Analyzing evolutionary patterns in amniote embryonic development*

Jonathan E. Jeffery,[a,1] Olaf R. P. Bininda-Emonds,[b,2] Michael I. Coates,[c] and Michael K. Richardson[a,1]

[a]Department of Anatomy and Developmental Biology, St. George's Hospital Medical School, Tooting, London SW17 ORE, UK; [b]Institute of Evolutionary and Ecological Sciences, Leiden University, Kaiserstraat 63, P.O. Box 9516, 2300 RA Leiden, The Netherlands; [c]Department of Organismal Biology & Anatomy, University of Chicago,1027 East 57th Street, Chicago, IL 60637-1508, USA; [1]Current address: Institute of Evolutionary and Ecological Sciences, Leiden University, Kaiserstraat 63, P.O. Box 9516, 2300 RA Leiden, The Netherlands; [2]Current address: Lehrstuhl für Tierucht, Technical University of Munich, D-85354 Fresing-Weihenstephan, Germany

[†]Correspondence (email: jeffery@rulsfb.leidenuniv.nl)

SUMMARY Heterochrony (differences in developmental timing between species) is a major mechanism of evolutionary change. However, the dynamic nature of development and the lack of a universal time frame makes heterochrony difficult to analyze. This has important repercussions in any developmental study that compares patterns of morphogenesis and gene expression across species. We describe a method that makes it possible to quantify timing shifts in embryonic development and to map their evolutionary history. By removing a direct dependence on traditional staging series, through the use of a relative time frame, it allows the analysis of developmental sequences across species boundaries. Applying our method to published data on vertebrate development, we identified clear patterns of heterochrony. For example, an early onset of various heart characters occurs throughout amniote evolution. This suggests that advanced (precocious) heart development arose in evolutionary history before endothermy. Our approach can be adapted to analyze other forms of comparative dynamic data, including patterns of developmental gene expression.

INTRODUCTION

Heterochrony is a change in developmental timing during evolution (Gould 1977; McKinney and McNamara 1991). It may be manifest as species differences in growth patterns (Gould 2000), in temporal patterns of gene expression (Ferkowicz and Raff 2001), or in the sequence in which the constituent parts of the embryo develop (i.e., the "developmental sequence" (Alberch 1985).

Developmental sequence heterochrony has been known for almost 150 years (Haeckel 1866). The *Normal Tables* of Oppel (1891) and Keibel (1897–1938) showed that sequence heterochronies were widespread within the vertebrates (see also Richardson 1995, 1999; Richardson et al. 1997; Hirata et al. 1998). However, it is only within the last few years that methods have been developed to study the phenomenon in a quantitative phylogenetic framework (e.g., Mabee and Trendler 1996; Smith 1997; Velhagen 1997; Schlosser 2001). Indeed, sequence heterochronies were often considered to be noise, obscuring a supposedly universal or conserved pattern. For example, Haeckel (1866) believed that sequence heterochronies disrupted the ability to interpret phylogeny on the basis of his Biogenetic Law.

With the recent revived interest in the relationship between evolution and development (Holland 1999; Raff et al. 1999), the value of sequence heterochrony as a data source has been reassessed (e.g., Mabee and Trendler 1996; Smith 2001a). First, as heritable traits, the heterochronies themselves should be phylogenetically informative. Second, heterochrony is of great intrinsic interest as a factor in the evolution of developmental mechanisms (Knoetgen et al. 1999; Smith 2001b). Keibel (1912) discussed sequence heterochronies as evidence of the autonomy of organ anlagen. Interest in the hierarchical nature of developmental integration has recently reemerged, in discussions of developmental "modularity" (e.g., Wagner 1996; von Dassow and Munro 1999; Bolker 2000; Schlosser 2001; Winther 2001). In the last few years, the genetic basis for specific heterochronic changes have been identified in the nematode *Caenorhabditis elegans* (Pasquinelli et al. 2000; Reinhart and Ruvkun 2001). Finally, sequence heterochronies have important practical implications for the way in which researchers compare development in different model species. For example, similar organ systems may arise at different times (and therefore in different embryonic contexts) in different species (Richardson 1995).

*The authors will distribute additional data used in this article upon request, free of charge via post or email. Limits of space prevented the inclusion of the data here. The additional data were made available to the reviewers.

Unfortunately, despite the recognition of sequence heterochrony as an important evolutionary and developmental phenomenon, there have been surprisingly few detailed studies. This is because its value as a data source has often been undermined by analytical difficulties (Smith 2001a; Jeffery et al. 2002 in press; see below).

Here we analyze sequence heterochrony in amniotes using a modified form of the event-pairing method (Mabee and Trendler 1996; Smith 1997; Velhagen 1997; Jeffery et al. 2002 in press). We concentrate on the amniotes because of their historical association with studies of sequence heterochrony (e.g., Haeckel 1866; Oppel 1891; Keibel 1897–1938, 1912). They are also a well-studied group in terms of their development and evolutionary relationships. Another factor is that data are available for a range of outgroup species. We analyze heterochrony in the mid-embryonic (organogenetic) period because it is the period when most organ primordia are laid down and when the body plan is specified under the control of regulatory genes (Slack et al. 1993). In addition to the classical morphological examples of sequence heterochronies, there is also evidence for heterochronies in developmental gene expression (Blanco et al. 1998). These may, in turn, affect pattern formation (Dolle et al. 1993; Duboule 1994).

MATERIALS AND METHODS

Techniques

To judge whether, during evolution, a particular event has shifted earlier or later in development, a universal standard (i.e., one applicable to all members of the taxon) would ideally be required for measuring developmental time. This may not be possible, however. For example, absolute chronological age is an unsatisfactory metric in comparative studies, because of interspecific variation in developmental rates (Hall and Miyake 1997). Methods for "normalizing" chronological ages among species (e.g., Dettlaff and Dettlaff 1961) can be difficult to use in practice, because they involve a continuous study of embryos in vivo. Also, they may only be applicable to closely related taxa. Most often, developmental biologists avoid a reliance on chronological age by subdividing development into "stages" based on morphology. The stages of Hamburger and Hamilton (1951) for the chick are a well-known example of this.

Basing the stages on the appearance of a single "landmark feature" or "key character" (e.g., "tailbud stage," "limbbud stage") makes it possible to identify the stage in any species possessing that feature. However, this carries with it the assumption that the landmark itself does not show heterochronic change. This assumption is not necessarily valid. A conspicuous example is the wide difference in the stage of maturity at which limbs develop in different species (Richardson 1995). Alternatively, a set of morphological characters may be used to define a stage. Although this gives a more precise definition of maturity, it can be difficult to apply that definition to more than one species. This is because sequence heterochronies can alter the set of morphological characters appearing at any one time (Fig. 1). Ironically, therefore, heterochrony can erase the common staging landmarks needed to study it.

Despite this problem, comparisons can be among species by using, not individual events, but the whole developmental sequences as a single entity. One approach is to use statistical techniques that can compare the similarity of two or more sequences (e.g., Kendall's concordance coefficient; Kendall 1970). It is not the timing of individual events that is important but rather their rank in the overall sequence (first, second, third, etc.). Such methods provide an empirical framework for comparison but also suffer from two drawbacks. First, because the events in the sequence must be ranked, a complete set of observations must be available for every species. In practice, missing data may not be a problem in species that can be bred in large numbers in captivity. However, material collected from the wild or rare species may often be incomplete—either due to a lack of sufficient numbers of embryos or poor preservation conditions. Also, in evolutionary studies, researchers may wish to include data on structures or events that are an evolutionary novelty, present only in a subset of the study group. The second more fundamental shortcoming is that the comparisons are phenetic—they measure the degree of variance of the rankings rather than identifying primitive and derived features. They are therefore unable to give any information on the patterns of heterochrony across phylogeny.

A more flexible method for comparing developmental sequences, known as "event pairing," was developed in the 1990s (Mabee and Trendler 1996; Smith 1996, 1997; Velhagen 1997). This method describes the developmental sequence by a series of statements on the relative timing of pairs of events. A pair of events, A and B, can have one of three possible timing relationships. Either event A can occur before event B in development, events A and B can occur simultaneously, or event A can occur after event B. In event-pair analysis, every possible pair of events from the sequence is examined and their relative timing relationship scored (using the values 0, 1, or 2, respectively, to represent the three possible relationships; cf. Smith 1996). For a data set of n events there are $1/2(n^2 - n)$ possible nonredundant event pairs. These scores can be compared among species using standard phylogenetic software (e.g., PAUP* [Swofford 2001] and MacClade [Maddison and Maddison 2001]) to highlight patterns of timing change (e.g., Smith 1997). These patterns can be either shared derived changes (synapomorphies) along the individual branches of an evolutionary tree or similar, independently acquired changes along separate branches of the tree (convergence, a form of homoplasy).

Like the statistical methods, examining the overall sequence removes the problems of absolute timing. However, rather than making a phenetic statement on similarity, event pairing can highlight individual timing shifts. It can also accommodate missing data. If an event cannot be observed in a particular species, its timing relationships may be scored as uncertain for that species (i.e., event pairs involving that event are scored as "?"). The phylogenetic software will estimate the most probable score, based on the position of the species on the tree. Scoring for inapplicable events (i.e., events that do not occur in a particular species) is more problematic; using the "uncertain" score can lead to illogical character optimizations in certain extreme cases. However, various alternative scoring schemes proposed to accommodate inapplicable events all have problems associated with them (see Forey and Kitching 2000 for a recent review). We therefore used the "uncertain" score but restricted ourselves as far as possible to events that occur in all the species under

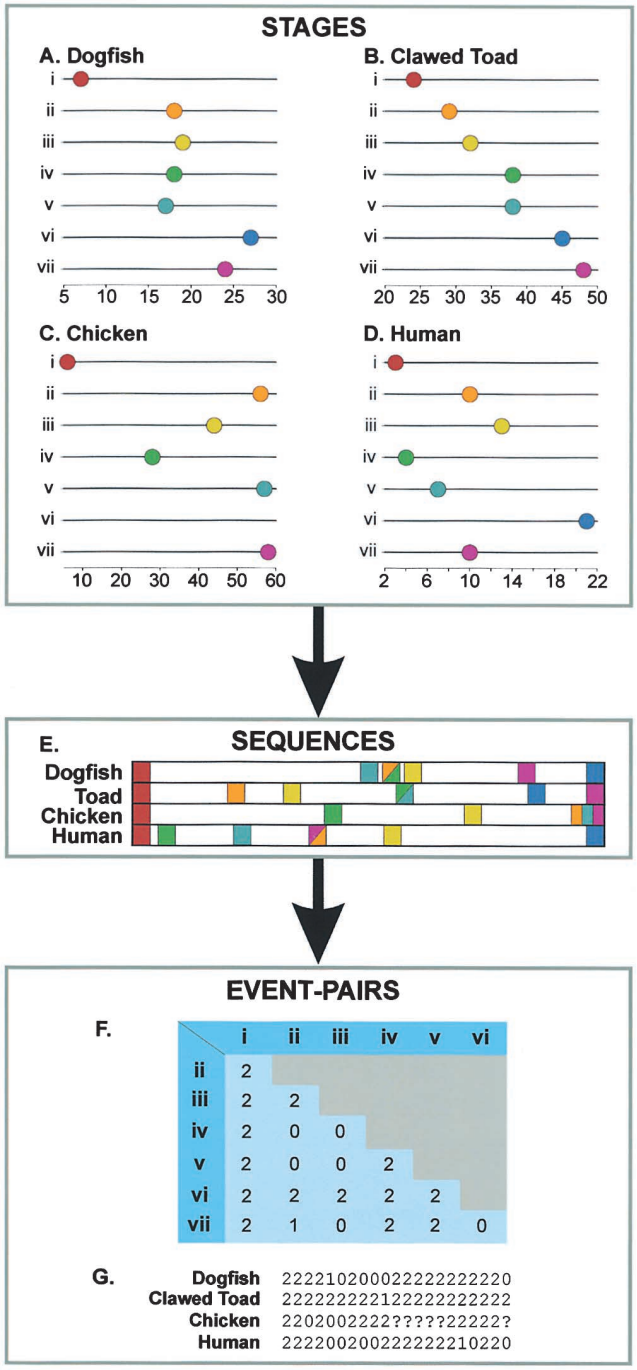

Fig. 1. The relationship between developmental stage, developmental sequence, and heterochrony, as used in this study. (A–D) The development of seven embryonic structures in four vertebrate species. The first appearance of each structure is indicated by the position of a bead on the horizontal line of an "abacus" diagram (Richardson 1999). The X axes (developmental time) are measured in "stages," but the published stages for each species cannot be directly compared. The structures are (i) first somite, (ii) nasal placode, (iii) optic cup, (iv) heart loop, (v) thyroid depression, (vi) spleen anlage, and (vii) forelimb buds. (E) Developmental sequences derived from the stages. These also cannot be directly compared, because the alignment of the sequences (i.e., which events are chosen to be constant) is arbitrary. (F) Event-pair matrix for the human derived from the seven events. This encodes the relative timing of events in a series of 21 event-pair statements, overcoming the reliance on an a particular staging scheme. (G) The 21 event-pair scores for each of the four species, assembled ready for phylogenetic analysis. Data after Keibel (1897–1938) and Nieuwkoop and Faber (1994).

study (Table 1). Even if it has not been observed, the occurrence of an event can be inferred by the presence of the structure concerned later in ontogeny (e.g., the event "first appearance of liver diverticulum" can be inferred by the presence of a liver in the adult).

Despite its advantages, interpretation of the results of event-pair analyses can be complicated. Each synapomorphy (a shared change of event-pair score) is informative about the timing relationship between only two events within the developmental sequence. Without further information, it is not possible to determine the magnitude of the change and whether it involved one or both events in the event pair. We therefore developed a procedure for interpreting the synapomorphies along a given branch en bloc to determine the shifts of individual events relative to all other developmental events surveyed. The method is discussed in detail elsewhere (see Jeffery et al. 2002 in press). Briefly, it is a heuristic to find the smallest set of event movements at any given node that could account for the observed synapomorphic event-pair changes. It does this by identifying any events that are constituent members of more event-pair synapomorphies than average over all events for that node. By assuming that these events have moved, most of the observed change can be accounted for. For example, say a particular node has 10 synapomorphic event-pair changes. If one particular event (X) is a constituent member all 10 event pairs, we could account for all the observed changes by assuming that event X has moved, changing its relative timing relationship with 10 other events. Any alternative explanation would be less parsimonious, because it would require a greater number of events to move, changing their timing relationship only with X and no other event. Once a minimum set of moving events has been established, the polarity of each movement can be determined—whether it was earlier or later in development (referred to as "advance" and "delay," respectively). The magnitude of each movement can also be noted by recording how many other events the shift was relative to. The method is conservative in that it requires an event to change its relative timing to several other events before its movement can be successfully identified. For example, if no event is a constituent member of more than one synapomorphic event pair, then no underlying pattern of movement can be determined. There is no absolute rule about how many synapomorphic event pairs are required before an underlying pattern can be established. In practice, it was typically 20 or more synapomorphic event pairs in our study.

Species

We used data from 14 vertebrate species to study heterochronic changes within nine amniote species (one lizard, three birds, and five placental mammals); five nonamniote species (one toad, three salamanders, and one shark) were used as an outgroup to infer the

Table 1. List of developmental events used in the analysis

Event Label	Event
1. Axial A	First Somite
2. Cardiovascular A	Endocardial anlage (single or paired rudiments)
3. Cardiovascular B	Endocardial tubes start to fuse
4. Cardiovascular C	Heart looping
5. Cardiovascular D	Atrioventricular canal indicated by constriction (or atrium distinct from ventricle)
6. Cardiovascular E	First aortic arch formed
7. Cardiovascular F	Septum primum of atrium
8. Cardiovascular G	Endocardial cushions of atrioventricular canal
9. Cardiovascular H	Outflow tract cushions
10. Cardiovascular I	Trabeculae carneae in ventricles
11. Cardiovascular J	Outflow tract valves
12. Intestinal A	Anterior intestinal portal begins as a diverticulum (or archenteron reaches head fold)
13. Intestinal B	Liver diverticulum
14. Intestinal C	Dorsal pancreas as a diverticulum
15. Intestinal D	Gallbladder as a diverticulum
16. Intestinal E	Liver cords
17. Intestinal F	Ventral pancreas anlage(n)
18. Intestinal G	Spleen anlage
19. Kidney A	Mesonephric duct anlagen
20. Kidney B	Paramesonephric duct anlagen
21. Kidney C	Mesonephric ducts open into cloaca
22. Limb A	Forelimb (or pectoral fin) bud
23. Neural A	Neural folds begin to fuse
24. Olfactory A	Nasal placodes appear as ectodermal thickenings
25. Olfactory B	Nasal placodes depressed (formation of olfactory pit)
26. Optic A	Optic vesicle as lateral evagination from neural tube
27. Optic B	Lens placode
28. Optic C	Optic vesicle starts to invaginate to form optic cup (secondary optic vesicle)
29. Optic D	Lens placode depressed (formation of optic pit)
30. Optic E	Lens vesicle pinches off from surface ectoderm
31. Optic F	Secondary retinal pigmentation
32. Otic A	Otic placode
33. Otic B	Otic placode depressed (formation of otic pit)
34. Otic C	Otocyst closed but still connected with surface ectoderm
35. Otic D	Otocyst detached from ectoderm
36. Otic E	Endolymphatic appendage
37. Pharyngeal A	Second visceral pouch contacts ectoderm (formation of hyoid arch)
38. Pharyngeal B	Thyroid anlage
39. Pharyngeal C	Third visceral pouch contacts ectoderm (formation of first branchial arch)
40. Pharyngeal D	Hypophysis anlage
41. Pharyngeal E	Lung buds as distinct paired evaginations

The events describe the first stage at which the particular structure or transformation is observed. The events have been grouped by organ system for ease of reference, although no a priori assumptions of modularity were made in the analysis. All but two of the events occur in the development of all the species we studied (even if timing data are not available). Two events—the first appearance of the septum primum in the atrium of the heart (cardiovascular F) and of the lung buds (pharyngeal E)—do not occur in the spiny dogfish, the most primitive member of our outgroup. However, they were included because they occur in the other members of the outgroup (making illogical optimizations less likely) and because they involve key tetrapod structures. Nevertheless, inferred changes involving these two events were scrutinized for the likely effects of illogical optimizations.

primitive condition for the amniotes. For Linnean and trivial names, see Fig. 2. Besides representing amniote diversity, an important criterion for selecting species for study was the availability and quality of published data on their development. The classical literature contains extremely detailed descriptions of developmental stages, notably Keibel's *Normal Tables* (Keibel 1897–1938). Data for the budgerigar were taken from Abraham (1901), for the toad from Nieuwkoop and Faber (1994), and for all other species from Keibel's *Normal Tables* (Keibel 1897–1938). Unfortunately, no detailed published information is available for some key groups (e.g., ray-finned fishes, turtles, crocodilians, and marsupial mammals), often for practical reasons. For example, because crocodilian eggs are laid at a relatively advanced stage of development (Ferguson 1985), gathering data for this group requires killing pregnant females; this in turn raises practical and ethical objections. Equally, staging series for species such as that for the zebrafish (Kimmel et al. 1995) are based only on external morphology rather than internal anatomy. This is because they are intended for rapid staging of embryos in the laboratory rather than as detailed embryological studies.

Events

Forty-one developmental events were selected, with no a priori consideration of whether they would show heterochrony. The events involved a range of embryonic structures (Table 1). Data on the se-

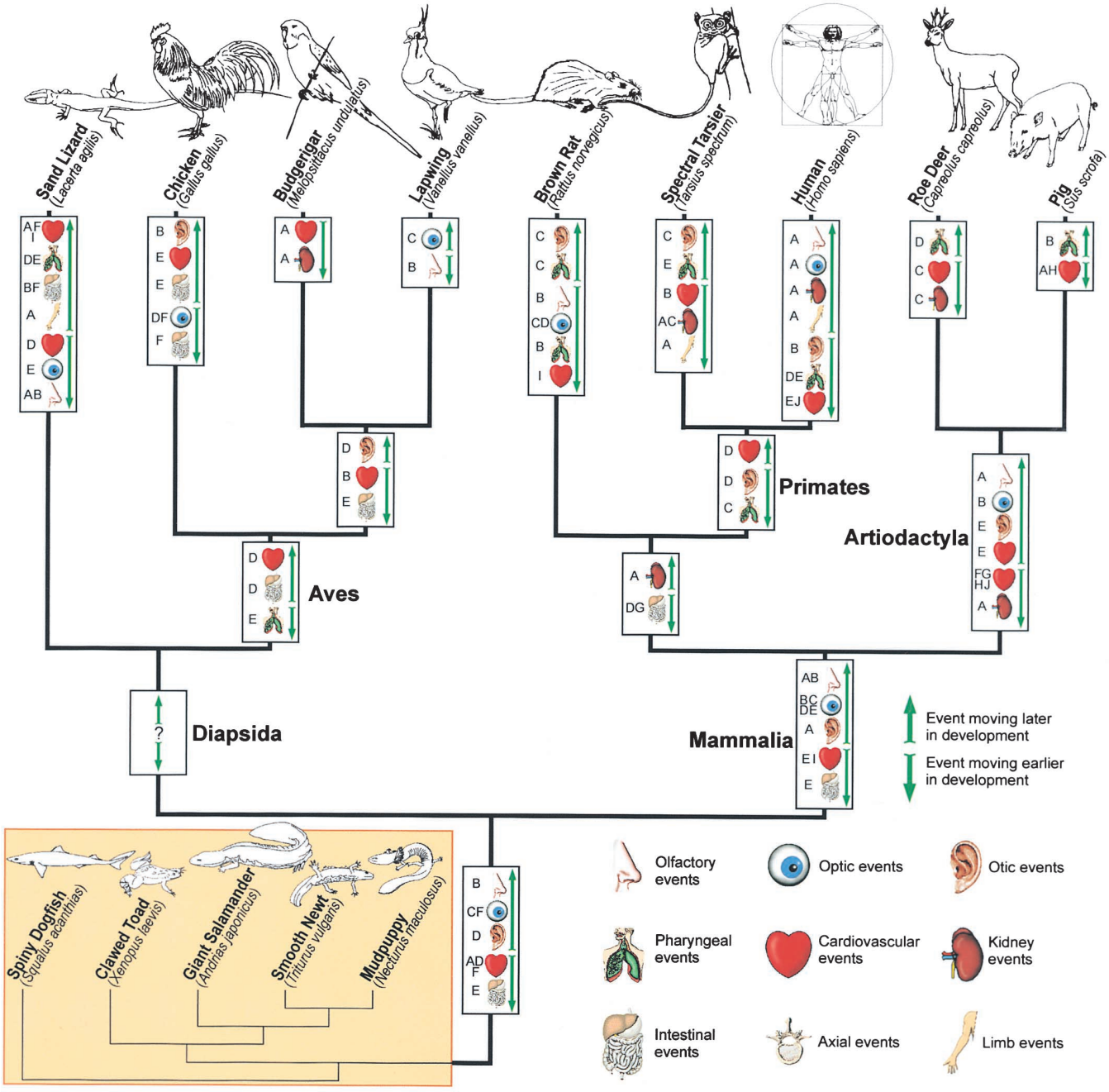

Fig. 2. Heterochrony in amniote evolution. Event-paired data were mapped onto a reference cladogram representing current opinion of vertebrate phylogeny (see Materials and Methods), and the implied synapomorphies were analyzed for patterns of shifts at each node. The events have been grouped by organ system for ease of reference, although no a priori assumptions of modularity (sensu Wagner 1996) were made in the analysis. For simplicity, the magnitudes of the shifts are not shown (see the Additional Data sheet, available by email, for a complete list of the direction and magnitude of all the shifts at each node).

lected events had to be available in all or most of our study species. The events we examined consisted of developmental transformations, including morphogenetic (e.g., neural fold fusion) and differentiation events (e.g., retinal pigmentation appearing). They encompassed the whole organogenetic period, ranging from early transformations such as the appearance of the first somite to late ones such as the development of the paramesonephric duct.

For each species, we recorded the earliest stage at which each of the 41 events occurred. The developmental sequence thus obtained was then subjected to a pair-wise comparison to generate a second-

ary data set consisting of 820 event pairs for each species (Fig. 1). Three event-pair scores were defined as above: 0 (character A occurs before B), 1 (A and B simultaneous), and 2 (A occurs after B).

Analysis

The character sequences for all species were compiled into a single data file and analyzed in two ways to examine the distributions of inferred synapomorphies and homoplasies. The first method of analysis was to use MacClade v4.0 phylogenetic analysis software (Maddison and Maddison 2001) to map the data onto reference cladograms (cf. Smith 1997) representing the current consensus opinion of vertebrate phylogeny. The topologies of the reference trees were based on data from the *Tree of Life Homepage* (http://phylogeny.arizona.edu/tree/phylogeny.html) and references therein. There is a general consensus on the interrelationships of all our study species, with the exception of the rat. A number of recent molecular studies, although not focusing specifically on rodents, recovered two alternative relationships: either (1) rodents and primates form a clade to the exclusion of artiodactyls or (2) primates and artiodactyls form a clade to the exclusion of rodents. The topology recovered depends on the data set and in some cases the consensus method used (e.g., see Waddell et al. 1999; Liu and Miyamoto 1999; Liu et al. 2001; Murphy et al. 2001; Madsen et al. 2001). In view of these differences, we examined the effects of both alternatives.

The second method of analysis was to perform a branch-and-bound parsimony analysis using PAUP* v4.0b8 (Swofford 2001) to determine whether heterochrony data contain a phylogenetic signal. The use of event-paired developmental timing data in a parsimony analysis has not yet been investigated in detail. We agree with previous authors (e.g., Smith 1997; Nunn and Smith 1998) who expressed concern with the internal nonindependence of an event-paired data set (see Discussion). For this reason, we preferred to use the reference trees, rather than trees produced by our parsimony analysis, to examine patterns of heterochrony within amniotes. However, differences between the reference trees and our most parsimonious trees were used to highlight areas of interest.

In both methods of analysis, the event-pair scores were used unordered and unweighted. The resulting event-pair synapomorphies were analyzed en bloc to infer which events had actively shifted position in evolution (see below).

RESULTS

Limitations of space prevents the publication of the full results of the analysis. The following sections (and Figs. 2–4) summarize the results, before the discussion. However, to allow scrutiny of our full data we produced an Additional Data sheet, available on request in a variety of formats from the corresponding author. This gives a complete listing of the sequence data obtained for each species, the event-pair data derived from it, and synapomorphic changes inferred from the analysis.

Mapping of data

The results of mapping our data are summarized in Figs. 2 and 3. In both cases, details of changes could not be determined at one node (the diapsids), because there were an insufficient number of changes to establish a background pattern (14 and 13 synapomorphies, respectively).

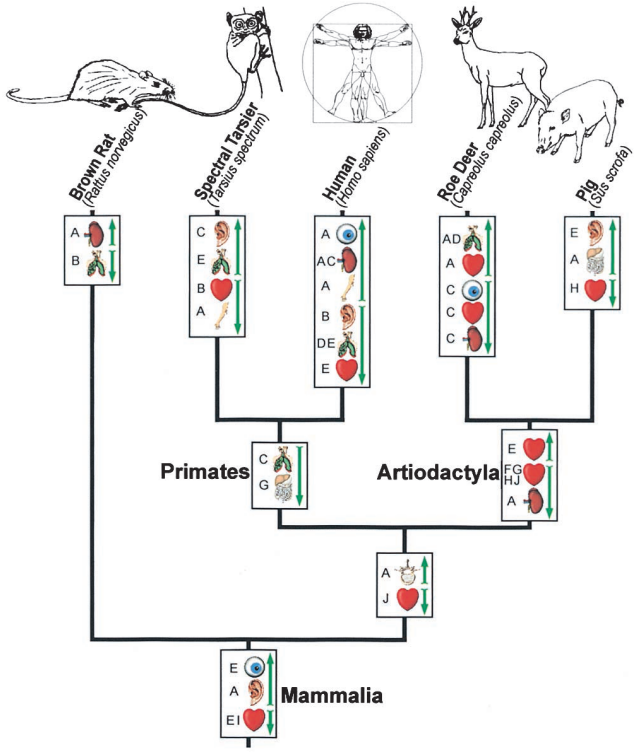

Fig. 3. Heterochrony in mammalian evolution. Detail of Fig. 2 showing the effects of an alternative position of the rat (see Materials and Methods). The rest of the tree was unaffected, with the exception of the birds (not shown), where a shared delay in the first appearance of the atrioventricular canal (cardiovascular D) was no longer recovered. Symbols are the same as in Fig. 2.

The data were first mapped onto the reference tree shown in Fig. 2, with the rat and the primates forming a clade. Of a total of 89 shifts identified within the amniotes, 26 relate to cardiovascular events. The proportion of shifts involving cardiovascular events (29.2%) is similar to the proportion of cardiovascular events in the data set (24.4%). However, the shifts show a large bias toward advances rather than delays (7 delays, 19 advances), especially within the mammals. Only a few of these shifts (1 delay, 5 advances) represent a change in the internal sequence of heart development. For example, mammals share an advance in the formation of the first aortic arch (cardiovascular E) with respect to the fusion of the endocardial tubes (cardiovascular B), as well as an advance of the ventricular trabeculae (cardiovascular I) with respect to the endocardial cushions of atrioventricular canal (cardiovascular G) and outflow tract cushions (cardiovascular H). One apparent change (an advance of the endocardial cushions of atrioventricular canal (cardiovascular F) with re-

spect to the formation of the outflow tract cushions (cardiovascular G) and ventricular trabeculae (cardiovascular H)) may be an artifact caused by the inapplicability of cardiovascular F to the spiny dogfish. The remainder of the shifts involving the cardiovascular system reflects the fact that the span of cardiovascular development is becoming "condensed" relative to noncardiac events.

Another interesting pattern is seen in the development of the lens and optic cup. Primitively in amniotes the optic vesicle starts to invaginate to form the optic cup (optic C) after the first appearance of the lens placode (optic B). This is a reversal of the order inferred to be primitive for tetrapods. Mammals retain the primitive amniote sequence but show a delay in four events ancestrally, namely the first appearance of the lens placode (optic B, relative to otic C, pharyngeal B, C, and D), the start of optic cup formation (optic C, relative to otic C, pharyngeal B and C), the depression of the lens placode to form the lens vesicle (optic D, relative to kidney C, pharyngeal E), and the final detachment of the lens vesicle from the surface ectoderm (optic E, relative to cardiovascular G and H, kidney C, optic F). Note that the internal sequence of events in mammalian lens formation remains unchanged. This is not surprising, because the events form a dependent sequence (cf. Alberch 1985)—for example, the depression of the lens placode (optic D) requires that the lens placode itself (optic B) has already formed. However, shifts identified in sand lizard, in the chicken, and at the artiodactyl node show that these events are also able to shift individually to some degree, relative to nonlens events. It is therefore interesting that the four events are delayed in concert (see Discussion). The rat shows a partial reversal of this phenomenon, with the start of optic cup formation (optic C) and the depression of the lens placode to form the lens vesicle (optic D) advancing relative to the condition primitive for mammals. Again, this does not alter the sequence of eye development; rather, it is a change in the spacing of eye developmental events relative to the rest of development.

Optimizing the data onto the tree in which primates and artiodactyls formed a clade to the exclusion of the rat (Fig. 3) changed many of the inferred synapomorphic changes within the mammals. Eighteen shifts were no longer recovered, 7 shifts were optimized to a deeper or higher node on the tree, and 3 new shifts were recovered. This is not surprising, because the movement of the rat to a basal position would affect the inferred primitive conditions of the mammals and thus the interpretation of convergence and synapomorphy. For example, when the rat and primates formed a clade, an advance in the appearance of the mesonephric duct (kidney A) was optimized as a shared character of artiodactyls, with a convergent advance in the tarsier. However, when primates and artiodactyls form a clade, the optimization changes, such that humans and the rat now show a convergent *delay* in kidney A.

Interestingly, only few of the shifts in cardiovascular development are affected by the change in topology. An advance in the formation of the outflow tract valves (cardiovascular J) is interpreted as a synapomorphy of primates + artiodactyls rather than an autapomorphy of humans. An advance in ventricular trabeculae formation (cardiovascular I) is unchanged as a synapomorphy of the mammals, but an additional advance of the same event is no longer recovered for the rat (now on an adjacent node). A delay of the appearance of the atrioventricular canal (cardiovascular D) is no longer recovered as a primate synapomorphy. A convergent shift of cardiovascular D, forming a synapomorphy of the birds, is also lost (this is the only difference between the optimizations among the nonmammalian amniotes).

In contrast, the inferred delays in eye development are heavily dependent on the position of the rat. Only a delay in the final detachment of the lens vesicle from the surface ectoderm (optic E) remains as a synapomorphy of the mammals. Clearly, the advances in optic cup formation (optic C) and the depression of the lens placode to form the lens vesicle (optic D), recovered as autapomorphies of the rat in the previous tree, affect the optimizations when the rat is the sister taxon to primates + artiodactyls. The extent of the difference can be investigated by examining the position (or rank) of these events within the developmental sequence. The rank of each event (optic B, C, D, and E) in the rat can be compared with its mean rank for diapsids and for artiodactyls + primates (the mean must be corrected for phylogenetic relatedness; Harvey and Pagel 1991). Although optic B, C, and D appear earlier in the sequence of the rat than the average for the remaining mammals, they are also much later than the average for diapsids: optic E in the rat has a similar rank to the average of the remaining mammals (data not shown).

Parsimony analysis

We recovered a single most parsimonious tree (Fig. 4). The monophyly of our ingroup (the amniotes) was corroborated, but only two of the clades present on our reference trees were recovered (mammals, artiodactyls). The diapsids were found to be polyphyletic, with the sand lizard as the sister taxon of the remaining amniotes. The rat was found to be the sister taxon of the other mammals (in agreement with the molecular studies cited above), but the primates were found to be polyphyletic. Templeton and winning-sites tests (Templeton 1983) applied as a one-tailed test (Goldman et al. 2000) show that this tree is a significantly better fit of the data ($P = 0.0001$) than either reference tree. The total number of changes identified within the amniotes (76) was slightly less than for either reference tree.

An interesting difference between the reference trees and the most parsimonious tree is the clustering of birds and mammals to the exclusion of the sand lizard. This clade is supported by six shifts, including an advance in the first ap-

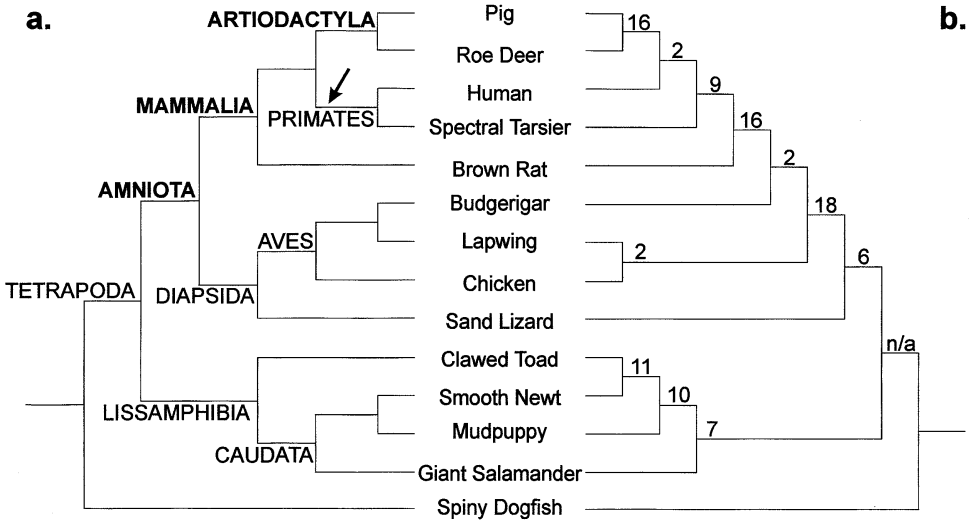

Fig. 4. Comparison of phylogenies. (A) Reference cladogram representing current phylogenetic opinion. Higher taxa listed in uppercase letters; ingroup taxa in bold type were recognized as monophyletic in our parsimony analysis. Arrow indicates alternative position of the rat. (B) Single most parsimonious tree of 1117 steps (CI excluding uninformative characters 0.570 [expected CI for 14 taxa is 0.634 {Sanderson and Donoghue 1989}]; RI 0.562; RC 0.351) generated using a branch and bound search strategy. The topology of the outgroup (spiny dogfish + four amphibians) was unconstrained, although we have rooted the most parsimonious tree in this figure using the spiny dogfish. This species is likely to be the most primitive member of our outgroup, and using it as the root produces a monophyletic lissamphibian clade. Strong hierarchical signal in the data set was revealed by a permutation tail probability test (Faith and Cranston 1991; $P = 0.0001$) and skewness test (Hillis and Huelsenbeck 1992; $g_1 = -0.756$, $P << 0.01$). Numbers on branches are Bremer decay index values (Bremer 1988) indicating relative levels of support.

pearance of the ventricular trabeculae (cardiovascular I) and a delay in the appearance and depression of the nasal placodes (olfactory A and B). On the reference trees, these three events were inferred to have shifted in the opposite direction in the sand lizard. Also, delays in olfactory A and B were inferred to be a synapomorphy of the mammals in the reference tree in which the rat and primates formed a clade. Thus, as might be expected, many shifts that mapped as convergent phenomena on the reference cladograms (homoplasies) now behave as shared derived characters (synapomorphies) at different nodes. One shift (advance in the first appearance of septum primum of atrium; cardiovascular F) is probably an artifact, caused by the inapplicability of this event in the spiny dogfish. However, the distribution of cardiovascular shifts in the rest of the tree is similar to that of the reference tree in which primates and artiodactyls form a clade.

DISCUSSION

Mapping of data

Our results show cardiovascular development has been subject to heterochronic changes throughout amniote evolution. Most of the change involves cardiovascular events shifting earlier with respect to noncardiovascular events. There has been little change in the internal sequence of events within cardiac development. This probably reflects the degree of internal dependence of the sequence (cf. lens formation, above).

Advanced development of the heart in birds and mammals was first linked to the independent evolution of high metabolic rates in these groups (culminating in adult endothermy) by McCrady (1938, p. 89). Our method allows a rigorous test of this hypothesis. It is clear from our analysis that advances in heart development are more widespread within amniotes than occurrences of adult endothermy; for example, amniotes, including the ectothermic sand lizard, share an advance in the first appearance of the endocardial anlage and the atrioventricular canal (cardiovascular A and D, respectively; Fig. 2). This implies that heart development was already moving earlier in ontogeny before the modern groups of amniotes diverged. Thus the first shift in heart development was likely to have been a response to demands, other than raised metabolic rates, that were shared at the broader amniote level. One possible demand might have been the evolution of the "closed" amniote egg. The timing of the onset of heartbeat and convective blood flow has been linked to angiogenesis (Burggren et al. 2000), an important factor in amniote embryonic respiration and food uptake (Luckett 1977; Seymour and Bradford 1995). Vascular development has also been proposed as a key factor in heat distribution within incubated eggs (Turner 1987).

Whatever their original cause, these shifts may have produced initial conditions that were subsequently "exapted" (sensu Gould and Vrba 1982) in the lineages leading to endothermic birds and mammals.

There is a concerted delay of four events in mammalian eye development (optic B, C, D, and E) under one reference tree. This correlates with the small size of embryonic eyes (and accompanying mid-brain structures) in mammals relative to other amniotes (Fig. 5). This correlation is consistent with the idea that the delayed onset of an organ results in its reduced size at later stages in embryonic development (Huxley 1932). However, the result is sensitive to the phylogenetic position of the rat, because this species does not share all the changes in sequence. If the rat is placed in a clade with the primates, its sequence is interpreted as a partial reversal of the primitive mammalian shifts. If it is placed as the sister taxon to the other mammals, then only a delay in the final detachment of the lens vesicle from the surface ectoderm (optic E) is retained as a synapomorphy. Clearly, this makes any hypothesis of mammalian optic heterochronies extremely tentative. However, it serves to illustrate a strength of the event-pair method: the robustness of specific hypotheses of heterochronic change can easily be tested by examining their sensitivity to perturbations in tree topology. In contrast to the optic heterochronies, for example, hypotheses of cardiovascular heterochrony within mammals are comparatively insensitive to the position of the rat. Once the sensitivity has been recognized, its causes can be identified (in this case, the unique developmental sequence of the rat). This in turn can prompt further research questions—for example, do other rodents show this sequence and can it be correlated with other biological phenomena?

Parsimony analysis

In our parsimony analysis, we found some degree of congruence with the reference cladograms (Fig. 4). This suggests that heterochrony data do contain a phylogenetic signal. However, there are several inconsistencies. These may reflect homoplasy caused by evolutionary convergence and insufficient taxon sampling. For example, it is possible that the shifts supporting the bird + mammal clade relate in some way to the independent evolution of endothermy in these groups (this phenomenon has been observed in molecular phylogenies based on globins; Bishop and Friday 1987, 1988). If so, this effect is likely to be exaggerated by the few species for which data are available (i.e., they represent a problem of taxon sampling). Including data from the ectothermic crocodilians could help resolve this issue. As the extant sister group of birds, crocodilians are likely to share much more of their developmental sequence with them, thereby overcoming any convergent development among endotherms. Unfortunately, sufficiently detailed developmental data are lacking for crocodilians (as discussed in Materials and Methods).

The inconsistencies may also stem from the inherent nonindependence of event-paired data (Smith 1997), which violates an assumption of phylogenetic analysis. This problem only affects the ability to infer a tree from event-paired data, not the interpretation of synapomorphies when event-paired data are mapped on to a tree. Although nonindependence is

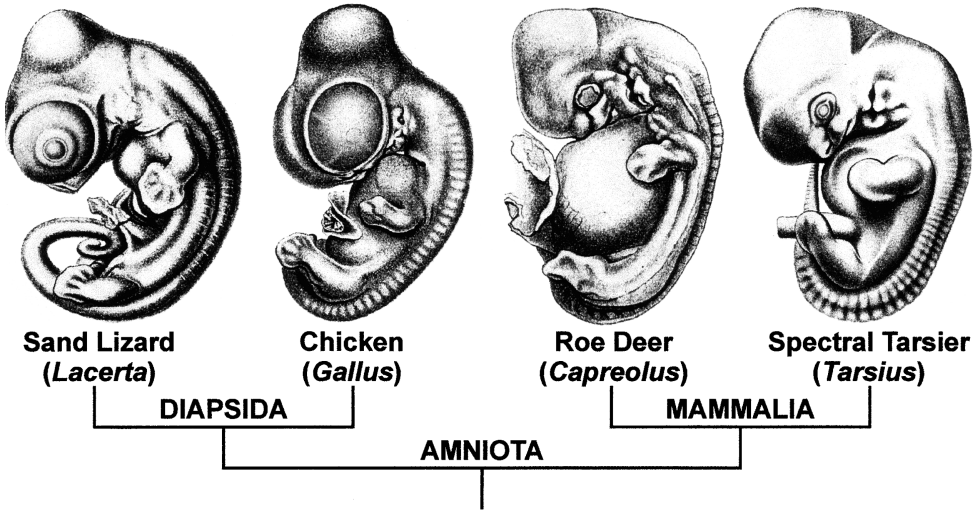

Fig. 5. Eye size in selected amniotes. The nocturnal tarsier has among the largest eye to body mass ratio of any mammal. Yet at embryonic stages its eyes are comparable in size with that of other mammals (e.g., roe deer) and much smaller than that of nonmammalian amniotes (e.g., sand lizard and chicken). If mammals share a late initiation of eye structures (see Discussion), the large eye of the tarsier must be achieved by other mechanisms (e.g., allometric heterochrony; Huxley 1932; Gould 1977). Drawings taken from Keibel (1897–1938).

a factor in most phylogenetic data sets, both morphological and molecular (e.g., Lovejoy et al. 1999) event pairing presents an additional unique category of nonindependence. This is because a cross-compared secondary data set (event pairs) is examined, not the primary data set (event times). As discussed above, event pairing is necessary to overcome the lack of a universal timing standard in development, but its effects on phylogenetic reconstruction are poorly understood. We are currently investigating any possible effects using computer simulations.

Applications of the method

We believe our method will be useful in the analysis of comparative data, both temporal and spatial. It is especially appropriate for developmental studies, ranging from embryonic anatomy to changing patterns of gene expression. It makes it possible to overcome a central problem in comparative developmental biology, namely the lack of common stages or a universal timing standard in different species.

Gene expression databases have been established to facilitate comparisons of developmental gene expression in different species. Unfortunately, the expression data are often not accompanied by timing data. For example, the organizers of the Mouse Gene Expression Database (Ringwald et al. 2000; http://www.informatics.jax.org/mgihome/GXD/aboutGXD.shtml) found that approximately two-thirds of publications on gene expression in the developmental literature do not report staging criteria for the expression patterns (David Hill, personal communication). In the future, a useful approach would be for such databases to report a group of developmental events that coincide with the expression pattern. This could allow the expression pattern to be positioned within a morphological developmental sequence and analyzed between species using the techniques demonstrated above.

The utility of event pairing for directly reconstructing evolutionary relationships is unclear because of the inherent nonindependence of event-paired characters (Smith 1997). However, when event-paired data are mapped onto existing phylogenies, the inferred states at internal nodes deliver hypotheses of ancestral developmental sequences (cf. Velhagen 1997). As we have shown, these sequences can be used to test theories linking adult conditions to embryonic heterochronies (e.g., the suggested link between endothermy and advanced heart development) and to suggest possible future lines of research.

Acknowledgments

We are grateful to Dr. H. Larsson and an anonymous reviewer for their comments on an earlier version of this manuscript. We also thank Profs. K. K. Smith and A. Minelli for helpful discussions. J. E. J. is supported by the Leverhulme Trust and the van der Leeuw Fonds, O. B.-E. by an NSERC postdoctoral fellowship, M. I. C. by a BBSRC Advanced fellowship, and M. K. R. by the van der Leeuw Fonds and Leiden University.

REFERENCES

Abraham, K. 1901. Beiträge zur Entwicklungsgeschichte des Wellensittichs (*Melopsittacus undulatus*). *Anat. Hefte* 17: 589–669.

Alberch, P. 1985. Problems with the interpretation of developmental sequences. *Syst. Zool.* 34: 46–58.

Bishop, M. J., and Friday, A. E. 1987. Tetrapod relationships: the molecular evidence. In C. Patterson (ed.). *Molecules and Morphology in Evolution: Conflict or Compromise?* Cambridge University Press, Cambridge, pp. 123–139.

Bishop, M. J., and Friday, A. E. 1988. Estimating the interrelationships of tetrapod groups on the basis of molecular sequence data. In M. J. Benton (ed.). *The Phylogeny and Classification of the Tetrapods, Volume 1: Amphibians, Reptiles, Birds.* Clarendon Press, Oxford, pp. 33–58.

Blanco, M. J., Misof, B. Y., and Wagner, G. P. 1998. Heterochronic differences of *Hoxa-11* expression in *Xenopus* fore- and hind limb development: evidence for lower limb identity of the anuran ankle bones. *Dev. Genes. Evol.* 208: 175–187.

Bolker, J. A. 2000. Modularity in development and why it matters to evo-devo. *Am. Zool.* 40: 770–776.

Bremer, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42: 795–803.

Burggren, W. W., Warburton, S. J., and Slivkoff, M. D. 2000. Interruption of cardiac output does not affect short-term growth and metabolic rate in day 3 and 4 chick embryos. *J. Exp. Biol.* 203: 3831–3838.

Dettlaff, T. A., and Dettlaff, A. A. 1961. On relative dimensionless characteristics of the development duration in embryology. *Arch. Biol.* 72: 1–16.

Dolle, P., Dierich, A., LeMeur, M., Schimmang, T., Schuhbaur, B., Chambon, P., and Duboule, D. 1993. Disruption of the Hoxd-13 gene induces localized heterochrony leading to mice with neotenic limbs. *Cell* 75: 431–441.

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.) 120: 135–142.

Faith, D. P., and Cranston, P. S. 1991. Could a cladogram this short have arisen by chance alone? On permutation tests for cladistic structure. *Cladistics* 7: 1–28.

Ferguson, M. W. J. 1985. Reproductive biology and embryology of the crocodilians. In C. Gans, F. Billett, and P. F. A. Maderson (eds.). *Biology of the Reptilia. Volume 14, Development A.* John Wiley & Sons, New York, pp. 329–491.

Ferkowicz, M. J., and Raff, R. A. 2001. Wnt gene expression in sea urchin development: Heterochronies associated with the evolution of developmental mode. *Evol. Dev.* 3: 24–33.

Forey, P. L., and Kitching, I. J. 2000. Experiments in coding multistate characters. In R. Scotland and T. Pennington (eds.). *Homology and Systematics. Coding Characters for Phylogenetic Analysis.* Taylor & Francis, London, pp. 54–80.

Goldman, N., Anderson, J. P., and Rodrigo, A. G. 2000. Likelihood-based tests of topologies in phylogenetics. *Syst. Biol.* 49: 652–670.

Gould, S. J. 1977. *Ontogeny and Phylogeny.* Belknap Press, Cambridge, Massachusetts.

Gould, S. J. 2000. Of coiled oysters and big brains: how to rescue the terminology of heterochrony, now gone astray. *Evol. Dev.* 2: 241–248.

Gould, S. J., and Vrba, E. S. 1982. Exaptation—a missing term in the science of form. *Paleobiology* 8: 4–15.

Haeckel, E. 1866. *General Morphologie der Organismen.* Georg Reimer, Berlin.

Hall, B. K., and Miyake, T. 1997. How do embryos measure time? In K. J. McNamara (ed.). *Evolutionary Change and Heterochrony.* Wiley, New York, pp. 3–20.

Hamburger, V., and Hamilton, H. L. 1951. A series of normal stages in the development of the chick embryo. *J. Morphol.* 88: 49–92.

Harvey, P. H., and Pagel, M. D. 1991. *The Comparative Method in Evolutionary Biology.* Oxford University Press, Oxford.

Hillis, D. M., and Huelsenbeck, J. P. 1992. Signal, noise, and reliability in molecular phylogenetic analyses. *J. Hered.* 83: 189–195.

Hirata, M., Ito, K., and Tsuneki, K. 1998. Significance of heterochronic differences in neural crest cell migration and sclerotomal development in evolution of the trunk body organization from agnathans to gnathostomes. *Zool. Sci.* 15: 903–912.

Holland, P. W. 1999. The future of evolutionary developmental biology. *Nature* 402 (suppl.): 41–44.

Huxley, J. S. 1932. *Problems of Relative Growth.* Methuen, London.

Jeffery, J. E., Bininda-Emonds, O. B. E., Coates, M. I., and Richardson, M. K. 2002. Analysing developmental sequences within a developmental framework. *Syst. Biol.* 51: In Press.

Keibel, F. 1897–1938. *Normentafeln zur Entwicklungsgeschichte der Wirbelthiere.* Verlag von Gustav Fischer, Jena.

Keibel, F. 1912. The interdependence of the various developmental processes. In F. Keibel and F. P. Mall (eds.). *Manual of Human Embryology.* J. B. Lippincott, Philadelphia & London, pp. 980–1002.

Kendall, M. G. 1970. *Rank Correlation Methods.* Griffin, London.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., and Schilling, T. F. 1995. Stages of embryonic development of the zebrafish. *Dev. Dynam.* 203: 253–310.

Knoetgen, H., Viebahn, C., and Kessel, M. 1999. Head induction in the chick by primitive endoderm of mammalian, but not avian origin. *Development* 126: 815–825.

Liu, F.-G. R., and Miyamoto, M. M. 1999. Phylogenetic assessment of molecular and morphological data for eutherian mammals. *Syst. Biol.* 48: 54–64.

Liu, F.-G. R., Miyamoto, M. M., Freire, N. P., Ong, P. Q., Tennant, M. R., Young, T. S., and Gugel, K. F. 2001. Molecular and morphological supertrees for eutherian (placental) mammals. *Science* 291: 1786–1789.

Lovejoy, C. O., Cohn, M. J., and White, T. D. 1999. Morphological analysis of the mammalian postcranium: a developmental perspective. *Proc. Natl. Acad. Sci. USA* 96: 13247–13252.

Luckett, W. P. 1977. Ontogeny of amniote fetal membranes and their application to phylogeny. In M. K. Hecht, P. C. Goody, and B. M. Hecht (eds.). *Major Patterns in Vertebrate Evolution.* Plenum Press, New York, pp. 439–516.

Mabee, P. M., and Trendler, T. A. 1996. Development of the cranium and paired fins in *Betta splendens* (Teleosti: Percomporpha): intraspecific variation and interspecific comparisons. *J. Morphol.* 227: 249–287.

Maddison, W. P., and Maddison, D. R. 2001. *MacClade: Analysis of Phylogeny and Character Evolution*, version 4.0. Sinauer Associates, Sunderland, Massachusetts.

Madsen, O., Scally, M., Douady, C. J., Kao, D. J., DeBry, R. W., Adkins, R., Amrine, H. M., Stanhope, M. J., de Jong, W. W., and Springer, M. S. 2001. Parallel adaptive radiations in two major clades of placental mammals. *Nature* 409: 610–614.

McCrady, E. 1938. Embryology of the opossum. *Am. Anat. Mem.* 16: 1–233.

McKinney, M. L., and McNamara, K. J. 1991. *Heterochrony, the Evolution of Ontogeny.* Plenum, New York.

Murphy, W. J., Eizirik, E., Johnson, W. E., Zhang, Y. P., Ryder, O. A., and O'Brien, S. J. 2001. Molecular phylogenetics and the origins of placental mammals. *Nature* 409: 614–618.

Nieuwkoop, P. D., and Faber, J. 1994. *Normal Table of Xenopus laevis (Daudin)*, 2 Ed. Garland Publishing, London.

Nunn, C. L., and Smith, K. K. 1998. Statistical analyses of developmental sequences: the craniofacial region in marsupial and placental mammals. *Am. Nat.* 152: 82–101.

Oppel, A. 1891. *Vergleichung des Entwicklungsgrades der Organe.* Fischer, Jena.

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., Hayward, D. C., Ball, E. E., Degnan, B., Müller, P., Spring, J., Srinivasan, A., Fishman, M., Finnerty, J., Corbo, J., Levine, M., Leahy, P., Davidson, E., and Ruvkun, G. 2000. Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA. *Nature* 408: 86–89.

Raff, R. A., Arthur, W., Carroll, S. B., Coates, M. I., and Wray, G. 1999. Chronicling the birth of a discipline. *Evol. Dev.* 1: 1–2.

Reinhart, B. J., and Ruvkun, G. 2001. Isoform-specific mutations in the *Caenorhabditis elegans* heterochronic gene *lin-14* affect stage-specific patterning. *Genetics* 157: 199–209.

Richardson, M. K. 1995. Heterochrony and the phylotypic period. *Dev. Biol.* 172: 412–421.

Richardson, M. K., Hanken, J., Gooneratne, M. L., Pieau, C., Raynaud, A., Selwood, L., and Wright, G. M. 1997. There is no highly conserved embryonic stage in the vertebrates: implications for current theories of evolution and development. *Anat. Embryol.* 196: 91–106.

Richardson, M. K. 1999. Vertebrate evolution: the developmental origins of adult variation. *BioEssays* 21: 604–613.

Ringwald, M., Eppig, J. T., Kadin, J. A., and Richardson, J. E. 2000. GXD: a gene expression database for the laboratory mouse: current status and recent enhancements. *Nucleic Acids Res.* 28: 115–119.

Sanderson, M. J., and Donoghue, M. J. 1989. Patterns of variation in levels of homoplasy. *Evolution* 43: 1781–1795.

Schlosser, G. 2001. Using heterochrony plots to detect the dissociated coevolution of characters. *J. Exp. Zool. (Mol. Dev. Evol.)* 291: 282–304.

Seymour, R. S., and Bradford, D. F. 1995. Respiration of amphibian eggs. *Physiol. Zool.* 68: 1–25.

Slack, J. M., Holland, P. W., and Graham, C. F. 1993. The zootype and the phylotypic stage. *Nature* 361: 490–492.

Smith, K. K. 1996. Integration of craniofacial structures during development in mammals. *Am. Zool.* 36: 70–79.

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

Smith, K. K. 2001a. Heterochrony revisited: the evolution of developmental sequences. *Biol. J. Linn. Soc.* 73: 169–186.

Smith, K. K. 2001b. Early development of the neural plate, neural crest and facial region of marsupials. *J. Anat.* 199: 121–131.

Swofford, D. L. 2001. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*, version 4.0b8. Sinauer Associates, Sunderland, Massachusetts.

Templeton, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution* 37: 221–244.

Turner, J. S. 1987. Blood circulation and the flows of heat in an incubated egg. *J. Exp. Zool* 1 (suppl.):99–104.

Velhagen, W. A. 1997. Analyzing developmental sequences using sequence units. *Syst. Biol.* 46: 204–210.

von Dassow, G., and Munro, E. 1999. Modularity in animal development and evolution: elements of a conceptual framework for EvoDevo. *J. Exp. Zool.* 285: 307–325.

Waddell, P. J., Okada, N., and Hasegawa, M. 1999. Towards resolving the interordinal relationship of placental mammals. *Syst. Biol.* 48: 1–5.

Wagner, G. P. 1996. Homologues, natural kinds and the evolution of modularity. *Am. Zool.* 36: 36–43.

Winther, R. G. 2001. Varieties of modules: kinds, levels, origins, and behaviors. *J. Exp. Zool.* 291: 116–129.