Metabolic Rate and Directional Nucleotide Substitution in Animal Mitochondrial DNA

Andrew P. Martin
Department of Biological Sciences, University of Nevada, Las Vegas

There is marked heterogeneity of nucleotide composition in mitochondrial DNA across divergent animals. Differences in nucleotide composition presumably reflect differences in directional nucleotide substitution for A+T or G+C nucleotides. In mitochondrial DNA, there is A+T directional nucleotide substitution in most (if not all) animals surveyed, and the magnitude of directional A+T nucleotide substitution differs greatly within and among groups. Differences in directional nucleotide substitution among lineages of mammals can be explained by changes in metabolic physiology. This relationship is thought to be mediated by the effect of oxygen radicals because these toxic compounds are by-products of aerobic metabolism and are known mutagens. Association between metabolism and nucleotide composition provides additional evidence in favor of the hypothesis that rates and patterns of nucleotide substitution in mitochondrial DNA can be influenced by factors that impinge on rates of endogenous DNA damage.

Introduction

There is marked heterogeneity in the nucleotide composition of mitochondrial genomes across taxa. Disparity among taxa is most evident at sites free to vary (i.e., four-fold degenerate codon positions of protein coding genes). The proportion of G+C nucleotides at four-fold degenerate sites varies from less than 0.05 in some insects to almost 0.50 in human (fig. 1). In all cases, G+C nucleotide content at sites free to vary is less than 50%. The G+C content of DNA is thought to be an effect of rates of nucleotide substitutions in two directions: from AT to GC base pairs and from GC to AT base pairs. Sueoka (1988) defined “directional mutation pressure” (μΔ) as the rate of mutation from AT to GC base pairs divided by the sum of the two mutation pathways (AT to GC and GC to AT). In general, G+C contents less than 50% indicate that rates of GC to AT mutation are greater than rates of AT to GC mutation. Thus, differences in nucleotide composition among taxa are thought to reflect differences in the relative rates of AT to GC and GC to AT mutation pathways.

Various alternative hypotheses have been proposed to explain differences in nucleotide composition across taxa. One hypothesis posits that directional mutation pressure is driven by the availability of nucleotides during DNA synthesis (Jukes and Blushan 1986; Wolfe 1991). Although there is a large literature showing that polymerase preference can be altered by changing the relative concentration of available nucleotides (Hibner and Alberts 1980), the skew of nucleotide abundance necessary to observe polymerase selectivity for specific nucleotides is typically several-fold or orders of magnitude greater (Hibner and Alberts 1980; Kunz 1982; Kaguni et al. 1986) than differences observed in cellular or organelle nucleotide pools (Skoug and Bjursell 1974; Bestwick and Mathews 1982; Bestwick et al. 1982; Kaguni et al. 1986). Moreover, Drosophila mitochondrial polymerases do not show a distinct kinetic preference for A or T nucleotides relative to G or C (Kaguni et al. 1986).

Correlation between G+C content and temperature led Bernardi and Bernardi (1986) to hypothesize that GC base pairs are selected for because they ensure “better protection against DNA breathing and mutability” (p. 6); in other words, GC base pairs increase the stability of DNA relative to AT base pairs, presumably because of the additional hydrogen bond. Although this is an intriguing hypothesis, it fails to explain why organisms which experience similar temperature regimes (e.g., different primates) have substantially different nucleotide compositions.

Other hypotheses invoked to explain compositional heterogeneity (e.g., mutator genes [Strazewski 1990], structural association with histone genes and differences
in chromatin state [Sueoka 1988], existence of ancient repetitive DNA sequences [Ohno 1988], codon-anticodon interactions [Grantham et al. 1981], differences in patterns of cytosine methylation [Quigley et al. 1989]) have little relevance for mtDNA.

In this report, I consider the hypothesis that differences in nucleotide composition (GC content) of mitochondrial genomes results from differences in rates of DNA damage across taxa. It is important to remember that differences in composition among taxa reflect heterogeneity in the probabilities of different nucleotide substitutions and that the probabilities of particular nucleotide substitutions are a function of the molecular process of mutation and population-level mechanisms that eventually result in nucleotide substitution as inferred from comparative analysis of individual sequences. Therefore, differences in the proportion of GC or AT nucleotides among sequences reflect differences in directionality of nucleotide substitution, not necessarily directional mutation pressure. Nevertheless, it is likely that nucleotide substitution at silent (degenerate) codon positions approximates mutation, and this has led many researchers to refer to directional nucleotide substitution as directional mutation pressure (Sueoka 1988; Jermiin et al. 1994).

Recent analysis of directional nucleotide substitution in mitochondrial DNA revealed marked heterogeneity among groups of taxa (Jermiin et al. 1994), indicating that probabilities of particular nucleotide substitutions differ across taxonomic groups. Accumulating evidence indicates that nucleotide substitution rates in mitochondrial DNA are heterogeneous across taxa (Avise et al. 1992; Martin et al. 1992; Martin and Palumbi 1993; Cantatore et al. 1994). Observed nucleotide composition is a product of the relative probabilities of the different substitution pathways; therefore, substitution rate variation and differences in nucleotide composition among taxa are probably effects of a common mechanism. Recent reports indicate that much of the variation in nucleotide substitution rates can be accounted for by differences in cellular metabolism (Adachi et al. 1993; Martin and Palumbi 1993; Cantatore et al. 1994). In particular, rates of weight-specific metabolism are positively related to nucleotide substitution rates (Martin and Palumbi 1993). Association between metabolic rate and nucleotide substitution is thought to be mediated by the mutagenic effects of highly reactive oxygen radicals that are abundant by-products of aerobic respiration (Richter et al. 1988; Lindahl 1993; Shigenaga et al. 1994).

Recognition that oxidation is a major source of DNA damage, increasing evidence for a link between rates of DNA damage and the accumulation of mutations, and the inextricable association between the probability of nucleotide substitution and composition, suggests that nucleotide composition heterogeneity reflects differences in rates of DNA damage across taxa. To explain the association between DNA damage and nucleotide composition, it is necessary to review identified mutation pathways. Biochemical studies indicate that increasing rates of DNA damage can shift directional mutation pressure in favor of AT nucleotides (Randall et al. 1987; Storz et al. 1987; Cheng et al. 1992; Lindahl 1993; Thomas and Kunkel 1993; Feig et al. 1994). Wagner et al. (1992, pp. 3383–3384) state that “a number of recent studies suggest that G-C damage is more likely to lead to base changes of which the majority are G-C to A-T transitions . . . . Damage to dC residues appears to be the principal cause of G-C to A-T transitions since damage to dG leads to G-C to T-A transversions.” One reason for the shift from GC to AT nucleotides is the preference of polymerases to insert dATP opposite abasic sites (Schaaper et al. 1983; Randall et al. 1987). In the mitochondria, oxidatively damaged bases are removed by nuclear-encoded repair enzymes which are imported from the cytoplasm (Myers et al. 1988; Sato et al. 1988; Tomkinson et al. 1988), providing opportunity for incorporation of dATP. In addition, lability of the N-glycosyl bond that anchors the base to the sugar-phosphate background can result in appreciable rates of spontaneous release of bases (mainly purines) from DNA (Lindahl 1993). If the damaged or missing base is a G or C, the result is a C to A or G to T transversion, respectively. By contrast, damage to an A or T results in AT to TA transversion. A second reason

---

**Fig. 1.**—Percentage of G+C nucleotides at four-fold degenerate sites of mitochondrially encoded protein-coding genes from various taxa (from Perna and Kocher 1995b). In cases where there were more than one taxon, values are means, and the standard deviation is indicated.
for the shift toward AT with increasing rates of DNA damage is hydrolytic deamination which can convert cytosine to uracil—at rates similar to depurination—resulting in a G:C to A:T transition (Lindahl 1993). Conversely, deamination of adenine occurs at 2%—3% of the rate of cytosine deamination (Lindahl 1993). Thus, a combination of increased susceptibility of dGTP and dCTP to damage and the preferential incorporation of dATP opposite abasic sites by polymerase probably explains most of the shift toward AT base pairs in the face of accelerating rates of DNA damage.

In contrast to the accumulation of A+T nucleotides associated with DNA damage mutation pathways, mismatch mutation pathways do not exhibit directionality. Models and empirical evidence indicates that base mismatches can and do occur. Although non-Watson-Crick base pairings are orders of magnitude less stable than standard Watson-Crick pairing (i.e., A:T and G:C), G:T and A:C pairs can, and do, occur (Topal and Fresco 1976). In the absence of mismatch repair, G:T and A:C mismatches result in transition mutations with symmetric probabilities (i.e., A to G = G to A, T to C = T to C). In addition, A:G and A:A mispairing have been inferred (Ferhst and Knill-Jones 1981), which will result in C:G to A:T and A:T to C:G transversions in the former case and A:T to T:A transversions in the latter case. Thus, DNA damage mutation pathways result in an asymmetric mutation matrix (directional mutation from GC to AT base pairs), whereas mismatch mutation pathways produce symmetric probabilities of change (fig. 2). If mutation is driven by mismatch mutation pathways (and therefore a symmetric mutation matrix), then nucleotide composition will evolve to an equilibrium of equal frequency for all nucleotides. By contrast, if we alter the probabilities of GC to TA and GC to AT mutations relative to AT to GC and AT to CG mutations, then it is possible to change the equilibrium frequency of G+C nucleotides in the direction expected with increasing rates of DNA damage (fig. 3).

There is a strong positive correlation between weight-specific metabolic rate (SMR) and the production of highly reactive oxygen radicals (Sohal et al. 1990). Furthermore, there is a significant positive correlation between SMR and rates of DNA damage and mutation (Adelman et al. 1988; Cortopassi et al. 1992), presumably as a result of greater flux of oxygen radicals. These observations suggest that change in SMR along a lineage, either as a result of body size evolution or metabolic adaptation, should be accompanied by change in directional nucleotide substitution, although the connection between metabolism and nucleotide composition is indirect and potentially influenced by a large number of variables. Increase in SMR should lead to a reduction in G+C nucleotides and vice versa. Association between acceleration or deceleration in SMR and directional nucleotide substitution will most likely be evident for mitochondrial DNA because of its close proximity to sites of oxygen radical production. In addition, there are dra-
matic differences among taxa in nucleotide composition at the third codon positions of mitochondrial genes. Therefore, attention is focused on mitochondrial DNA.

Material and Methods

Mammals were chosen to test for an association between change in SMR and directional nucleotide substitution. Examination of directional nucleotide substitution was done for third positions of protein-coding genes because these sites conform most closely with expectations of neutral sites (i.e., the probability of substitution at synonymous sites is identical across the gene; Irwin et al. 1991). Estimates of directional mutation pressure (= directional nucleotide substitution) at third codon positions were available for the cytochrome b gene from mammals (Jermiin et al. 1994). Estimates of directional nucleotide substitution for primates were for the third codon position of the cytochrome oxidase II gene (Adkins and Honeycutt 1994). Directional nucleotide substitution for this gene was estimated by applying the regression equation for the relationship of directional "mutation pressure", $\mu_D$ (defined as the rate of AT to GC mutation divided by the sum of the rate of AT to GC and GC to AT mutation; Sueoka 1988) and the proportion of G+C nucleotides at third codon positions for the cytochrome b gene in animals ($\mu_D = 1.94 \text{GC} - 0.414, r = 0.984, P = 0.0001; \text{from Jermiin et al. 1994}$). Estimates of SMR (in milliliters of oxygen consumed per gram of tissue per hour) were taken from the literature for mammals (Altman and Ditmer 1968) and estimated for primates from female body size using the well-known allometric relation between mass and specific metabolic rate ($\text{VO}_2/M_b = 0.676 M_b^{0.75}$; Schmidt-Nielsen 1979; p. 185). For each group, nodal values for given topologies were estimated for directional nucleotide substitution and SMR using the continuous trace option in MacClade with a rooted tree (Maddison 1991; Maddison and Maddison 1992). In this way, change in SMR and directional nucleotide substitution could be estimated for each lineage. Both standard linear regression and Kendall's nonparametric correlation statistic (Sokal and Rohlf 1981; p. 601) were used to test for association between variables.

Results

For mammals (fig. 4) and primates (fig. 5) there is a significant negative correlation between change in SMR and change in directional nucleotide substitution along lineages. (Similar results were obtained for a few species of insects; data not shown.) Increasing SMR is associated with increasing directional nucleotide substitution in favor of A+T nucleotides. It is important to remember that all of the lineages experience A+T di-

![Fig. 4.—A, Phylogenetic hypothesis for lineages of mammals and estimates of mutation pressure ($\mu_D$) from Jermiin et al. 1994) and specific metabolic rate (in milliliters oxygen consumed per gram per hour) (from Altman and Ditmer 1968 and Schmidt-Nielsen 1979). Values at the nodes were determined using the continuous trace option in MacClade and are as follows ($\mu_D$, SMR): 1, 0.392, 0.308; 2, 0.391, 0.284; 3, 0.412, 0.244; 4, 0.415, 0.276; 5, 0.422, 0.368; 6, 0.412, 0.230; 7, 0.418, 0.291; 8, 0.421, 0.223; 9, 0.422, 0.420; 10, 0.419, 0.387; 11, 0.410, 1.072; 12, 0.437, 0.745; 13, 0.430, 0.582. B, Bivariate plot of $\Delta$ SMR and $\Delta$ mutation pressure for lineages of mammals. Kendall's nonparameter correlation test: $N = 24$, tau = −0.292, $P = 0.0456$. Regression: $r^2 = 0.558$, $P = 0.0001$, $\mu_D = 0.077 \text{SMR} - 0.002$. Rectangular nucleotide substitution overall (estimates of directional nucleotide substitution at third codon positions are less than 0.50; Jermiin et al. 1994); thus, figures 4 and 5 provide an assessment of change in directional nucleotide substitution along a lineage relative to another in the context of changes in metabolism.

Mammals and primates exhibit qualitatively similar relationships between change in directional nucleotide substitution and change in SMR; nevertheless, the slope of the relationship is significantly greater for primates than for mammals (fig. 6). It is also interesting that differences in directional nucleotide substitution are larger in primates than in mammals, whereas differences in metabolic physiology are greater among mammals (i.e., between mouse and whale) than among primates.
increased rates of DNA damage. Over time this event may become obscured as the accumulation mismatch mutations equalizes nucleotide composition, and the effect of accelerated evolution of metabolism is gradually erased from the genome.

Discussion

There is a negative association between increase in specific metabolic rate and the abundance of G+C nucleotides at silent sites in mitochondrial DNA. It is important to keep in mind the indirect association between SMR and nucleotide composition (i.e., SMR \(\rightarrow\) free radicals \(\rightarrow\) DNA damage \(\rightarrow\) mutation \(\rightarrow\) DNA substitution \(\rightarrow\) %GC); thus, factors that impinge on each of the steps in the cascade can modify nucleotide composition. Nevertheless, the results obtained are consistent with the predictions of biochemical studies of mutation and metabolism and suggest that the rate at which organisms live (measured by oxygen consumption) can influence rates and patterns of mutation.

Significant correlation between directional nucleotide substitution and rates of DNA damage (estimated by SMR) provides additional evidence that the metabolic physiology of animals may play an important role in mtDNA evolution (Martin and Palumbi, 1993). Predictions of the “physiological clock” hypothesis can be tested by examining rates and patterns of DNA evolution for genomes subject to different rates of DNA damage. For instance, the nucleus is anoxic relative to mitochondria and chloroplasts. In animals cells, greater than 90% of the oxygen is consumed by the mitochondria (Richter et al., 1988). A difference in the flux of oxygen radicals between organelles and the nucleus should lead to a difference in nucleotide composition. In plants, AT content of genes in close proximity to sites of oxygen radical production via photosynthesis (e.g., cyanelle and

These results underscore the fact that SMR is not the only causative factor driving DNA damage. Flux of oxygen radicals is a product of specific metabolic rate + the efficiency of oxygen use + the ability of cells to detoxify oxygen radicals + the efficiency and activity of DNA repair. Each of these components of endogenous DNA damage are themselves under the control of multiple genes. Differences in time for evolutionary change between groups may also partially explain the difference in slope. Rapid evolution of SMR may be accompanied by a rapid shift in nucleotide composition as a result of

Fig. 5.—A. Phylogenetic hypothesis for lineages of primates and estimates of mutation pressure (\(\mu_D\)), proportional G+C nucleotides at third codon positions (from the COII gene; Adkins and Honeycutt, 1994) and SMR (in milliliters oxygen consumed per gram per hour). Values at the nodes were determined using the continuous trace option in MacClade and are as follows (\(\mu_D\), G+C, SMR): 1, 0.447, 0.443, 0.260; 2, 0.439, 0.439, 0.300; 3, 0.390, 0.414, 0.379; 4, 0.363, 0.410, 0.409; 5, 0.368, 0.407, 0.511; 6, 0.365, 0.482, 0.403; 7, 0.350, 0.394, 0.446; 8, 0.323, 0.380, 0.446; 9, 0.333, 0.385, 0.468; 10, 0.302, 0.369, 0.483; 11, 0.288, 0.361, 0.579; 12, 0.168, 0.300, 0.992; 13, 0.228, 0.330, 0.785. B. Bivariate plot of \(\Delta\) SMR and \(\Delta\) mutation pressure for lineages of primates. Kendall’s nonparametric correlation test: N = 24, tau = -0.448, P = 0.0022. Linear regression: \(r^2 = 0.44, P = 0.0004, \mu_D = -0.376\)SMR + 0.002. Similar results were obtained when G+C was used as the dependent variable.

Fig. 6.—Bivariate plot of \(\Delta\) SMR and \(\Delta\) mutation pressure for lineages of mammals and primates (from figs. 4 and 5) showing a difference in slopes between the two groups.
chloroplast genes) and oxidative respiration (e.g., mitochondrial genes) are greater than the AT content of eubacterial and plant nuclear genes (Lockhart et al. 1992). The similarity of AT content of cyanelle and chloroplasts genes is thought to reflect similar directional nucleotide substitution. In addition, there is evidence that chloroplast genes of green plants that transferred to the nucleus a long time ago exhibit less A+T bias than recently transferred genes (Campbell and Gowri 1990; Lockhart et al. 1992). These results suggest that there is directional A+T nucleotide substitution on genes in oxygenic environments (cyanelle, chloroplasts) relative to nuclear-encoded genes in plants.

There are a number of cases of endosymbiosis in which the endosymbiont can potentially alter DNA damage rates of host genomes. The observation that rates of oxygen production by endosymbiotic algae control the level of superoxide dismutase activity of their host tissue (Dyken and Shick 1982; Lesser and Shick 1989) suggests that rates of DNA damage in cnidarians and other hosts of endosymbiotic photoautotrophs may be set by the endosymbionts. If this is true, then we should expect cnidarians which do not harbor algae to have slower rates of DNA substitution and higher proportions of G+C nucleotides than those species which do harbor algae. A recent report of the phylogenetic relationships of deep-water corals revealed that the branch length for a shallow-water coral with photosynthetic endosymbionts is roughly five-fold longer than in a deep-water sister species (Agenbroad et al. 1995). In addition, percent G+C at third codon position in the anemone Metridium senile is 32.8, roughly similar to the composition of high-metabolic rate taxa like mouse and nematodes (Pont-Kingdon et al. 1994). The low G+C content of M. senile may be the result of oxidative damage from photoautotrophic endosymbionts. There are other examples of shifts in nucleotide composition associated with endosymbiosis. The genomes of many prokaryotic parasites such as mycoplasmas, rickettsiae, and chlamydiae have high A+T contents (Winkler and Wood 1988). Ohtaka and Ishikawa (1993) discovered that an intracellular symbiont of aphids has been subjected to AT-biased mutation pressure.

Of course, oxidative damage resulting from by-products of metabolism is only one source of mutation. Ultraviolet and other types of radiation can be powerful mutagens. Taxa vary in the degree that germ line cell lineages are protected against the damaging effects of UV radiation. The oocytes of mammals, for instance, are shielded from radiation behind multiple tissue layers. Recent studies of humans indicate that endogenous processes account for most germ line mutations and that environmental (exogenous) mutagens play a very minor role (Sommer 1995). By contrast, many fish and amphibians lay clear, gelatinous eggs that are exposed to toxic UV radiation. Among these groups, there is considerable variation in the actual exposure of eggs to UV light. Some amphibians, for instance, bury their eggs in soil. Among species which lay eggs exposed to UV light, there are significant differences in levels of repair enzyme activities that afford protection against UV damage (Blaustein et al. 1994). Studies of UV-induced mutagenesis indicate that many mutants result from C to T transitions at dipyrimidine sites (Thomas and Kunzel 1993). Irradiation also is known to increase rates of DNA lesions and oxidative damage (Schulte-Frohlinde and von Sonntag 1985). The higher percentage AT in the mitochondrial genome of the frog Xenopus than predicted from consideration of its SMR may reflect relatively high rates of DNA damage from UV irradiation. Differences in physical and biochemical insulin of germ line cells from environmental mutagens may, in part, explain differences in DNA substitution rate and nucleotide composition among taxa.

It is clear that acceleration in rate of oxygen consumption is associated with an increase in A+T nucleotides (figs. 4–6) and in overall rates of nucleotide substitution (Martin and Palumbi 1993). Association between cellular metabolism and mitochondrial DNA evolution is probably mediated by DNA damage from oxygen radicals. Increasing evidence for a link between physiology and mtDNA evolution, the observation that substitution matrices may be very asymmetric (Peron and Kocher 1995a), and an indication that some substitution types may be forbidden (i.e., GC to CG transversions) have important implications in molecular systematics. First, current models of DNA evolution for all methods of phylogenetic inference assume compositional stationarity (i.e., a constant mutation matrix). It is difficult to predict what sorts of error will result when trying to infer topology and branch lengths for a group of taxa in which mutation matrices change across the tree as a result of evolution in characters linked to aerobic metabolism. Second, marked asymmetry of substitution matrices, a situation likely to be particularly acute for taxa with high SMR, requires adopting more complex models of DNA evolution to accurately estimate evolutionary history. Finally, inferences of mutation matrices for mtDNA are characterized by low numbers of GC to CG transversions. Many authors have taken this as evidence that these changes occur, are rare, and can therefore be weighted heavily in order to choose among alternative topologies (see, e.g., Knight and Mindell 1993). However, in light of the fact that GC to CG transversions rarely, if ever, occur, it is more likely that inferred GC to CG changes are the result of multiple
substitutions at a site (one example is a G to A to C change in one lineage, and no change in a second).

Recognition that molecular evolution may be highly variable within and among taxonomic groups—at least for mtDNA—should not be construed as evidence against the ability to infer phylogenies and estimate rates of evolution. Instead, if we can quantify lineage effects (e.g., factors that drive DNA damage rates), it should be possible to more accurately model the process of base substitution. If this can be successfully accomplished, it will be possible to attach a more realistic assessment of confidence to estimates of evolutionary history.

Acknowledgments

Thanks to T. Kocher for sending copies of manuscripts in press. Two anonymous reviewers provided constructive criticism.

LITERATURE CITED


NAOYUKI TAKAHTA, reviewing editor

Received March 27, 1995

Accepted June 28, 1995
18S rRNA Data Indicate That Aschelminthes Are Polyphyletic in Origin and Consist of at Least Three Distinct Clades

Birgitta Winnepenningckx, * Thierry Backeljau, † Laura Y. Mackey, ‡ Jacqueline M. Brooks, ‡ Rupert De Wachter, * Sudhir Kumar, § and James R. Garey ‡

*Departement Biochemie, Universiteit Antwerpen (UIA); †Koninklijk Belgisch Instituut voor Natuurwetenschappen; ‡Department of Biological Sciences, Duquesne University; §Institute of Molecular Evolutionary Genetics, Pennsylvania State University

The Aschelminthes is a collection of at least eight animal phyla, historically grouped together because the absence of a true body cavity was perceived as a pseudocoelom. Analyses of 18S rRNA sequences from six Aschelminth phyla (including four previously unpublished sequences) support polyphyly for the Aschelminthes. At least three distinct groups of Aschelminthes were detected: the Priapulida among the protostomes, the Rotifera-Acanthocephala as a sister group to the protostomes, and the Nematoda as a basal group to the triploblastic Eumetazoa.

Introduction

High-level phylogenetic relationships among animals have been based upon several characters, including the number of embryonic tissue layers, early embryonic cleavage patterns, larval morphology, body symmetry, and the type of body cavity present. Three body cavity conditions are commonly recognized among the triploblastic animals: the absence of a body cavity (acoelomate), the presence of a false body cavity that appears as a persistent blastocoel (pseudocoelom), and the true body cavity that arises from within the mesoderm (eucelom). The bilateral ancestor to the triploblastic Metazoa is often presented as a "flatworm-like" acoelomate from which the pseudocoelomate and eucelomate animals arose following the development of the pseudocoelom and eucelom. The body cavity presumably allowed better internal organization that enabled animals to become larger and more active (reviewed in Brusca and Brusca 1990). The eucelomates led to the "mainstream" line of evolution that split into the protostomes and the deuterostomes. Most of the pseudocoelomate animals have been relegated to a diverse group of at least eight phyla known as the Aschelminthes (reviewed in Hyman 1951; Marcus 1958; Clark 1979; Brusca and Brusca 1990).

Ultrastructural studies have suggested that the boundaries between the three body cavity types are not well defined. These studies suggest that the body cavity conditions perceived as pseudocoelomate and acoelomate may be the result of a reduced or modified eucelom (Rieger 1985; Ruppert 1991a; Rieger et al. 1991). For example, the nemerteans are traditionally allied with the acoelomate Platyhelminthes, but molecular, morphological, and embryonic studies consistently suggest that they are actually protostome coelomates (Turberville et al. 1992). Recently, the free-living marine nematode Anoplostoma vivipara was described as lacking a pseudocoel (Ehlers 1994), and the author concluded that the acoelomate condition is ancestral to the nematodes. However, if acoelomates are the ancestors of modern triploblastic Metazoa, one would expect that a novel structure such as the pseudocoelom would most likely evolve only once. If so, a putative pseudocoelomate ancestor would have evolved into the modern aschelminth phyla, and the Aschelminthes would be a valid monophyletic taxon. Conversely, if eucelomates were the ancestors of modern triploblastic Metazoa, one would expect that the pseudocoelom could have evolved any number of times by modification or partial loss of the eucelom. This could have occurred under evolutionary pressures that favored small animals in which the eucelom was not an advantage (Ruppert 1991a), in which case the aschelminth phyla would not necessarily be related to one another, and the Aschelminthes could not be considered a valid taxon.

The evolutionary relationships between animal phyla have been examined many times by a number of authors using a variety of morphological characters (for recent examples, see Brusca and Brusca 1990; Schram 1991; Eernisse et al. 1992; Backeljau et al. 1993; Nielsen 1995) with varying results concerning the polyphyly of the Aschelminthes. However, the most explicit cladistic
analysis that focused specifically on aschelminth evolutionary relationships is that of Lorenzen (1985). He considered the phyla Rotifera, Acanthocephala, Nematoda, Nematomorpha, Gastrotricha, Kinorhyncha, and Priapulida and concluded that the pseudocoelomates are polyphyletic, forming a number of distinct clades (Rotifera-Acanthocephala, Nematoda-Nematomorpha-Gastrotricha, and Priapulida-Kinorhyncha).

Molecular data can complement morphological data and are useful when comparing distantly related organisms that have few common characters. The small ribosomal subunit RNA gene (18S rRNA) has proven useful with distantly related organisms because it is reasonably large (about 1,800 nucleotides in length), highly conserved, and data are available from a large number of organisms (Hillis and Dixon 1991; Larsen et al. 1993; Neefs et al. 1993). The 18S rRNA gene has been used to investigate the origin of the animal kingdom (Field et al. 1988; Wainright et al. 1993) and several other problems in animal phylogeny (see, e.g., Nadler 1992; Turbeville et al. 1992; Telford and Holland 1993; Halanych et al. 1995).

We have chosen to test the hypothesis that the Aschelminthes is made up of several clades and to determine the location of these clades among the Eumetazoa by analysis of the 18S rRNA gene. This analysis includes new unpublished sequences from one representative of the phyla Gastrotricha, Nematomorpha, Priapulida, and Rotifera (Monogononta) as well as previously published sequences of representatives of the phyla Acanthocephala and Nematoda (Ellis et al. 1986; Telford and Holland 1993; Fitch et al. 1995) that have historically been included in the Aschelminthes.

Material and Methods

Rotifer (Brachionus plicatilis) cysts were obtained commercially (Aquaculture Supply; Dade City, Fla.) and cultured in sea water (Snell et al. 1987). Rotifers were separated from feeder algae and concentrated by sieving through plankton netting and starved for several hours to clear algae from the gut. Gastrotrichs (Lepidodermella squammatata) were purchased from a commercial supplier (Carolina Biological Supply Company; Burlington, N.C.), and several hundred individuals were isolated from the culture with a mouth micropipette and starved for several hours. DNA was prepared (Hempstead et al. 1990) and 0.1 µg used as template for polymerase chain reaction (PCR) amplification. Four primers were used that resulted in two overlapping fragments representing a nearly complete fragment of the 18S rRNA gene corresponding to nucleotides 130-1,965 of the human sequence (GenBank accession M10098). PCR amplification was carried out for 35 cycles with 30 s at 94°C denaturing, 90 s at 55°C annealing, and 120 s at 72°C extension (Ausbuhl et al. 1995). The primers contained restriction endonuclease site containing tails for cloning purposes (Garey et al. 1992). The PCR products were cloned into M13 mp18 nondirectionally using an appropriate restriction enzyme. M13 clones containing inserts in opposite orientation were identified by complement testing. DNA sequencing of single cloned fragments was carried out completely in both directions from the M13 templates with the chain termination method using Sequenase (US Biochemical; Cleveland, Ohio), commercial M13 primers, and conserved internal primers. Additional sequencing reactions were carried out using inosine mixes as needed to resolve some sequencing artifacts. Nematomorphs (Gordius aquaticus) were collected in the Pyrenees (France). A priapulid (Priapulus caudatus) was found in the coastal waters of Kristineberg (Sweden). DNA was extracted (Winne-penninkx et al. 1993) from a single nematomorph specimen and from the skin tissue of the priapulid. The 18S rRNA genes were PCR amplified in two overlapping fragments using two primers complementary to the 5' and 3' ends of the 18S rRNA gene and two primers complementary to a conserved part of the 18S rRNA gene and the 5' end of the 28S rRNA gene. PCR amplification was carried out on 10 ng DNA template for 30 cycles of 60 s at 94°C, 60 s at 55°C, and 120 s at 72°C. PCR fragments were ligated into T-tailed PSK+ vector (BioRad; Richmond, Calif.), and DNA from a pool of 10 clones was sequenced using a variety of primers (Winne-penninkx et al. 1994). Recently, an unpublished sequence of the 18S rRNA gene of Priapulus caudatus, appeared in GenBank (GenBank accession number Z38009) and differs from the sequence reported here at 15 different nucleotide sites. It did not affect the topology of our trees when substituted for the priapulid sequence reported here.

Aschelminth sequences were aligned with those of other animals and yeast (see below) according to a secondary structure model (Neefs et al. 1993). Sites containing gaps were excluded from phylogenetic analyses to reduce systematic errors. Alignments were analyzed with the MEGA program (Kumar et al. 1994) to produce neighbor-joining (NJ) trees using the Kimura two-parameter model in which substitution rates follow a gamma distribution with shape parameter a = 0.72 to correct for multiple substitutions at the same site (Jin and Nei 1990). The gamma parameter of 0.72 was estimated from the distribution of the number of nucleotide substitutions across different sites as obtained in a parsimony analysis of 202 diverse eukaryotic 18S rRNA sequences (S. Kumar and A. Rzhetsky, unpublished data). Confidence in NJ trees was determined by analyzing 1,000 bootstrap replicates using the MEGA pro-