Molecular Clocks Do Not Support the Cambrian Explosion

Jaime E. Blair and S. Blair Hedges

NASA Astrobiology Institute and Department of Biology, The Pennsylvania State University, University Park

The fossil record has long supported the view that most animal phyla originated during a brief period approximately 520 MYA known as the Cambrian explosion. However, molecular data analyses over the past 3 decades have found deeper divergences among animals (~800 to 1,200 MYA), with and without the assumption of a global molecular clock. Recently, two studies have instead reported time estimates apparently consistent with the fossil record. Here, we demonstrate that methodological problems in these studies cast doubt on the accuracy and interpretations of the results obtained. In the study by Peterson et al., young time estimates were obtained because fossil calibrations were used as maximum limits rather than as minimum limits, and not because invertebrate calibrations were used. In the study by Aris-Brosou and Yang, young time estimates were obtained because of problems with rate models and other methods specific to the study, and not because Bayesian methods were used. This also led to many anomalous findings in their study, including a primate-rodent divergence at 320 MYA. With these results aside, molecular clocks continue to support a long period of animal evolution before the Cambrian explosion of fossils.

Introduction

Numerous studies using molecular clocks have timed divergences among animal phyla since the early 1970s (e.g., Brown et al. 1972; Wray, Levinton, and Shapiro 1996; Hedges et al. 2004). Until recently, such analyses have indicated deep origins for animal phyla (~800 to 1,200 MYA), much earlier than predicted by the Cambrian explosion, a period when many animal phyla first appear in the fossil record (~520 MYA). Now, two studies have proposed molecular time estimates that are apparently consistent with the Cambrian explosion. We show here that young times in the first study (Peterson et al. 2004) were obtained primarily because fossil-based calibration points were improperly assigned to be maximum bounds, without justification. In the other study (Aris-Brosou and Yang 2003), major inconsistencies between their results and well-established aspects of the eukaryote fossil record show that their model used for describing the evolution of substitution rates among lineages is flawed.

All molecular clock studies rely on calibrations to establish evolutionary rates throughout the phylogenetic tree. Fossils are the most common calibrations, and they are widely accepted as being underestimates of the true divergence time between two lineages. However, if fossils are gross underestimates, they can substantially miscalculate divergence times and lead to misinterpretations of evolutionary history. Therefore, it is desirable to have robust calibrations for estimating divergence times (Hedges et al. 1996; Hedges and Kumar 2004).

Some methods of time estimation permit fossil calibrations to be used as either minimum or maximum constraints, or both (Kishino, Thorne, and Bruno 2001; Sanderson 2003). The use of a fossil as a minimum constraint is relatively uncontroversial, especially if the identity and date of the fossil are accurate; the divergence must have occurred before that time. However, the use of a fossil as a maximum constraint is fundamentally different. The maximum limit on a calibration is almost never known with certainty. Rare transitions in the fossil record, such as the series of fossils corresponding to the colonization of land by tetrapods (Benton 2000), provide some justification for assigning a maximum limit (Hedges et al. 1996; Kumar and Hedges 1998; Hedges and Kumar 2004). Nonetheless, even in those cases, the limits and probability distribution of the fossil time range can be debated (Hedges and Kumar 2004; Ruta and Coates 2004). An additional problem with the use of maximum limits on calibrations, relevant here, is that the primary conclusions of a study may hinge upon the use of such a limit. Thus, the justification for the maximum limit becomes the key factor requiring scrutiny.

Results and Discussion

A recent study reported molecular divergence times for animal phyla that were apparently consistent with the Cambrian explosion (Peterson et al. 2004). Sequences from seven proteins were used with a Langley-Fitch global clock method (Sanderson 2003). Maximum time limits were set to equal minimum time limits for fossil calibrations (i.e., calibrations were "fixed"), and only invertebrates were used to calibrate. The resulting estimates for the arthropod-deuterostome divergence were 573 to 656 MYA, with the older time obtained when a gamma distribution was used to model rate heterogeneity among sites. The difference between these time estimates compared with much older times for the arthropod-deuterostome divergence found in most other studies was ascribed to the use of invertebrate calibrations instead of vertebrate calibrations used in other studies.

However, two other recent molecular clock analyses addressing the same arthropod-deuterostome divergence, using sequences from 5 to 43 proteins, also used invertebrate calibrations (61 to 151 proteins were used with vertebrate calibrations) and did not find a bias with the use of vertebrate calibrations (Hedges et al. 2004; Pisani et al. 2004). Therefore, we conducted a reanalysis of the data set of Peterson et al. (2004) to determine the explanation for their relatively young time estimates. We treated fossil times as both fixed calibrations (maximum = minimum) and as only minimum constraints, and we used the same method of analysis as Peterson et al. (2004). We also reanalyzed the data using maximum-likelihood estimates

Key words: animals, Metazoa, time estimation, evolution, fossil record.

E-mail: jeb322@psu.edu.

Mol. Biol. Evol. 22(3):387–390. 2004
doi:10.1093/molbev/msi039
Advance Access publication November 10, 2004

Table 1
Molecular Divergence Times between Arthropods and Deuterostomes

	Fixed Calibrations		Fossil Minimums	
Distance Branch Lengths	$\alpha = \infty$	$\alpha = 0.28$	$\alpha = \infty$	$\alpha = 0.28$
Langley-Fitch	578 (559, 599)[a]	665 (643, 695)[a]	777 (613, 1048)	851 (722, 927)
Penalized likelihood	582 (564, 608)	676 (651, 705)	1072 (969, 1113)	1014 (968, 1049)
	Fixed Calibrations		Fossil Minimums	
ML Branch Lengths	$\alpha = \infty$	$\alpha = 0.4$	$\alpha = \infty$	$\alpha = 0.4$
Langley-Fitch	671 (643, 699)	666 (638, 694)	941 (895, 987)	933 (881, 985)
Penalized likelihood	748 (708, 788)	706 (668, 744)	882 (834, 930)	904 (844, 964)

NOTE.—Divergence times are recalculated from Peterson et al. (2004). Numbers are means followed by the 95% confidence interval. Distance-based branch lengths use the curvature of the likelihood surface to estimate the 95% confidence interval (Sanderson 2003). Maximum- likelihood (ML) branch lengths use 100 bootstrap replicates to estimate error.

[a] Similar to times presented in Peterson et al. (2004).

of the branch lengths (Yang 1997) and bootstrapping to estimate error. We obtained similar (young) time estimates for the arthropod-deuterostome divergence using distance-based branch lengths, both with and without a gamma correction, and a global clock (table 1).

In contrast, when we treated fossil times as minimum constraints, the same data set recovered older dates, estimating the arthropod-deuterostome divergence as 777 to 851 MYA. The use of a local clock method (penalized likelihood) (Sanderson 2003) also showed discordance between those times estimated with fixed calibrations and those where the fossils were used as minimum constraints. The use of maximum likelihood to estimate branch lengths resulted in older divergence times under a global clock and produced a similar pattern as distance-based branch lengths when fossils were treated either as fixed, resulting in younger time estimates, or as minimums, resulting in older time estimates. In addition, we estimated the divergence times for the calibration nodes when fossils were treated as minimum constraints (table 2); the resulting molecular times were, on average, 68.6% older than the fossil times.

Therefore, our reanalysis of the data of Peterson et al. (2004) shows that their main conclusion—young time estimates for divergences among animal phyla—was the result of incorrectly assuming that fossils provide maximum limits on calibration points. As we have shown, treating fossil times as minimum estimates for both global and local clock analyses recovers time estimates that are similar to previous molecular clock studies, indicating that arthropods and deuterostomes diverged 300 to 400 million years before the Cambrian explosion.

The viewpoint that the Cambrian explosion represented a time when many groups of animals evolved and diverged from one another is the same as the assumption that the maximum limits of those Cambrian phyla correspond to their minimum limits in the fossil record. Therefore, if molecular clocks are to properly test this model, one should not include maximum limits on fossil calibrations unless the resulting time estimates are treated implicitly as minimums. This was not the case with the Peterson et al. (2004) study, where they concluded that their results "support the view that the Cambrian explosion reflects, in part, the diversification of bilaterian phyla."

A second recent molecular clock study claiming support for the Cambrian explosion used a Bayesian approach to accommodate rate variation across lineages, with nucleotide sequence data from 22 genes (Aris-Brosou and

Table 2
Molecular Divergence Times for Calibration Nodes

		Distance Branch Lengths[a]		ML Branch Lengths	
Calibration Node[b]	Fossil Minimum[c]	Langley-Fitch	Penalized Likelihood	Langley-Fitch	Penalized Likelihood
1	50	108	108	123 (97, 149)	100 (76, 124)
2	190	241	296 (261, 333)	319 (277, 361)	271 (229, 313)
3	260	282	354 (307, 394)	386 (340, 432)	333 (285, 381)
4	475	511	695	635 (575, 695)	582 (510, 654)
5	485	572 (552, 592)	773 (726, 820)	730 (672, 788)	681 (611, 751)
6	20	79	122 (92, 161)	70 (48, 92)	67 (45, 89)
7	325	325	481 (425, 545)	345 (315, 375)	332 (306, 358)
8	485	485	691 (627, 745)	586 (540, 632)	554 (502, 606)
9	120	125	174	178 (144, 212)	169 (135, 203)
10	235	283	419 (372, 471)	369 (327, 411)	347 (303, 391)
11	325	369	543 (494, 593)	509 (461, 557)	481 (427, 535)

NOTE.—Divergence times are recalculated from Peterson et al. (2004). Confidence intervals are as described in table 1. Branch lengths were estimated with a gamma correction (results without a gamma correction were similar, data not shown).

[a] Nodes where there was no convergence on a solution for the 95% confidence interval are left blank.

[b] Node numbers from Peterson et al. (2004).

[c] Fossil minimums from Peterson et al. (2004).

Table 3
Molecular Divergence Times

Divergence	Number of Genes	Time Estimates (MYA)		Fossil Minimum
		Clock	OUP	
Protostome-Deuterostome	21	1117 (789, 1428)	581 (557, 610)	543
Echinoderm-Chordate	20	747 (641, 848)	536 (524, 544)	530
Mammal-Bird	17	368 (343, 389)	398 (386, 423)	310
Cnidaria-Bilateria	16	1331 (964, 1591)	611 (569, 643)	600
Mollusca-Arthropoda	16	1037 (822, 1323)	556 (493, 576)	543
Annelida-Arthropoda	15	854 (665, 1289)	528 (465, 542)	543
Osteichthys-Dipnoi/Tetrapod	13	389 (378, 395)	454 (441, 462)	425
Drosophila-Anopheles/Aedes	13	184 (150, 204)	141 (77, 241)	235
Asteroidea-Echinoidea	12	447 (353, 549)	455 (383, 486)	485
Arachnida-Merostomata	12	447 (422, 500)	344 (200, 406)	480
Agnatha-Gnathostomata	12	555 (504, 600)	500 (492, 509)	495
Coelacanth-Dipnoi/Tetrapod	12	378 (343, 382)	429 (419, 438)	410
Perissodactyl-Cetartiodactyl	12	148 (124, 195)	213 (169, 242)	53
Monotreme-Placental mammals	11	256 (242, 285)	336 (320, 353)	115
Primate-Rodent	8	243 (168, 359)	320 (176, 355)	60
Amphibia-Amniote	5	405 (356, 637)	396 (356, 462)	370
Primate-Artiodactyl	5	178 (154, 328)	286 (227, 305)	60
Animals-Plants	18	n/a	671 (591, 722)	1200

NOTE.—Divergence times are from Aris-Brosou and Yang (2003). Numbers are the median values, followed by the first and third quartiles of the median. Fossil minimums from various sources (Benton 1993, 2000; Butterfield 2000; Shu et al. 2001; Gaunt and Miles 2002; Chen et al. 2004; Donoghue, Smith, and Sansom 2004; Pisani et al. 2004). n/a = divergence time not calculated. OUP = Ornstein-Uhlenbeck process.

Yang 2003). By allowing rates to vary through time and modeling rate variation with a complex probability distribution (Ornstein-Uhlenbeck process), the authors obtained time estimates for the arthropod-deuterostome and echinoderm-chordate divergences that were apparently consistent with the fossil record (Cambrian explosion). The authors used the posterior Bayes factor to show support for the use of a complex probability distribution over other, simpler distributions (Aris-Brosou and Yang 2003).

We have reanalyzed the data set of Aris-Brosou and Yang (2003), but we first point out inconsistencies reported in their paper (including Supplementary Material online) that were not discussed. These inconsistencies (table 3) violate widely held views of animal evolution and indicate fundamental problems with their methods. For example, some estimated times of divergence were younger than known fossils, such as the divergences between Arachnida and Merostomata and between Annelida and Arthropoda, or the fossil minimum was within the extreme upper boundary of the interquartile range, such as Brachycera (*Drosophila*) versus Nematocera (*Anopheles/Aedes*) and Asteroidea versus Echinoidea (Benton 1993, 2000; Shu et al. 2001; Chen et al. 2004; Donoghue, Smith, and Sansom 2004; Pisani et al. 2004). There are also cases within the vertebrates where the molecular times significantly overestimate the divergences, such as with monotremes versus placental mammals (336 MYA), perissodactyls versus cetartiodactyls (213 MYA), primates versus rodents (320 MYA), and primates versus artiodactyls (286 MYA). Also, their time estimate for the split of mammals and birds (386 to 423 MYA) is earlier than the fossil evidence for the colonization of land by tetrapods (Benton 2000). Molecular time estimates should always be older than corresponding fossils, but most paleontologists would find those molecular times to be unrealistic because of the presence of transitional stages in the vertebrate fossil record. Such anomalies were also present in a previous study by these authors (Aris-Brosou and Yang 2002), where, for example, their estimate (~15 MYA) for the divergence of great apes and Old World monkeys was considerably younger than the fossil evidence for that divergence (Benton 2000). Other molecular clock studies of mammals have not found such gross inconsistencies with the fossil record (Kumar and Hedges 1998; Springer et al. 2003).

We also reanalyzed the data set of Aris-Brosou and Yang (2003) with an additional outgroup to estimate the animal-plant divergence time, using the same methods as outlined in the original study. As shown in table 3, the estimated animal-plant divergence time (671 MYA) violates the fossil minimum (1,200 MYA) for this divergence (Butterfield 2000). We also analyzed the data using a different Bayesian method, Divtime5b (Kishino, Thorne, and Bruno 2001), and recovered deep Precambrian times consistent with previous clock studies (arthropod-deuterostome divergence, approximately 1,125 MYA; echinoderm-chordate divergence, approximately 985 MYA; and animal-plant divergence, approximately 1,530 MYA) and not with the Bayesian study of Aris-Brosou and Yang (2003). This disagrees with the conclusion of Aris-Brosou and Yang (2003) that the Bayesian method itself is responsible for obtaining time estimates that support the Cambrian explosion model of animal evolution.

The complex methods used in this study (Aris-Brosou and Yang 2003) to accommodate rate variation do not appear to be producing reliable divergence times across the phylogenetic tree. Therefore, the authors cannot justify selecting some divergence times to support the Cambrian explosion model while ignoring many other time estimates that are grossly inconsistent with general knowledge of animal evolution.

Conclusion

Here we have shown that two recent studies supporting the paleontological interpretation of the Cambrian explosion suffer from methodological problems that place into question their results. In both cases, the reasons given by those authors for why their time estimates were younger than those of previous studies were incorrect. In the first case (Peterson et al. 2004), it was not a rate bias in the vertebrate fossil record, but rather the imposition (unjustified) of maximum time limits on fossil calibrations. In the second case (Aris-Brosou and Yang 2003), it was not the Bayesian method of analysis, but rather the complex rate models and methods specific to their analysis. In addition, a third recent molecular clock study (Douzery et al. 2004) reported relatively young divergence times among animals (642 to 761 MYA) using a Bayesian method of analysis. However, this study also violates the eukaryote fossil record by underestimating divergence times for red algae (~925 MYA versus 1,200 MYA fossil), stramenopiles (~800 MYA versus 1,000 MYA fossil), and cholorophytan green algae (~730 MYA versus 1,000 MYA fossil) (Woods, Knoll, and German 1998; Butterfield 2000; Kumar 2001).

Elsewhere (Hedges et al. 2004), we have analyzed larger sequence data sets with all available time estimation methods, including Bayesian, and found deep Precambrian divergences among animal phyla, consistent with clock studies over the past 3 decades. Specifically, those time estimates are 976 (786 to 1,166) MYA for arthropods versus deuterostomes, 1,351 (1,116 to 1,586) MYA for Porifera versus Eumetazoa, and 1,513 (1,384 to 1,642) MYA for animals versus fungi. Therefore, the viewpoint that the fossil record is missing a long period in the early history of animals remains alive and well.

Acknowledgments

We thank K. J. Peterson for kindly providing alignments and S. Aris-Brosou for helpful information concerning his program, PhyBayes. We also thank Sudhir Kumar for valuable comments on this manuscript. This work was supported by grants to S.B.H. from the National Science Foundation and the NASA Astrobiology Institute.

Literature Cited

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.

———. 2003. Bayesian models of episodic evolution support a late Precambrian explosive diversification of the Metazoa. Mol. Biol. Evol. **20**:1947–1954.

Benton, M. J. 1993. The fossil record, Vol. 2. Chapman & Hall, New York.

———. 2000. Vertebrate paleontology. Chapman & Hall, New York.

Brown, R. H., M. Richardson, D. Boulter, J. A. Ramshaw, and R. P. Jefferies. 1972. The amino acid sequence of cytochrome c from *Helix aspersa* Muller (garden snail). Biochem. J. **128**: 971–974.

Butterfield, N. J. 2000. Bangiomorpha pubescens n. gen., n. sp.: Implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. Paleobiology **26**:386.

Chen, J.-Y., D. J. Bottjer, P. Oliveri, S. Q. Dornbos, F. Gao, S. Ruffins, H. Chi, C.-W. Li, and E. H. Davidson. 2004. Small bilaterian fossils from 40 to 55 million years before the Cambrian. Science 1099213.

Donoghue, P. C. J., M. P. Smith, and I. J. Sansom. 2004. The origin and early evolution of chordates: molecular clocks and the fossil record. Pp. 190–223 *in* P. C. J. Donoghue and M. P. Smith, eds. Telling the evolutionary time: molecular clocks and the fossil record. CRC Press, New York.

Douzery, E. J. P., E. A. Snell, E. Bapteste, F. Delsuc, and H. Philippe. 2004. The timing of eukaryote evolution: Does a relaxed molecular clock reconcile proteins and fossils? Proc. Natl. Acad. Sci. USA.

Gaunt, M. W., and M. A. Miles. 2002. An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. Mol. Biol. Evol. **19**:748–761.

Hedges, S. B., J. E. Blair, M. L. Venturi, and J. L. Shoe. 2004. A molecular timescale of eukaryote evolution and the rise of complex multicellular life. BMC Evol. Biol. **4**:2.

Hedges, S. B., and S. Kumar. 2004. Precision of molecular time estimates. Trends Genet. **20**:242–247.

Hedges, S. B., P. H. Parker, C. G. Sibley, and S. Kumar. 1996. Continental breakup and the ordinal diversification of birds and mammals. Nature **381**:226–229.

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. 2001. Mesoproterozoic megafossil Chuaria-Tawuia association may represent parts of a multicellular plant, Vindhyan Supergroup, Central India. Precambrian Res. **106**:187–211.

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

Peterson, K. J., J. B. Lyons, K. S. Nowak, C. M. Takacs, M. J. Wargo, and M. A. McPeek. 2004. Estimating metazoan divergence times with a molecular clock. Proc. Natl. Acad. Sci. USA **101**:6536–6541.

Pisani, D., L. L. Poling, M. Lyons-Weiler, and S. B. Hedges. 2004. The colonization of land by animals: molecular phylogeny and divergence times among arthropods. BMC Biol **2**:1.

Ruta, M., and M. I. Coates. 2004. Bones, molecules, and crown-tetrapod origins. Pp. 224–262 *in* P. C. J. Donoghue and M. P. Smith, eds. Telling the evolutionary time: molecular clocks and the fossil record. CRC Press, New York.

Sanderson, M. J. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics **19**:301–302.

Shu, D. G., L. Chen, J. Han, and X. L. Zhang. 2001. An early Cambrian tunicate from China. Nature **411**:472–473.

Springer, M. S., W. J. Murphy, E. Eizirik, and S. J. O'Brien. 2003. Placental mammal diversification and the Cretaceous-Tertiary boundary. Proc. Natl. Acad. Sci. USA **100**:1056–1061.

Woods, K. N., A. H. Knoll, and T. N. German. 1998. Xanthophyte algae from the Mesoproterozoic/Neoproterozoic transition: confirmation and evolutionary implications. G.S.A. Abstr. Programs **30**:A232.

Wray, G. A., J. S. Levinton, and L. H. Shapiro. 1996. Molecular evidence for deep Precambrian divergences among metazoan phyla. Science **274**:568–573.

Yang, Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. **13**: 555–556.

Brian Golding, Associate Editor

Accepted November 2, 2004