Biological Journal of the Linnean Society (2001), *73:* 169–186. With 7 figures
doi:10.1006/bijl.2001.0535, available online at http://www.idealibrary.com on IDEAL®

Heterochrony revisited: the evolution of developmental sequences

KATHLEEN K. SMITH

Departments of Biology and Biological Anthropology and Anatomy, Box 90338, Duke University, Durham, NC 27708, USA

Received 17 July 2000; revised and accepted for publication 6 March 2001

The concept of heterochrony is a persistent component of discussions about the way that evolution and development interact. Since the late 1970s heterochrony has been defined largely as developmental changes in the relationship of size and shape. This approach to heterochrony, here termed growth heterochrony, is limited in the way it can analyse change in the relative timing of developmental events in a number of respects. In particular, analytical techniques do not readily allow the study of changes in developmental events not characterized by size and shape parameters, or of many kinds of events in many taxa. I discuss here an alternative approach to heterochrony, termed sequence heterochrony, in which a developmental trajectory is conceptualized as a series of discrete events. Heterochrony is demonstrated when the sequence position of an event changes relative to other events in that sequence. I summarize several analytical techniques that allow the investigation of sequence heterochrony in phylogenetic contexts and also quantitatively. Finally, several examples of how this approach may be used to test hypotheses on the way development evolves are summarized.

ADDITIONAL KEY WORDS: development – craniofacial – mammal – marsupial.

INTRODUCTION

The concept of heterochrony has accompanied attempts to link evolution and development for well over a century. In the broadest sense, heterochrony refers to a change in the relative timing of developmental events in one species relative to an ancestral species. In practice, heterochrony is almost always studied as change in the relative timing of events in two related taxa, as true ancestral-descendent comparisons are rarely available. The term was first defined by Haeckel to describe cases in which an ontogenetic sequence of events did not recapitulate the phylogenetic sequence (Russell, 1916). The general issues of recapitulation and heterochrony were examined by a number of workers in the early part of the 20th century, including Garstang, Sewertzoff and others (see Gould, 1977; Russell, 1916). Heterochrony began to acquire its modern definition in the work of de Beer who, in a series of books (1930, 1940, 1951, 1958), attempted to bring development into the evolutionary synthesis of the mid-20th century. De Beer's aims were to separate the concept of heterochrony from the idea of recapitulation and to demonstrate how many different kinds of changes in developmental timing could produce evolutionary change. Two elements were particularly important in de Beer's discussions. First, he argued that evolutionary innovations could occur at any stage in an organism's life history. Likewise, changes in the timing of events could also occur at any point in ontogeny. These observations refuted recapitulation and also broadened the context for heterochrony. Second, de Beer discussed potential mechanisms for heterochronic changes arising out of the increasing understanding of genetics and developmental biology. He therefore attempted to move heterochrony from a static description to a means to link developmental and evolutionary mechanisms (Ridley, 1985).

For the most part, the concept of heterochrony entered the active vocabulary of modern evolutionary biology in 1977 with S. J. Gould's book *Ontogeny and Phylogeny*. Like de Beer, much of Gould's focus was on the concept of recapitulation but, unlike de Beer who aimed to refute recapitulation, Gould was interested in exploring the reasons for its persistence. He reviewed the history of the concept of heterochrony and

E-mail: kksmith@duke.edu

0024–4066/01/060169+18 $35.00/0

proposed a streamlined terminology for various kinds of heterochrony. Gould's treatment of heterochrony was enormously influential and has been authoritative in defining the way the concept is used today. Several aspects of Gould's conception of heterochrony have become fixed in current usage even though in many ways they represent departures from previous views.

First, Gould solidifies a fundamental change in Haeckel's original definition of heterochrony – a "curious odyssey" in Gould's words. For Haeckel, heterochrony was one type of pattern in which the ontogenetic sequence was not parallel to the phylogenetic sequence. He termed this lack of recapitulation false history or "caenogenesis". (The other kind of false history defined by Haeckel was heterotopy, referring to a shift in the germ layer from which a set of cells originates.) Later workers, including de Beer, separated heterochrony from its recapitulatory implications and defined heterochrony as any change in the timing of a developmental event relative to other events. Heterochronies could occur at any time in development and heterochrony had no necessary relationship to adult morphology or the phylogenetic sequence. Gould discarded many of de Beer's categories and in so doing applied the concept in a much narrower context. Gould, in distinct contrast to de Beer, only considered a shift in the timing of an event to be heterochrony if it produced a parallel between ontogeny and phylogeny. This view limits the cases of heterochrony to recapitulation, when the sequence of events in ontogeny is directly parallel to the sequence of characters in phylogeny, and reverse recapitulation (paedomorphosis) when the "ontogeny of the most remote ancestor goes through the same stages as a phylogeny of adult stages read in the reverse order" (Gould, 1977: 215). Gould discards other cases of timing shifts as not producing parallels and, therefore, not qualifying as heterochrony. For Gould, heterochrony is the mechanism that produces parallels between ontogeny and phylogeny.

The second major change in the concept of heterochrony arising from Gould's treatment concerns the types of timing shifts considered. Haeckel viewed heterochrony largely in terms of the sequence of developmental events or stages. De Beer likewise focused on the order of structural changes or events and stated, "the strengths of the internal factors of development can vary and exert their effects at different rates with the result that the time of appearance of a structure can be altered. To this shifting along the time-scale the term heterochrony is applied. It is thus possible for two organs to reverse the order of their appearance in successive ontogenies" (1958: 34). To de Beer, allometric growth, termed heterogony, was only one subset of the kinds of heterochrony possible. Gould however, focused almost entirely on rates of relative growth. In particular, and unique to Gould, is the definition of allometric growth not as growth, but as differentiation. Gould cites Needham's (1933) arguments on dissociation and argues that heterochrony is the dissociation of fundamental processes in ontogeny. Needham developed his concept of dissociation around three fundamental processes of development: growth (increase in spatial dimensions and weight), differentiation (increase in complexity and organization) and metabolism (chemical changes). Gould modifies these fundamental ontogenetic processes to growth (i.e. size), maturation (i.e. sexual maturation) and development (differentiation and allometric growth). Gould, by and large, ignores differentiation and defines allometric growth as the third fundamental process "because it allows me to formulate a compellingly simple clock model for heterochrony" (1977: Fig. 1).

Gould therefore shifted the focus on heterochrony from the relative timing of developmental events to changes in size and shape (Table 1). The almost exclusive focus on size and shape changes as the important heterochronic phenomenon was a significant redirection of the concept by Gould and is almost universally accepted today. Heterochrony has become virtually synonymous with allometry. "Heterochrony was born because patterns of covariant growth permeate ontogeny and phylogeny" (McKinney, 1999: 149).

Gould's approach was further generalized and extended in the influential paper 'Size and shape in ontogeny and phylogeny' by Alberch *et al.* (1979). This paper begins with a discussion of the importance of shifts in the sequence of discrete events but, as with Gould, the analysis restricts itself to relative growth, i.e. systems whose "functions are characterized by their morphological appearance. For example, we diagnose reproductive maturity by the size and shape of the reproductive organs" (1979: 298). They go on to define the ontogenetic trajectory as the trajectory traced by size and shape changes because "it is a complete record of the physical appearance of the system" (1979: 300). However, unlike Gould, who presents "an essentially static" model of heterochrony, Alberch *et al.* focus on growth processes. Heterochrony is defined in terms of shifts in specific processes such as change in onset, cessation, or rate of growth, rather than results (Fig. 1; Table 1; see Klingenberg, 1998 for an excellent review of the consequences of these different approaches).

The work by Gould (1977) and by Alberch *et al.* (1979) generated an enormous new interest in the relation of ontogeny and phylogeny and led to significant increases in our understanding of the developmental and evolutionary patterns and process involved in the change in relative body size and shape.

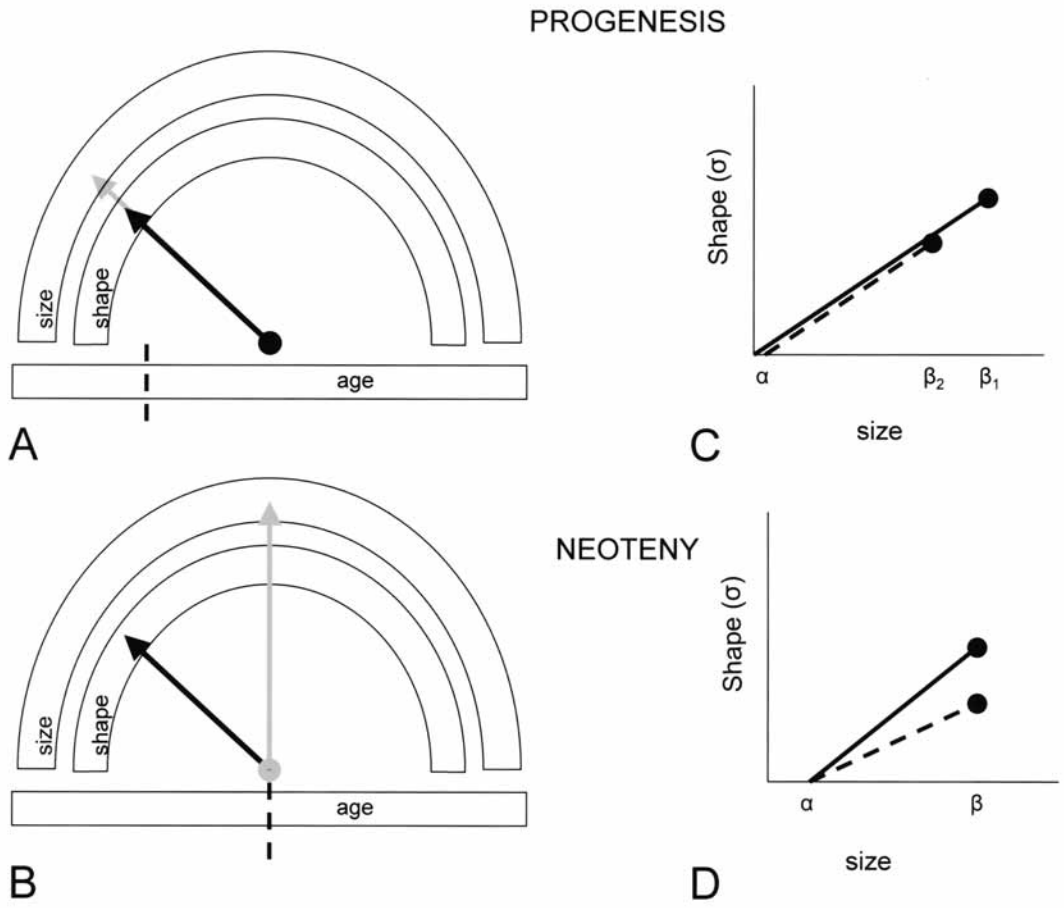

Figure 1. Models of growth heterochrony, showing shifts in the relation of size, shape and time. A & B from Gould (1977); C & D from Alberch *et al*. (1979). A & B, heterochrony is characterized as a dissociation of one of three parameters: size, shape and age at which a reference point is achieved, in this case sexual maturity. In A the age of maturation is early, while the relation of size and shape are constant. In B the age of and size at maturation are unchanged, but shape changes more slowly. C & D, heterochrony expressed as change in one of three control parameters: α, onset of growth; β, offset of growth; k, rate of growth. In C the descendent (dotted line) shows early offset of growth (β_2) relative to ancestor (β_1) although the relation of size and shape (slope of the line) remain unchanged. In D the rate of shape growth relative to size growth (slope of dotted line) in descendent is decreased, although the time of onset and offset are unchanged. A & C both show a phenomenon referred to as progenesis; B & D show a phenomenon referred to as neoteny.

Although Gould's definition of heterochrony, with emphasis on relative size and shape, was a significant change from previous conceptions, it rapidly became fixed as the context in which to view heterochrony and therefore the way to link ontogeny and phylogeny. Numerous papers following this approach appeared in the 1980s and 1990s (for recent reviews see Gould, 2000; Hall, 1992, 1999; Klingenberg, 1998; McKinney, 1988, 1999; McNamara, 1995, 1997; Raff, 1996; Raff & Wray, 1989; Zelditch & Fink, 1996).

This approach to heterochrony, here termed growth heterochrony (called deBeerian heterochrony by Raff & Wray, 1989), is limited as a general way to view changes in the timing of development. First, because of the emphasis on changing size and shape, size is often taken as a surrogate for time. In part this substitution is due to practical considerations, as many such studies have no access to data on time. Therefore, many studies of 'heterochrony' are not comparisons of shifts in timing but instead purely allometric studies. In many cases size is an appropriate surrogate for age, but there are cases in which this substitution obscures patterns or is theoretically questionable (e.g. Blackstone, 1987a,b; Emerson, 1986; Godfrey & Sutherland, 1995a,b; Klingenberg, 1998; Klingenberg & Spence, 1993; Roth, 1984; Snow, Tam & McLaren, 1981). Size, rate of development, and shape may evolve independently. Unless the appropriateness of size as a

Table 1. The terminology of growth heterochrony as presented by Gould (1977) and Alberch *et al.* (1979). Paedomorphosis refers to cases in which the outcome is less relative growth; peramorphosis to cases in which the outcome is more relative growth; cases of giantism and dwarfism do not include relative growth. See also Figure 1 for examples of graphic representation of these phenomena. Note the explicit characterization of all types of heterochrony as either recapulatory or reverse recapitulatory phenomena, a concept that arose with Gould, and was included in the characterization of Alberch *et al.* (1979). Note also that only limited types of changes are characterized by growth heterochrony approaches. The terms used to describe growth heterochrony (e.g. neoteny, hypermorphosis, etc.) are by and large global terms, not particularly useful for the kinds of issues discussed in sequence analysis. These terms are not discussed in this paper

Heterochronic phenomenon	Gould (1977) characterization	Alberch *et al.* (1979) control parameter	Relative size change	Phylogenetic effect
Progenesis	Size/shape relation constant, early maturation	Early growth offset	Paedomorphosis	Reverse recapitulation
Neoteny	Shape slowed relative to size and maturation	Decrease shape growth rate	Paedomorphosis	Reverse recapitulation
Postdisplacement		Later growth onset	Paedomorphosis	Reverse recapitulation
Proportional dwarfism	Size slowed relative to shape and maturation	Decrease size growth rate		Reverse recapitulation
Hypermorphosis	Size/shape relation constant, late maturation	Later offset growth	Peramorphosis	Recapitulation
Acceleration	Shape increased relative to size and maturation	Increases shape growth rate	Peramorphosis	Recapitulation
Predisplacement		Early growth onset	Peramorphosis	Recapitulation
Proportionate giantism	Size increased relative to shape and maturation	Increase size growth rate		Recapitulation

substitute for age is explicitly tested in a given case, it is not certain that a comparative study of size and shape is in fact a study of heterochrony – a change in the timing of developmental events.

Second, the emphasis on size and shape and the analytical approaches of growth heterochrony have limited the focus of studies of heterochrony to global (whole body) events and relatively late processes (Hall, 1992, 1999; Raff & Wray, 1989). Although many changes between closely related species arise through patterns of relative growth, it can be argued that the most critical events in development occur early when size and shape may not be the appropriate reference points. These early events include such things as the initial differentiation and patterning of the major elements of the body, appearance of segmental and regional identity, patterns of regulatory gene expression, induction and signalling cascades, cell and tissue specification and differentiation, and the differentiation of skeletal elements and organ systems. Because these events do not rely on size and shape parameters they are by definition excluded from the kinds of analyses presented by Gould (1977) and Alberch *et al.* (1979). These are increasingly the kinds of events examined in studies of evolution and development (e.g. Collazo, 1994; Cubbage & Mabee, 1996; Hall, 1984a; Jeffery & Swalla, 1992; Langille & Hall, 1989; Richardson, 1995; Slack & Ruvkun, 1997; Smith, 1995; Swalla *et al.*, 1994; Velhagen, 1997; Wray, 1995; Wray & McClay, 1989). Changes in the relative timing of such events are likely to be critical in producing evolutionary change.

Furthermore, the analytical approaches provided for growth heterochrony are primarily useful for analysis of "one or a few body parts at a time" (Alberch *et al.*, 1979: 300). Neither Gould's clock model nor the growth models of Alberch *et al.* provide methods to analyse changes in the interaction of multiple events in multiple taxa. Therefore it is difficult to study important questions concerning modularity or integration in a broad phylogenetic context.

In order to apply the concept of heterochrony to a wider range of developmental processes, approaches must be developed that: first, provide for an appropriate means of standardization not dependent on external size or time measures and, second, offer analytical approaches that allow the examination of shifts in timing of a variety of kinds of developmental events

and processes. In this paper I present an alternative route to the analysis of heterochrony, which I refer to as developmental sequence analysis. The focus on sequence does not imply that different processes or mechanism are at work, but instead emphasizes an alternative way to describe developmental events and standardize and analyse developmental trajectories. Previous authors have examined changes in developmental sequence (e.g. Hanken & Hall, 1984; Larsson, 1998; Mabee & Trendler, 1996; Richardson, 1995; Velhagen, 1997). However, by and large, these analyses have lacked any kind of unifying theoretical or analytical approach. Here I describe analytical tools that allow the consideration of different data sets and different kinds of questions than possible with growth heterochrony studies. Below I will discuss the basic approach of developmental sequence analysis, provide brief summaries of examples where the approach has been applied, discuss newly developed analytical techniques and probe the types of questions open to this kind of analysis.

DEVELOPMENTAL SEQUENCES

As pointed out by de Beer, changes in developmental sequence arise from the same processes examined by growth heterochrony – change in onset, offset or rate of process. A study of sequence heterochrony differs from an analysis of growth heterochrony in two important ways. The first difference concerns the way embryonic time is conceptualized and the second the kinds of events that are analysed. One problem with previous studies of heterochrony arises from the lack of an appropriate measure for interspecific comparison of developmental time (e.g. Blackstone, 1987a,b; Hall & Miyake, 1995; Raff & Wray, 1989; Reiss, 1989; Roth, 1984). Hall & Miyake (1995) list four potential criteria to standardize comparisons across species: (1) discrete landmarks of maturation (e.g. metamorphic stages); (2) measures of growth, (e.g. size or rate variables); (3) measures of chronological age (either absolute or scaled) or (4) attainment of morphological stages. Many of these criteria are inadequate for intraspecific comparisons. All present significant theoretical and practical difficulties as measures for interspecific comparisons (Hall & Miyake, 1995, and references therein). A growing literature has sought to elucidate how embryos measure time and properly schedule developmental events (e.g. Ambros & Horvitz, 1984, 1987; Cooke & Smith, 1990; Dale & Pourquié, 2000; Gorodilov, 1992; Hall & Miyake, 1995; Howe *et al.*, 1995; Johnson & Day, 2000; McClung, Fox & Dunlap, 1989; Palmeirim *et al.*, 1997; Pourquié, 1998; Power & Tam, 1993; Reiss, 1989; Satoh, 1982; Stern & Vasiliauskas, 1998; Yasuda & Schubiger, 1992). This literature demonstrates that different organisms, at different stages in their life history, track developmental time using many different types of measures, including number of cell cycles, time oscillations, DNA replication cycles, cell volumes and 'lists' of scheduling events. Many developmental events are by nature dependent on a sequence of prior events as found with induction, signalling cascades, or expression of regulatory genes. No data support any single external metric (e.g. clock time or size) for interspecific comparison.

I propose that the developmental trajectory be modelled as any sequence of morphogenetic events. Instead of an external size or time standard, the beginning of the trajectory is when event a occurs and the end is when event n occurs. Ontogeny in multiple organisms is standardized as this series of events from a to n. Heterochrony is recognized when the sequence position of an event changes relative to the other events. In effect, this approach assumes that the most critical 'clock' for the embryo is not an external or internal time base, but the completion of a series of morphogenetic events and processes. Using sequence as a method of standardization proposes that an embryo in part recognizes it is 'time' for event n when events k, l and m have been completed.

Focus on events, rather than size and shape, means that any part of the developmental trajectory and changes in the timing of any kind of process or event can be examined by sequence heterochrony techniques. Examples of the kinds of events that can be analysed include the onset of expression of specific genes at specific sites, the differentiation of specific tissue types, the establishment of specific connections or interactions, the appearance of distinct morphological elements, numerical or quantitative landmarks, or the attainment of specific stages of morphological differentiation. Furthermore, multiple kinds of events may be incorporated and integrated in the single analysis. Thus the analysis of sequence can easily incorporate elements located on all of the axes on the model of development presented by Raff & Wray (1989): local–global, early–late, and molecular–morphological. These are increasingly the kinds of events of greatest interest to evolutionary developmental biologists.

The use of sequence as a basis for standardizing interspecific comparison rather than size or absolute time is useful for at least three reasons. First, it is a biologically important and functionally relevant method to "compare the development of different embryos on the basis of the timing of decision making . . . such a coupling of heterochrony to the epigenetic, hierarchical organization of embryonic development would put us on the path to understanding heterochrony as a process" (Hall, 1992: 212). Traditional methods use one point of reference and attempt to scale other events relative to an external base. Using

a developmental sequence as a criterion of standardization substitutes multiple reference points for the time base. "Although all developmental events are subject to temporal change in evolution, every system for classifying and analyzing heterochronies assumes that polarities of individual heterochronies can be determined. The most convincing support for polarity is to show a consistent offset in the timing of the altered developmental event relative to the timing of several other reference events. In general, the more such reference events, the more strongly corroborated the polarity of the heterochrony becomes" (Raff & Wray, 1989: 421–422).

Second, analytical techniques have recently been developed to investigate sequence changes that allow the consideration of types of data and questions not possible with current methods of growth heterochrony (see below). Sequence approaches broaden the applicability of the concept of heterochrony. Finally, the developmental sequence, and changes to that sequence, have been important parts of the concept of heterochrony from its origins in Haeckel to its more modern formulation in de Beer. Alberch *et al.* (1979: 298) state, "the importance of sequence in developmental events cannot be overemphasized". Developmental sequences have lost their place as a central component of the concept of heterochrony only over the past 20 years with the current, almost exclusive, focus on size and shape.

AN EXAMPLE: DEVELOPMENT IN THERIAN MAMMALS

Several authors have attempted to analyse heterochronies in developmental sequence (e.g. Dunlap & Sanchiz, 1996; Hanken & Hall, 1984; Hufford, 1995, 1996; Irish, 1989; Larsson, 1998; Mabee, 1993; Mabee & Trendler, 1996; Richardson, 1995; Strauss, 1990; Velhagen, 1995, 1997; Wake & Hanken, 1982). I provide here an extended example from my own work to illustrate some of the above points. Marsupial and placental mammals possess fundamentally different reproductive and life history strategies. Marsupial young are born after an extremely short intrauterine period and are characterized by a particularly short period of organogenesis. Most maternal investment occurs during an extended postnatal period via lactation. Eutherians have relatively long intrauterine gestation periods and even the most altricial eutherian is well developed relative to any marsupial newborn. However, in marsupials the highly altricial neonate must independently find and recognize the teat, attach and feed. Therefore, a number of specific elements, including the forelimbs and oral apparatus are developmentally advanced relative to the general embryonic state of the neonate. The phenomenon of developmental acceleration and deceleration (heterochrony) of some organs in marsupials has been appreciated for well over a century, with focus generally on the forearms, tongue and facial regions (e.g. Clark & Smith, 1993; Cockburn, 1989; Filan, 1991; Gemmell & Selwood, 1994; Hill & Hill, 1955; Hughes & Hall, 1988; Klima, 1987; Maier, 1987; Nelson, 1987; Smith, 1994; Tyndale-Biscoe & Renfree, 1987, and references therein). However there have been few detailed analyses of the specific heterochronies of multiple events across therian mammals. The approaches of growth heterochrony cannot be usefully applied to analyse the specific shifts in development for a number of reasons. First, there is a clear mosaic of processes – some are accelerated and others are delayed. Second, most of the most interesting patterns involve shifts in the early differentiation of structures and not size and shape changes. Third, overall development in marsupials and placentals is so different that no useful criterion of standardization has been defined to compare development across these clades. Finally, the most important questions involve the interaction of elements, which cannot be addressed by existing growth heterochrony methods.

Smith (1996, 1997, 2000; Nunn & Smith, 1998) examined changes in the sequence of events in craniofacial development in marsupials and placentals in the period between the early differentiation of the forebrain to the onset of ossification of the last bone in the cranium. These studies document the relative timing of 28 different developmental events in the central nervous system, muscular system and cranial skeleton. Complete developmental series of four marsupials and six placentals were examined. The criterion of standardization was the sequence in which the 28 events occurred. These studies introduced two techniques (see below) to examine shifts in sequence of multiple events in multiple taxa. Three overarching questions were the foci of these studies. First, which specific elements appear relatively accelerated in marsupials as a consequence of the necessity for independent function at an embryonic state? Second, how does the overall pattern of craniofacial development differ in these animals, e.g. are elements relatively delayed as a result of the advancement of some structures? Finally, what does the study of differential acceleration and delay of craniofacial elements reveal about the integration of craniofacial structures during development?

The studies showed that overall the developmental sequence of craniofacial structures was largely conserved within a set of elements. There was, for example, little difference in the order in which bones began ossification, either within or between the major groups; there were no differences in the order of events within

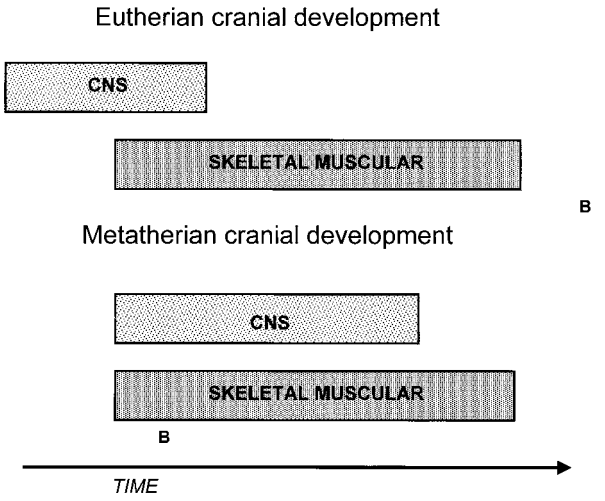

Figure 2. A summary of the relative timing of development of craniofacial features in metatherians and eutherians. The upper box (light stipppling) for each clade represents events in central nervous system (CNS) development; the lower box (dark stippling) represents events in the development of skeletal–muscular systems. The arrow at the bottom represents time. The letter B represents the approximate time of birth in each group. For comparison, the two taxa were scaled and aligned relative to the timing of the events of the skeletal–muscular system; the CNS in each was plotted relative to that. This illustration demonstrates that even though time was not a component of the original analysis, the relative timing and rate of developmental events may be studied if sequence is used as the basis for standardization. See Smith (1997) for more detail on the characters and the analysis.

the central nervous system. However the analysis revealed a major shift in the timing of differentiation of cranial systems. In marsupials, the development of a number of bones and muscles are advanced relative to the central nervous system. This heterochrony, or shift in relative timing of the differentiation of two major systems, is revealed by the fact that these events occur earlier in the sequence of development in marsupials relative to eutherians. Specifically, in all eutherians, several events of CNS differentiation were always the first events in the sequence. In marsupials a number of events of the musculo-skeletal system appeared early in the sequence and occurred before or at the same time as the first events examined in the CNS (Fig. 2).

METHODS TO STUDY CHANGES IN DEVELOPMENTAL SEQUENCE

In the absence of specific analytical methods, most studies of sequence heterochrony have been graphic.

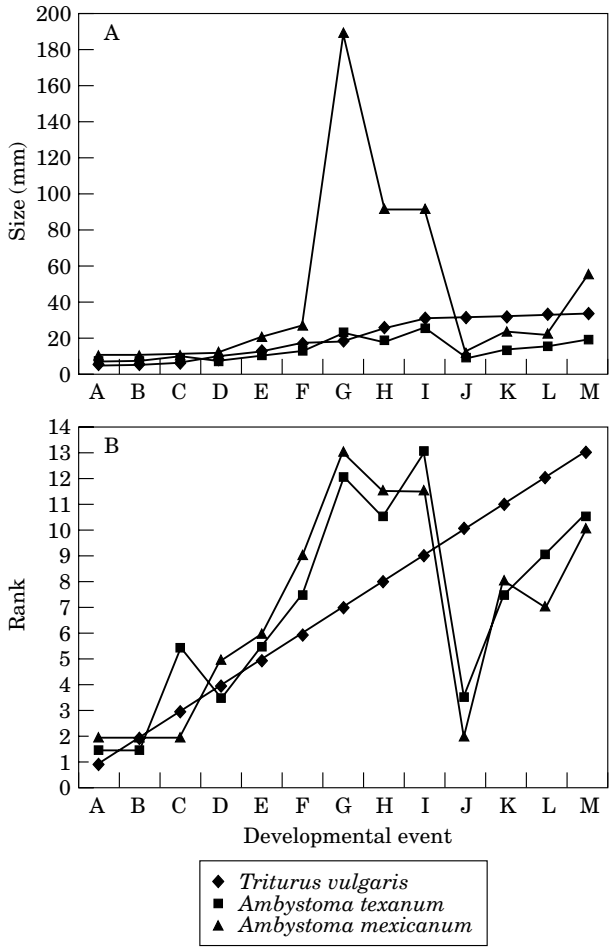

Figure 3. Developmental sequence evolution in three salamanders. Redrawn from Alberch *et al.* (1979). A, drawing of the events (ossification of cranial bones) relative to size. B, drawing of events relative to rank (sequence order). In B it is clear that the sequence of ossification is quite similar in the two species of *Ambystoma*, a pattern that was somewhat obscured by the plot of events by size. Further the specific events that occur early and occur late in *Ambystoma*, relative to *Triturus vulgaris*, are easy to discern. The data used to construct these plots are in Table 2.

For example, Alberch *et al.* (1979) plot a series of events relative to size in three species of salamanders (Fig. 3A). Similar types of illustrations appear in a number of analyses such as those of Wake (1989), Hanken & Hall (1984), Strauss (1990), Dunlap & Sanchiz (1996) and Richardson (1995). While graphical illustrations can demonstrate simple patterns, they are not particularly useful when the taxa being compared develop at different rates or sizes, or many taxa or many traits are analysed. Recently several new types of analytical approaches have been used by a number of authors.

Table 2. Size and rank data for salamander species; data from Alberch *et al.* (1979). Event rank demonstrates the conversion of raw size data to rank data. Events are aligned in the order they occur and given a rank between 1 and *N*, where *N* equals the total number of events in the sequence (in this case 13). Ties (events that occur at same time) are given the mean rank of the events at that time. For example in *A. texanum*, events A & B are the first two events and occur together. They are given the rank (1+2)/2, which equals 1.5. In *A. mexicanum* events A, B, C and J occur first, at size 11, and are given the rank (1+2+3+4)/4, which equals 2. A large number of ties indicates that the sequence is not sufficiently resolved

	Event size												
Data set/taxon	A	B	C	D	E	F	G	H	I	J	K	L	M
Event size													
Triturus vulgaris	5	5.5	7	10	12	17	18	25	30	31	31	32	33
Ambystoma texanum	7	7	10	8	10	12	22	18	25	8	12	15	18
Ambystoma mexicanum	11	11	11	12	20	26	188	91	91	11	23	22	55
Event rank													
Triturus vulgaris	1	2	3	4	5	6	7	8	9	10	11	12	13
Ambystoma texanum	1.5	1.5	5.5	3.5	5.5	7.5	12	10.5	13	3.5	7.5	9	10.5
Ambystoma mexicanum	2	2	2	5	6	9	13	11.5	11.5	2	8	7	10

CONSTRUCTION OF A DEVELOPMENTAL SEQUENCE

In all the analytical approaches discussed here, the first step is to construct a developmental sequence. For each taxon in the analysis there must be an ordered set of embryos in which a series of comparable development events can be defined. Within each taxon the embryos can be ordered by age, size, stage or some combination, and each set can be ordered by a different criterion. In taxa with considerable intraspecific variability, a conserved or representative sequence must be constructed. The specific events in the sequence will vary depending on the specific hypothesis to be tested but may include events from many different kinds of processes. For example, one might test the hypothesis that an evolutionary innovation is due to the early development of a series of morphological elements. In this case the sequence would include the elements hypothesized to develop early, as well as a suite of independent elements that occur before, during and after the elements of interest. This hypothesis must be tested in the context of a phylogeny that includes taxa with and without the evolutionary innovation.

From an ordered set of embryos, a sequence of developmental events can be tabulated. For some of the analyses discussed below each event in the sequence is given a rank order, from $1-N$ (where N is the total number of events in the sequence). Table 2 shows the data from Alberch *et al.* converted into ranks. These data are plotted in Figure 3B. A comparison of Figures 3A and 3B highlight the kinds of information revealed by a sequence analysis. Figure 3B demonstrates clearly that the developmental sequence in the two *Ambystoma* species is virtually identical, an observation that is somewhat masked in Figure 3A, which focuses on relative size. The sequence analysis further reveals that the major differences in these taxa relative to Triturus include the late occurrence of events G, H and I, and the early occurrence of events J, K, L and M. Again, this pattern is not immediately obvious from Figure 3A. The plots together provide a complete view on the important differences and similarities in development among these species.

PHYLOGENETIC APPROACHES

The first analytical procedure was designed to compare developmental sequences in an explicit phylogenetic framework and was independently developed by Mabee (Mabee & Trendler, 1996), Smith (Smith, 1996, 1997) and Velhagen (1995, 1997). The aim of this method is to identify developmental events that shift relative timing in a hierarchically structured phylogenetic array, a cladogram. The technique converts sequence data into characters, which are assigned various character states that represent changes in the sequence. These character states may then be plotted on independently derived phylogenies using software such as MacClade (Maddison & Maddison, 1992) to examine phylogenetic patterns of change in developmental timing.

Briefly, the method of analysis is as follows. Once the developmental events have been standardized in a sequence, the data are converted to a form that can be analysed phylogenetically. Smith (1996, 1997) and Velhagen (1995, 1997) constructed a matrix so that the timing of each event was compared with every other event in that species. For example, if six events,

	EVENT PAIR														
TAXON	EF	EG	EH	EI	EJ	FG	FH	FI	FJ	GH	GI	GJ	HI	HJ	IJ
Triturus vulgaris	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ambystoma texanum	0	0	0	0	2	0	0	0	2	2	2	2	0	2	2
Ambystoma mexicanum	0	0	0	0	2	0	0	0	2	2	2	2	1	2	2

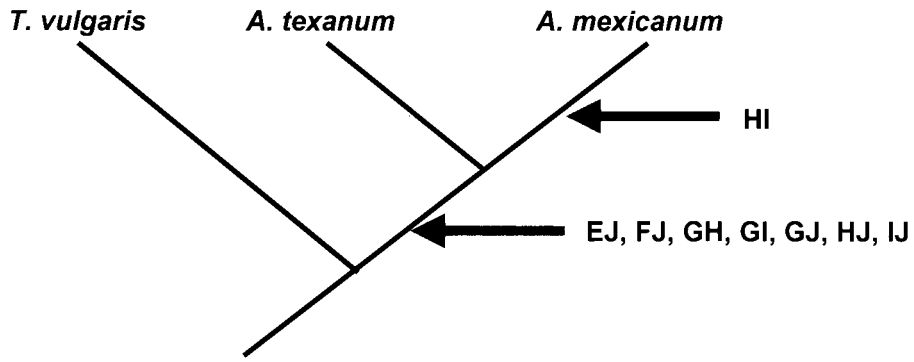

Figure 4. Demonstration of the event pair analysis. A, the sequence units (event pairs) of events, E, F, G, H, I and J. Character state 0 indicates the first event in the pair occurs before the second event in the pair. Character state 1 indicates the two events occur at same size (the surrogate for time in this study). Character state 2 indicates first event in pair occurs after second event in the pair. The overall sequence is ranked according to the sequence in *Triturus vulgaris*, therefore, all event pairs have character state 0 in that taxon. Character states EJ, FJ, GH, GI, HJ, and IJ are shifted in *Ambystoma* species; specifically event J is shifted anteriorly in the sequence (occurs early relative to E, F, G, H and I) and event G is shifted later in the sequence (occurs late relative to H, I and J). In *A. mexicanum* event I may occur early (or event H late), however, the occurrence 'same time' may be due to insufficient resolution. See Smith (1997) for more detail on the method and interpretation of the analysis.

E, F, G, H, I and J, were studied, the pairs would be EF, EG, EH, EI, EJ, FG, FH, and so forth (Fig. 4A). These pairs were called sequence units by Velhagen (1997). Each sequence unit (pair of events) is then assigned one of three character states that represent the relative timing of these two events in a taxon. For example, for sequence unit EF the following character states would be used (Smith, 1997): character state 0 if event E occurred before event F; character state 1 if E and F occurred at the same time, and character state 2 if E occurred after F. These three character states allow the depiction of shifts in timing of each developmental event relative to every other event. The characters (relative timing within a sequence unit or pairs of event) can be plotted onto a phylogeny (Fig. 4B).

In the study of marsupial and placental mammal craniofacial development discussed above (Smith, 1996, 1997) the developmental sequence of 28 events in four marsupials and five placentals was examined. In each species, these 28 events were converted to a matrix of 378 sequence units (event pairs); each event pair was assigned one of the three character states above. The character states were plotted on a phylogeny of marsupial and placental mammals in order to determine which event pair had character state distributions that: (1) were conserved across Theria (marsupials + placentals); (2) distinguished marsupials and placentals (i.e. had a consistent state within each clade that differentiated the two clades), (3) distinguished individual taxa within either clade, and (4) had no discernible phylogenetic pattern. In this manner specific heterochronies characterizing the two groups were revealed. Of the 378 event pairs, 65% were essentially uniform across the Theria, 15% distinguished eutherian and metatherian mammals (i.e. 56 event pairs had one character state in eutherians and another, essentially non-overlapping, character state distribution in metatherians) and 20% had no discernable pattern. Of the 56 event pairs that distinguished the groups, 80% involved advancement in marsupials of a musculo-skeletal element relative to an element of the central nervous system. The remaining event pairs represented relative ad-

vancement of certain musculo-skeletal elements that are functional at birth, relative to other elements of the musculo-skeletal system.

QUANTITATIVE APPROACHES

The second major kind of analytical approach applied to developmental sequences is quantitative. The first step in the quantitative approach is to convert the sequence of events into a rank ordering of events so that the data are numerically equivalent. The earliest event is given the number 1; the last the number *N*, equal to the total number of events in the series. Events that occur at the same time are given the mean rank of all the events occurring at that time (or stage) (Table 2; see Nunn & Smith, 1998, for further discussion). Heterochrony is demonstrated by analysis of the numerical changes in the rank position of specific events. Again, this method allows comparison of data sets for which no external time standard can be established, and also allows analysis of any kind of developmental event. While this method does not allow phylogenetic analysis to the same extent as the previous approach, non-parametric statistical tests may be used to test hypotheses about differences in timing of specific events among groups.

Several authors (e.g. Mabee & Trendler, 1996; Strauss, 1990) used rank correlation coefficients to attempt to assess overall sequence conservation in various groups. Nunn & Smith (1998) used a similar measure, Kendall's coefficient of concordance (*W*), which was designed to test the similarity of ranked series. However, it is somewhat difficult to interpret the significance of overall similarity of developmental sequence (rank order) for a number of reasons (discussed by Nunn & Smith, 1998). For example, it is not clear what null hypothesis should be established. Statistically rejecting null hypotheses of either zero association (two sets of developmental sequences are completely different) or complete similarity (two sets of developmental sequences are identical) are not likely to be biologically informative. If the taxa are related, then the appropriate null hypothesis is some degree of similarity in developmental sequence, however, how much overall change or similarity should be expected is not obvious.

Larsson (1998) used a different correlation approach to test hypotheses on the relationship between the sequence of events in phylogeny and ontogeny. Larsson's method compares the sequence of appearance of events in ontogeny and phylogeny, as represented by the fossil record, and consists of the following steps. First, a phylogenetic hypothesis is generated to represent the history of a particular study taxon. Fossil taxa are included. This tree is collapsed into a linear series of taxa, the phylogenetic trajectory, which represents the ancestry of the specific terminal taxon of interest. The diagnostic characters and character states of the taxa in the phylogenetic trajectory are assigned a rank order from 1 to *N*. This ranking represents the phylogenetic sequence of the acquisition of characters and character states. After the phylogenetic sequence has been defined, an ontogenetic series of the study taxon is obtained. The appearance of each of the phylogenetically diagnostic characters or character states is mapped in the ontogeny to produce an ontogenetic sequence. This ontogenetic sequence is also ranked from 1 to *N*. Bivariate plots and Spearmann rank coefficients are used to test the association of the ontogenetic and phylogenetic sequences (see Larsson, 1998, for details). If the sequence in ontogeny and phylogeny are conserved, the ranks of the specific characters in the two sequences will by highly correlated.

Larsson used this method to study the appearance of characters in ontogeny and phylogeny in the evolution of the crocodilian secondary palate. His aim was to identify features that were developmentally integrated, predicting that such characters would retain particular patterns of association in phylogeny and ontogeny (e.g. Wimsatt, 1986). His data suggested that the premaxilla and maxilla are developmentally independent from a complex that includes the palatine, pterygoid and ectopterygoid bones. The characters of the maxilla and premaxilla exhibited little correlation in ontogeny and phylogeny, while the pterygoid–palatine complex and a complex of features involved with the choanae exhibited a positive, statistically significant correlation. As pointed out by Larsson, this method provides a precise means to test hypotheses such as von Baer's law, Wimsatt's (1986) hypothesis of generative entrenchment, or more general hypotheses on the relation between phylogenetic appearance and ontogenetic appearance of characters (e.g. Nelson, 1978, see below).

Another kind of quantitative approach uses ANOVA rather than correlation methods to examine changes in the rank order (relative timing) of individual elements in a developmental sequence. This method, proposed by Nunn & Smith (1998), divides taxa into groups (in this case marsupials and placentals) and applies ANOVA to investigate which characters exhibit a greater difference in rank position between groups than within groups. The results of the analysis on the data set of therian mammals discussed above demonstrated that 11 of the 28 elements differ statistically in marsupials and placentals (Fig. 5). This approach uses a common statistical method to identify specific elements that are either advanced or delayed in one group relative to another. It may be applied to both intraspecific comparisons (i.e. to experimental manipulations, different litters or broods, etc.) or to

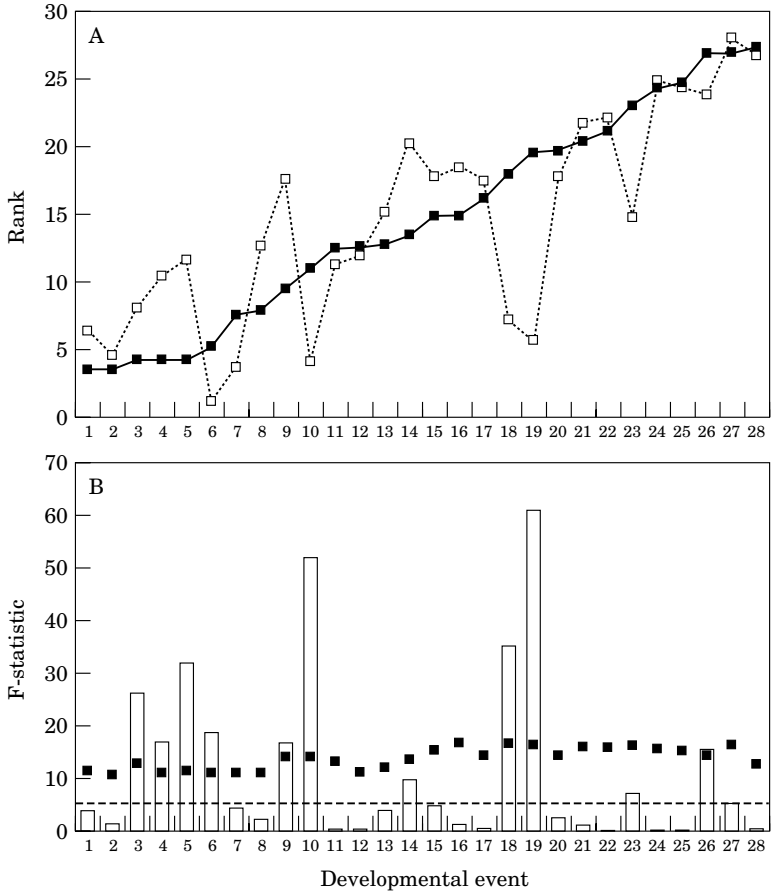

Figure 5. Results of the analysis of sequence shifts in marsupials and placentals. A, mean rank of four marsupials (■) and six eutherians (□). The order of events is arranged by the mean rank of marsupials. B, results of ANOVA. (□) *F*-statistic; (– – –) the statistical calculation of $P<0.05$; (■) $P<0.05$ resulting from simulation to correct for phylogenetic non-independence. See Nunn and Smith (1998) for discussion of methods. The combination of the two charts allows the identification of which events are significantly different in the two clades, and also the polarity of the shift. It is important to note that, even though the mean rank may be shifted (i.e. events 1, 7, 8, 15), the difference between the two groups may not be statistically significant, given the variance within the groups. Key to events: 1, cartilage in the basicranium; 2, alignment of myoblasts in the tongue; 3, ossification in the dentary; 4, ossification in premaxilla; 5, ossification in maxilla; 6, evagination of telencephalon; 7, pigment in retina; 8, striations in muscles; 9, secondary palate closes; 10, olfactory nerve contacts bulb; 11, tooth buds; 12, cartilage on condyle; 13, frontal bone ossifies; 14, exoccipital ossifies; 15, jugal ossifies; 16, craniofacial muscles organized; 17, squamosal ossifies; 18, primary lens cells fill lens vesicle; 19, thalamus and hypothalamus; 20, parietal ossifies; 21, alisphenoid ossifies; 22, basioccipital ossifies; 23, layering in cortex; 24, basisphenoid ossifies; 25, malleus and incus separate from Meckel's cartilage; 26, membrane bones meet over skull roof; 27, periotic bone begins ossification; 28, joint capsule forms. See Smith (1997) for more detail on the events.

interspecific comparisons. In the latter case methods that assess significance levels corrected for phylogenetic non-independence should be applied (Nunn & Smith, 1998). The phylogenetic and the ANOVA approaches to the Smith marsupial/placental data set provided congruent results. For example, in the ANOVA, eleven characters were significantly different in rank order. These 11 characters were either one or both of the elements in 55 of the 56 event-pairs that distinguished marsupials and placentals in the event-pair analysis.

To summarize, the above methods allow examination of patterns of heterochrony not accessible by techniques of growth heterochrony. They have a number of features in common. First, aligning the ontogeny by sequence allows comparison of timing of events in development across taxa for which an external means of standardization cannot be established because it is

unknown or incomparable. Second, the methods allow comparison of the timing of events that are not characterized by size or shape criteria and are designed to allow assessment of multiple events in multiple taxa. These kinds of analyses could easily incorporate events at the cellular, molecular or genetic level. The three different approaches are each best suited for different kinds of data sets and questions. The phylogenetic event-pair mapping approach provides a means to analyse changes in many taxa for which a well-established phylogeny exists. Further it allows the comparison of patterns of change at multiple levels within the phylogeny. The correlation approach presented by Larsson best compares two developmental sequences and is ideal for testing the conservation of a given series of events in ontogeny and phylogeny. Finally, the ANOVA approach is best for testing changes in developmental sequence in two groups of organisms with multiple members. These may be two clades with multiple species, or two groups within a single species (such as treatment groups in an experiment, different litters, different populations and so forth). This approach can provide statistical tests of hypotheses of patterns of change of specific elements, or groups of elements (i.e. developmental modules), relative to the rest of the developmental trajectory.

DISCUSSION

MISCONCEPTIONS ABOUT DEVELOPMENTAL SEQUENCE

The approaches of growth heterochrony discussed in the introduction of this paper document changes in the onset, cessation, or rate of growth relative to a standardized metric such as size or time. In contrast, sequence heterochrony analyses study changes in the timing of any kind of events relative to other morphogenetic events. This is not a new concept, as de Beer (1930, 1940, 1951, 1958), Gould (1977) and Alberch *et al*. (1979) all discussed the relevance of sequence changes to evolution. However the popularity of the concept of heterochrony in the last 20 years has centred almost entirely on Gould's focus on size and shape. Relative to the mass of literature on growth heterochrony, sequence heterochronies have received little or no attention. In large part this may be because, until very recently, methods to analyse sequence heterochrony, or to compare developmental sequences across taxa, have not existed. But perhaps more importantly, discussions of sequence heterochrony have been dominated by preconceptions and misunderstandings that have diverted attention from the issue of developmental sequence evolution. Two major issues have clouded our understanding.

First, a sequence of developmental events is commonly confused with a sequence of developmental stages. The notion that development proceeds as a series of discrete, conserved stages in which an embryo possesses a number of specific characters is long standing and pervasive. The idea that stages are directly comparable across taxa has its origins in Haeckel and also in the pre-Darwinian comparative embryologists, and is demonstrably false. Using stages as a criterion of standardization does avoid problems of determining absolute age and comparing animals with different rates of development and is perhaps the best metric for intraspecific comparisons. However, stages can only roughly be compared across taxa, and virtually all detailed comparative developmental studies show that a regular progression of stages with detailed equivalence across taxa at higher levels simply does not exist (e.g. Richardson, 1995; Richardson *et al*., 1997, 1998). Developmental sequence analysis compares individual events and makes no a priori assumption about the linkage of events in conserved stages.

Second, perhaps because sequence was the focus of Haeckel's attempts to parallel ontogeny and phylogeny, most discussion about the evolution of developmental sequences has focused on whether developmental sequences are, or should be, recapitulatory (see Alberch, 1985; Raff & Wray, 1989, for discussion). As pointed out by Gould, the idea of recapitulation has been pervasive. Recently, there has been a great deal of attention paid to the use of ontogeny in phylogenetics and, in particular, the use of ontogenetic data to determine character polarity in phylogeny reconstruction (e.g. Alberch, 1985; de Queiroz, 1985; Fink, 1982; Kluge, 1985, 1988; Kluge & Strauss, 1985; Mabee, 1993, 1996; Meier, 1997; Nelson, 1978; O'Grady, 1985; Patterson, 1996; Rieppel, 1990, and references therein). With some exceptions (e.g. Mabee, 1993 and Meier, 1997) this literature largely discusses the issue in principle, or by providing general examples, with little detailed testing of the data. As proposed here, analysis of developmental sequence makes no a priori assumptions on the conservation of developmental sequences, the frequency of terminal additions, or about how sequences should evolve. Instead an analysis of developmental sequence data provides means to empirically test such hypotheses on sequence conservation or change. Through these means the longstanding debates (or misunderstandings in the words of Alberch, 1985, and Raff & Wray, 1989) may actually be resolved.

TIME AND SIZE DO MATTER

A focus on developmental sequences does not necessarily eliminate size or time from the analysis of heterochrony. However, neither size nor time are the basis for comparison, so that if data are not available or are incomparable, analysis can proceed. If data on

absolute or relative size or time are available, they may also be examined in the context of developmental sequences. The absolute timing of events is critical, for example, in life history analysis and many ecological situations (e.g. Emerson, 1986; Gould, 1977). Likewise, the absolute or relative size at which an event occurs may have important effects on morphogenetic processes (e.g. Alberch *et al.*, 1979; Hanken & Wake, 1993). For example in Figure 3 the same events (from Alberch *et al.*, 1979) are plotted by size and by sequence rank. Figure 3B highlights the essential sequence similarity of the two *Ambyostoma* species; Figure 3A highlights the fact that the latest events occur at very large sizes in *A. mexicanum*. There are likely important biological consequences of both sets of relations. In this case size data are available and comparable, but in many other cases size data are not available. However, when examining the sequence of events, even if data on absolute time is not available or comparable, the relative percentage of the sequence occupied by any subset of events can be examined, which provides information on relative rates of development (Fig. 2).

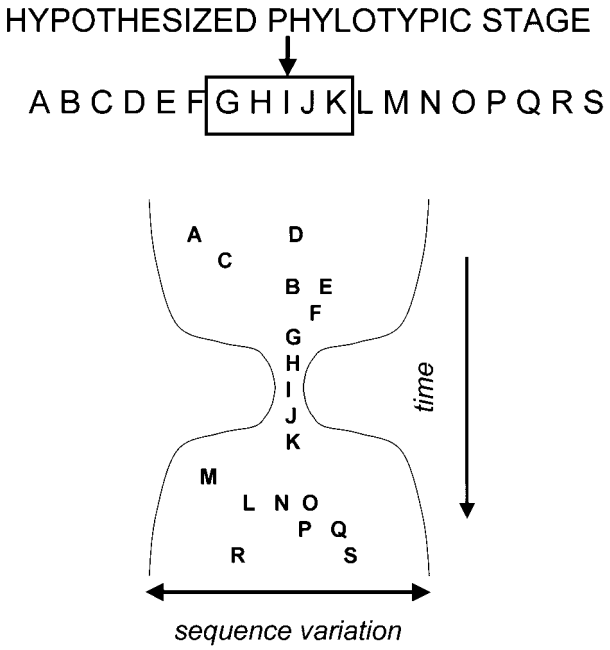

Figure 6. Testing the hypothesis of the hourglass model of the evolution of development. If the phylotypic stage is hypothesized to encompass events G–K, then the sequence of these events should be conserved in a broad taxonomic study. Specifically, events G–K should retain a conserved sequence: (a) relative to each other, and (b) relative to the sequence as a whole. The hourglass model would predict that events before and after the hypothesized phylotypic stage should show significantly more variation than those within the phylotypic stage.

THE CONSERVATION OF EARLY DEVELOPMENT

My argument throughout this paper has been that a focus on developmental sequence heterochronies does not imply different mechanisms or processes are in place, nor is it intended to replace more traditional growth heterochrony methods. Rather, it is a way to extend the concept and analysis of heterochrony so that additional data sets and different kinds of questions may be examined. In this final section I briefly discuss a few of the types of hypotheses that may be tested with the above methods.

One of the most persistent hypotheses about evolution and development is that, in general, early development is conserved relative to late development. The roots of this hypothesis lie in the notion that given the hierarchical nature of development, changes in early events will have far-reaching effects while changes in later events will be much more localized in their impact. Numerous empirical studies have shown that early development can be quite diverse, with little evidence of conservation, or impact on adult morphology (see Raff, 1996, for review). There still exists, however a pervasive assumption that within a group, such as vertebrates, development goes through a relatively conserved phase often called the phylotypic stage. The concept of a conserved stage is often modelled as an hour glass, in which early and late development are variable, but the stages near the establishment of a body axis are quite conserved (e.g. Duboule, 1994; Raff, 1996; Slack, Holland & Graham, 1993). Richardson and colleagues (1995; Richardson *et al.*, 1997, 1998) have provided empirical data to help move the debate from the general (somewhat exaggerated) concept of a uniform phylotypic stage to a specific analysis of the patterns of vertebrate development. Richardson (1995) has provided quite convincing graphical and descriptive evidence that there is enormous underlying variability in the supposed conservative phylotypic stage.

The evaluation of these issues could be aided by the analytical approaches discussed above. In the case of the hourglass model, the test would require a broad range of developmental events spanning early, prephylotypic, phylotypic and later stages over a broad phylogenetic sample. The conservation hypothesis would predict that early events would exhibit little sequence variation across taxa while variation in later events would be significant. The hourglass hypothesis would predict that the events immediately surrounding and including the phylotypic period exhibit little sequence variation, while earlier and later events exhibit significant variation (Fig. 6). It would be particularly useful to include both morphological characters and data on the timing of gene expression in such an analysis to test the concept of the zootype

(Slack *et al.*, 1993). Either the phylogenetic or quantitative approaches discussed above could provide such an analysis. The former would be useful if there were many taxa at multiple levels in the phylogenetic hierarchy; the latter would be more useful if the taxa were arrayed in a few major clades. Such studies are in progress by Richardson and colleagues (Richardson, pers. comm.).

In connection with this hypothesis, the events examined in Smith (1996, 1997) began at or near the 'phylotypic' stage (e.g. differentiation of the neural tube and of facial structures), and extended well into the period of emergence of species-specific characters. However, there was no evidence that early events were conserved, instead events that differ significantly extend through almost the entire developmental period (see Fig. 5).

The debate on the usefulness of ontogeny to polarize characters or character states in phylogeny is simply a subset of this issue, stating that characters and character states should appear in a hypothesized phylogeny in the order in which they appear in ontogeny. The types of analyses discussed here such as the phylogenetic event pair mapping approach or the correlation method used by Larsson (1998) would allow empirical tests of this hypothesis.

DEVELOPMENTAL INTEGRATION AND DEVELOPMENTAL MODULES

One issue that has long been tied to discussions on evolution and development is the degree to which elements may be dissociated during development (e.g. Needham, 1933; Gould, 1977). Most recently Raff (e.g. Raff & Raff, 2000; Raff & Wray, 1989; Raff, 1996) has cast this question in terms of developmental modules. The general argument is that some degree of integration in development is expected, but if certain sets of elements (modules) are dissociated from the rest of the sequence, then those events may undergo differential selection and evolve independently. As Raff states "without dissociation, there would be no way for a novel element to be introduced into a developmental pathway, regardless of its source. No feature can be subtracted from or added to an ontogeny with too much linkage to allow some dissociation to occur. Those features of development that cannot be dissociated may be the features that define the phylotypic stage, and thus the conserved elements of the body plan". (1996: 337). Developmentally independent modules would be predicted to exhibit change in timing in the sequence across phylogeny.

The hypothesis of modularity lies within the more general hypothesis that development is highly integrated. These issues of integration and dissociation are simply alternatives of the same basic questions.

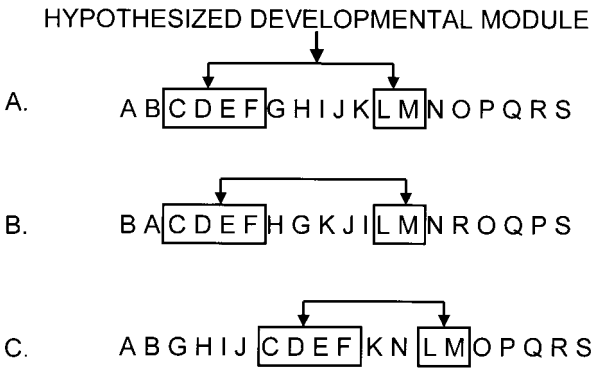

Figure 7. Testing hypothesis of developmental modules. In this case the hypothesized module consists of events C, D, E, F, L and M. Note that a module need not contain events that are always adjacent in the sequence, as a sequence may be defined to contain many events involved in multiple processes. The prediction is that the events of the module will retain a conserved internal sequence despite sequence changes in surrounding elements (B), and will shift in concert relative to the entire sequence (C).

Across phylogeny can sets of events in development be shown either to: (1) change the timing of their development independent of other events (modularity, dissociation) or (2) exhibit exceptional conservation despite changes in timing of surrounding events (integration). Both hypotheses can be investigated by study of changes in developmental sequence in large numbers of characters in large numbers of taxa (Fig. 7). Sequence analysis can test whether a specific set of events (the hypothesized developmental module) retain conserved relations as a unit despite their dissociation from the rest of the developmental sequence. Elements in a developmental module need not necessarily be contiguous within a series of events as the series of events in a study can be any set defined by an investigator. However, if events are linked by some degree of causality or integrative mechanism (Alberch, 1985), then they should retain a conserved sequence relative to each other, despite changes in the sequence among surrounding events or shifts in the timing of the module as a whole.

Smith (1996) explored patterns of craniofacial integration and dissociation through the analysis of the marsupial/placental data set. In this study specific hypotheses on three types of developmental integration were proposed. The first hypothesis was that elements that were spatially adjacent would be expected to be integrated because they might be subject to similar mechanical or inductive influences. The second hypothesis was that elements that are part of a single organ system might be under integrated control mechanisms. The third explored the notion that some elements are

functionally or evolutionarily integrated and therefore they may be also developmentally integrated (Alberch, 1980; Kay, 1986).

This study revealed many different levels of integration and dissociation in the head of therian mammals. For example, ossification of cranial bones was clearly dissociated and under local control. There were no data to suggest that the skeletal system was integrated as a system; instead, it appeared to develop and evolve as independent modules. Onset of ossification of bones of the facial region was accelerated in marsupials and clearly independent of the bones of the calvaria. There was significant evidence of integration between the central nervous system and the skeletal elements forming the neurocranium, as shifts in timing of these sets of structures appeared to be coordinated. Such coordinated changes would be expected given all current evidence on mechanistic relations between CNS differentiation and cranial ossification (e.g. Hall, 1984b, 1987; Hanken & Thorogood, 1993; Herring, 1993; Moss, 1968; Schöwing, 1968a,b; Thorogood, 1988; Thorogood, Bee & Mark, 1986; Tyler, 1983; Wood *et al.*, 1991). A single bone of the braincase, the exoccipital, apparently was not controlled by the same influences as other bones of the neurocranium, but instead appeared to be largely influenced by cervical musculature, as it is markedly advanced in marsupials relative to other bones of the neurocranium. This observation confirmed the apparent dissociability of individual cranial bones in development. In contrast, the musculature of the head showed little regional differentiation, or heterochrony, despite changes in spatially or functionally adjacent elements, and despite significant changes in the timing of when a particular muscle would be used. Instead cranial musculature appeared to develop as a unified system under some overriding set of influences. For example, tongue and pharyngeal muscles, which must be functional at birth, matured only a few days before eye muscles, which would not be functional until eye opening, at 35 days after birth. Finally, elements of the first arch, which have been hypothesized to be both evolutionarily (Alberch, 1980) and developmentally (Kay, 1986) integrated, gave no evidence of developmental integration, but instead showed a mixed pattern of advancement and delay of development.

SUMMARY AND CONCLUSIONS

Over the past 20 years the concept of heterochrony has received tremendous attention by those attempting to link changes in development with changes in evolution. During this period heterochrony has been viewed almost entirely as a phenomenon of relative growth. Further, it has been applied so non-specifically by some that the usefulness of the concept has been questioned. In this paper I present an alternative conceptual and methodological approach to heterochrony. The methods proposed here broaden the applicability of heterochrony rather than replace the more traditional growth heterochrony approach. I re-focus the analysis on the relative timing of developmental events, and view heterochrony as the change in the sequence of events. I argue that this approach has several distinct advantages. First, it provides a means to standardize development across taxa independently of external criteria (age, stage or size), all of which are theoretically and empirically problematical for broad cross-taxa comparison. Second, it permits analysis of changes in developmental timing of events not characterized by size and shape parameters while also allowing the analysis of multiple types of events (e.g. morphological, molecular or genetic). Third, it provides means to simultaneously analyse many elements in many taxa – providing the opportunity to test hypotheses of conservation of developmental sequences, the degree to which ontogeny parallels phylogeny, and the existence of developmental modules, developmental dissociation or developmental integration.

ACKNOWLEDGEMENTS

I am grateful to P. Mabee, B. Velhagen and C. Nunn for discussions on approaches to the analysis of developmental sequences, to H. F. Nijhout, J. Vaglia, A. van Nievelt and G. Wray for comments on earlier drafts of this paper and the NSF (IBN 9407616 and IBN 9816985) for financial support.

REFERENCES

Alberch P. 1980. Ontogenesis and morphological diversification. *American Zoologist* **20:** 653–667.

Alberch P. 1985. Problems with the interpretation of developmental sequences. *Systematic Zoology* **34:** 46–58.

Alberch P, Gould SJ, Oster GF, Wake DB. 1979. Size and shape in ontogeny and phylogeny. *Paleobiology* **5:** 296–317.

Ambros V, Horvitz HR. 1984. Heterochronic mutants of the nematode *Caenorhabditis elegans. Science* **226:** 409–416.

Ambros V, Horvitz HR. 1987. The *lin-14* locus of *Caenorhabditis elegans* controls the time of expression of specific postembryonic developmental events. *Genes and Development* **1:** 398–414.

Blackstone NW. 1987a. Allometry and relative growth: pattern and process in evolutionary studies. *Systematic Zoology* **36:** 76–78.

Blackstone NW. 1987b. Size and time. *Systematic Zoology* **36:** 211–215.

Clark CT, Smith KK. 1993. Cranial osteogenesis in *Monodelphis domestica* (Didelphidae) and *Macropus eugenii* (Macropodidae). *Journal of Morphology* **215:** 119–149.

Cockburn A. 1989. Adaptive patterns in marsupial reproduction. *Trends in Ecology and Evolution* **4:** 126–130.

Collazo A. 1994. Molecular heterochrony in the pattern of fibronectin expression during gastrulation in amphibians. *Evolution* **48:** 2037–2045.

Cooke J, Smith JC. 1990. Measurement of developmental time by cells of early embryos. *Cell* **60:** 891–894.

Cubbage CC, Mabee PM. 1996. Development of the cranium and paired fins in the zebrafish *Danio rerio* (Ostariophysi, Cyprinidae). *Journal of Morphology* **229:** 121–160.

Dale KJ, Pourquié O. 2000. A clock-work somite. *BioEssays* **22:** 72–83.

de Beer GR. 1930. *Embryology and evolution.* Oxford: Clarendon Press.

de Beer GR. 1940. *Embryos and ancestors.* Oxford: Clarendon Press.

de Beer GR. 1951. *Embryos and ancestors,* revised edition. Oxford: Clarendon Press.

de Beer GR. 1958. *Embryos and ancestors,* third edition. Oxford: Oxford University Press.

de Queiroz K. 1985. The ontogenetic method for determining character polarity and its relevance to phylogenetic systematics. *Systematic Zoology* **34:** 280–299.

Duboule D. 1994. Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Development* **Suppl:** 135–142.

Dunlap KD, Sanchiz BA. 1996. Temporal dissociation between the development of the cranial and appendicular skeletons in *Bufo bufo* (Amphibia: Bufonidae). *Journal of Herpetology* **30:** 506–513.

Emerson SB. 1986. Heterochrony and frogs: the relationship of a life history trait to morphological form. *American Naturalist* **127:** 167–183.

Filan SL. 1991. Development of the middle ear region in *Monodelphis domestica* (Marsupialia, Didelphidae): marsupial solutions to early birth. *Journal of Zoology, London* **225:** 577–588.

Fink WL. 1982. The conceptual relationship between ontogeny and phylogeny. *Paleobiology* **8:** 254–264.

Gemmell RT, Selwood L. 1994. Structural development in the newborn marsupial, the stripe-faced dunnart, *Sminthopsis macroura. Acta Anatomica* **149:** 1–12.

Godfrey LR, Sutherland MR. 1995a. Flawed inference: why size-based tests of heterochronic processes do not work. *Journal of Theoretical Biology* **172:** 43–61.

Godfrey LR, Sutherland MR. 1995b. What's growth got to do with it? process and product in the evolution of ontogeny. *Journal of Human Evolution* **29:** 405–431.

Gorodilov YN. 1992. Rhythmic processes in lower vertebrate embryogenesis and their role for developmental control. *Zoological Science* **9:** 1101–1111.

Gould SJ. 1977. *Ontogeny and phylogeny.* Cambridge, MA: Belknap Press of Harvard University Press.

Gould SJ. 2000. Of coiled oysters and big brains: how to rescue the terminology of heterochrony, now gone astray. *Evolution and Development* **2:** 241–248.

Hall BK. 1984a. Developmental processes underlying heterochrony as an evolutionary mechanism. *Canadian Journal of Zoology* **62:** 1–7.

Hall BK. 1984b. Genetic and epigenetic control of connective tissues in the craniofacial structures. *Birth Defects: Original Article Series* **20:** 1–17.

Hall BK. 1987. Tissue interactions in the development and evolution of the vertebrate head. In: Maderson PFA, ed. *Developmental and evolutionary aspects of the neural crest.* New York: Wiley-Interscience, 215–259.

Hall BK. 1992. *Evolutionary developmental biology.* London: Chapman and Hall.

Hall BK. 1999. *Evolutionary developmental biology,* 2nd edition. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Hall BK, Miyake T. 1995. How do embryos measure time? In: McNamara KJ, ed. *Evolutionary change and heterochrony.* New York: John Wiley & Sons, 3–20.

Hanken J, Hall BK. 1984. Variation and timing of the cranial ossification sequence of the Oriental fire-bellied toad, *Bombina orientalis* (Amphibia, Discoglossidae). *Journal of Morphology* **182:** 245–255.

Hanken J, Thorogood P. 1993. Evolution and development of the vertebrate skull: the role of pattern formation. *Trends in Ecology and Evolution* **8:** 9–15.

Hanken J, Wake DB. 1993. Miniaturization of body size: organismal consequences and evolutionary significance. *Annual Review of Ecology and Systematics* **24:** 501–519.

Herring SW. 1993. Epigenetic and functional influences on skull growth. In: Hanken J, Hall BK, eds. *The Skull.* Chicago: University of Chicago Press, 153–206.

Hill JP, Hill WCO. 1955. The growth stages of the pouch young of the native cat (*Dasyurus viverrinus*) together with observations on the anatomy of the new born young. *Transactions of the Zoological Society, London* **28:** 349–453.

Howe JA, Howell M, Hunt T, Newport JW. 1995. Identification of a developmental timer regulating the stability of embryonic cyclin A and a new somatic A-type cyclin at gastrulation. *Genes and Development* **9:** 1164–1176.

Hufford L. 1995. Patterns of ontogenetic evolution in perianth diversification of *Besseya* (Scrophulariaceae). *American Journal of Botany* **82:** 655–680.

Hufford L. 1996. Ontogenetic evolution, clade diversification, and homoplasy. In: Sanderson MJ, Hufford L, eds. *Homoplasy: the reoccurrence of similarity in evolution.* New York: Academic Press, 271–301.

Hughes RL, Hall LS. 1988. Structural adaptations of the newborn marsupial. In: Tyndale-Biscoe CH, Janssens PA, eds. *The developing marsupial: models for biomedical research.* Berlin: Springer, 8–27.

Irish FJ. 1989. The role of heterochrony in the origin of a novel bauplan: evolution of the ophidian skull. *Geobios* **12 (memoire special):** 227–233.

Jeffery WR, Swalla BJ. 1992. Evolution of alternate modes of development in ascidians. *BioEssays* **14:** 219–226.

Johnson MH, Day ML. 2000. Egg timers: How is developmental time measured in the early vertebrate embryo? *BioEssays* **22:** 57–63.

Kay E. 1986. The phenotypic interdependence of the musculoskeletal characters of the mandibular arch in mice. *Journal of Embryology and Experimental Morphology* **98:** 123–136.

Klima M. 1987. *Early development of the shoulder girdle and sternum in marsupials (Mammalia: Metatheria).* Berlin: Springer.

Klingenberg CP. 1998. Heterochrony and allometry: the analysis of evolutionary change in ontogeny. *Biological Reviews* **73:** 79–123.

Klingenberg CP, Spence JR. 1993. Heterochrony and allometry: lessons from the water strider genus *Limnoporus*. *Evolution* **47:** 1834–1853.

Kluge AG. 1985. Ontogeny and phylogenetic systematics. *Cladistics* **1:** 13–27.

Kluge AG. 1988. The characterization of ontogeny. In: Humphries CJ, ed. *Ontogeny and systematics*. New York: Columbia University Press, 57–81.

Kluge AG, Strauss RE. 1985. Ontogeny and systematics. *Annual Review of Ecology and Systematics* **16:** 247–268.

Langille RM, Hall BK. 1989. Developmental processes, developmental sequences and early vertebrate phylogeny. *Biological Reviews* **64:** 73–91.

Larsson HCE. 1998. A new method for comparing ontogenetic and phylogenetic data and its application to the evolution of the crocodilian secondary palate. *Neues Jahrbuch für Geologie und Paläontologie, Abhandlung* **210:** 345–368.

Mabee PM. 1993. Phylogenetic interpretation of ontogenetic change: sorting out the actual and artefactual in an empirical case study of centrarchid fishes. *Zoological Journal of the Linnean Society* **107:** 175–291.

Mabee PM. 1996. Reassessing the ontogenetic criterion: a response to Patterson. *Cladistics* **12:** 169–176.

Mabee PM, Trendler TA. 1996. Development of the cranium and paired fins in *Betta splendens* (Teleostei: Percomorpha): Intraspecific variation and interspecific comparisons. *Journal of Morphology* **227:** 249–287.

Maddison WP, Maddison DR. 1992. *MacClade, version 3.* Sinauer: Sunderland, MA, USA.

Maier W. 1987. The ontogenetic development of the orbitotemporal region in the skull of *Monodelphis domestica* (Didelphidae, Marsupialia), and the problem of the mammalian alisphenoid. In: Kuhn H-J, Zeller U, eds. *Morphogenesis of the mammalian skull*. Hamburg: Verlag Paul Pary.

McClung CR, Fox BA, Dunlap JC. 1989. The *Neurospora* clock gene *frequency* shares a sequence element with the *Drosophila* clock gene *period*. *Nature* **339:** 558–562.

McKinney ML. 1988. *Heterochrony in evolution.* New York: Plenum.

McKinney ML. 1999. Heterochrony: beyond words. *Paleobiology* **25:** 149–153.

McNamara KJ. 1995. *Evolutionary change and heterochrony.* New York: John Wiley & Sons.

McNamara KJ. 1997. *Shapes of time: the evolution of growth and development.* Baltimore: Johns Hopkins University Press.

Meier R. 1997. A test and review of the empirical performance of the ontogenetic criterion. *Systematic Biology* **46:** 699–721.

Moss ML. 1968. A theoretical analysis of the functional matrix. *Acta Biotheoretica* **18:** 195–202.

Needham J. 1933. On the dissociability of the fundamental processes in ontogenesis. *Biological Reviews* **8:** 180–223.

Nelson G. 1978. Ontogeny, phylogeny, paleontology, and the biogenetic law. *Systematic Zoology* **27:** 324–345.

Nelson JE. 1987. The early development of the eye of the pouch-young of the marsupial *Dasyurus hallucatus*. *Acta Embryologica* **175:** 387–398.

Nunn CL, Smith KK. 1998. Statistical analysis of developmental sequences: The craniofacial region in marsupial and placental mammals. *American Naturalist* **152:** 82–101.

O'Grady RT. 1985. Ontogenetic sequences and the phylogenetics of parasitic flatworm life cycles. *Cladistics* **1:** 159–170.

Palmeirim I, Henrique D, Ish-Horowicz D, Pourquié O. 1997. Avian *hairy* gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell* **91:** 639–648.

Patterson C. 1996. Comments on Mabee's "empirical rejection of the ontogenetic polarity criterion". *Cladistics* **12:** 147–167.

Pourquié. 1998. Clocks regulating developmental processes. *Current Opinion in Neurobiology* **8:** 665–670.

Power M-A, Tam PPL. 1993. Onset of gastrulation, morphogenesis and somitogenesis in mouse embryos displaying compensatory growth. *Anatomy and Embryology* **187:** 493–504.

Raff RA. 1996. *The shape of life.* Chicago: University of Chicago Press.

Raff EC, Raff RA. 2000. Dissociability, modularity, evolvability. *Evolution and Development* **2:** 235–237.

Raff RA, Wray GA. 1989. Heterochrony: developmental mechanisms and evolutionary results. *Journal of Evolutionary Biology* **2:** 409–434.

Reiss JO. 1989. The meaning of developmental time: a metric for comparative embryology. *American Naturalist* **134:** 170–189.

Richardson MK. 1995. Heterochrony and the phylotypic period. *Developmental Biology* **172:** 412–421.

Richardson MK, Allen SP, Wright GM, Raynaud A, Hanken J. 1998. Somite number and vertebrate evolution. *Development* **125:** 151–160.

Richardson MK, Hanken J, Gooneratne ML, Pieau C, Raynaud A, Selwood L, Wright GM. 1997. There is no highly conserved embryonic stage in the vertebrates: implications for current theories of evolution and development. *Anatomy and Embryology* **196:** 91–106.

Ridley M. 1985. Embryology and classical zoology in Great Britain. In: Horder TJ, Witkowski JA, Wylie CC, eds. *A history of embryology.* Cambridge: Cambridge University Press, 35–67.

Rieppel O. 1990. Ontogeny – a way forward for systematics, a way backward for phylogeny. *Biological Journal of the Linnean Society* **39:** 177–191.

Roth VL. 1984. How elephants grow: heterochrony and the calibration of developmental stages in some living and fossil species. *Journal of Vertebrate Paleontology* **4:** 126–145.

Russell ES. 1916. *Form and function.* London: John Murray Ltd.

Satoh N. 1982. Timing mechanisms in early embryonic development. *Differentiation* **22:** 156–163.

Schöwing J. 1968a. Influence inductrice de l'encéphale embryonnaire sur le développement du crâne chez le poulet. *Journal of Embryology and Experimental Morphology* **19:** 23–32.

Schöwing J. 1968b. Mise en évidence du rôle inducteur de l'encéphale dans l'ostéogenese du crâne embryonnaire du poulet. *Journal of Embryology and Experimental Morphology* **19:** 83–94.

Slack F, Ruvkun G. 1997. Temporal pattern formation by heterochronic genes. *Annual Review of Genetics* **31:** 611–634.

Slack JMW, Holland PWH, Graham CF. 1993. The zootype and the phylotypic stage. *Nature, London* **361:** 490–492.

Smith KK. 1994. The development of craniofacial musculature in *Monodelphis domestica* (Didelphidae, Marsupialia). *Journal of Morphology* **222:** 149–173.

Smith KK. 1996. Integration of craniofacial structures during development in mammals. *American Zoologist* **36:** 70–79.

Smith KK. 1997. Comparative patterns of craniofacial development in eutherian and metatherian mammals. *Evolution* **51:** 1663–1678.

Smith KK. 2000. The evolution of mammalian development. *Bulletin of the Museum of Comparative Zoology* **155:** 603–619.

Smith MM. 1995. Heterochrony in the evolution of enamel in vertebrates. In: McNamara KJ, ed. *Evolutionary change and heterochrony.* New York: John Wiley & Sons, 125–150.

Snow MHL, Tam PPL, McLaren A. 1981. On the control and regulation of size and morphogenesis in mammalian embryos. In: Subtelny S, Abbott UK, eds. *Levels of genetic control in development.* New York: Alan R. Liss, 201–217.

Stern CD, Vasiliauskas D. 1998. Clocked gene expression in somite formation. *BioEssays* **20:** 528–531.

Strauss RE. 1990. Heterochronic variation in the developmental timing of cranial ossifications in poeciliid fishes (Cyprinodontiformes). *Evolution* **44:** 1558–1567.

Swalla BJ, White ME, Zhou J, Jeffery WR. 1994. Heterochronic expression of an adult muscle actin gene during ascidian larval development. *Developmental Genetics* **15:** 51–63.

Thorogood P. 1988. The developmental specification of the vertebrate skull. *Development* **103 Supplement:** 141–153.

Thorogood P, Bee J, von der Mark K. 1986. Transient expression of collagen type II at epitheliomesenchymal interfaces during morphogenesis of the cartilaginous neurocranium. *Developmental Biology* **116:** 497–509.

Tyler MS. 1983. Development of the frontal bone and cranial meninges in the embryonic chick: an experimental study tissue interactions. *The Anatomical Record* **206:** 61–70.

Tyndale-Biscoe H, Renfree M. 1987. *Reproductive physiology of marsupials.* Cambridge: Cambridge University Press.

Velhagen WA. 1995. A comparative study of cranial development in the thamnophiine snakes (Serpentes: Colubridae). Unpublished Ph.D. Thesis, Duke University, Durham, NC, USA.

Velhagen WA. 1997. Analyzing developmental sequences using sequence units. *Systematic Biology* **46:** 204–210.

Wake DB. 1989. Phylogenetic implications of ontogenetic data. *Geobios* **12 memoire special:** 369–378.

Wake MH, Hanken J. 1982. Development of the skull of *Dermophis mexicanus* (Amphibia: Gymnophiona) with comments on skull kinesis and amphibian relationships. *Journal of Morphology* **173:** 203–223.

Wimsatt WC. 1986. Developmental constraints, generative entrenchment and the innate-acquired distinction. In: Bechtel W, ed. *Integrating scientific disciplines.* Dordrecht, The Netherlands: Martinus Nijhoff Publishers, 185–208.

Wood A, Ashhurst DE, Corbett A, Thorogood P. 1991. The transient expression of type II collagen at tissue interfaces during mammalian craniofacial development. *Development* **111:** 955–968.

Wray GA. 1995. Causes and consequences of heterochrony in early echinoderm development. In: McNamara KJ, ed. *Evolutionary change and heterochrony.* New York: John Wiley & Sons, 197–223.

Wray GA, McClay DR. 1989. Molecular heterochronies and heterotopies in early echnoid development. *Evolution* **43:** 803–813.

Yasuda GK, Schubiger G. 1992. Temporal regulation in the early embryo: is MBT too good to be true? *Trends in Genetics* **8:** 124–127.

Zelditch ML, Fink WM. 1996. Heterochrony and heterotopy: stability and innovation in the evolution of form. *Paleobiology* **22:** 241–254.