics is not just neutral with
respect to evolutionary theory but is actually at odds with it. Two points are made in this context. First, there is the transformed cladists’ claim that unique apomorphies, or unique sets of apomorphies, are necessary and sufficient for determining group membership. The laws of von Baer, in part, are claimed to form the basis for the assertion that group-defining characters are never lost but are simply added onto. Contrary to the set-theoretical argument of Nelson & Platnick (70, pp. 157–9), such defining characters cannot lose or replace any of their defining properties and still retain their diagnostic content. Thus, defining properties cannot themselves evolve. To the phylogenetic systematist, genealogical groups may be described in terms of a hierarchy of apomorphies, but the ultimate criterion of group identity remains the unique historical origin (51, pp. 79–80). The second point is that von Baer’s law, qua “law,” is anti- evolutionary. Just as there are no permanent results of evolution, even the rules of evolution must be allowed to change (57). For example, our most cherished Mendel’s laws are now known to depend on “normal” meiosis, and evolutionary contingencies are required to save their generality (J. Beatty, personal communication).

PHYLOGENETIC SYSTEMATICS

Phylogenetic systematists assume that any particular aspect of an organism’s life history can evolve. Six basic ontogenetic patterns are possible, depending on whether the change occurs terminally or nonterminally in ontogeny and whether the type of transformation involved is addition, deletion, or substitution (23). The best-documented categories of transformation are terminal addition (recapitulation), terminal deletion (paedomorphosis), and nonterminal addition or substitution (caenogenesis).

In phylogenetic inference, it is essential that evidence for sistergroup relationships, apomorphy, be distinguished from nonevidence, plesiomorphy, and for this reason character state transformations must be polarized, i.e. the direction of evolution (plesiomorphy to apomorphy) must be determined. A great deal has been written about polarization, and the outgroup and ontogeny criteria appear to provide practical guidelines, with few assumptions, for inferring the direction of character evolution (23, 33, 37, 57, 63, 84, 92, 103, 104). The former criterion specifies that of two states exhibited by the taxa under study, the state occurring throughout the related groups is plesiomorphic. [Maddison et al (63) discuss how to treat characters that vary among the related groups.] The outgroup criterion is explicitly evolutionary because it requires a hypothesis of phylogeny (particular outgroups, plus the ingroup), and it is intended to provide an inference about the state of the outgroup node [i.e. the common ancestor of the outgroup(s) + ingroup; see (63) for further explanation]. The ontogeny criterion has been stated in terms of either developmental precedence or commonality.
Hennig (51, pp. 95–96) emphasized developmental precedence—the earlier appearing state in ontogeny is plesiomorphic. According to Hennig, this form of the ontogeny criterion can be traced to Naef’s proposition (67; see above), which in turn might have had its conceptual origin in von Baer’s second law and Haeckel’s biogenetic law. In order for this form of the ontogeny criterion to work, one must be able to observe the differentiation of one state into a different state and two ontogenies must be compared. By itself, an ontogeny is usually not hierarchical. It only takes on such a perspective when two or more different ontogenies are compared, or when cell lines and tissues are observed to differentiate into two or more derivatives. The precedence form of the ontogeny criterion does not avoid the problems associated with delimiting comparable preadult states in two or more ontogenies, nor does it avoid the assumptions that dedifferentiation (57) and nonterminal changes do not occur (51; see below). Delimiting the “same” state may be more difficult than expected because the continuous, and usually gradual, nature of ontogeny does not always allow it to be cut in some nonarbitrary manner (32, 53). Even when discrete stages are present, as in arthropods, it may be difficult to identify the same state in different stages (46, 97, 98, 99). Moreover, there is the problem of how to code the same state when it appears more than once in the same ontogeny. For example, early in the ontogenies of most lower tetrapods, and some amniotes, the pharyngeal region becomes serially segmented but no visceral pouches or clefts are present. Later in ontogeny pouches form, which in turn become perforate and form gill clefts or slits. Later still, the slits close as the organism metamorphoses into the more terrestrial and air-breathing phase of its life cycle. Of course, one might argue that the absence of gill slits in the adult is not the same state as their absence in early ontogeny (85), the former being a loss, the latter truly never having been there. However, making such an a priori assumption necessarily shields the precedence form of the ontogeny criterion from exceptions and reduces it to a tautology of no particular interest to scientists.

Patterson (78, p. 25) concluded that “von Baer’s second law—less general characters are developed from the most general, and so on—is not presented as a guide to sequence in the whole organism, but to sequence in structural relations,” and as such is not known to be violated in vertebrates. While dedifferentiation may or may not occur naturally among vertebrates, it seems to be a particularly common phenomenon among acariform mites (46, 97, 98, 99), and occasionally it may be observed in other nonvertebrate animals and in plants as well (57).

Nelson’s law—the more general character is primitive and the less general, advanced—is an example of the commonality form of the ontogeny criterion, which can be rendered in terms of frequency of either particular developmental states or whole ontogenetic sequences (102). The relative importance of these two formulations in discerning plesiomorphic from apomorphic has yet to be fully explored. Nelson’s law casts von Baer’s first and second laws then in
evolutionary terms (Figure 1). Such formulations of the ontogeny criterion are not theory neutral (70) but are explicitly evolutionary (general equals primitive). Thus, Nelson’s effort (68, p. 325) to distinguish the ontogeny criterion from all other protocols, in not requiring an “a priori assumption of relationship (or a higher level phylogeny),” pertains only to particular propositions about history, not to the existence of a species genealogy and some more general theories of evolution and speciation. Patterson (78, p. 25) has pointed out that “the amnion and allantois, characterising Amniota, a subgroup of Tetrapoda, appear before any feature of the latter group, and even before features of Gnathostomata such as jaws and limb girdle,” and in so doing demonstrate that von Baer’s first law “fails as a sequential guide to hierarchical ranking of features.” In fact, even von Baer (100) interpreted the chick’s yolk sac as a nonterminal addition.

There is considerable disagreement as to the relative merits of the outgroup and ontogeny criteria in phylogenetic inference. For example, Nelson (68) argued that the ontogeny criterion was to be preferred because the transformation of one state into another can be directly observed, while the outgroup utilizes transformations that are unobservable, and consequently he denigrated it as an “indirect method.” The quality of direct observation of ontogenetic transformation is rarely realized in practice because of the technical difficulties of studying living systems, particularly over those periods when significant transformation occurs. Moreover, in some instances the transformation is simply masked by blastema or suppression of gene expression (e.g. enameloid teeth in living birds). In addition, one is forced to delimit discrete states arbitrarily in applying the ontogeny criterion and to infer their comparability on the basis of the overall similarity of the whole organisms (see the section on homology below). The ontogeny criterion, as formulated by Nelson, was also claimed to be falsifiable, but unfalsified, and to require fewer assumptions than the outgroup criterion, thereby being least protected from falsification by ad hoc propositions. These claims by Nelson and others have been critically evaluated and found deficient (23, 34, 57). Phylogeneticists who assume a law-like quality for the ontogeny criterion must claim, by prior knowledge, that nonrecapitulatory evolution does not occur (57).

While both the outgroup and ontogeny criteria correctly interpret the

---

**Figure 1** The sense of the ontogeny criterion. The more general state appears earlier in ontogeny than does the less general (or the more specialized), and it is relatively more primitive.
phylogenetic information in a case of terminal addition (Figure 2, $H_1$), the two protocols do not make the same predictions for the five nonrecapitulatory modes of evolution (84). Figures 2–6 illustrate the circumstances under which use of the ontogenetic criterion produces incongruent (terminal deletion) or less informative phylogenetic hypotheses (terminal substitution; and nonterminal addition, deletion, and substitution). Equally important, Figures 2–6 illustrate that the ontogeny criterion is insensitive to nonrecapitulatory modes of evolution and the outgroup criterion is required if heterochronic events are to be interpreted correctly (39; see below). Nelson (68, p. 344) also recognized these effects, particularly in the case of neoteny (that is, paedomorphism; see also 50, pp. 107–8). However, he asserted that neoteny was also a potential falsifier of the outgroup criterion. In a further attempt to protect the ontogeny criterion, he proposed that no character is truly lost, rather it is transformed—“Apparent loss may be an indication that the characters and transformations are merely poorly understood and, consequently, wrongly defined.” Brooks & Wiley (23) have concluded that because of this insensitivity and because the ontogeny criterion does not provide any unique information on character evolution or phylogeny,
it is less decisive than the outgroup criterion. Further, Maddison et al (63, p. 83) have argued that application of the outgroup criterion ensures a better global fit to data, because “parsimony is measured simultaneously over the ingroup and outgroups.” Their conclusion appears to hold even though it has been shown that the outgroup criterion goes beyond the weak assumptions of the likelihood justification for parsimony (89, 90; E. Sober, unpublished information).

Over the years, Haeckel’s (48, Vol. 2, p. 300) biogenetic law has been given special consideration in evolution (41, 52). The biogenetic law assumes that evolution occurs by terminal addition—new adult forms are added to the preexisting ontogeny, and, with further comparable evolution, are pushed back in descendants and become nonterminal (Figure 2, H1). In effect, Haeckel simply revived the Meckel-Serres concept of recapitulation, “grafted [it] onto a system of divergent development [sensu von Baer],” and interpreted its cause in terms of Darwinian evolution (74, p. 6). There are certain difficulties with the biogenetic law. Since it pertains to whole organisms, it can only recapitulate an unbranching part of phylogeny, because a given individual’s ontogeny is a strictly linear sequence. Viewed in this light, the biogenetic law predicts an array of common ancestral forms, but any particular ancestor of two or more taxa can only be deduced from the congruent pattern of two or more recapitulated ontogenies (78, p. 23). Further, it has long been recognized, even by

Figure 4  Given only the ontogenies of the ingroup taxa (1–3), H1 is the best fitting phylogenetic hypothesis. However, if the outgroup criterion is employed (i.e. taxon X is taken into account) phylogenetic hypothesis H2 is preferred. The latter suggests evolution by nonterminal addition. See Figures 2–3 for further explanation.

Figure 5  Given only the ontogenies of the ingroup taxa (1–3), H1 is the best fitting phylogenetic hypothesis. However, if the outgroup criterion is employed, i.e. taxon X is taken into account, phylogenetic hypothesis H2 is preferred. The latter suggests evolution by nonterminal deletion. See Figures 2–3 for further explanation.
Haeckel (48), that paedomorphosis and nonterminal events occur and may disrupt the perfect parallelism between ontogeny and phylogeny indicated by recapitulation (30, 32, 41).

In their study of the bivalve *Chesapeake*, Miyazaki & Mickevich (65) claimed to have found Haeckelian recapitulation in several characters, as well as paedomorphosis in other features. These authors concluded (p. 397) that the occurrences of paedomorphosis “are not so abundant as to overwhelm the predictive power of Haeckel’s law. . . . The overall morphology of later *Chesapeake* taxa is more adult in form than the ancestral taxa.” In contrast, Brooks et al (21, 22) found little evidence of recapitulation in their quantitative study of larval digenean platyhelminths, most of the evolution being nonterminal additions. While a high degree of congruence existed between the larval and adult characters of these parasitic flatworms (which we interpret to mean that one type of data has no primacy over the other), neither the larval nor the adult characters were capable of fully resolving the phylogeny. While such studies are too few to permit generalizations on the relative frequency of the different modes of evolution, it is clear that one must investigate the entire life cycle in order to obtain maximum historical information. A study of adults can only sample terminal novelties.

HOMOLOGY

Only homology justifies the presumption of common ancestry and sistergroup relationships, and only synapomorphy ranks as a hypothesis of homology. According to Stevens (93), a homology is a similarity that delimits a group of organisms alone and that was found in their common ancestor. In this definition, similarity is a 1:1 correspondence or identity. Conclusions as to homology usually seem to be preceded by two distinctly different evaluations—preliminary and final testing (76). Preliminary testing involves judgments as to conjunction and overall similarity (topographic, compositional, and ontogenetic). The test of conjunction is failed when supposed homologues are found in
the same organism. Congruence is the final test, and in this context it means correlation among synapomorphies at a particular level of taxonomic generality. This last form of evaluation is especially important because the various types of nonhomology (homoplasy) may pass the tests of conjunction and/or similarity but finally fail the test of congruence. In phylogenetic systematics, maximum congruence is achieved by applying parsimony algorithms to data that have been pretested and passed, namely synapomorphies. While all homologies are group-delimiting synapomorphies, not all synapomorphies need be homologies.

Nelson (68, p. 335) claimed that "the mode of development itself is the most important criterion of homology," but more recently [69, p. 43; see also Roth, (86)] he stated that homology is directly presented by ontogeny, "vertical relations," whereas it is only indirectly presented by outgroup comparisons, "horizontal relations." This vertical-horizontal dichotomy serves to emphasize that homology can be conceptualized in two ways, transformationally and taxically, respectively. The taxic view is that presented above (76, 93). Transformational homology is rather different. For example, ontogeny is considered an explanation of homology, instead of a criterion involved in its recognition (64), and the vertical relation presupposes that the transformation will not be so great as to remove confidence in the hypothesis of homology (see also 70, p. 161). Stevens (93, p. 402) also underscored the fact that the transformation may remain unclear, even with a well-corroborated phylogenetic hypothesis and detailed developmental evidence. Patterson (76) noted that while transformational homology may be important and interesting in its own right, it is not necessary in understanding taxic relationships. However, in the following sections we will consider how descriptions of "continuous" ontogenetic changes in form, based upon hypotheses of topographic and transformational similarity, may be used to infer taxic relationships.

HETEROCHRONY

Gould (43) was the first to emphasize the surprising lack of effort directed toward a synthesis of two fundamental and longstanding concepts in evolutionary morphology: heterochrony, the change in relative timing of expression as a function of age, and allometry, the change in relative shape as a function of size. Since the publication of Gould’s landmark volume, a number of important steps have been taken to unify these concepts in terms of ontogenetic and evolutionary development.

Heterochrony is an alteration in the timing of ontogeny such that the relative time of appearance or rate of development of a feature is either accelerated or retarded (43). De Beer (31, 32) and Zimmermann (105), among others, had attempted to categorize such changes qualitatively and in terms of correspond-
ing stages in ancestors and descendants. Alberch et al (8), in an important step towards a unified view of developmental biology and evolution [but see Alberch (6)], developed a formalism intended to structure current ideas on the transformation of continuous size and shape. Their model was based on ancestor-descendant relationships and quantified heterochronic phenomena in terms of the three variables that Gould (43) had thought important in recapitulation and paedomorphosis: size, shape, and age. The model, based on the study of "growth trajectories," has been refined in several ways, discussed here and in the following section.

In the formalism of Alberch et al (8), the growth of an organism is described as an ontogenetic trajectory, a record of the physical appearance of the organism from inception to maturity or death (Figure 7A). This trajectory is an idealization of the growth of an individual, but in practice characterizes the mean developmental path of a population or species. The trajectory is purely descriptive in that the organism's size and shape are represented by a collection of mensural traits visualized as a moving point in a geometric "size-shape" space. When the trajectories for a descendant taxon and its ancestor are viewed simultaneously within the same morphological space (Figure 7B), variation in three controlling parameters of development—time of onset of growth (α), time of cessation of growth (β), and growth rate (k)—can account for two of the classic categories of heterochronic phenomena: paedomorphosis, including neoteny, progenesis, and postdisplacement; and peramorphosis (recapitulation), including acceleration, hypermorphosis, and predispacement. Alone or in combination, and without the necessity of identifying discrete, comparable

![Figure 7](image-url)

**Figure 7** A model of heterochronic comparisons. (A) The growth trajectory of an individual described as a function of three parameters: α, the onset of growth; β, the cessation of growth; and k, growth rate. At any point on the trajectory the individual is described by vectors of size and shape traits. (B) A comparison of two trajectories, for an ancestral taxon (t_a) and its descendant (t_d), showing simultaneous changes in all three parameters. (C) A comparison of three trajectories for two sister taxa (t_x and t_y) and an outgroup (t_o). The sister taxa share a synapomorphy, the forward displacement of time at offset, resulting in an extended developmental period (hypermorphosis). Taxon t_x is distinguished by an apomorphy, a decrease in developmental rate (neoteny) resulting in an adult shape convergent with that of the outgroup.
developmental "stages", these morphological changes can account for the
terminal modifications often inferred when comparing ontogenetic sequences
to hypothesized higher-order phylogenetic relationships (45).

Fink (39) refined the operational value of this model to phylogenetic studies
by emphasizing two key points. First, because hypotheses of heterochrony are
comparative statements about evolutionary processes, some phylogenetic con-
text, however tentative, is a fundamental prerequisite (35). Because ancestors
cannot be identified and studied directly, hypotheses of heterochrony are not
two-taxon statements (ancestor-descendant) but rather, minimally three-taxon
statements [the taxon of interest, a putative sister taxon, and appropriate
outgroup(s) (Figure 7C); but see (26)]. Combining the trajectory-comparison
model with Fink's phylogenetic context, the six "pure" cases of recapitulation
and paedomorphosis (8) can theoretically be identified. For example, if the
ontogenetic trajectory of a species is compared to those of its sister taxon and
one or more outgroups, hypermorphosis appears as the extrapolation of the
plesiomorphic trajectory (t_y in Figure 7C), producing a preterminal phenotype
similar to the terminal phenotype of the outgroup (t_o). Neoteny (t_x in Figure 7C)
appears as a decrease in mean developmental rate, as compared to t_y, over the
same duration. The remaining pure cases can be described in similar terms (8,
39). More realistic mixed cases, in which two or three of the parameters change
simultaneously, are more difficult to describe as simple discrete transformation
series but may be used in quantitative descriptions of ontogenetic change (as in
the following section).

Fink's (39) second major point is that the features of organisms are
heterochronic, not the organisms themselves (32, p. 8). Thus, some characters
or character complexes may display one form of heterochrony, while other
characters may display no heterochrony at all or even a different form. While
begging the familiar question of what constitutes a taxonomic character, this
point also emphasizes the care with which mensural descriptors of size and
shape must be defined and treated if they are to be biologically informative (17,
18, 94).

GROWTH, ALLOMETRY, AND HISTORICAL
INFERENC

The ontogenetic formalism of Alberch et al (8; see also 4), with its emphasis on
comparisons of growth trajectories, made explicit an aspect of development of
potential value for systematics: the description of ontogeny as a unit phenom-
enon (14, 15, 96). This approach contrasts with the more traditional methods of
using ontogenetic information in phylogenetic inference, in which the units of
comparison are the "character" and the "stage." In the first case, discrete
taxonomic characters identified in juveniles or adults are polarized for
phylogenetic inference (from plesiomorphic to apomorphic) by observing their state-transformations during development or determining their generality, according to the “ontogeny criterion” discussed above, or on the basis of the outgroup criterion. In the second case, ontogenies are divided into discrete stages, usually defined by the onset or cessation of particular developmental events, and are then compared in terms of the presence, absence, and relative positions of stages among taxa (7, 32, 38, 56, 66). Observed differences among ontogenetic sequences are characterized by additions, deletions, and substitutions of stages.

The trajectory model (3, 8) is much more general than traditional methods in the amount of information it subsumes about the organism because it can incorporate information about specific “growth laws” and the timing of differentiation events with respect to size and age (60, 96). The dynamic life cycle is taken to be the unit of study and is characterized simultaneously by two types of developmental changes: relative and absolute growth rates of various morphological features (change with age), and covariation among morphological traits, both continuous and discrete (change with size). Additional information is provided by the mapping of particular developmental events (e.g. ossifications, physiological/behavioral/ecological shifts, timing of onset of particular structures) onto growth curves (3). Recent attempts to expand or make more specific the original formulation have been basically along two lines: (a) replacing the single-parameter notions of size and shape with multivariate size-factors and shape-factors in order to describe and distinguish the two aspects of form more carefully; and (b) implementing particular quantitative models of change in size and shape over time, in order to describe in specific terms the changes in morphological form of individuals or populations with time.

Growth (Figure 8A) may be described by any of several quantitative models (83, 88). The use of such models accomplishes more than merely “simplifying certain practical aspects of curve fitting” (8); such models provide the basic descriptive variables with which comparisons among taxa can be made. What is required for this purpose is a growth model that is sufficiently precise to characterize individual trajectories, yet sufficiently flexible to facilitate interspecific comparisons. Creighton & Strauss (27), for example, reformulated a generalized model of negative exponential growth (101) so that the controlling parameters in the heterochronic trajectory model (α, β, and k) could be explicitly and uniquely defined for each character examined, along with a number of other variables. For phylogenetic comparisons among cricetine rodents they defined α to be age at birth or at some extrapolated prenatal body size, β as the age at which a fixed proportion of the adult size is attained, and k the mean growth rate from α to β. Although other (e.g. sigmoidal) growth models may be more exact when extrapolated into early ontogeny (58, 59), the
simpler negative exponential model describes postnatal growth extremely well and, more importantly, can be uniquely represented by $\alpha$, $\beta$, and $k$ with no loss of information. Thus, it leads naturally to comparisons of linear trajectories (Figure 7B, C), although other linearizations are possible (8, 25). The data base for an analysis of heterochrony would be a set of such curves, one for each metric trait for each taxon, from which are derived corresponding values of $\alpha$, $\beta$, and $k$. We may superimpose upon these any additional information of potential phylogenetic value, such as samples of age-specific sizes or growth rates or the points at which particular developmental events occur (Figure 8A; 3, 4, 27).

Use of such trajectories allows specification of relative onsets, offsets, and rates among all descriptors of form, and may be employed to assess heterochrony trait by trait. If the mensural characters adequately describe body form, then together they will also describe change in shape over time by their relative allometries (18, 24, 94). Alberch et al (8) discussed change in form in terms of quantitative but unspecified shape traits ($\sigma$). Character ratios (7) and bivariate allometric relationships (47) have been used to study heterochronic shape-changes. Bookstein et al (18) described how the various shape-changes summarized by $\sigma$ may be made most informative by contrasting patterns of regional shape-change over the form via continuous biorthogonal analyses, which allow quantitative comparisons of form by revealing the principal directions of shape-change with respect to corresponding anatomical landmarks, in the manner of Thompson (96). They also outlined a methodology for comparing series of forms from two or more ontogenies without respect to explicit time correspondences by examining shape-changes between stages in the two taxa independently. By using suitable sets of metric traits selected on the basis of

---

**Figure 8**  (A) A model of age-specific growth of a mensural character $x$. Open circles indicate sizes at ages 1, 2, \ldots (arbitrary scale). Closed circles indicate timing of developmental events. (B) The allometric relationship of the same character with general size. Circles represent the corresponding age-specific sizes (open) and developmental events (closed) from the growth curve of A.
topographic similarity (e.g. “trusses,” 94, 95), we can approximate this continuous description of form-change with multivariate size-factors and shape-factors (17). Any available information about age-specific growth (Figure 8B) or the occurrence of developmental events may be superimposed on this plot so that it portrays both allometry and timing. The information on this trait may also be summarized by two algebraic elements: an allometric scalar and a vector of age-specific coefficients. Such descriptions hold even if ontogenetic development is diphasic, triphasic, or discontinuous (Figure 9), to the extent that the hypothetical transformational homologies defining the trait are corroborated by available anatomical evidence. Indeed, the presence of such discontinuities, especially if subtle, may lead to the discovery of caenogenetic changes. The composite of age- and size-dependent changes across all traits, added to information about the timing of discrete events, comprises a reasonably complete description of the ontogeny (an interrelated set of ontogenetic trajectories) in terms of size-change and shape-change as a function of age. The description may be portrayed geometrically as a series of linear plots, or algebraically as a vector of allometric coefficients and matrixes of growth and timing coefficients.

Thus, by extending the basic formalism of Alberch et al (8) with specific growth models and with geometric descriptors of size and shape, we can study both heterochrony and morphological integration (18, 43, 73) in terms of dissociable character suites at the levels of ontogeny and phylogeny. Also, within the same context, it should be possible, using geometric descriptions of form, to ask what specific kinds of ontogenetic trajectories would give rise to unobserved (theoretical) morphologies (18, 82). Such prior analyses of change remain largely unexplored (5) but might be used, for example, to distinguish cases in which theoretically possible forms are unobserved due to historical or

---

**Figure 9** Allometric relationships of a mensural character with general size, showing (A) one (diphasic) and (B) two (triphasic) allometric shifts during development. Open circles represent corresponding age-specific sizes derived from growth curves.
sampling accident, from those in which specific sets of developmental trajectories result in physically impossible morphological configurations.

For historical inference, developmental patterns can be compared using discrete or continuous characters. In the same way that we examine adult organisms for relatively discrete morphological differences, we may find that ontogenetic trajectories possess particular traits that vary "significantly" among taxa in relation to variation within, traits that can be polarized with respect to outgroups and used in phylogenetic comparisons and inferences. Ontogenetic characters in this sense are descriptors of the ontogenetic program as a whole [rather than arbitrarily derived discrete stages (6)] and are selected after description of ontogeny rather than before. Alternatively, quantitative data on growth and allometry may be analyzed as continuous characters using conventional parsimony algorithms (36). The expectation in either case is that phylogenetic hypotheses based on ontogeny will be consistent with, and more informative than, corresponding hypotheses based on adult morphology alone (65), and differences between the two should tell us something about heterochrony. Creighton & Strauss (27) found that a Wagner tree for cricetine rodents based on age-specific growth and allometry for the first eight weeks of age corresponded in major topological outline with results from previous studies based on adults. Similarly, growth patterns before and after larval transformation in a group of fishes (95) yielded a Wagner network consistent with previous phylogenetic hypotheses based on adults, but more highly resolved. In both cases the ontogenetic data give a "dynamic" assessment of homoplasy, because cases in which similar adult morphologies may have been derived from different ontogenetic pathways can be detected (87). An understanding of the processes leading to developmental change depends on knowing the ways in which particular ontogenies differ and, ultimately, the ways in which their different developmental programs are mediated (9, 16, 44, 49, 81, 85).

If hierarchical patterns of organization among taxa are resolvable by ontogeny it is because changes in development become irreversible at some level. But, to be phylogenetically useful, patterns of ontogenetic development need not be any more "conservative" (6) than any other conventional taxonomic characters. In fact, the question of where in the course of evolution ontogenetic changes in growth and time become hierarchic and indicative of relationship subsumes the problem of when and in what way morphological traits become irreversible. A more intensive comparative study of development will be needed before we can know which aspects of ontogeny are historically constrained and which are free to vary, and to what extent patterns of constraint are similar in even closely related groups (42, 81). If parallels between ontogeny and phylogeny are of major significance, it is because heterochrony and allometry are pervasive phenomena in evolution (43, 81).
ACKNOWLEDGMENTS

John Beatty, David Hull, and Elliott Sober provided thoughtful comments on philosophical issues. The University of Michigan Systematics Discussion Group and Morphometrics Study Group served as the forums in which many of our ideas were first developed, and we thank the participants. Sara Fink helped to clarify several of our points of view. The National Science Foundation, grants BSR-83-04581 (to A. G. Kluge) and BSR-83-07719 (to R. E. Strauss and F. L. Bookstein), provided financial assistance.

Literature Cited

growth and form in animals. Q. Rev. Biol. 41:131–90
50. Laird, A. K., Barton, A. D., Tyler, S. A.


64. Mason, H. L. 1957. The concept of the flower and the theory of homology. *Madroño* 14:81–95


78. Patterson, C. 1983. How does phylogeny differ from ontogeny? See Ref. 42, pp. 1–31


