THE ROYAL SOCIETY

Can fast early rates reconcile molecular dates with the Cambrian explosion?

Lindell D. Bromham[1]* **and Michael D. Hendy**[2]

[1]*Department of Zoology, University of Queensland, Brisbane 4072, Australia*
[2]*Institute of Fundamental Sciences, Massey University, PO Box 11–222, Palmerston North, New Zealand*

Molecular dates consistently place the divergence of major metazoan lineages in the Precambrian, leading to the suggestion that the 'Cambrian explosion' is an artefact of preservation which left earlier forms unrecorded in the fossil record. While criticisms of molecular analyses for failing to deal with variation in the rate of molecular evolution adequately have been countered by analyses which allow both site-to-site and lineage-specific rate variation, no analysis to date has allowed the rates to vary temporally. If the rates of molecular evolution were much higher early in the metazoan radiation, molecular dates could consistently overestimate the divergence times of lineages. Here, we use a new method which uses multiple calibration dates and an empirically determined range of possible substitution rates to place bounds on the basal date of divergence of lineages in order to ask whether faster rates of molecular evolution early in the metazoan radiation could possibly account for the discrepancy between molecular and palaeontological date estimates. We find that allowing basal (interphylum) lineages the fastest observed substitution rate brings the minimum possible divergence date (586 million years ago) to the Vendian period, just before the first multicellular animal fossils, but excludes divergence of the major metazoan lineages in a Cambrian explosion.

Keywords: Cambrian explosion; molecular clock; substitution rate; Metazoa

1. INTRODUCTION

The use of a 'molecular clock' for estimating the dates of lineage divergence has proved one of the most useful aspects of molecular phylogenetic analyses. However, the results of many studies have been controversial because molecular date estimates are, in many cases, dramatically older than the divergence times estimated from palaeontological data alone. The timing of the origin of animal phyla is perhaps the best known example. Molecular dates consistently place the divergence of major metazoan lineages in the Precambrian (table 1), at least 100 million years (Myr) before the first multicellular animal body fossils in the Vendian (600–545 Myr ago) and long before the first undisputed members of modern metazoan phyla appeared in an 'explosion' of fossils in the early Cambrian (544–520 Myr ago) (Bowring *et al*. 1993). These results have been used to suggest that the 'Cambrian explosion', the inferred rapid evolutionary radiation of virtually all animal phyla in as little as 5–10 Myr, is an artefact of preservation which left earlier forms unrecorded in the fossil record.

A similar story can be told for other rapid evolutionary radiations inferred from palaeontological evidence, most notably for the ordinal diversification of birds and mammals. Most modern orders of birds and mammals appear in the fossil record not long after the Cretaceous–Tertiary (K–T) boundary. This has been interpreted as the signal of a massive adaptive radiation following the final extinction of the dinosaurs (e.g. Feduccia 1995; Foote *et al*. 1999). However, molecular dates for the ordinal diversification of birds and mammals are commonly almost twice as old as the earliest fossil evidence, suggesting a hidden Cretaceous radiation of birds and mammals (e.g. Cooper & Penny 1997; Kumar & Hedges 1998).

There are three plausible explanations for the dramatic discrepancy between the molecular and palaeontological date estimates. First, molecular and palaeontological dates may mark different events which are separated in time. The divergence of evolutionary lineages marked by molecular dates is expected to precede the development of the defining morphological features of a taxon marked by fossil evidence. An extensive lag between lineage divergence and the development of the phylum-specific characters could allow both Precambrian molecular dates and Cambrian palaeontological dates (Cooper & Fortey 1998; Archibald 1999; Bromham *et al*. 1999*a*,*b*). Second, palaeontological dates may be too young if systematic biases in the fossil record obscure early metazoan history. The Precambrian fossil record is notably disjunct, with a discontinuity between the Ediacaran assemblages of the Vendian period and the metazoan faunas of the early Cambrian. However, the continuity of taxa, as demonstrated by the presence of metazoans in the Vendian and Ediacarans in the Cambrian (Conway Morris 1993*a*,*b*; Gehling & Rigby 1996; Jensen *et al*. 1998; Li *et al*. 1998), suggests that some macroscopic multicellular metazoans persisted through the latest Proterozoic, despite the lack of appropriate body fossils from this period. This raises the question of how to place the beginning of the animal kingdom in time when we cannot guarantee having the first animals represented as readily identifiable fossils.

Third, molecular dates may be too old if the molecular clock does not hold reliably over time. Molecular studies have been criticized for failing to account for variation in the rate of molecular evolution, both across sites and between lineages, or for relying too heavily on few questionable calibration dates (Conway Morris 1997; Ayala *et al*. 1998; Lee 1999*a*). Variation in the rate of molecular

*Author for correspondence (lbromham@zoology.uq.edu.au).

Received 18 January 2000 *Accepted* 14 February 2000

Table 1. *Recent molecular date estimates of the protostome–deuterostome split*

date estimate	sequence data	calibration date (Myr ago)	study
1200	seven protein coding (mitochondrial and nuclear), 18S rRNA	> 15 dates (mostly vertebrate)	Wray *et al*. (1996)
730, 850	21 protein coding (mitochondrial and nuclear)	six dates (vertebrate: 100–450)	Feng *et al*. (1997)
670, 736	18 protein coding (mitochondrial and nuclear)	more than eight dates	Ayala *et al*. (1998)
> 680	11 protein coding (mitochondrial), 18S rRNA	12 dates (protostome and deuterostome: 240–530)	Bromham *et al*. (1998)
830	22 protein coding (nuclear)	three dates (primate–rodent: 100; mammal–bird: 310; animal–fungus: 1100)	Gu (1998)
630	ten protein coding (mitochondrial)	one date (fish–tetrapod: 430)	Lynch (1999)
993	50 genes	three dates (primate–rodent: 100; mammal–bird: 310; animal–fungus 1100)	Wang *et al*. (1999)

evolution between lineages is a serious problem for the accuracy of molecular dates. The tests used to exclude rate-variable sequences often have unacceptably low power and this could result in consistent overestimation of dates of divergence (Bromham *et al*. 1999*a*, 2000). Because this bias would be common to many studies, congruence of estimates could lend false confidence for the accuracy of molecular date estimates. Establishing rate constancy by linear regression of distance through time is also problematic due to the frequent use of non-independent data points (Ayala *et al*. 1998; Bromham *et al*. 1999*a*; Lynch 1999). However, these factors alone cannot be wholly responsible for Precambrian molecular date estimates, because an analysis which included a large number of taxa and a range of calibration dates (Bromham *et al*. 1998) and allowed for both site-to-site and lineage-specific rate variation (Rambaut & Bromham 1998) also produced Precambrian divergence dates. But, the problem of rate variation has not been vanquished, because no analysis to date has dealt adequately with the problem of temporal variation in the rate of molecular evolution.

A corollary of the observation of lineage-specific rates of molecular evolution is that the rates must evolve along lineages, giving rise to temporal patterns in the rates. It has been suggested that the discrepancy between molecular and palaeontological date estimates is a result of faster rates of molecular evolution early in the metazoan radiation (Vermeij 1996; Conway Morris 1998; Valentine *et al*. 1999). Fast early rates (perhaps due to rapid diversification, bursts of adaptive change or smaller body size accelerating generation turnover) would cause consistent overestimation of the dates of divergence because any calibration rate, whether estimated from a tip lineage or averaged over the phylogeny, would underestimate the true rate at the base of the radiation and, therefore, overestimate the age of lineages. However, this hypothesis is difficult to test because it is not possible to directly estimate rates on the internodes of a phylogeny without knowing the basal date of divergence. This is because a faster rate is most parsimoniously reconstructed as a longer branch with the same substitution rate as other

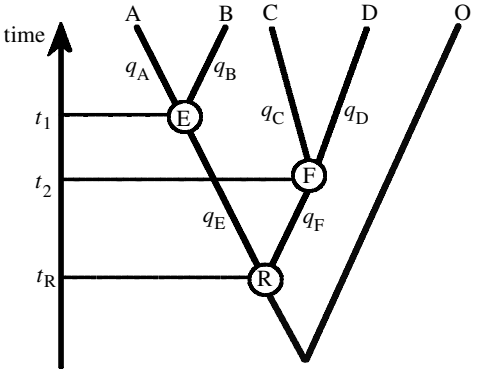

Figure 1. A quintet is formed from two monophyletic pairs (A–B and C–D) each of which has an independently derived date of origin (t_1 and t_2). t_R is the unknown date of divergence which we wish to estimate given the range of possible rates on the internodes q_E and q_F.

branches. Conventional 'clock tests', such as the relative rates test, can detect only departures from 'parallel rate equality' (difference in total branch length) (Gingerich 1986) and cannot detect concerted patterns in the rate which affect all lineages equally. Without a means of directly testing for fast early rates and with no clear evidence of a mechanism which could cause a concerted deceleration independently in all lineages, it is difficult to assess whether the rates of molecular evolution were faster early in the metazoan radiation. But, we can ask whether such a pattern could account for the discrepancy between the molecular and palaeontological dates of the origins of animal phyla.

In order to estimate the rates of evolution for internal branches of a molecular phylogeny without knowing the basal date of divergence, we extend the quartet analysis (Cooper & Penny 1997; Rambaut & Bromham 1998) by constructing a 'quintet' of two dated pairs of lineages (quartet) and an outgroup (figure 1). The quintet method is designed to express the limits of confidence of molecular estimates of divergence dates given a range of possible substitution rates on the interphylum lineages.

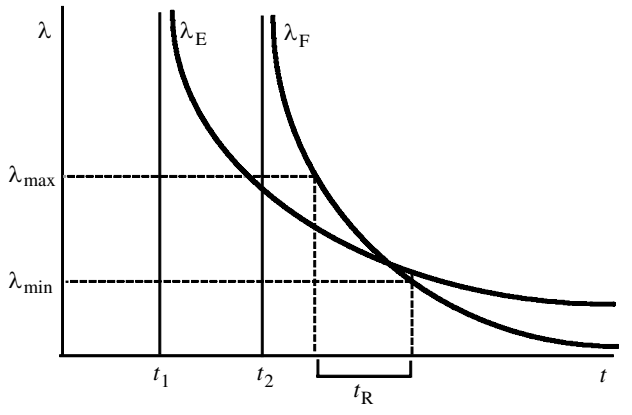

Figure 2. The internode rates (λ_E and λ_F) bounded by the maximum and minimum values for the substitution rate (λ_{max} and λ_{min}) are solved to give a range of possible values of the unknown date t_R (see figure 1).

Because the range of possible dates does not include the early Cambrian, we conclude that even allowing all early lineages the maximum observed substitution rate does not make the molecular data compatible with the origination of these lineages in the early Cambrian.

2. METHODS

If we have a quartet consisting of two monophyletic pairs (A–B and C–D) (figure 1) each with independently derived dates of origin t_1 and t_2, then we can estimate the expected number of substitutions (q_A, q_B, q_C and q_D) and substitution rates (λ_A, λ_B, λ_C and λ_D) directly for each of the four tips from the pairwise distances between them. Let E and F be the nodes representing the points of divergence of the pairs A–B and C–D, respectively and let R be the root of the quartet. Given the dates t_1 of E and t_2 of F, we wish to determine t_R, the date of R. To do this we add an outgroup, O, to form a quintet. This allows us to use the pairwise distances between the five sequences (A, B, C, D and O) in order to estimate the expected number of substitutions q_E and q_F on the branches from R to E and F. The average rates of change λ_E and λ_F on these branches depend on t_R. We can provide a range of possible values of t_R if we can put an upper and lower bound on λ_E and λ_F (figure 2).

Let λ_A and λ_B be the average rates on the branches from E to A and B and let λ_C and λ_D be the average rates on the branches from F to C and D. The average rates on the branches from root R to the bifurcations at E (of A–B) and F (of C–D) are λ_E and λ_F. We now show how the bounds on the rates λ_E and λ_F give the bounds on the time t_R. The quantities $q_X = \lambda_X t_X$, where $t_A = t_B = t_1$, $t_C = t_D = t_2$, $t_E = t_R - t_1$ and $t_F = t_R - t_2$, represent the expected number of substitutions (branch lengths) on the corresponding branches. From the pairwise distances, we can estimate q_A, q_B, q_C and q_D and the sum $q_E + q_F$. The distances from A, B, C and D to an outgroup O (d_{A-O}, d_{B-O}, d_{C-O}, and d_{D-O}) allow us to distinguish the values q_E and q_F. Let q_O represent the number of substitutions between R and O. Then the seven quantities $q_A, \ldots, q_F$ and q_O can be derived from the ten pairwise distances between the five taxa. By summing the q_X-values along the paths we find $\boldsymbol{d} = A\boldsymbol{q}$, where $\boldsymbol{d} = (d_{A-B}, d_{A-C}, d_{A-D}, d_{A-O}, d_{B-C}, d_{B-D}, d_{B-O}, d_{C-D}, d_{C-O}, d_{D-O})^t$ is the vector of the distances, A is the 10×7 matrix

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix},$$

and $\boldsymbol{q} = (q_A, q_B, q_C, q_D, q_E, q_F, q_O)^t$ is the vector of the branch lengths. The least squares best estimate for $\boldsymbol{q}$ is $\boldsymbol{q} = A^+ d$ (Noble 1969), where

$$A^+ = (A^t A)^{-1} A^t$$

$$= 1/24 \begin{bmatrix} 12 & 4 & 4 & 4 & -4 & -4 & -4 & 0 & 0 & 0 \\ 12 & -4 & -4 & -4 & 4 & 4 & 4 & 0 & 0 & 0 \\ 0 & 4 & -4 & 0 & 4 & -4 & 0 & 12 & 4 & -4 \\ 0 & -4 & 4 & 0 & -4 & 4 & 0 & 12 & -4 & 4 \\ -12 & 3 & 3 & 6 & 3 & 3 & 6 & 0 & -6 & -6 \\ 0 & 3 & 3 & -6 & 3 & 3 & -6 & -12 & 6 & 6 \\ 0 & -3 & -3 & 6 & -3 & -3 & 6 & 0 & 6 & 6 \end{bmatrix},$$

is a generalized inverse of A (A^t is the transpose of A).

Hence, in particular,

$$q_A = (3d_{A-B} + d_{A-C} + d_{A-D} + d_{A-O} - d_{B-C} - d_{B-D} - d_{B-O})/6, \quad (1)$$

$$q_E = (-4d_{A-B} + d_{A-C} + 2d_{A-D} + 2d_{A-O} + d_{B-C} + d_{B-D} + 2d_{B-O} - 2d_{C-O} - 2d_{D-O})/8, \quad (2)$$

and

$$q_F = (d_{A-C} + d_{A-D} - 2d_{A-O} + d_{B-C} + d_{B-D} - 2d_{B-O} - 4d_{C-D} + 2d_{C-O} + 2d_{D-O})/8. \quad (3)$$

Now $q_E = \lambda_E(t_R - t_1)$ and $q_F = \lambda_F(t_R - t_2)$, so if we know the rates λ_E and λ_F we have two estimates for t_R:

$$t_R = t_1 + q_E/\lambda_E, \quad (4)$$

and

$$t_R = t_2 + q_F/\lambda_F. \quad (5)$$

However, if we can only assume a range of values which bound these rates, then this gives us a range of times for t_R. Thus, we see

$$\max(t_1 + q_E/\lambda_{max},\ t_2 + q_F/\lambda_{max}) \leqslant t_R \leqslant \min(t_1 + q_E/\lambda_{min},\ t_2 + q_F/\lambda_{min}), \quad (8)$$

provides the lower and upper limits on the time t_R.

3. DATA

We used the sequences, calibration dates and estimated branch lengths from a previous study (Bromham *et al.* 1998): 1710 bp of 18S rRNA and 5676 bp of mitochondrial protein-coding genes (the first and second codon positions only). We set the maximum and minimum substitution rates (λ_{max} and λ_{min}) to encompass the estimates of the substitution rates for a range of taxa (figure 3). For 23 18S sequences with calibration dates, λ_{max} and λ_{min} were set to 4.0×10^{-4} and 0.4×10^{-4} substitutions per site per million years (substitutions site^{-1} Myr^{-1}), respectively. For 16

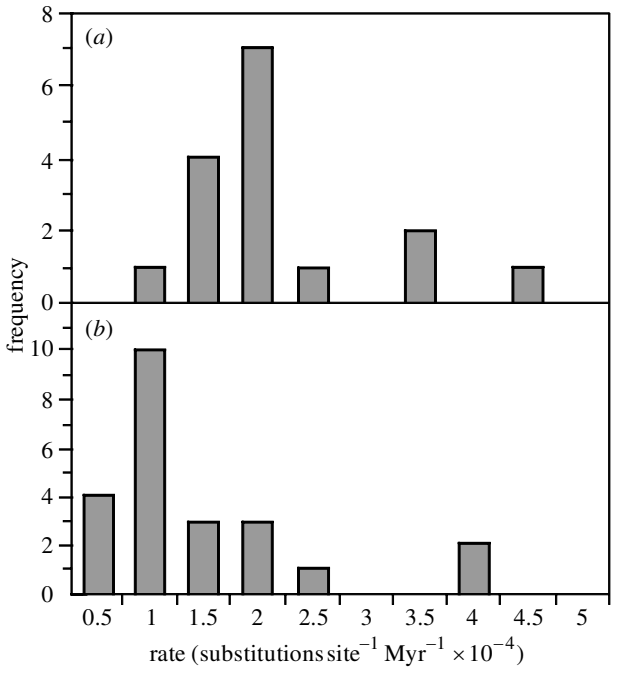

Figure 3. The range of observed substitution rates (substitutions site^{-1} Myr^{-1}) estimated from the maximum likelihood (HKY + Γ) branch lengths and fossil calibration dates (see Bromham *et al.* 1998). (*a*) Mitochondrial proteins, and (*b*) 18S rRNA.

mitochondrial sequences, λ_{max} and λ_{min} were 5.0×10^{-4} and 0.5×10^{-4} substitutions site^{-1} Myr^{-1}, respectively. For 18S rRNA, 21 quintets were formed with deuterostome and arthropod pairs (outgroup Cnidaria) and 20 Echinodermata–Chordata quintets (outgroups from Cnidaria, Chelicerata, Mollusca and Annelida). For the mitochondrial alignment, 164 quintets were constructed from Echinodermata and Chordata pairs (outgroups Arthropoda and Annelida). The pairwise distances between the sequences were calculated using HKY85 distance (Swofford 1999), with a gamma parameter and transition–transversion values from the maximum-likelihood tree of all taxa (18S $t_i/t_v = 1.74$ and $\alpha = 0.38$ and mitochondrial $t_i/t_v = 1.21$ and $\alpha = 0.38$) (see Bromham *et al.* 1998).

4. RESULTS

We report the overall minimum t_R—the youngest of all minimum t_R-values from all quintets tested—in order to test the compatibility of the molecular data with early Cambrian divergences. Because the quintets are not phylogenetically independent, they cannot be combined statistically to give an average minimum t_R. The estimates of the minimum t_R for 164 mitochondrial quintets were all older than 589 Myr ago: 161 were older than the Vendian period (> 600 Myr ago). The overall minimum t_R for 18S quintets was 588 Myr ago for the protostome–deuterostome split and 586 Myr ago for the within-deuterostome split (table 2). The results also give a minimum date of origin for any sister groups to these lineages, such as Porifera, Cnidaria, Ctenophora and Placozoa (Brusca & Brusca 1990). We conclude that, given the range of possible rates observed for the metazoan phylogeny, the molecular data points to a Precambrian (Early Vendian or earlier) origin of at least half a dozen metazoan lineages, which predates their first fossil appearance.

Table 2. *Minimum possible divergence dates given the maximum substitution rates on the internodes*

node	overall minimum t_R (youngest confidence interval of all quintets (Myr ago))	
protostome–deuterostome	18S rRNA	588
echinoderm–chordate	18S rRNA	586
	mitochondrial	589

This analysis is conservative for testing the Cambrian explosion hypothesis (biased towards younger dates), because the overall minimum t_R is not representative of the average outcome of the analysis but is the extreme case (e.g. 98% of mitochondrial quintets gave pre-Vendian estimates of the minimum t_R). Furthermore, the fossil calibration dates must postdate the genetic divergence of the lineages, which will cause overestimation of the observed substitution rate and, thus, upward bias of the maximum substitution rate, thereby producing more recent estimates of the minimum t_R.

5. DISCUSSION

While molecular dating has the potential for revolutionizing many fields of biology, it is important that the accuracy and precision of molecular date estimates are not overstated. The accuracy and precision are limited by, amongst other things, the 'sloppiness' of the clock (uneven tick rate), the ability to estimate genetic distances accurately and the difficulty establishing a calibration rate which can be extrapolated from one branch of a phylogeny to another. Because of these combined sources of error, molecular data are less suited to the production of point estimates of lineage divergence dates and are most useful when presented with realistic confidence intervals which allow the testing of biological hypotheses by asking whether the molecular data is compatible or not with a hypothesized date of divergence.

If the rates vary between the tips of a phylogeny then we should expect them to vary over any branches including the internodes. This presents a problem: we cannot estimate the divergence dates without knowing the internode rates, but we cannot estimate the internode rates without knowing the basal date of divergence. The quintet method is designed for expressing the limits of confidence of molecular date estimates given a range of possible substitution rates on the internodes. An alternative use of this method is to estimate the maximum and minimum rates implied by a given value of t_R in order to test the implications for the molecular evolution of a given evolutionary hypothesis.

We have demonstrated that molecular data do not allow early Cambrian divergences and can only be made compatible with a Vendian origin of major metazoan lineages by assuming that the rates of substitution were universally higher in the Proterozoic than they have been throughout the Phanerozoic. Invoking fast early rates

requires a genetic mechanism which affects all of the 50 or more genes which have been used to date the origin of metazoan lineages (table 1) and must affect both mitochondrial and nuclear genomes. Examination of possible genetic mechanisms for genome-wide fast early substitution rates is essential to assess whether fast early rates provide a plausible explanation for the discrepancy between molecular and palaeontological dates. Here we consider bias in the observed rates, morphological diversification, selective constraints, DNA mutation rates and speciation rate. Since none of these hypothesized mechanisms has empirical support, much more research into the patterns of rate variation is needed before convincing arguments for fast early rates can be raised.

The body size trend in molecular evolution rate in tetrapod vertebrates (Martin & Palumbi 1993; Mooers & Harvey 1994; Bromham *et al*. 1996) has prompted the suggestion that the presumed small size of the earliest metazoans generated faster rates early in the metazoan radiation (Conway Morris 1998). If this was true and since we do not estimate substitution rates for soft-bodied lineages or microscopic taxa, we might have underestimated the maximum substitution rates. But, there is currently no evidence to suggest a body size relationship in the rate of molecular evolution for invertebrates, and branch lengths of small soft-bodied taxa are not consistently longer than their nearest larger biomineralized relatives.

A proposed link between rapid morphological change and the rate of molecular evolution has been invoked to suggest higher substitution rates during the metazoan radiation, particularly given Omland's (1997) observation of an association between the rates of molecular change and rates of morphological change for a number of phylogenies (Vermeij 1996; Conway Morris 1998; Knoll & Carroll 1999; Lee 1999*b*; Valentine *et al*. 1999). Omland's (1997) relationship is surprising because sequence evolution is generally held to be largely unlinked to phenotypic change. It has long been recognized that the relationship between molecular change and morphological change is, at very least, not a strict one. For example, bradytelic 'living fossil' taxa such as the coelacanth and the horseshoe crab do not have appreciably shorter molecular branch lengths than horotelic taxa. Conversely, phylum-level lineages which show a high level of morphological disparity between their constituent taxa, such as molluscs and echinoderms, do not appear to accumulate more molecular changes in genes used for phylogenetic analysis than less morphologically diverse clades. Experimental evidence of the disassociation of molecular and phenotypic evolution has been recently described for long-running experiments with bacteria, where initial rapid rates of adaptive change had no impact on the rates of accumulation of genetic changes (Papadopoulos *et al*. 1999). Because a close relationship between morphological and molecular rates of change is not expected, possible mechanisms which could generate such an association need to be examined before such claims can be supported.

Strong selection for novel traits during a Cambrian explosion could directly accelerate the rates of substitution in a relatively small number of target genes, such as genes involved in body-plan diversification. While non-target mutations in neighbouring genes might be indirectly promoted, this 'hitchhiking' will be most effective for non-recombining genomes (Papadopoulos *et al*. 1999) and could be limited in effect for the multiple independently segregating and recombining metazoan chromosomes. It is difficult to see that hitchhiking could produce a sufficiently great increase in the substitution rates across the entire nuclear and mitochondrial genomes for all early metazoan lineages, including not only those lineages which produced high diversity in the Cambrian (e.g. arthropods) but also those whose major radiation occurred subsequent to the Cambrian explosion (e.g. echinoderms).

Alternatively, the substitution rate could have been higher early in the metazoan radiation if selection was somehow less stringent. If the genomes of all early metazoan taxa were subject to universally lower selective constraints before and during the Cambrian explosion, this might make more mutations effectively neutral (Kimura 1983), speeding up the substitution rate relative to the more complex and selectively constrained post-explosion genetic architecture. However, there is no obvious mechanism for such universally lowered constraints early in the metazoan radiation, nor is there any evidence that pre-explosion metazoan lineages (such as the diploblastic taxa Cnidaria and Porifera) had higher substitution rates. Similarly, if DNA repair efficiency increased dramatically in association with the rise in metazoan complexity, then perhaps the post-Cambrian explosion mutation rates could have been lower than the Proterozoic mutation rates. However, consideration of metazoan sister groups and the earliest metazoans suggests that sophisticated DNA replication and repair systems must have been present in the earliest stem lineages of Metazoa.

Population size can influence genome-wide substitution rates, as the rate of fixation of nearly neutral mutations should be higher in small populations (Kimura 1983; Ohta 1993). Rapid speciation could result in repeated population subdivision (as incipient species become reproductively isolated from each other), potentially creating a lower average population size in rapidly radiating lineages. This might explain the observation of faster rates of molecular evolution in more speciose plant clades (Barraclough *et al*. 1996; Savolainen & Goudet 1998). However, the opposite prediction is made for the rate of fixation of adaptive alleles, which is expected to be higher in larger populations or populations undergoing exponential growth (Otto & Whitlock 1997). Therefore, it is necessary to examine whether the proposed fast early rates would be due to an increase in the neutral (or nearly-neutral) substitution rate or increased fixation of adaptive alleles. This hypothesis would be strengthened if empirical evidence of an association between speciation rate and rate of molecular evolution was found for metazoans, and requires modelling of the change in population size between Proterozoic and Phanerozoic metazoans needed to produce a sufficient decrease in the substitution rate.

While further investigation into the correlates of variation in the rate of molecular evolution may reveal plausible scenarios for faster rates early in the metazoan radiation, there is currently no compelling reason to

suppose that the rates of molecular evolution were universally higher before and during the Cambrian explosion and that the rates have dramatically decelerated independently in all major metazoan lineages since then. Even allowing the highest observable rate on the basal internodes still implicates a substantial Precambrian history which is currently hidden from the fossil record and so cannot wholly account for the discrepancy between the molecular and palaeontological dates for the metazoan radiation.

Nor does the 'compromise' view (Conway Morris 1997)—Precambrian genesis plus Cambrian explosion—which suggests that, although metazoan lineages may have originated in the Late Proterozoic, they underwent a massive adaptive radiation in the early Cambrian (Conway Morris 1998; Knoll & Carroll 1999; Valentine *et al*. 1999). This hypothesis suggests that complex macroscopic metazoans persisted throughout the latest Precambrian, despite the paucity of metazoan body fossils immediately preceding the Cambrian explosion. This absence of fossils cannot be entirely explained by presumed small size and possible interstitial habitat of the earliest metazoans (Vermeij 1996; Fortey *et al*. 1997), as Precambrian stem metazoans must have had reasonably sophisticated developmental pathways and a degree of morphological complexity approaching that of the Cambrian metazoans (e.g. Knoll & Carroll 1999), consistent with the interpretation of some Ediacaran taxa as mobile triploblastic metazoans of a grade of organization equivalent to some Cambrian taxa (Conway Morris 1993*b*; Fedonkin & Waggoner 1997). The Precambrian genesis plus Cambrian explosion scenario also implies a perplexing view of a major evolutionary radiation: a long period of cryptic speciation with increasing developmental complexity and anatomical sophistication, yet without substantial behavioural or ecological modification or diversification, followed by a sudden increase in morphological and ecological innovation. This suggests that the degree to which lineage divergence and morphological diversification is disassociated is a critical feature of evolutionary radiations requiring attention (Cooper & Fortey 1998; Bromham *et al*. 1999*b*). Clearly no type of evidence is more important than the fossil record, yet resolution of the tempo and mode of the metazoan radiation profits from a multidisciplinary approach (Jablonski 1999). Molecular analyses, though imperfect (as alas is any means of investigating evolutionary events in deep time), can potentially make a valuable contribution.

6. CONCLUSIONS

The potential contribution of molecular clocks to understand the patterns of evolution in deep time cannot be overstated, but it is critical that we develop this technique with an honest appreciation of the margins of error involved in molecular date estimates. In particular, we must recognize that, if the rates of molecular evolution can vary between taxa, then the rates must also vary over time. Concerted temporal patterns in the rate of molecular evolution are difficult to detect, but could result in biased molecular date estimates. To test the Cambrian explosion hypothesis, we have demonstrated that giving interphylum lineages the fastest observed substitution rate can bring molecular date estimates forward by at least 100 Myr, allowing molecular data to nearly coincide with the first multicellular animal fossils in the Vendian period. However, even with universally fast rates on the internodes of the metazoan tree, we cannot reconcile the molecular data to an origin of all animal phyla in the early Cambrian. Although we demonstrate that molecular dates could be overestimated if the rates were universally higher on the interphylum lineages, this is not equivalent to demonstrating that previous estimates of metazoan divergence times are incorrect. We currently have no compelling reason to suppose that the rates of molecular evolution were universally faster in the Proterozoic. This requires further investigation of the causes and correlates of variation in the rate of molecular evolution across metazoan lineages.

Thanks are due to Richard Fortey, Kathi Huber, Michael Langton, David Penny, Bernie Degnan and Mike Lee for their help and to Mike Steel for organizing the Kaikoura '99 meeting.

REFERENCES

Archibald, J. D. 1999 Molecular dates and the mammalian radiation. *Trends Ecol. Evol.* **14**, 278.

Ayala, F. J., Rzhetsky, A. & Ayala, F. J. 1998 Origin of the metazoan phyla: molecular clocks confirm palaeontological estimates. *Proc. Natl Acad. Sci. USA* **95**, 606–611.

Barraclough, T. G., Harvey, P. H. & Nee, S. 1996 Rate of *rbcL* gene sequence evolution and species diversification in flowering plants (Angiosperms). *Proc. R. Soc. Lond.* B **263**, 589–591.

Bowring, S. A., Grotzinger, J. P., Isachsen, C. E., Knoll, A. H., Pelechaty, S. M. & Kolosov, P. 1993 Calibrating rates of Early Cambrian evolution. *Science* **261**, 1293–1298.

Bromham, L., Rambaut, A. & Harvey, P. H. 1996 Determinants of rate variation in mammalian DNA sequence evolution. *J. Mol. Evol.* **43**, 610–621.

Bromham, L., Rambaut, A., Fortey, R., Cooper, A. & Penny, D. 1998 Testing the Cambrian explosion hypothesis by using a molecular dating technique. *Proc. Natl Acad. Sci. USA* **95**, 12 386–12 389.

Bromham, L. D., Phillips, M. J. & Penny, D. 1999*a* Growing up with dinosaurs: molecular dates and the mammalian radiation. *Trends Ecol. Evol.* **14**, 113–118.

Bromham, L. D., Penny, D. & Phillips, M. J. 1999*b* Molecular dates and the mammalian radiation. *Trends Ecol. Evol.* **14**, 278.

Bromham, L. D., Hendy, M. D., Penny, D. & Rambaut, A. E. 2000 The power of relative rates tests depends on the data. *J. Mol. Evol.* **50**, 296–301.

Brusca, R. C. & Brusca, G. L. 1990 *Invertebrates*. Sunderland, MA: Sinauer Associates.

Conway Morris, S. 1993*a* Ediacaran-like fossils in Cambrian Burgess Shale-type faunas of North America. *Palaeontology* **36**, 593–635.

Conway Morris, S. 1993*b* The fossil record and the early evolution of the Metazoa. *Nature* **361**, 219–225.

Conway Morris, S. 1997 Defusing the Cambrian explosion? *Curr. Biol.* **7**, R71–R74.

Conway Morris, S. 1998 Early metazoan evolution: reconciling paleontology and molecular biology. *Am. Zool.* **38**, 867–877.

Cooper, A. & Fortey, R. 1998 Evolutionary explosions and the phylogenetic fuse. *Trends Ecol. Evol.* **13**, 151–156.

Cooper, A. & Penny, D. 1997 Mass survival of birds across the Cretaceous–Tertiary boundary: molecular evidence. *Science* **275**, 1109–1113.

Fedonkin, M. A. & Waggoner, B. M. 1997 The Late Precambrian fossil *Kimberella* is a mollusc-like bilaterian. *Nature* **388**, 868–871.

Feduccia, A. 1995 Explosive evolution in tertiary birds and mammals. *Science* **267**, 637–638.

Foote, M., Hunter, J. P., Janis, C. M. & Sepkoski Jr, J. J. 1999 Evolutionary and preservational constraints on origins of biologic groups:divergence times of eutherian mammals. *Science* **283**, 1310–1314.

Fortey, R. A., Briggs, D. E. G. & Wills, M. A. 1997 The Cambrian evolutionary 'explosion' recalibrated. *BioEssays* **19**, 429–434.

Gehling, J. G. & Rigby, J. K. 1996 Long expected sponges from the Neoproterozoic fauna of South Australia. *J. Palaeontol.* **70**, 185–195.

Gingerich, P. D. 1986 Temporal scaling of molecular evolution in primates and other mammals. *Mol. Biol. Evol.* **3**, 205–221.

Jablonski, D. 1999 The future of the fossil record. *Science* **284**, 2114–2116.

Jensen, S., Gehling, J. G. & Droser, M. L. 1998 Ediacara-type fossils in Cambrian sediments. *Nature* **393**, 567–569.

Kimura, M. 1983 *The neutral theory of molecular evolution*. Cambridge University Press.

Knoll, A. H. & Carroll, S. B. 1999 Early animal evolution: emerging views from comparative biology. *Science* **284**, 2129–2137.

Kumar, S. & Hedges, S. B. 1998 A molecular timescale for vertebrate evolution. *Nature* **392**, 917–920.

Lee, M. S. Y. 1999*a* Molecular clock calibrations and metazoan divergence dates. *J. Mol. Evol.* **49**, 385–391.

Lee, M. S. Y. 1999*b* Shortening the phylogenetic fuse. *Trends Ecol. Evol.* **13**, 323–323.

Li, C.-W., Chen, J.-Y. & Hua, T.-E. 1998 Precambrian sponges with cellular structures. *Science* **279**, 879–882.

Lynch, M. 1999 The age and relationships of the major animal phyla. *Evolution* **53**, 319–325.

Martin, A. P. & Palumbi, S. R. 1993 Body size, metabolic rate, generation time and the molecular clock. *Proc. Natl Acad. Sci. USA* **90**, 4087–4091.

Mooers, A. O. & Harvey, P. H. 1994 Metabolic rate, generation time and the rate of molecular evolution in birds. *Mol. Phylogenet. Evol.* **3**, 344–350.

Noble, B. 1969 *Applied linear algebra*. Englewood Cliffs, NJ: Prentice Hall.

Ohta, T. 1993 An examination of the generation time effect on molecular evolution. *Proc. Natl Acad. Sci. USA* **90**, 10 676–10 680.

Omland, K. E. 1997 Correlated rates of molecular and morphological evolution. *Evolution* **51**, 1381–1393.

Otto, S. P. & Whitlock, M. C. 1997 The probability of fixation in populations with changing size. *Genetics* **146**, 723–733.

Papadopoulos, D., Schneider, D., Meier-Eiss, J., Arber, W., Lenski, R. E. & Blot, M. 1999 Genomic evolution during a 10,000-generation experiment with bacteria. *Proc. Natl Acad. Sci. USA* **96**, 3807–3812.

Rambaut, A. & Bromham, L. 1998 Estimating divergence dates from molecular sequences. *Mol. Biol. Evol.* **15**, 442–448.

Savolainen, V. & Goudet, J. 1998 Rate of gene sequence evolution and species diversification in flowering plants: a re-evaluation. *Proc. R. Soc. Lond.* B **265**, 603–607.

Swofford, D. L. 1999 *PAUP*. Phylogenetic analysis using parsimony (*and other methods)*, v. 4. Sunderland, MA: Sinauer Associates.

Valentine, J., Jablonski, D. & Erwin, D. 1999 Fossils, molecules and embryos: new perspectives on the Cambrian explosion. *Development* **126**, 851–859.

Vermeij, G. 1996 Animal origins. *Science* **274**, 525–526.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.