

What can DNA Tell us About the Cambrian Explosion?¹

LINDELL BROMHAM²

Centre for the Study of Evolution, School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG, UK

SYNOPSIS. Molecular data is ideal for exploring deep evolutionary history because of its universality, stochasticity and abundance. These features provide a means of exploring the evolutionary history of all organisms (including those that do not tend to leave fossils), independently of morphological evolution, and within a statistical framework that allows testing of evolutionary hypotheses. In particular, molecular data have an important role to play in examining hypotheses concerning the tempo and mode of evolution of animal body plans. Examples are given where molecular phylogenies have led to a re-examination of some fundamental assumptions in metazoan evolution, such as the immutability of early developmental characters, and the evolvability of bauplan characters. Molecular data is also providing a new and controversial time-scale for the evolution of animal phyla, pushing the major divisions of the animal kingdom deep into the Precambrian. There have been many reasons to question the accuracy and precision of molecular date estimates, such as the failure to account for lineage-specific rate variation and unreliable estimation of rates of molecular evolution. While these criticisms have been largely countered by recent studies, one problem has remained a challenge: could temporal variation in the rate of molecular evolution, perhaps associated with “explosive” adaptive radiations, cause overestimation of diversification dates? Empirical evidence for an effect of speciation rate, morphological evolution or ecological diversification on rates of molecular evolution is examined, and the potential for rate-variable methods for molecular dating are discussed.

The questions asked by George Gaylord Simpson in his 1944, “The Tempo and Mode of Evolution,” concerning the size of mutations, the pace of morphological change and the apparent discontinuous origins of taxa in the fossil record, are far from resolved. Indeed, they are being debated more strongly than ever, because of the growing conviction amongst many biologists that observations from developmental biology and palaeontology are inconsistent with the Neodarwinian hypothesis championed by Simpson. The origin of the animal phyla has been a key case study in the tempo and mode of evolution.

The near-simultaneous appearance in the early Cambrian of the first recognizable members of many animal phyla, at high diversity and with few clear precursors, has been attributed by many researchers an incomplete fossil record (*e.g.*, Darwin, 1859). However, the description of late Precambrian body fossils and traces from around the world, from which members of diverse Cambrian phyla such as arthropods are conspicuously absent, has made arguments for a missing history less convincing. More broadly, analyses of fossil ranges across a wide range of taxa have led to increasing faith in the fossil record as an accurate record of tempo and mode of evolution (*e.g.*, Benton *et al.*, 2000). Such analyses have led to the suggestion that rather than a steady accumulation of biological diversity, the global biota has been fundamentally shaped by a series of major events, mass extinctions of taxa followed by explosive radiations of new taxa. The first explosive radiation—the Cambrian explo-

sion—is particularly important, as it has been argued that the diversity of animal body plans was achieved soon after their origin, and that virtually no new body plans have evolved since.

The apparently sudden origin of animal phyla has contributed to the view that phyla represent a fundamental level of organization. In particular, phyla are held to be defined primarily by a set of discrete body plan characters, rather than simply differing at a large number of continuously varying traits. Just as species are considered by many to represent a fundamental evolutionary unit, not just an arbitrary division of biological diversity, phyla are considered by many not simply as a taxonomic division, but as a reflection of an underlying biological structure.

The origin of animal phyla (both their sudden appearance in the fossil record and the discontinuous variation in bodyplan traits between phyla) has also been the focus of research in the relatively young field of evolutionary development (Evo-Devo). The characterization of homeotic mutants, that alter the development of whole structures such as limbs or eyes, has led some researchers to suggest that the differences in body plan between animal phyla could have arisen through relatively few genetic changes. In particular, the Hox genes (and related genes) have been implicated as controllers of body plan characters, and the differences between phyla have been attributed to variation in the number and expression of Hox-like genes. Furthermore, the Cambrian explosion has been interpreted as the point at which Hox clusters formed, and were subsequently canalized so that no new body plans formed after that time (*e.g.*, Valentine *et al.*, 1999).

So these three observations—the sudden appearance of animal phyla in the fossil record, the discontinuity between animal body plans, and mutations that bring

¹ From the Symposium *The Cambrian Explosion: Putting the Pieces Together* presented at the Annual Meeting of the Society for Integrative and Comparative Biology, 2–6 January 2002, at Anaheim, California.

² E-mail: L.D.Bromham@sussex.ac.uk

about discrete body-plan-like changes—have been combined in the hypothesis that the evolution of animal body plans did not occur by the gradual accumulation of small genetic differences, but by relatively few developmental changes with large phenotypic effects. Because this hypothesis offers a direct challenge to the Neodarwinian view of evolution (evolutionary change by the gradual accumulation of adaptive changes), the importance of these interpretations of the Cambrian explosion go beyond explaining a single evolutionary radiation.

The Neodarwinian hypothesis does not assume constant rates of evolutionary change, but it does assume uniformity of process—that all evolutionary change arises by the same basic mechanism. Darwin's (1859) hypothesis provided a unifying theory for biological diversity because his mechanism for evolution was explicitly uniformitarian. He followed the revolutionary works of Charles Lyell (1830) in which the dramatic geological changes of the past—such as mountain building or sea level changes—were explained in terms of the cumulative effect of processes witnessed in operation today, such as sedimentation or uplift. Darwin explained the extraordinary changes in species over time using a simple mechanism operating at the level of populations: the same process that causes relatively modest differences over observable periods produces major changes over much longer timescales. In other words, macroevolution (formation of higher taxa) is simply the cumulative effect of microevolution (population genetic changes). The new view of evolution challenges this assumption, by proposing that the evolution of animal body-plans occurred by a different process than that which modifies existing body parts. If the differences between animal phyla must be explained by a process that operated in the early Cambrian, but not before or since, then this is a serious challenge to the uniformitarian principles upon which evolutionary biology has been based for the past century and a half. It is therefore very important that the hypothesis is thoroughly tested.

Molecular data can contribute to this debate by providing a means of inferring phylogeny independently of morphological characters, and a timescale independent of the fossil record. The key features of molecular data that make it so useful for assessing tempo and mode of evolution are universality, abundance and stochasticity. DNA data is directly comparable across all extant organisms, including those that do not fossilize well such as soft bodied taxa, and contains information for the entire history of every lineage. DNA sequence data provides an abundant data source, with thousands of independently evolving characters in even the smallest genome. DNA sequences evolve in a predominantly stochastic manner because, although every heritable change to the phenotype must be associated with a change in the genome, the reverse is not true: most changes to the genome are apparently unconnected to phenotypic change. Many changes occur in non-translated DNA sequences, or at sites that do not make a

functional change to a gene product. Because most DNA substitutions will have little effect on fitness, the rate of accumulation of genetic change should increase as a function of time. These features of molecular data provide the basis for independent tests of hypotheses concerning the tempo and mode of metazoan evolution generated from morphological phylogeny and fossil evidence.

EVOLUTION OF ANIMAL BODY PLANS

The hypothesis that body plans were formed in the Cambrian Explosion then canalized predicts that there are critical characters that not only define phyla, but were instrumental in their formation (*e.g.*, Carroll, 1995; Erwin, 1993; Valentine *et al.*, 1999). This contrasts to the Neodarwinian perspective that the apparent origin of phyla in the Cambrian is simply a reflection of the amount of time needed to accumulate sufficient differences to be considered separate phyla (*e.g.*, Williams, 1992). The claim that critical body plan characters formed in the Cambrian, but not since, can be tested directly from fossil evidence, which provides evidence of observable bodyplan characters present before, during and after the Cambrian explosion (*e.g.*, Budd and Jensen, 2000), or indirectly by inference of the relative timing of character evolution from phylogenies. However, assessing the relative timing of evolution of body plan characters from phylogenies constructed from morphological data risks a circular argument. If one of the assumptions on which the phylogeny is based is that a set of characters are so fundamental that they will evolve only once, and therefore define major groups, then the resulting phylogeny cannot be used to test the same assumption. The stochastic nature of molecular change makes it ideal for providing an independent source of phylogenetic information to test these claims about the tempo and mode of morphological evolution.

The deuterostomes provide an illustrative example of the use of molecular phylogenies to examine the evolution of body plan characters. The superphylum Deuterostomia contains the diverse phyla Chordata and Echinodermata and the minor phyla Hemichordata and Urochordata. Although united by early developmental characteristics, echinoderms and chordates have strikingly different adult body plans. The head-and-tail chordate body plan is characterized by a pharynx, dorsal nerve cord and post-anal tail. The echinoderm body plan has pentameral symmetry, a water vascular system and a ring-shaped nervous system. Hemichordates have traditionally been allied with the chordates (Fig. 1a) on the basis of shared adult body plan features, particularly the pharynx and dorsal nerve chord. Molecular phylogenies (Fig. 1b) have suggested that hemichordates are in fact a sister lineage to the echinoderms (*e.g.*, Castresana *et al.*, 1998; Holland *et al.*, 1991). The morphological phylogeny gave few clues of the body plan of the ancestor to all deuterostome phylogeny (Fig. 1a). But if hemichordates are the sister group to echinoderms, then the deu-

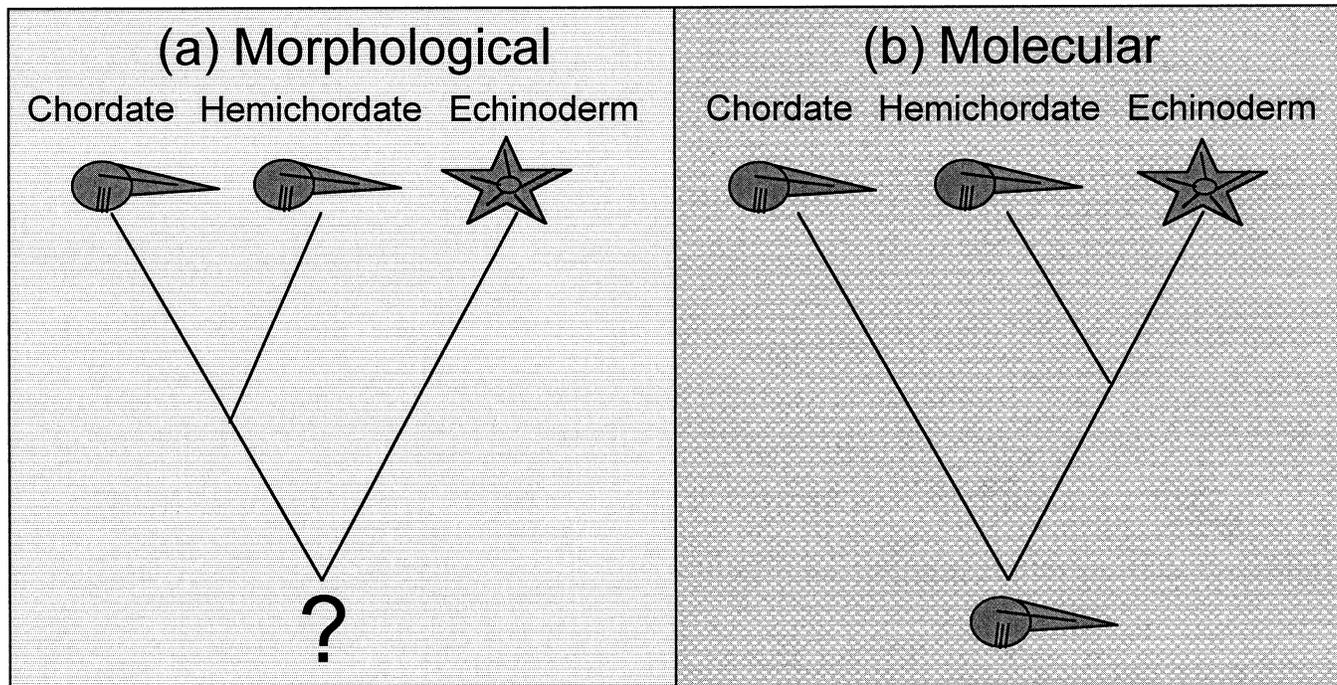


FIG. 1. Hemichordates were previously allied to chordates on the basis of the shared body plan characters of a pharynx and dorsal nerve cord. Molecular phylogenies have suggested that hemichordates are more closely related to echinoderms, a grouping that implies the deuterostome ancestor had the chordate body plan features of a pharynx and dorsal nerve cord, which were subsequently lost in the echinoderm lineage.

terostome ancestor must have had chordate body plan characters including a pharynx and dorsal nerve cord, because these characters are present in both major branches of the deuterostomes (Bromham and Degnan, 1999). So molecular data suggest that the echinoderm lineage, with its distinctive pentamerous body plan, evolved from an ancestor that had the key features of the head-and-tail chordate body plan. It also implies that aspects of the chordate body plan did not evolve as a single complex. Instead, the postanal muscularized tail of chordates evolved after the pharynx (Bromham and Degnan, 1999; Hinman and Degnan, 2000). Whatever timescale this change in body plan characters occurred on, it suggests that body plans of metazoan phyla were not formed once then canalized, but were able to evolve in stepwise fashion from one body plan to another. Animal body plans are therefore not immutable.

Early developmental characters have also been considered to be relatively immutable. The assumption that the earlier a character is in the developmental hierarchy, the less it may be modified in evolution, has provided the basis for constructing phylogenies of the animal kingdom (*e.g.*, Haeckel, 1879). By providing phylogenies independent of developmental characters, DNA sequence data have allowed these assumptions to be re-examined. In some cases, this has supported the phylogenetic significance of developmental characters, such as the superphylum Spiralia which is characterized by a pattern of spiral cleavage in the early embryo. In other cases, molecular phylogenies suggest

that some developmental characters are less constrained than is often assumed. The minor phyla of Spiralia may provide an illustrative case. The phylum Sipuncula (acorn worms) has been considered allied to the molluscs on the basis of a pattern of cells in early spiral cleavage of the embryo, known as the molluscan cross. But some molecular phylogenies (*e.g.*, Peterson and Eernisse, 2001; Boore and Staton, 2002) place Sipuncula with Annelida, a grouping that implies that molluscan cross pattern of early spiral cleavage evolved several times independently. The phylogenetic position of the Sipuncula is still unclear, but whether or not the alliance between Sipuncula and annelids is confirmed by future phylogenetic analyses (both molecular and morphological), this initial result should lead to careful scrutiny of character coding. The perception that early developmental characters are fundamental may have led to bias in coding some developmental patterns, which may form a continuous variation in early cell sizes and patterns, rather than a number of discrete, distinctive developmental patterns (see Jenner, 2003).

Molecular phylogenies of Metazoa suggest that even fundamental developmental and body plan characters can evolve along the metazoan phylogeny, rather than being fixed at the base of the radiation. Clearly the phylogenetic hypotheses generated from molecular data need to be thoroughly tested, particularly as at present most metazoan phylogenies are based on a single gene, 18S rRNA. It may be that some of these phylogenetic hypotheses are incorrect. But by provid-

ing a source of phylogenetic information independent of morphological and developmental characters, molecular phylogenies offer a way of testing traditional assumptions about the evolutionary lability of traits, and will in many cases lead to a reexamination of the characters on which systematic divisions of Metazoa have been based.

MOLECULAR CLOCKS

The stochastic nature of DNA sequence evolution leads to the prediction that substitutions should increase as a function of time. This prediction is broadly borne out: in general, the more distantly related two species are, the more sites in a DNA or protein sequence differ between them. In many cases, if genetic distance is plotted against time, a linear relationship is revealed (*e.g.*, Fleischer *et al.*, 1998; Runnegar, 1982; Zuckerkandl and Pauling, 1965), suggesting that genetic data can be used to predict divergence times. But molecular clock estimates for the origin of metazoan lineages are at odds with the timing of appearance of metazoan phyla in the fossil record. While molecular clock estimates vary greatly, all estimates for the major divisions of the Metazoa (*e.g.*, between bilaterians and deuterostomes) are all at least 55 Ma before the first unambiguous metazoan body fossils in the early Cambrian, and more than 25 Ma before the first multicellular animals in the Vendian (see Bromham and Hendy, 2000). Most molecular dates are one hundred million years or more before the Cambrian explosion.

There are a number of possible explanations of the discrepancy between molecular and morphological dates of the metazoan radiation. Firstly, the gap between the molecular dates and the first fossils may be due to incompleteness of the fossil record, such that the early evolutionary history of the metazoans is obscured. This is difficult to reconcile with the apparent increase in complexity of body and trace fossils across the latest Proterozoic and earliest Phanerozoic. Secondly, molecular and palaeontological dates may be effectively measuring different things, if lineage origination is disconnected from the evolution of morphological divergence. It has been suggested that the basal splits of the metazoan tree occurred in the late Proterozoic, but these lineages persisted in relatively low diversity until the early Cambrian. In this case, it may be that the predominance of molecular date estimates for basal splits (such as the protostome–deuterostome split) is responsible for the huge disparity between molecular and fossil dates, and that dates for “shallower” splits will be closer to the Cambrian. This view raises important questions about the tempo and mode of evolution, as it requires that the early metazoan lineages persisted in some kind of evolutionary stasis even though they had all of the elements of the body plan “toolkit” that have been considered by many to have been key innovations of the metazoan radiation. Thirdly, the discrepancy between the molecular and palaeontological dates might be due to a systematic bias in molecular clock estimates that results in consistent

overestimation of the date of divergence of animal phyla.

There are a number of reasons to be worried about the accuracy and precision of molecular dates for the metazoan radiation:

1. **Failure to account for rate variation:** Most molecular dating techniques assume that the rate of molecular evolution is approximately the same in all taxa (*e.g.*, Lynch, 1999; Nei *et al.*, 2001). Since lineages can vary in rate of molecular evolution, various “clock tests” have been used to identify genes for which rates do not vary between lineages (Wang *et al.*, 1999), such as the relative rates test (Wu and Li, 1985) or the Tajima test (Tajima, 1993). These tests have low power for the type of sequences typically used in molecular clock studies (Bromham *et al.*, 2000). Using saturated sequences or distant outgroups (such fungi as an outgroup to Metazoa) also reduces the power of relative rates tests to detect rate variation (*e.g.*, Robinson *et al.*, 1998). Even relatively high levels of rate variation between lineages may not be detected by clock tests, leading to an erroneous impression of rate constancy. Undetected rate variation can lead to consistently overestimated dates of divergence (Bromham *et al.*, 2000).
2. **Poor estimation of branch length:** selection of substitution model is a critical part of molecular dating. For example, failure to account for variation in substitution rate between sites in the sequence will result in inaccurate estimation of branch length. In the case of internal calibrations (where the calibration date is younger than the node for which the molecular date is estimated), this will generally shorten branches and thus make dates too early (Ayala *et al.*, 1998): the opposite is likely to be true for external calibrations (where the calibration date is older than the estimated node). Similarly, exclusion of invariant sites (Lynch, 1999) is likely to bias estimates of both substitution rates and branch lengths.
3. **Misleading confidence intervals:** Molecular date estimates are sometimes presented with confidence intervals that represent the standard deviation of a number of estimates (commonly from different genes). This is a reflection of difference between estimates, not of the accuracy of the estimates per se. For example, erroneous estimates that all share the same measurement bias might be represented by a mean with narrow confidence intervals. In this way, collating large numbers of genes can give a false sense of confidence in molecular date estimates. Giving equal weight to all genes also falsely reduces confidence intervals, because it does not account for the effect of sequence length, substitution rate, and degree of rate variation across sites on the precision of date estimation. In addition, care must be taken to consider whether the estimates are truly independent before mean estimates are cal-

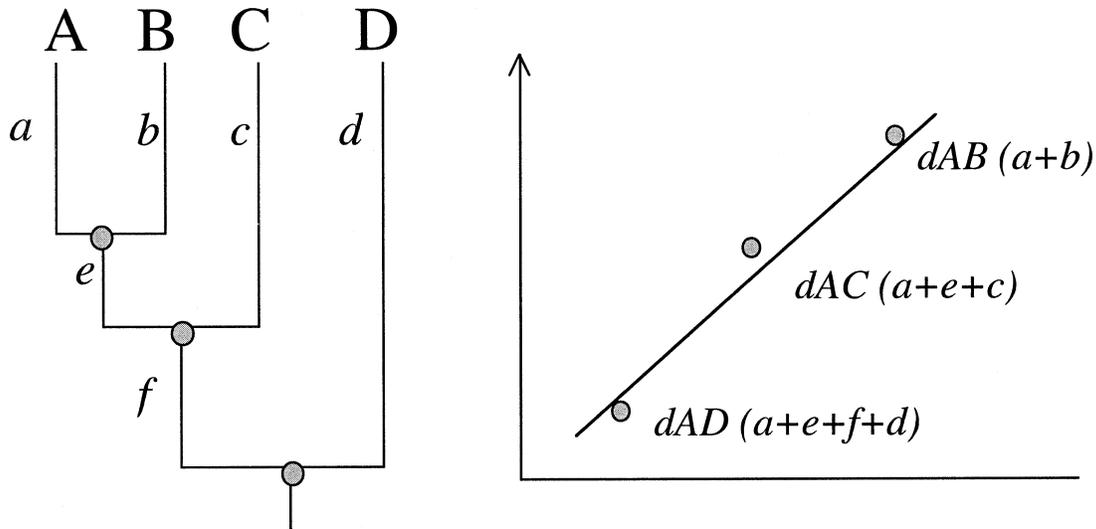


FIG. 2. The use of nested calibration dates can lead to the use of non-independent datapoints when using regression to estimate rates of molecular evolution. In this example, each of the datapoints contains the branch a , which connects node e and species A. The inclusion of a in each datapoint results in a decrease in the degrees of freedom of the regression, and may lead to an inflation of the apparent association between time and molecular divergence.

culated (Bromham *et al.*, 1998). Ideally, confidence intervals should reflect not only the difference between estimates obtained from different genes or different calibration dates, but should also reflect lineage-specific rate variation and the “sloppiness” of the molecular clock (the random distribution of the time interval between substitutions), both of which add a great deal of imprecision to molecular date estimates (see Rambaut and Bromham, 1998).

4. **Poor calibration dates:** While some studies have aimed to use as many fossil calibration dates as possible, an alternative strategy has been use a single calibration date deemed to be well supported. The accuracy of one commonly used calibration date—310 Myr for the split between reptiles and mammals (Gu, 1998; Wang *et al.*, 1999)—has been questioned (Lee, 1998). More broadly, the strategy of using rates calculated for a specific, potentially unrepresentative, group—particularly vertebrates—to date the rest of the metazoan tree (Feng *et al.*, 1997; Wang *et al.*, 1999) has been criticized (*e.g.*, Lynch, 1999). Even “local” molecular clocks should be rigorously tested, as closely related lineages can vary substantially in substitution rate (*e.g.*, Bromham, 2002; Bromham *et al.*, 1996; Mooers and Harvey, 1994). Calibrating rates on other molecular clock estimates (*e.g.*, using estimate of primate-rodent at 100 Myr to date metazoan divergences: Gu, 1998; Wang *et al.*, 1999) is an ill-advised strategy, potentially compounding the error of earlier molecular clock estimates.
5. **Regression through non-independent points:** In some studies, calibration rates are calculated by taking the molecular distance between a number of pairs of taxa, each with a different calibration point, and extrapolating a linear relationship between ge-

netic distance and time (*e.g.*, Doolittle *et al.*, 1996; Wray *et al.*, 1996). While these regressions of distance against time have the advantage of providing an instinctive test of the molecular clock, they are usually flawed due to the inclusion of non-independent points (Lynch, 1999). The inclusion of the same branch-lengths in many datapoints (Fig. 2) is likely to inflate the apparent association between branch-length and time, therefore giving false confidence in the “clock-like” evolution of a sequence.

6. **Reliance on few sequences:** 18s rRNA has dominated attempts to reconstruct molecular phylogenies for metazoans, and so has been used in many molecular dating studies (*e.g.*, Bromham *et al.*, 1998; Wray *et al.*, 1996). If this sequence gives misleading results, then it could influence the many studies that rely on 18s rRNA sequences (Abouheif *et al.*, 1998). Phylogenetic inference from rRNA genes is complicated by the secondary structure of the RNA product: for example, no substitution model used in molecular phylogenetics accounts for paired substitutions in stem sites, and so the number of independent molecular changes may be overestimated for these sites, potentially producing molecular dates that are too old (Bromham *et al.*, 1998).

However, none of these criticisms is sufficient to explain the Precambrian molecular dates for major metazoan divergences, because studies that overcame these problems have produced similar results. The accuracy of branch length estimation has been improved by using long sequences and methods that allow for variation in substitution rates across sites and between lineages (Rambaut and Bromham, 1998). Biases due to specific genes or calibration dates have been countered by studies that use a range mitochondrial and

nuclear protein-coding genes and many non-vertebrate calibrations (Ayala *et al.*, 1998; Bromham *et al.*, 1998; Feng *et al.*, 1997; Gu, 1998; Wang *et al.*, 1999). And yet all of these analyses point to Precambrian divergences of major metazoan lineages.

MOLECULAR DATES AND EXPLOSIVE RADIATIONS

One problem with molecular clocks that is more difficult to address is the potential for concerted changes in rate of molecular evolution over time. If lineages can have consistently different average substitution rates, this implies that rates evolve along phylogenies. This could create temporal patterns of substitution rates that would be difficult to detect, and hard to model. This is particularly worrying given that substitution rates can be correlated with species traits, such as life history or population dynamics (*e.g.*, Bromham, 2002; Bromham *et al.*, 1996; Johnson and Seger, 2001; Martin and Palumbi, 1993; Mooers and Harvey, 1994). A concerted evolutionary change in such species traits could generate temporal patterns in rates across many lineages, which could produce consistently misleading molecular date estimates (Bromham, 2003). This problem can be illustrated with an example taken from another “explosive” radiation which mirrors the controversy over molecular dates for the metazoan radiation.

Modern mammal orders appear suddenly in the early Tertiary fossil record, but molecular date estimates put the the major mammalian divergences deep into the Cretaceous, long before the final extinction of the dinosaurs (Kumar and Hedges, 1998; Madsen *et al.*, 2001). So molecular dates have been used to challenge the notion that the extinction of the dinosaurs was the key determinant of the radiation of modern mammals and birds, as they filled the niches left vacant by dinosaurs. This same pattern of early Tertiary fossils but Cretaceous molecular dates is repeated for the evolution of modern bird orders (Cooper and Fortey, 1998; Cooper and Penny, 1997). Is this repeated pattern of molecular dates twice as old as an explosion of diversity in the fossil record due to systematic biases in the fossil record, or is there some aspect of explosive radiations that could lead to consistent overestimation of molecular dates?

Rates of molecular evolution can vary substantially between mammalian lineages (*e.g.*, Gillespie, 1991; Li *et al.*, 1996; Yang and Nielsen, 1998). This variation will not always be detected by “clock tests” such as the relative rates test, likelihood ratio tests, or the Tajima test (Bromham *et al.*, 2000; Robinson *et al.*, 1998; Scherer, 1989). Undetected lineage-specific rate variation can result in consistent overestimation of molecular date estimates (Bromham *et al.*, 2000). Furthermore, variation in rate of molecular evolution appears to be associated with body size in mammals (Bromham *et al.*, 1996). This could be important for the accuracy of molecular dates of the mammalian radiation because virtually all modern mammal orders increased in average body size from their first appearance in the early Tertiary to the more diverse members

of the order in the Eocene. If smaller species have faster rates of molecular evolution, then an increase in average body size during the mammalian radiation could have led to a concerted slowdown in rates of molecular evolution across many lineages. This slowdown would be difficult to detect, but could result in consistent overestimation of molecular dates for divergences between mammalian orders (Bromham, 2003).

It is important to note that this line of reasoning does not prove that mammalian molecular dates are wrong. The molecular dates for mammalian divergences may be broadly correct, and the discrepancy with fossil data could be due to geographical biases in the late Cretaceous terrestrial record, a hypothesis supported by molecular phylogenetic analyses that place the origins of modern mammals in Gondwana (*e.g.*, Madsen *et al.*, 2001; Penny *et al.*, 1999). It is intended simply as an illustration of the possibility of consistent bias in molecular date estimates arising from patterns of rate of molecular evolution across lineages. Could a similar pattern account for the discrepancy between paleontological and molecular dates for other explosive radiations? The body size pattern is unlikely to provide a general explanation—for example, the radiation of modern birds was not obviously marked by an increase in average body size. The pattern of body size evolution at the base of the metazoan radiation is unknown, and there have been no systematic studies of correlates of rate variation in invertebrates, so it is not known if life history influences molecular evolution in the majority of metazoan lineages.

EXPLOSIVE RADIATIONS AND THE MOLECULAR CLOCK

Are there other determinants of rate of molecular evolution that could cause systematic errors in the molecular dates for the metazoan radiation? Adaptive radiations are characterized by several evolutionary processes that might influence rates of molecular evolution: increased speciation rate, rapid morphological change, and directional change in species characteristics.

Molecular evolution has been considered to be largely independent of the pace of phenotypic evolution, based on theoretical (Kimura, 1983) and empirical (Papadopoulos *et al.*, 1999) studies. Most substitutions are expected to be effectively neutral (Kimura, 1983; Ohta, 1993), their fate determined by chance rather than selection, and the number of genes affected by any given selection event will be vanishingly small compared to the total genome size. However, a recent study suggesting an association between molecular and morphological branch lengths (Omland, 1997) has been used to suggest that rapid phenotypic evolution associated with explosive radiations should be accompanied by faster molecular rates (Archibald, 1999; Conway Morris, 1998; Lee, 1998). If true, such a relationship would have serious implications for using molecular clocks to date adaptive radiations. However, this study (Omland, 1997) was limited by few datasets,

and statistical flaws in the experimental design that could have inflated the probability of an artefactual association between morphological and molecular branch lengths for a number of phylogenies. A study that used three new methods designed to overcome these biases on thirteen “total evidence” vertebrate datasets found no evidence of association between morphological and molecular rates of change (Bromham *et al.*, 2002). Clearly this relationship must be examined further before claims for rapid morphological evolution speeding the molecular clock in explosive radiations can be supported. Similarly, although an association between net speciation rate and substitution rate has been observed for three genes for angiosperms (Barraclough and Savolainen, 2001) and for DNA hybridization distances for passerine birds (Barraclough *et al.*, 1998), neither the generality of the relationship nor its underlying causes are known, so the relevance of this association to the Cambrian explosion is currently unknown.

The ideal way to test the effect of explosive radiations on the rate of molecular evolution would be to compare clades with a demonstrably rapid rate of diversification to similar lineages that have not undergone high rates of evolution. Island endemic radiations, characterized by high rates of speciation, adaptation into new ecological niches, and rapid morphological change, provide one way of making this comparison. These radiations are associated with many evolutionary processes that might effect the rate of molecular evolution: genetic bottlenecks as populations are initiated from a small number of colonists, rapid rate of phenotypic evolution as species are released from the constraints of the mainland ecosystems (such as predators), and novel adaptation as they evolve into a range of new niches. A study of thirteen island datasets, where the rate of molecular evolution for the island clades was compared to mainland relatives that showed no evidence of rapid evolution, revealed no consistent effect of rapid adaptive radiation on molecular evolution (unpublished data).

MOLECULAR DATING WITH VARIABLE RATES

So there is currently no empirical evidence to suggest that rates of molecular evolution would have been faster in the early evolution of the metazoan kingdom. However, it is possible that some unknown effect sped the molecular clock during the Cambrian explosion—could such unknown rate variation be allowed for in molecular clock analyses? One approach is to develop molecular phylogenetic methods that explicitly model variation in substitution rate across the phylogeny as part of the phylogenetic estimation process. There are a number of new methods that incorporate variable rates, either as a random walk of rates through time (Sanderson, 1997), or using a Bayesian framework (Kishino *et al.*, 2001).

For example, Aris-Brosou and Yang (2002) used a Bayesian approach (Thorne *et al.*, 1998; Kishino *et al.*, 2001) to allow rates of substitution of the 18S rRNA

gene to vary across the phylogeny of 39 species of metazoan. Under a clock model (constant rates across the phylogeny), they estimated divergence dates for major divisions of the metazoan tree deep in the Precambrian (1,062–1,567 million years ago, Mya). But under the rate-variable model, the date estimates were dramatically younger (516–619 Mya), coinciding neatly with the appearance of the earliest multicellular animal fossils in the Vendian. It should be noted that the estimated date of the plant-animal kingdom split was also surprisingly young at around 650 Mya, in contrast to the more common age estimates of the kingdom split of at least 1,400 Mya (however the kingdom split was not a focus of the study and only one plant species was included). The authors concluded that the 18S gene underwent an acceleration in evolutionary rate in all of the sampled metazoan lineages in the early Cambrian (550–500 Mya), after which substitution rates declined across all lineages (excepting a later burst in the chordate lineage). No mechanism for this co-ordinated acceleration and deceleration of substitution rates across many independent lineages was offered (Aris-Brosou and Yang, 2002).

It is important to note that, as with other phylogenetic methods, the rate-variable methods rely on a number of assumptions, some of which may be inappropriate for the data considered. For example, the Bayesian method employed by Aris-Brosou and Yang (2002) assumed a constant birth-death process over the phylogeny (lineages created by speciation and removed by extinction at an even rate throughout the history of the taxa under consideration). This presents a dilemma—although divergence date estimates around 550 Mya were interpreted to support the reality of a Cambrian explosion, the results were built on the assumption that no such explosion had occurred, because speciation rates were assumed to be constant throughout time (Aris-Brosou and Yang, 2002). Species sampling strategy must also be considered: rather than being a random sample, the data used by Aris-Brosou and Yang (2002) contained representatives chosen from major metazoan lineages for which reliable fossil dates were available (see Bromham *et al.*, 1998). It is not entirely clear how to sample metazoan sequences randomly in order to satisfy the assumptions of the underlying birth-death process: if each metazoan species should have an equal chance of being included, then the dataset would probably be characterized by a preponderance of beetles. More broadly, it is currently unclear whether stochastic methods, which allow random variation in rate, would be suited to systems where rate varied systematically, for example with body size or speciation rate. Rate-variable methods, such as the Bayesian approach employed by Aris-Brosou and Yang (2002), are a promising new approach to dating the metazoan radiation that may challenge the conclusions of many of the earlier molecular clock studies, but much work needs to be done to test their reliability.

A less sophisticated approach is to reflect the un-

certainty in the molecular clock due to the possibility of temporal rate variation. For example, the observed range in substitution rates over the metazoan phylogeny can be used to put bounds on the possible divergence dates between animal phyla. For 18S rRNA and mitochondrial protein coding genes, allowing all interphylum lineages the maximum observed rate of the metazoan tree could bring the molecular dates for the deep divisions of the metazoan phylogeny closer to the Cambrian, but the date estimates were still long before the appearance of the first undisputed bilaterian fossils (Bromham and Hendy, 2000). However, without a mechanism for universally faster early rates, these confidence limits are only useful to rule out some dates as being beyond reasonable inference given current understanding of rates of molecular evolution.

METAZOAN MOLECULAR DATES: WHERE TO NEXT?

The discrepancy between molecular and palaeontological dates for the origin of animal phyla remains unresolved. The accuracy and precision of molecular dates should be subject to scrutiny, as there are many aspects of molecular studies that could lead to error in molecular date estimates. The wide variation between estimates shows that the precision of molecular dates leaves much to be desired. More specifically, it has been suggested that processes associated with explosive radiations—accelerated morphological change, relaxed adaptive constraint or increased speciation rate—could speed the rate of molecular evolution, making molecular date estimates for the origin of metazoan phyla too old. There is a lack of empirical evidence for this hypothesis, partly because determinants of rates of molecular evolution have not been explored for invertebrates, but also because comparative tests have thus far revealed no association between morphological evolution or adaptive radiation and rates of molecular evolution. Furthermore, rates of molecular evolution at the base of the metazoan radiation would have to have been many times higher than throughout the remainder of the Phanerozoic in order to reconcile molecular data to a Cambrian origin of the major divisions of the metazoan kingdom. It may be that future work reveals an as yet unknown mechanism for such a dramatic deceleration in molecular rates that could account for the apparently deep molecular divergence between phyla. But until such a mechanism is uncovered, it will be difficult to reconcile the molecular data to a Cambrian explosion of all animal phyla.

The bottom line is that the molecular data need to be explained, not dismissed—if the deep molecular divergences were not produced by a long Precambrian history of metazoan lineages, then they must have been produced by some other process. The imperfections of the fossil record—such as temporal gaps in preservation, taxon bias and patchy geographical representation—do not make the fossil record worthless, but they do complicate the interpretation of palaeontological data. Similarly, variation in rates of molec-

ular evolution, imprecision of date estimates, and variation between different genes and taxa do not make the molecular data useless, but they do illustrate the danger of a simplistic interpretation of molecular distances. Molecular data must be telling us something about metazoan evolution, whether or not current molecular date estimates are correct. Much work remains to be done to arrive at an adequate explanation of the deep molecular divergence between animal phyla.

ACKNOWLEDGMENTS

I thank Andrew Rambaut for helpful discussions about rate-variable methods, and Graham Budd and Kevin Peterson for inviting me to the Cambrian Explosion symposium.

REFERENCES

- Abouheif, E., R. Zardoya, and A. Meyer. 1998. Limitations of metazoan 18S rRNA sequence data: Implications for reconstructing a phylogeny of the animal kingdom and inferring the reality of the Cambrian explosion. *J. Mol. Evol.* 47:394–405.
- Archibald, J. D. 1999. Pruning and grafting on the mammalian phylogenetic tree. *Acta Palaeontol. Pol.* 44:220–222.
- Aris-Brosou, S. and Z. Yang. 2002. Effects of models of rate evolution on estimation of divergence dates with special reference to the metazoan 18S ribosomal RNA phylogeny. *Syst. Biol.* 51: 703–714.
- Ayala, F. J., A. Rzhetsky, and F. J. Ayala. 1998. Origin of the metazoan phyla: Molecular clocks confirm palaeontological estimates. *Proc. Natl. Acad. Sci. U.S.A.* 95:606–611.
- Barracough, T. G. and V. Savolainen. 2001. Evolutionary rates and species diversity in flowering plants. *Evolution* 55:677–683.
- Barracough, T. G., A. P. Vogler, and P. H. Harvey. 1998. Revealing the factors that promote speciation. *Phil. Trans. R. Soc. London B* 353:241–249.
- Benton, M. J., M. A. Willis, and R. Hitchin. 2000. Quality of the fossil record through time. *Nature* 403:534–537.
- Boore, J. L. and J. L. Staton. 2002. The mitochondrial genome of the sipunculid *Phascolopsis gouldii* supports its association with Annelida rather than Mollusca. *Evolution* 19:127–137.
- Bromham, L. 2002. Molecular clocks in reptiles: Life history influences rate of molecular evolution. *Mol. Biol. Evol.* 19:302–309.
- Bromham, L. 2003. Molecular clocks and explosive radiations. *J. Mol. Evol.* (In press)
- Bromham, L., A. Rambaut, R. Fortey, A. Cooper, and D. Penny. 1998. Testing the Cambrian explosion hypothesis by using a molecular dating technique. *Proc. Natl. Acad. Sci. U.S.A.* 95: 12386–9.
- Bromham, L., A. Rambaut, and P. H. Harvey. 1996. Determinants of rate variation in mammalian DNA sequence evolution. *J. Mol. Evol.* 43:610–21.
- Bromham, L., M. R. Q. Woolfit, M. S. Y. Lee, and A. Rambaut. 2002. Testing the relationship between morphological and molecular rates of change along phylogenies. *Evolution* 56:1921–1930.
- Bromham, L. D. and B. M. Degnan. 1999. Hemichordates and deuterostome evolution: Robust molecular phylogenetic support for a hemichordate+echinoderm clade. *Evol. Dev.* 1:166–171.
- Bromham, L. D. and M. D. Hendy. 2000. Can fast early rates reconcile molecular dates to the Cambrian explosion? *Proc. R. Soc. London B* 267:1041–1047.
- Bromham, L. D., A. Rambaut, M. D. Hendy, and D. Penny. 2000. The power of relative rates tests depends on the data. *J. Mol. Evol.* 50:296–301.
- Budd, G. E. and S. Jensen. 2000. A critical reappraisal of the fossil record of the bilaterian phyla. *Biol. Rev.* 75:253–295.
- Carroll, S. B. 1995. Homeotic genes and the evolution of arthropods and chordates. *Nature* 376:479–485.
- Castresana, J., G. Feldmaier-Fuchs, S. Yokobori, N. Satoh, and S.

- Paabo. 1998. The mitochondrial genome of the hemichordate *Balanoglossus carnosus* and the evolution of deuterostome mitochondria. *Genetics* 150:1115–1123.
- Conway Morris, S. 1998. Early metazoan evolution: Reconciling paleontology and molecular biology. *Amer. Zool.* 38:867–877.
- Cooper, A. and R. Fortey. 1998. Evolutionary explosions and the phylogenetic fuse. *Trends Ecol. Evol.* 13:151–156.
- Cooper, A. and D. Penny. 1997. Mass survival of birds across the Cretaceous-Tertiary boundary: Molecular evidence. *Science* 275:1109–1113.
- Darwin, C. 1859. *The origin of species by means of natural selection*. John Murray, London, p. 286–289.
- Doolittle, R. F., D. F. Feng, S. Tsang, G. Cho, and E. Little. 1996. Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science* 271:470–477.
- Erwin, D. H. 1993. The origin of metazoan development: A paleobiological perspective. *Biol. J. Linn. Soc.* 50:255–274.
- Feng, D. F., G. Cho, and R. F. Doolittle. 1997. Determining divergence times with a protein clock: Update and reevaluation. *Proc. Natl. Acad. Sci. U.S.A.* 94:13028–33.
- Fleischer, R. C., C. E. McIntosh, and C. L. Tarr. 1998. Evolution on a volcanic conveyor belt: Using phylogeographic reconstructions and K-Ar based ages of the Hawaiian islands to estimate molecular evolutionary rates. *Mol. Ecol.* 7:533–545.
- Gillespie, J. H. 1991. *The causes of molecular evolution*. Oxford University Press, Oxford.
- Gu, X. 1998. Early metazoan divergence was about 830 million years ago. *J. Mol. Evol.* 47:369–371.
- Haeckel, E. 1879. *The evolution of man: A popular exposition of the principal points of human ontogeny and phylogeny*, London.
- Hinman, V. and B. D. Degnan. 2000. Retinoic acid perturbs *Otx* gene expression in the ascidian pharynx. *Dev. Genes Evol.* 210:129–139.
- Holland, P. W. H., A. M. Hacker, and N. A. Williams. 1991. A molecular analysis of the phylogenetic affinities of *Saccoglossus cambrensis* Brambell and Cole (H). *Phil. Trans. R. Soc. London B* 332:185–189.
- Jenner, R. A. 2003. Unleashing the force of cladistics? Metazoan phylogenetics and hypothesis testing. *Integrative and Comparative Biology* 43:000–000.
- Johnson, K. P. and J. Seger. 2001. Elevated rates of nonsynonymous substitution in island birds. *Mol. Biol. Evol.* 18:874–881.
- Kimura, M. 1983. *The neutral theory of molecular evolution*. Cambridge University Press, Cambridge.
- Kishino, H., J. L. Thorne, and W. J. Bruno. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. *Mol. Biol. Evol.* 18:352–361.
- Kumar, S. and S. B. Hedges. 1998. A molecular timescale for vertebrate evolution. *Nature* 392:917–920.
- Lee, M. S. Y. 1998. Molecular clock calibrations and metazoan divergence dates. *J. Mol. Evol.* 49:385–391.
- Lee, M. S. Y. 1999. Shortening the phylogenetic fuse. *Trends Ecol. Evol.* 13:323–323.
- Li, W.-H., D. L. Ellesworth, J. Krushkal, B. H.-J. Chang, and D. Hewett-Emmett. 1996. Rates of nucleotide substitution in primates and rodents and the generation-time effect hypothesis. *Mol. Phylog. Evol.* 5:182–187.
- Lyell, C. 1830. *Principles of geology*. London.
- Lynch, M. 1999. The age and relationships of the major animal phyla. *Evolution* 53:319–325.
- Madsen, O., M. Scally, C. J. Douady, D. J. Kao, R. W. DeBry, R. Adkins, H. M. Amrine, M. J. Stanhope, W. W. de Jong, and M. S. Springer. 2001. Parallel adaptive radiations in two major clades of placental mammals. *Nature* 409:610–614.
- Martin, A. P. and S. R. Palumbi. 1993. Body size, metabolic rate, generation time and the molecular clock. *Proc. Natl. Acad. Sci. U.S.A.* 90:4087–4091.
- Mooers, A. Ø. and P. H. Harvey. 1994. Metabolic rate, generation time and the rate of molecular evolution in birds. *Mol. Phylog. Evol.* 3:344–350.
- Nei, M., P. Xu, and G. Glazko. 2001. Estimation of divergence times from multiprotein sequences for a few mammalian species and several distantly related organisms. *Proc. Natl. Acad. Sci. U.S.A.* 98:2497–2502.
- Ohta, T. 1993. An examination of the generation time effect on molecular evolution. *Proc. Natl. Acad. Sci. U.S.A.* 90:10676–10680.
- Omland, K. E. 1997. Correlated rates of molecular and morphological evolution. *Evolution* 51:1381–1393.
- Papadopoulos, D., D. Schneider, J. Meier-Eiss, W. Arber, R. E. Lenski, and M. Blot. 1999. Genomic evolution during a 10,000-generation experiment with bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 96:3807–3812.
- Penny, D., M. Hasegawa, P. J. Waddell, and M. D. Hendy. 1999. Mammalian evolution: Timing and implications from using the LogDeterminant transform for proteins of differing amino acid composition. *Syst. Biol.* 48:76–93.
- Peterson, K. J. and D. J. Eernisse. 2001. Animal phylogeny and the ancestry of bilaterians: Inferences from morphology and 18S rDNA gene sequences. *Evol. Dev.* 3:170–2–5.
- Rambaut, A. and L. Bromham. 1998. Estimating divergence dates from molecular sequences. *Mol. Biol. Evol.* 15:442–8.
- Robinson, M., M. Gouy, C. Gautier, and D. Mouchirod. 1998. Sensitivity of relative rates tests to taxonomic sampling. *Mol. Biol. Evol.* 15:1091–1098.
- Runnegar, B. 1982. A molecular-clock date for the origin of the animal phyla. *Lethaia* 15:199–205.
- Sanderson, M. J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. *J. Mol. Evol.* 14:1218–1231.
- Scherer, S. 1989. The relative-rate test of the molecular clock hypothesis: A note of caution. *Mol. Biol. Evol.* 6:436–441.
- Simpson, G. G. 1944. *Tempo and mode in evolution*. Columbia University Press, New York.
- Tajima, F. 1993. Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* 135:599–607.
- Thorne, J. L., H. Kishino, and I. S. Painter. 1998. Estimating the rate of the rate of molecular evolution. *Mol. Biol. Evol.* 15:1647–1657.
- Valentine, J. W., D. Jablonski, and D. H. Erwin. 1999. Fossils, molecules and embryos: New perspectives on the Cambrian explosion. *Development*. 126:851–859.
- Wang, D. Y.-C., S. Kumar, and S. B. Hedges. 1999. Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. *Proc. R. Soc. London B*. 266:163–171.
- Williams, G. C. 1992. *Natural selection: Domains, levels and challenges*. Oxford University Press, Oxford.
- Wray, G. A., J. S. Levington, and L. H. Shapiro. 1996. Molecular evidence for deep Precambrian divergences among metazoan phyla. *Science* 274:568–573.
- Wu, C.-I. and W.-H. Li. 1985. Evidence for higher rates of nucleotide substitutions in rodents than in man. *Proc. Natl. Acad. Sci. U.S.A.* 82:1741–1745.
- Yang, Z. H. and R. Nielsen. 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *J. Mol. Evol.* 46:409–418.
- Zuckerkindl, E. and L. Pauling. 1965. Evolutionary divergence and convergence in proteins. In V. Bryson and H. J. Vogel (eds.), *Evolving genes and proteins*. Academic Press, New York.