Comp. Biochem. Physiol. Vol. 117**B**, No. 3, pp. 379–385, 1997

ELSEVIER

ISSN 0305-0491/97/$17.00
PII S0305-0491(96)00333-1

Molecular Phylogenetics of a Protein Repair Methyltransferase

Ron M. Kagan, Helen J. McFadden,† Philip N. McFadden,† Clare O'Connor‡ and Steven Clarke**

*Department of Chemistry and Biochemistry and the Molecular Biology Institute, University of California, Los Angeles, CA 90095-1569, U.S.A.; †Department of Biochemistry and Biophysics, Oregon State University, Corvallis, OR 97330, U.S.A. and ‡Department of Biology, Boston College, Chestnut Hill, MA 02167-3811, U.S.A.

ABSTRACT. Protein-L-isoaspartyl (D-aspartyl) *O*-methyltransferase (E.C. 2.1.1.77) is a well-conserved and widely distributed protein repair enzyme that methylates isomerized or racemized aspartyl residues in age-damaged proteins. We exploited the availability of protein sequences from 10 diverse animal, plant and bacterial taxa to construct a phylogenetic tree and determine the rates of amino acid substitution for this enzyme. We used a likelihood ratio test to show that this enzyme fulfills the conditions for a molecular clock. We found that the rate of substitution is 0.39 amino acid substitutions per site per 10^9 years and remains relatively constant from bacteria to humans. We argue that this degree of sequence conservation may result from the functional constraints necessitated by the requirement to specifically recognize altered aspartyl but not normal aspartyl residues in proteins. Relative rate analysis of the *Caenorhabditis elegans* sequence suggests that the amino acid substitution rate in the nematode lineage may be higher than that in other lineages and that the divergence of nematodes may have been a more recent event than suggested by previous analysis. COMP BIOCHEM PHYSIOL 117B;3:379–385, 1997. © 1997 Elsevier Science Inc.

KEY WORDS. L-isoaspartyl methyltransferase, *Caenorhabditis elegans*, evolution, phylogenetics, molecular clock, protein repair, protein sequence comparison, *Arabidopsis thaliana*

INTRODUCTION

Protein L-isoaspartate (D-aspartate) *O*-methyltransferase (EC 2.1.1.77) catalyzes the transfer of methyl groups from AdoMet to the free carboxyl groups of the isomerized and racemized residues derived from aspartyl and asparaginyl residues in polypeptides (1,11,22,26,28). The methylation of these abnormal residues, the products of spontaneous aging processes, can lead to their conversion to normal L-aspartyl residues and may function to minimize the effects of protein isomerization and racemization in cells (Fig. 1). The enzyme is particularly active on L-isoaspartyl residues, the product of isomerization reactions, and "repair" of this type of damage has been directly demonstrated *in vitro* (2,15,16,30). The most direct evidence for the biological relevance of this methylation pathway is the significantly diminished stationary phase survival of *Escherichia coli* mutants lacking this enzyme (24).

L-isoaspartyl methyltransferase activity has been detected in a broad spectrum of organisms, including gram-negative bacteria (23,34), the protozoan *Tetrahymena thermophila* (R. Kagan, unpublished data), fungi such as a common mushroom (12,17) and *Neurospora crassa* (R. Kagan, unpublished data); a number of plants (17,31,32), nematodes (19) and other invertebrates (17) and all vertebrates examined including humans (5,34). This enzyme has been found in almost every cell type examined so far, although it has not been detected in four types of gram-positive bacteria (23) or in the yeast *Schizosaccharomyces pombe* (R. Kagan, unpublished data). In addition, we find no evidence for a homologous gene in the recently completed *Saccharomyces cerevisiae* genome sequence (14). The widespread distribution of this enzyme suggests that its function in recognizing damaged proteins may be essential to most life forms.

In light of the sequence and functional conservation of the L-isoaspartyl methyltransferases from *E. coli* to humans, we were interested in examining the phylogenetic relationships of the 10 orthologous gene products for which amino acid sequences are available. In this work, we show that this enzyme fulfills the conditions for a molecular clock. We also derive the evolutionary rate constant expressed as amino acid substitutions per site per year of this methyltransferase and then use this rate to attempt to corroborate the early divergence date of 1000 Myr postulated for the nematode

Address reprint requests to: S. Clarke, Department of Chemistry and Biochemistry and the Molecular Biology Institute, University of California, Los Angeles, CA 90095-1569, U.S.A. Tel. (310) 825-8754; Fax (310) 825-1968; E-mail: clarke@ewald.mbi.ucla.edu.

Received 29 August 1996; accepted 26 November 1996.

Caenorhabditis elegans based on cytochrome *c* and globin evolutionary rates (41).

MATERIALS AND METHODS

Sources of L-Isoaspartyl Methyltransferase Sequence Data

Sequences included human (*Homo sapiens*) PCMT isozyme I (NBRF-PIR accession number A34489), bovine (*Bos taurus*) PIMT-1 isozyme I (NBRF-PIR accession number A34242), rat (*Rattus rattus*) PCMT-1 (Genbank accession number D11475), mouse (*Mus musculus*) PCMT-1 (NBRF-PIR accession number JC1248), *C. elegans* PCM-1 (Genbank accession number U09669), wheat (*Triticum aestivum*) PCM (Genbank accession number L07941), zebrafish (*Danio rerio*) PCMT (Genbank accession number U37434), fruit fly (*Drosophila melanogaster*) Pcmt (Genbank accession number U37432), *Arabidopsis thaliana* PCMT (Genbank accession number U31288) and *E. coli* PCM (Genbank accession number M63493).

Sequence Alignments and Phylogenetic Tree Construction

A multiple alignment of 10 L-isoaspartyl methyltransferase amino acid sequences was constructed using the CLUSTAL W algorithm (40) and the default parameter settings. Amino acid sequences were compared for the full length of the protein species, ranging from 207 residues in *E. coli* to 229 residues (excluding the initiator methionine) in wheat and *A. thaliana*. Subsequent analysis was carried out with the Phylip version 3.56c program package (10). The fraction of identical amino acids between pairs of proteins was converted into difference scores ($P_d = 1 -$ fraction of identical residues) and the corrected distances were calculated according to the formula $K_{aa} = -\ln(1 - P_d - 0.2\,P_d^2)$ (21) using the PROTDIST program. Phylogenetic trees were constructed from the corrected distances using the neighbor-joining method (37). Bootstrap values for the nodes of the tree were obtained by applying the SEQBOOT program to generate 100 data sets and then calculating the distances with the PROTDIST program. The neighbor-joining trees were then constructed with the NEIGHBOR program, and

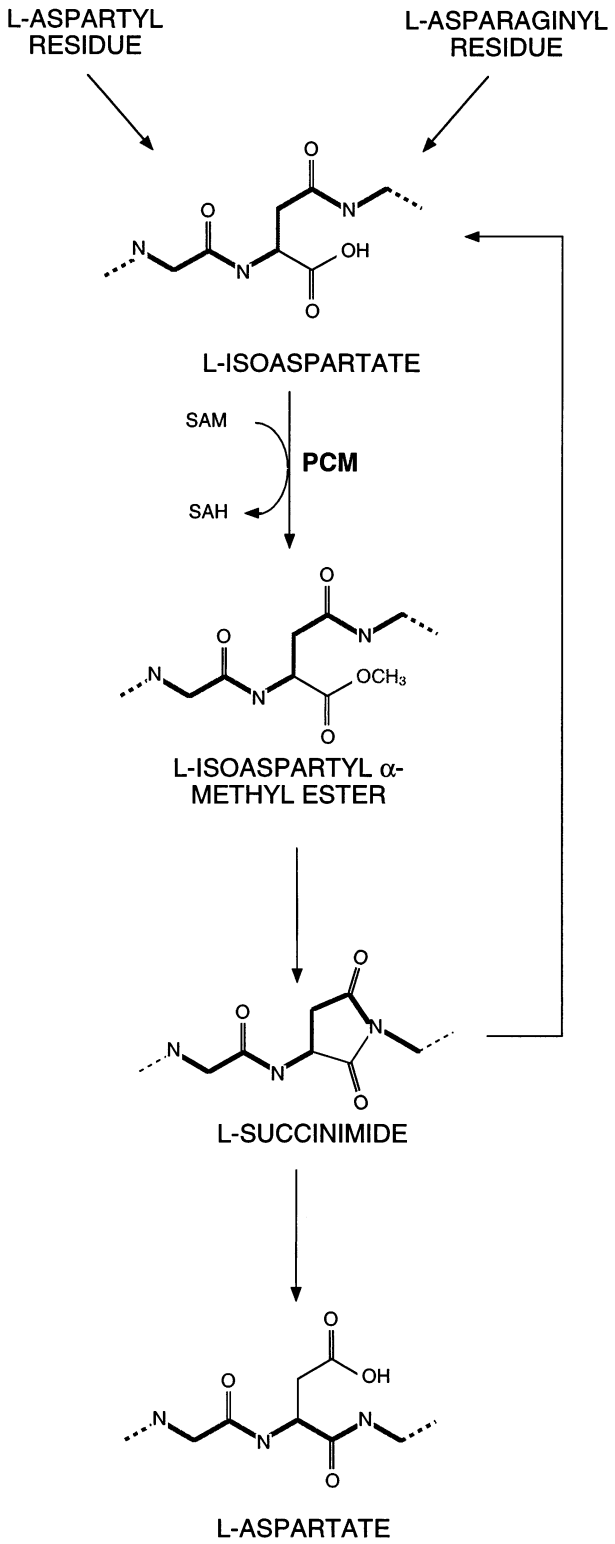

FIG. 1. Role of L-isoaspartyl methyltransferase in the repair of spontaneously damaged proteins. L-Isoaspartyl methyltransferase (PCM) can recognize the major L-isoaspartyl product of the non-enzymatic degradation of protein aspartyl and asparaginyl residues and initiate a process that leads to the conversion of isomerized aspartyl residues to normal L-aspartyl residues. Both the formation and the hydrolysis of the succinimide intermediate are spontaneous steps under physiological conditions. SAM, S-adenosyl-L-methionine; SAH, S-adenosyl-L-homocysteine; PCM, protein carboxyl methyltransferase.

the consensus tree and its bootstrap values were then found with the CONSENSE program. A maximum parsimony tree with the *E. coli* methyltransferase sequence chosen as the outgroup was obtained with the program PROTPARS and bootstrap values for 100 data sets were obtained in a manner similar to that described above, using SEQBOOT, PROTPARS and CONSENSE. Maximum likelihood analysis and likelihood ratio tests of the nine available L-isoaspartyl methyltransferase DNA sequences (bovine methyltransferase DNA sequence is not available) was carried out using DNAML and DNAMLK (10).

Evolutionary Rate Calculations

The rate of evolutionary change in terms of substitutions per site per 10^9 years (r) was calculated according to the formula $r = K_{aa}/2T$ (33), or one-half of the slope of the line of a plot of K_{aa} vs the divergence times for the aligned pairs of methyltransferases from various lineages. The values of T used for the divergence times of various lineages were as follows: eucaryote/procaryote, 1800 Myr (8); plant/animal, 1000 Myr (3); vertebrate/invertebrate, 600 Myr (8); zebrafish/mammal, 408 Myr (3); human/bovine, 75 Myr (36); mouse/rat, 29 Myr (35) and human/bovine vs mouse/rat, 80 Myr (35). Relative rate tests to compare the rates of amino acid substitution of lineages A and B in reference to a third species C that had branched off earlier were conducted according to the method of Sarich and Wilson (38,39) as described by Li and Grauer (25).

RESULTS AND DISCUSSION

The amino acid sequences of 10 L-isoaspartyl (D-aspartyl) *O*-methyltransferases from mammals (four sequences), fish (one sequence), insects (one sequence), nematodes (one sequence), plants (two sequences) and bacteria (one sequence) were aligned as shown in Fig. 2. All aligned proteins possess the three conserved S-adenosyl-L-methionine (AdoMet) and S-adenosyl-L-homocysteine (AdoHcy) motifs (designated I, II and III in the figure) characteristic of most AdoMet-dependent methyltransferases (18). The L-isoaspartyl (D-aspartyl) methyltransferases are also distinguished by two additional protein sequence motifs, designated pre-I and post-III in Fig. 2. These motifs have not yet been identified in any other type of methyltransferase; hence, they may be determinants of the substrate specificity for L-isoaspartyl or D-aspartyl residues. We find that amino acid substitutions among these enzymes do not cluster but are distributed throughout the sequence.

A phylogenetic tree was constructed from the distances corrected for multiple substitutions (21) by the neighbor-joining method (37) as shown in Fig. 3. The branching order of the tree follows accepted taxonomic relationships with *E. coli* appearing as the most divergent sequence. The same branching order was also supported at high values by 100 bootstrap replicates (10). A maximum parsimony tree constructed from the methyltransferase protein sequences produced the same tree with high bootstrap values except that in 58 of 100 trees the bovine methyltransferase sequence was clustered with the mouse and the rat sequences rather than with the human sequence. Maximum likelihood analysis (10) of the nine available L-isoaspartyl methyltransferase DNA sequences gave the same topology as the distance and parsimony trees (data not shown).

To determine the rate of amino acid substitution for the L-isoaspartyl methyltransferases, the number of amino acid substitutions per site were plotted against the divergence times for the respective taxons (Fig. 4, A and B). The *C. elegans* distance values were omitted from this analysis because we wished to use the amino acid substitution rate from the resulting linear fit to estimate the divergence date of this organism. The wheat vs *A. thaliana* distance was also omitted, as it was shown in relative rate tests not to follow rate constancy (see Table 2) and deviated substantially from the linear relationship established with the other 44 sequence pair distances in this plot. The data for all sequences or for eucaryotic sequences alone both gave a good linear fit with correlation coefficients (r^2) of 0.95 and 0.90, respectively.

The evolutionary rate for all protein L-isoaspartyl methyltransferase sequences was 0.39 substitutions/site/10^9 years (Fig. 4A), which corresponds to a unit evolutionary period of 12.8 Myr. When the *E. coli* sequence was omitted from the analysis to test whether the rate was different for eucaryotes, a slightly higher value of 0.42 substitutions/site/10^9 years was obtained (Fig. 4B). The rate of amino acid substitutions of the L-isoaspartyl methyltransferase is comparable with that of lactate dehydrogenase and cytochrome b_5 and about 77% greater than that of cytochrome *c*, an electron transport protein that is commonly used as a molecular clock (7,33).

Recent work has attempted to use molecular sequence data to establish the divergence time of the phylum Nematoda. Vanfleteren *et al.* (41) used nematode sequences for cytochrome *c* and globins to establish a divergence time based on the accepted unit evolutionary period of 21 Myr for cytochrome *c* and 5.0 Myr for globin. They obtained a balanced estimate for the nematode divergence time of 1047 ± 46 Myr (41). Table 1 shows the estimated divergence time of the Nematode phylum based on the amino acid substitution rate of the L-isoaspartyl methyltransferase in the nematode *C. elegans* and the rate constants determined in Fig. 4, A and B. The average divergence time obtained from the rate constant in Fig. 4A was 1019 ± 90 Myr and 940 ± 83 Myr for the eucaryote-only rate constant in Fig. 4B. These divergence times agree with the values obtained by Vanfleteren *et al.* (41).

A necessary assumption in obtaining divergence times from molecular sequence data is that the sequences in question obey the molecular clock. Gillespie (13) has argued

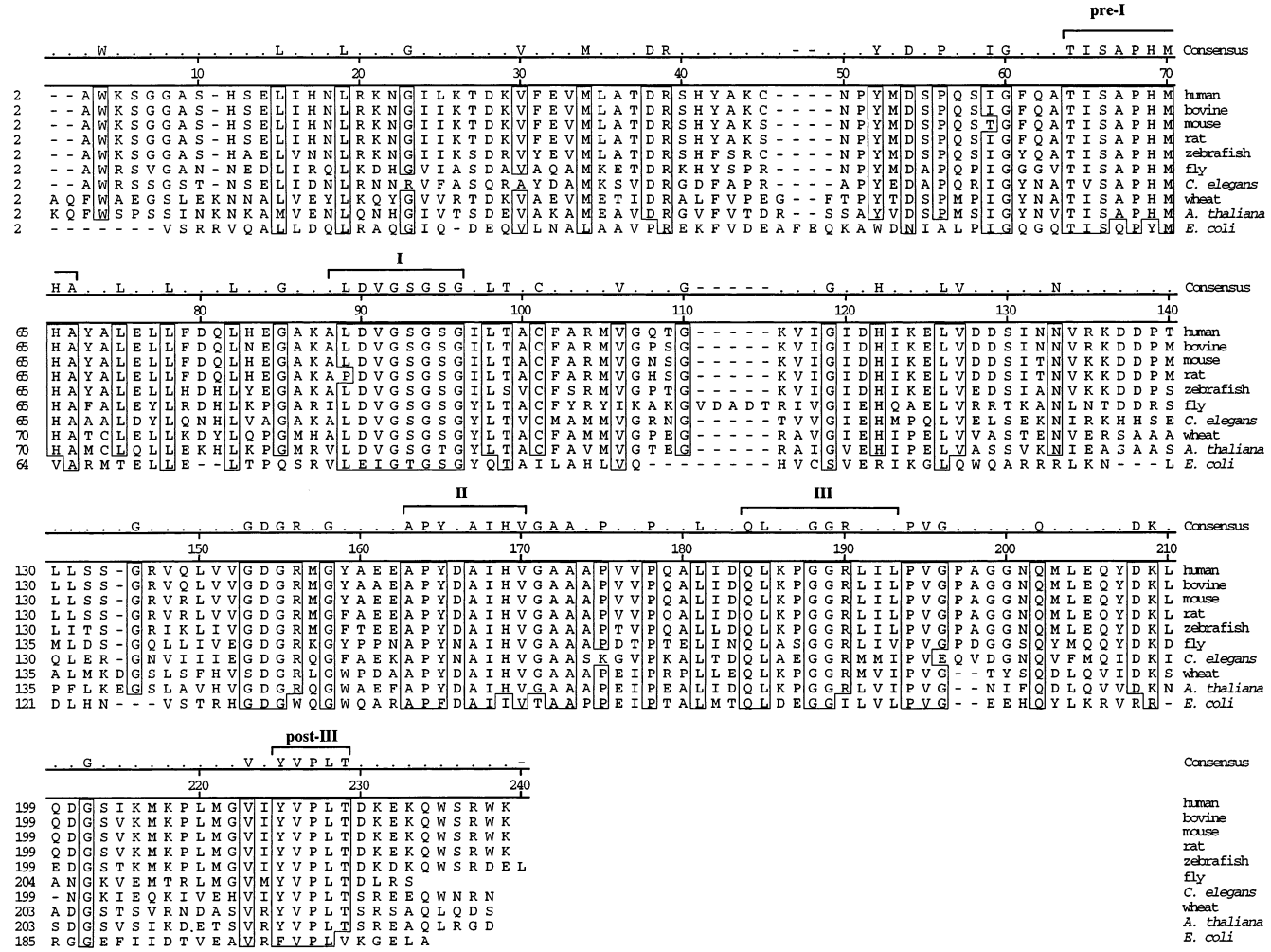

FIG. 2. Multiple alignment of L-isoaspartyl (D-aspartyl) methyltransferases. The methyltransferase protein sequences from 10 species were input into the CLUSTAL W multiple alignment program (40) and aligned with the default program parameters as described in Materials and Methods. "I," "II" and "III" designate the three AdoMet-binding motifs (18). The sequence motifs designated "pre-I" and "post-III" are described in the text.

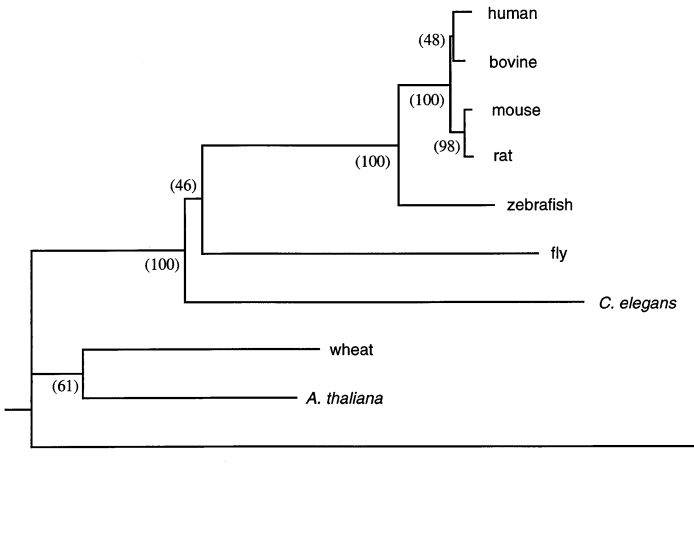

FIG. 3. Phylogenetic tree of the L-isoaspartyl methyltransferases. A neighbor-joining tree of the aligned sequences in Fig. 2 was constructed from 100 bootstrap replicates of the data. The numbers to the right of the tree nodes indicate the number of times that node appeared in 100 bootstrap samples. The branch lengths are proportional to the number of substitutions per site. 1 cm = 0.05 substitutions per site.

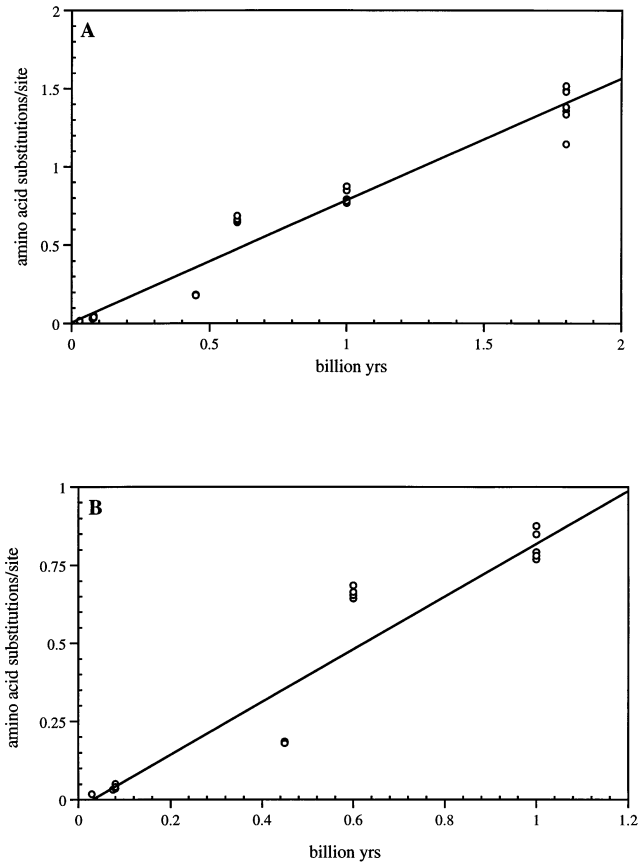

FIG. 4. Amino acid substitutions per site in the L-isoaspartyl methyltransferases plotted as a function of the estimated divergence times of the species used in this study. The divergence times for the lineages studied were taken from the sources listed in Materials and Methods. The substitutions per site between each pair of species was calculated according to Kimura's equation (21). The data points were fitted to a straight line by linear regression. (A) Plot of substitutions per site (K_{aa}) vs divergence time (T) in billions of years for all taxa used. The linear regression equation fitting the data was $K_{aa} = 0.778T + 0.009$ ($r^2 = 0.95$). (B) Plot of substitutions per site vs divergence time for eucaryotic taxa only. The linear regression equation fitting the data was $K_{aa} = 0.844T - 0.026$ ($r^2 = 0.90$).

that the molecular clock assumption may not be valid for many proteins. A maximum likelihood analysis of the DNA sequences of the L-isoaspartyl methyltransferases that we used to calibrate the molecular clock (Fig. 4, A and B) both with and without the assumption of a molecular clock allowed us to use the likelihood ratio test to ascertain the validity of the clock assumption for this enzyme (10). The log likelihood without the clock assumption was −4906.7, whereas it was −4912.5 with the clock assumption. The double of this difference is 11.6. Assuming $n - 2$ degrees of freedom and a significance level of $\alpha = 0.05$, this is less than the χ^2 (0.95) (6) value of 18.48; thus, we do not reject the null hypothesis that the molecular clock is obeyed.

We performed relative rate tests (25,38,39) to determine

TABLE 1. Estimation of the divergence date of *Caenorhabditis elegans* from L-isoaspartyl methyltransferase amino acid substitution rates

Nematode vs	K_{aa}	Myr (all)*	Myr (eucaryote)†
Human	0.730	938	865
Bovine	0.752	967	891
Mouse	0.763	981	905
Rat	0.752	967	891
Fish	0.741	952	878
Fly	0.817	1052	968
Wheat	0.928	1192	1099
Arabidopsis thaliana	0.861	1106	1020
Mean		1019 ± 34	940 ± 31

Divergence dates were calculated according to the equation $T = K_{aa}/2r$, where K_{aa} designates amino acid substitutions per site and r designates amino acids per site per 10^9 years.

*An r value of 0.39 substitutions/site/10^9 years was calculated from the pairwise distances of the L-isoaspartyl methyltransferases.

†An r value of 0.42 substitutions/site/10^9 years was calculated as described above from using the eucaryotic L-isoaspartyl methyltransferases alone.

whether the rate of substitution of the C. *elegans* methyltransferase sequence was constant with respect to the other lineages. Table 2 shows that the amino acid substitution rate for the C. *elegans* sequence is 38% greater than the mammalian rate (average for three reference groups shown in Table 2), 49% greater than the rate for wheat and as much as a 153% greater than the rate for A. *thaliana*. When the C. *elegans* DNA sequence was included in the maximum likelihood analysis, double the difference of the log likelihoods increased to 18, which is greater than the χ^2 (0.975) (7) value of 16.01, suggesting that the more rapid substitution rate in the C. *elegans* lineage is responsible for the failure of the molecular clock null hypothesis at the $\alpha = 0.025$ significance level in this second likelihood ratio test.

These data show that the rate of amino acid substitution in the C. *elegans pcm-1* gene is accelerated, and therefore the 1000 Myr estimate of the divergence date of nematodes may be too early. This conclusion may also be warranted by the cytochrome *c* substitution rates in the nematodes *Ascaris suum* and C. *elegans* that were also found to be high in relative rate tests with a protozoan outgroup (41). An anomalous mutation rate has also been discerned in C. *elegans* rRNA genes, where a small subunit rRNA evolutionary tree clusters C. *elegans* with the animals, whereas in the large subunit rRNA tree C. *elegans* branches as a protist (4). It would be useful to determine the amino acid sequences of *pcm-1* homologs from other nematode species to determine whether the evolutionary rate of C. *elegans* is anomalously high or whether the nematodes actually did diverge as early as 1000 Myr. The absence of Precambrian fossil forms of early metazoans cannot exclude the possibility of an earlier divergence date, because soft-bodied microscopic forms

TABLE 2. Relative rate tests of some L-isoaspartyl methyltransferases

Reference	K_{OA}	K_{OB}	$K_{OB} - K_{OA}$
	Mammal	Nematode	
Escherichia coli	0.304	0.445	0.141 ± 0.015
Wheat	0.303	0.447	0.144 ± 0.011
Arabidopsis thaliana	0.341	0.409	0.068 ± 0.000
	Wheat	Nematode	
Escherichia coli	0.372	0.555	0.182
	Arabidopsis thaliana	Nematode	
Escherichia coli	0.244	0.617	0.373
	Mammal	Fly	
Escherichia coli	0.276	0.380	0.103 ± 0.012
Wheat	0.296	0.361	0.065 ± 0.011
Arabidopsis thaliana	0.287	0.369	0.083 ± 0.000
	Arabidopsis thaliana	Mammal	
Escherichia coli	0.280	0.512	0.232 ± 0.015
	Wheat	Mammal	
Escherichia coli	0.371	0.413	0.042 ± 0.015
	Arabidopsis thaliana	Wheat	
Escherichia coli	0.151	0.341	0.191

Relative rate test of Sarich and Wilson (38,39) as described in Li and Grauer (25).

*Reference sequence to wich the amino acid substitutions per site of the two sequences were compared.

†To simplify te appearance of the table, average values obtained for human, bovine, mouse and rat sequences are given. The calculated $K_{OB} - K_{OA}$ values ± SE were obtained from the distance values for individual species.

would not have been preserved (6). However, a recent determination of the divergence times of major biological groupings derived from the amino acid substitution rates of 57 different enzymes has placed the schizocoelomate-pseudocoelomate divergence date at about 750 Myr (9), which is consistent with our argument against an earlier divergence date for nematodes. Recently, comparison of seven amino acid and nucleotide sequences from a number of chordate, mollusk and echinoderm species have provided evidence that metazoan phyla may have diverged at about 1000–1200 Myr (42). Nevertheless, even in the latter study, the authors found that relative rate tests for nematode sequences indicated consistently faster rates of sequence divergence, and these organisms were not included in the study (42).

The data in Table 2 also show that the amino acid substitution rate of the *A. thaliana* sequence is 45–56% slower than that of either mammals or wheat. Further, the corrected number of substitutions between *A. thaliana*, a dicot, and wheat, a monocot (0.492 substitutions per site), leads to the estimation of a monocot/dicot divergence date of ca. 600 Myr, which is incompatible with a 300 Myr divergence date estimated from plant nuclear and chloroplast sequence data (29). Kemmerer *et al.* (20) conducted phylogenetic analysis of *A. thaliana* cytochrome *c* and histone H3 genes and found that they clustered poorly with those of the other higher plants and clustered instead with fungal species raising the possibility of convergent evolution. Interestingly, the *A. thaliana* L-isoaspartyl methyltransferase amino acid sequence is the least divergent from *E. coli* (1.14 substitutions per site), contrasted to 1.34 substitutions per site for wheat/*E. coli* and 1.36–1.38 substitutions per site for mammals/*E. coli*. Because no L-isoaspartyl methyltransferase sequences are available from fungal species, it was not possible to test whether the sequence of this enzyme in *A. thaliana* also converged toward that of fungal sequences.

The high degree of sequence conservation of the L-isoaspartyl methyltransferase cannot be explained merely by the requirement to recognize the methyl donor AdoMet. AdoMet utilization is common to most methyltransferases with widely divergent sequences (18). A much greater constraint may be imposed on this enzyme by the requirement to distinguish the α-carboxyl from the β-carboxyl of aspartyl residues in proteins and selectively transfer a methyl group to the β-carboxyl of an L-isoaspartyl residue while leaving the α-carboxyl groups of normal L-aspartyl residues unmodified. The enzyme fails to recognize D-isoaspartyl substrates and only the mammalian enzymes appear to recognize D-aspartyl substrates (19,28). The enzyme also exhibits a high degree of discrimination among the sequence contexts of L-isoaspartyl residues in peptides and proteins, with a several thousand-fold variation in K_m values among peptide substrates and a thousand-fold or greater variation among protein substrates (27,28). Further structural studies, including X-ray crystallography of this enzyme, may be used in clarifying the structural constraints imposed by the function of this enzyme.

This work was supported by National Institutes of Health Grant GM 26020 to S.C., National Institutes of Health Grant AG 11087 to P.N.M. and National Institutes of Health Grant AG 08109 to C.O.

References

1. Aswad, D.W.; Johnson, B.A. The unusual substrate specificity of eukaryotic protein carboxyl methyltransferases. Trends Biochem. Sci. 12:155–158;1987.
2. Brennan, T.V.; Anderson, J.W.; Jia, Z.; Waygood, E.B.; Clarke, S. Repair of spontaneously deamidated HPr phosphocarrier protein catalyzed by the L-isoaspartate-(D-aspartate) O-methyltransferase. J. Biol. Chem. 269:24586–24595;1994.
3. Carroll, R.L. Vertebrate Paleontology and Evolution. New York: Freeman; 1988.
4. Cedergren, R.; Gray, M.W.; Abel, Y.; Sankoff, D. The evolutionary relationships among known life forms. J. Mol. Evol. 28:98–112;1988.
5. Clarke, S. Protein carboxyl methyltransferases: Two distinct classes of enzymes. Annu. Rev. Biochem. 54:479–506;1985.
6. Davidson, E.H.; Peterson, K.J.; Cameron, R.A. Origin of bilaterian body plans: Evolution of developmental regulatory mechanisms. Science 270:1319–1325;1995.

7. Dickerson, R.E. The structures of cytochrome c and the rates of molecular evolution. J. Mol. Evol. 1:26–45;1971.
8. Doolittle, R.F.; Anderson, K.L.; Feng, D.F. Estimating the prokaryote-eukaryote divergence time from protein sequences. In: Fernholm, B.; Bremer, K.; Jornvall, H. (eds). The Hierarchy of Life. Amsterdam: Elsevier Science Publishers B.V; 1989:73–85.
9. Doolittle, R.F.; Feng, D.-F.; Tsang, S.; Cho, G.; Little, E. Determining divergence times of the major kingdoms of living organisms with a protein clock. Science 271:470–476;1996.
10. Felsenstein, J. "PHYLIP (Phylogeny Inference Package)" version 3.56c. Distributed by the author, Dept. of Genetics, University of Washington, Seattle, WA; 1993.
11. Galletti, P.; Ingrosso, D.; Manna, C.; Clemente, G.; Zappia, V. Protein damage and methylation-mediated repair in the erythrocyte. Biochem. J. 306:313–325;1995.
12. Ghosh, A.K. Protein carboxyl methylation in the mushroom *Volvariella volvacea*. Phytochem. 32:1093–1096;1993.
13. Gillespie, J.H. The Causes of Molecular Evolution. New York: Oxford University Press; 1991.
14. Goffeau, A.; Barrell, B.G.; Bussey, H.; Davis, R.W.; Dujon, B.; Feldmann, H.; Galibert, F.; Hoheisel, J.D.; Jacq, C.; Johnston, M.; Louis, E.J.; Mewes, H.W.; Murakami, Y.; Philippsen, P.; Tettelin, H.; Oliver, S.G. Life with 6000 genes. Science 274:546–567;1996.
15. Johnson, B.A.; Langmack, E.L.; Aswad, D.W. Partial repair of deamidation-damaged calmodulin by protein carboxyl methyltransferase. J. Biol. Chem. 262:12283–12287;1987.
16. Johnson, B.A.; Murray, E.J., Jr.; Clarke, S.; Glass, D.B.; Aswad, D.W. Protein carboxyl methyltransferase facilitates conversion of atypical L-isoaspartyl peptides to normal L-aspartyl peptides. J. Biol. Chem. 262:5622–5629;1987.
17. Johnson, B.A.; Ngo, S.Q.; Aswad, D.W. Widespread phylogenetic distribution of a protein methyltransferase that modifies L-isoaspartyl residues. Biochem. Int. 24:841–847;1991.
18. Kagan, R.M.; Clarke, S. Widespread occurrence of three sequence motifs in diverse S-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes. Arch. Biochem. Biophys. 310:417–427;1994.
19. Kagan, R.M.; Clarke, S. Protein L-isoaspartyl methyltransferase from the nematode *Caenorhabditis elegans*: genomic structure and substrate specificity. Biochemistry 34:10794–10806; 1995.
20. Kemmerer, E.C.; Lei, M.; Wu, R. Structure and molecular evolutionary analysis of a plant cytochrome c gene: Surprising implications for *Arabidopsis thaliana*. J. Mol. Evol. 32:227–237; 1991.
21. Kimura, M. The Neutral Theory of Molecular Evolution. New York: Cambridge University Press; 1983.
22. Ladino, C.A.; O'Connor, C.M. Methylation of atypical protein aspartyl residues during the stress response of HeLa cells. J. Cell Physiol. 153:297–304;1992.
23. Li, C.; Clarke, S. Distribution of an L-isoaspartyl protein methyltransferase in eubacteria. J. Bacteriol. 174:355–361; 1992.
24. Li, C.; Clarke, S. A protein methyltransferase specific for altered aspartyl residues is important in *Escherichia coli* stationary-phase survival and heat-shock resistance. Proc. Natl. Acad. Sci. USA 89:9885–9889;1992.
25. Li, W.H.; Grauer, D. Fundamentals of Molecular Evolution. Sunderland, MA: Sinauer Associates; 1991.
26. Lindquist, J.A.; McFadden, P.N. Automethylation of protein (D-aspartyl/L-isoaspartyl) carboxyl methyltransferase, a response to enzyme aging. J. Protein Chem. 13:23–30;1994.
27. Lowenson, J.D.; Clarke, S. Structural elements affecting the recognition of L-isoaspartyl residues by the L-isoaspartyl/D-aspartyl protein methyltransferase. Implications for the repair hypothesis. J. Biol. Chem. 266:19396–19406;1991.
28. Lowenson, J.D.; Clarke, S. Recognition of D-aspartyl residues in polypeptides by the erythrocyte L-isoaspartyl/D-aspartyl protein methyltransferase. Implications for the repair hypothesis. J. Biol. Chem. 267:5985–5995;1992.
29. Martin, W.; Lydiate, D.; Brinkmann, H.; Forkmann, G.; Saedler, H.; Cerff, R. Molecular phylogenies in angiosperm evolution. Mol. Biol. Evol. 10:140–162;1993.
30. McFadden, P.N.; Clarke, S. Conversion of isoaspartyl peptides to normal peptides: Implications for the cellular repair of damaged proteins. Proc. Natl. Acad. Sci. USA 84:2595–2599; 1987.
31. Mudgett, M.B.; Clarke, S. Characterization of plant L-isoaspartyl methyltransferases that may be involved in seed survival: Purification, cloning, and sequence analysis of the wheat germ enzyme. Biochemistry 32:11100–11111;1993.
32. Mudgett, M.B.; Clarke, S. A distinctly regulated protein repair L-isoaspartyl methyltransferase from *Arabidopsis thaliana*. Plant Mol. Biol. 30:723–737;1996.
33. Nei, M. Molecular Evolutionary Genetics. New York: Columbia University Press; 1987.
34. O'Connor, C.M.; Clarke, S. Specific recognition of altered polypeptides by widely distributed methyltransferases. Biochem. Biophys. Res. Commun. 132:1144–1150;1985.
35. O'Huigin, C.; Li, W.H. The molecular clock ticks regularly in muroid rodents and hamsters. J. Mol. Evol. 35:377–384; 1992.
36. Saccone, C.; Lanave, C.; Pesole, G. Time and biosequences. J. Mol. Evol. 37:154–159;1993.
37. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406–425;1987.
38. Sarich, V.M.; Wilson, A.C. Rates of albumin evolution in primates. Proc. Natl. Acad. Sci. USA 58:142–148;1967.
39. Sarich, V.M.; Wilson, A.C. Generation time and genomic evolution in primates. Science 179:1144–1147;1973.
40. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673–4680;1994.
41. Vanfleteren, J.R.; Van de Peer, Y.; Blaxter, M.L.; Tweedie, S.A.; Trotman, C.; Lu, L.; Van Hauwaert, M.-L.; Moens, L. Molecular genealogy of some nematode taxa as based on cytochrome c and globin amino acid sequences. Mol. Phylogenet. Evol. 3:92–101;1994.
42. Wray, G.A.; Levinton, J.S.; Shapiro, L.H. Molecular evidence for deep Precambrian divergences among metazoan phyla. Science 274:568–573;1996.