Substitution Bias, Rapid Saturation, and the Use of mtDNA for Nematode Systematics

Michael S. Blouin, Charles A. Yowell,† Charles H. Courtney,† and John B. Dame†*

*Department of Zoology, Oregon State University; and †Department of Pathobiology, University of Florida

Only relatively recently have researchers turned to molecular methods for nematode phylogeny reconstruction. Thus, we lack the extensive literature on evolutionary patterns and phylogenetic usefulness of different DNA regions for nematodes that exists for other taxa. Here, we examine the usefulness of mtDNA for nematode phylogeny reconstruction and provide data that can be used for a priori character weighting or for parameter specification in models of sequence evolution. We estimated the substitution pattern for the mitochondrial ND4 gene from intraspecific comparisons in four species of parasitic nematodes from the family Trichostrongylidae (38–50 sequences per species). The resulting pattern suggests a strong mutational bias toward A and T, and a lower transition/transversion ratio than is typically observed in other taxa. We also present information on the relative rates of substitution at first, second, and third codon positions and on relative rates of saturation of different types of substitutions in comparisons ranging from intraspecific to interordinal. Silent sites saturate extremely quickly, presumably owing to the substitution bias and, perhaps, to an accelerated mutation rate. Results emphasize the importance of using only the most closely related sequences in order to infer patterns of substitution accurately for nematodes or for other taxa having strongly composition-biased DNA. ND4 also shows high amino acid polymorphism at both the intra- and interspecific levels, and in higher level comparisons, there is evidence of saturation at variable amino acid sites. In general, we recommend using mtDNA coding genes only for phylogenetics of relatively closely related nematode species and, even then, using only nonsynonymous substitutions and the more conserved mitochondrial genes (e.g., cytochrome oxidases). On the other hand, the high substitution rate in genes such as ND4 should make them excellent for population genetics studies, identifying cryptic species, and resolving relationships among closely related congeners when other markers show insufficient variation.

Introduction

Nematodes are very conservative in gross morphology, and even higher-level taxonomy in the phylum has been confounded by a lack of easily scored characters (e.g., Maggenti 1981; Blaxter et al. 1998). As researchers turn to molecular methods, mitochondrial DNA is being applied more frequently to questions of systematics and population genetics in nematodes (e.g., Blouin et al. 1992, 1995; Anderson, Romero, and Jaenike 1993; Powers, Harris, and Hyman 1993; Hugall et al. 1994; Joyce, Burnell, and Powers 1994; Anderson, Romero-Abal, and Jaenike 1995; Nadler 1995; Hyman and Whipple 1996; Anderson and Jaenike 1997; Hugall, Stanton, and Moritz 1997). However, we still know much less about the evolution of mtDNA in nematodes than we do about that in other taxa such as insects or vertebrates (Hyman and Slater 1990; Thomas and Wilson 1991; Okimoto et al. 1992; Okimoto, MacFarlane, and Wolstenholme 1994; Hyman and Azevedo 1996; Hugall, Stanton, and Moritz 1997). Understanding basic parameters such as patterns of nucleotide substitution and rate variation among sites is important for proper application of DNA sequence data to molecular systematics (Yang 1994; Lockhart et al. 1994; Jermiin, Graur, and Crozier 1995; Yang and Kumar 1996). In addition, the evolution of nematode mtDNA is itself interesting, because several lines of evidence suggest that it is unusual. For example, nematode mtDNA is highly A+T-rich (Thomas and Wilson 1991; Okimoto et al. 1992; Hyman and Azevedo 1996; Hugall, Stanton, and Moritz 1997), rivaled only by that of some insects (e.g., Clary and Wolstenholme 1985; Simon et al. 1994). It may also be prone to recombination and gene rearrangement (Hyman and Azevedo 1996; Lunt and Hyman 1997), and circumstantial evidence suggests that it evolves more quickly than mtDNA of other taxa (Blouin et al. 1995; Anderson, Blouin, and Beech 1998).

Here, we describe patterns of base substitution within and among species, genera, and orders of nematodes for the 3′ end of the mitochondrial ND4 gene. The specific goals of this study are (1) to describe the pattern of substitution, (2) to describe the relative rates at which different types of substitution saturate, and (3) to assess the appropriateness of the ND4 region for phylogenetic analyses at different taxonomic levels in the phylum nematoda. Thus, we provide data that others can use to decide whether mtDNA is appropriate for their research question and to choose values for a priori character weighting or for parameter specification in models of sequence evolution.

Materials and Methods

Sequences Used

In a study designed to compare patterns of gene flow in four species of parasitic nematodes in the family Trichostrongylidae, Blouin et al. (1995) sequenced 463 bp of the 3′ ends of the mitochondrial ND4 genes of individuals from each of four species: *Haemonchus placei* (40 individuals), *Haemonchus contortus* (40 individuals), *Teladorsagia circumcincta* (40 individuals), and

Key words: mitochondrial DNA, mtDNA, nematode, Trichostrongylidae, ND4.

Address for correspondence and reprints: Michael S. Blouin, Depatment of Zoology, Oregon State University, Corvallis, Oregon 97331-2914. E-mail: blouinm@bcc.orst.edu.

Mol. Biol. Evol. 15(12):1719–1727. 1998

Table 1
Taxonomic Relationships Among the Species Included in this Study

Subclass Spiruria
Order Ascaridida
Family Ascaridae
Ascaris suum
Subclass Rhabditia
Order Rhabditia
Family Rhabditidae
Caenorhabditis elegans
Order Strongylida
Family Trichostrongylidae
Mazamastrongylus odocoilei
Teladorsagia circumcincta
Ostertagia ostertagi
Haemonchus contortus
Haemonchus placei

Table 2
Number of Substitutions at Each Codon Position, Inferred by Tree Analysis for Each Species

Species	Codon Position		
	First	Second	Third
Haemonchus contortus	16	6	146
Haemonchus placei	11	4	68
Teladorsagia circumcincta	29	9	118
Mazamastrongylus odocoilei	30	7	134
Total	86	26	466
Ratio	3.3	1	17.9

Mazamastrongylus odocoilei (51 individuals). Here, we use 459 bp (153 complete codons, not including the stop codon) of those sequences and an additional 10 *H. contortus* sequences to describe substitution patterns in the mitochondrial ND4 gene. The 51 *Mazamastrongylus* sequences clustered into two very distinct groups of 38 and 12 sequences that may represent cryptic species (called groups A and B, respectively; see fig. 5 in Blouin et al. 1995). In order to be sure that we were using only intraspecific comparisons to infer the substitution matrix, we used only the 38 group A sequences. For comparisons above the intraspecific level, we included *Ostertagia ostertagi,* another trichostrongylid parasite of ruminants for which we have a single ND4 sequence (from the *O. ostertagi* mtDNA clone described in Tarrant et al. 1992), and the sequences from *Caenorhabditis elegans* (Rhabditidae) and *Ascaris suum* (Ascarididae) that were published in Okimoto et al. (1992). The phylogenetic relationships among the three families are well established by both morphological and molecular criteria (Blaxter et al. 1998), with the trichostrongylids and *Caenorhabditis* being more closely related than either is to *Ascaris* (table 1).

Estimating Substitution Patterns

In order to obtain the best estimate of the underlying pattern of substitution, we confined our analysis to fourfold degenerate sites and intraspecific comparisons. There were 47 fourfold degenerate sites in the *T. circumcincta* data set, 51 in the *M. odocoilei* data set, 44 in the *H. contortus* data set, and 44 in the *H. placei* data set. We inferred the pattern of substitution at these sites using Yang and Kumar's (1996) pairwise sequence comparison method as implemented in pamp (see also Yang 1995). This method produces a matrix Q which shows the expected instantaneous rate of substitution of i to nucleotide j per unit time.

The relative rates at which first, second, and third codon position sites accumulate substitutions were also estimated using only intraspecific comparisons. For each of the four species, a neighbor-joining tree was constructed from a matrix of Jukes-Cantor distances (in PHYLIP; Felsenstein 1993), and the number of substitutions at each site in the tree counted using MacClade (Madison and Madison 1992).

Relative Saturation Rates of Different Sites and Types of Substitution

In order to use coding sequences to accurately reconstruct phylogenies, it is important to adjust for the relative rates at which different codon sites and different types of substitution (e.g., transitions vs. transversions) saturate. We compared the relative rates at which transitions and transversions saturate across taxonomic categories by counting substitutions in all pairwise comparisons between sequences from the following categories: within species, between species within a genus, between genera within a family, and between orders. For intraspecific comparisons, we used the four species mentioned previously. We arbitrarily chose a single sequence from each of those four species to use in higher-level comparisons. *Haemonchus placei* versus *H. contortus* served as the comparison between species within a genus. For genera within a family, we compared the five trichostrongylid species. For interordinal comparisons, we compared *Caenorhabditis elegans, Ascaris suum,* and the trichostrongylids. Because we have no independent method of dating divergence times between nematode species, we cannot equate these taxonomic categories with true time. Therefore, for third positions, we plotted the numbers of transitions and transversions in each comparison against the number of substitutions accumulated at first plus second position sites. We also plotted first-, second-, and third-position substitutions and amino acid distance (percentage difference, hereinafter ''aa distance'') against 18S rDNA sequence distance (percentage difference over entire 18S rDNA subunit) for the species for which 18S sequences are available in GenBank (*O. ostertagi, H. contortus, H. placei, C. elegans,* and *A. suum*; note that *H. placei* and *H. contortus* have identical 18S sequences). In the appendix, we list, for each pairwise comparison, the number of each of the six types of substitution (C↔T, A↔G, and so on) and the aa distance calculated using the entire 459-bp data set.

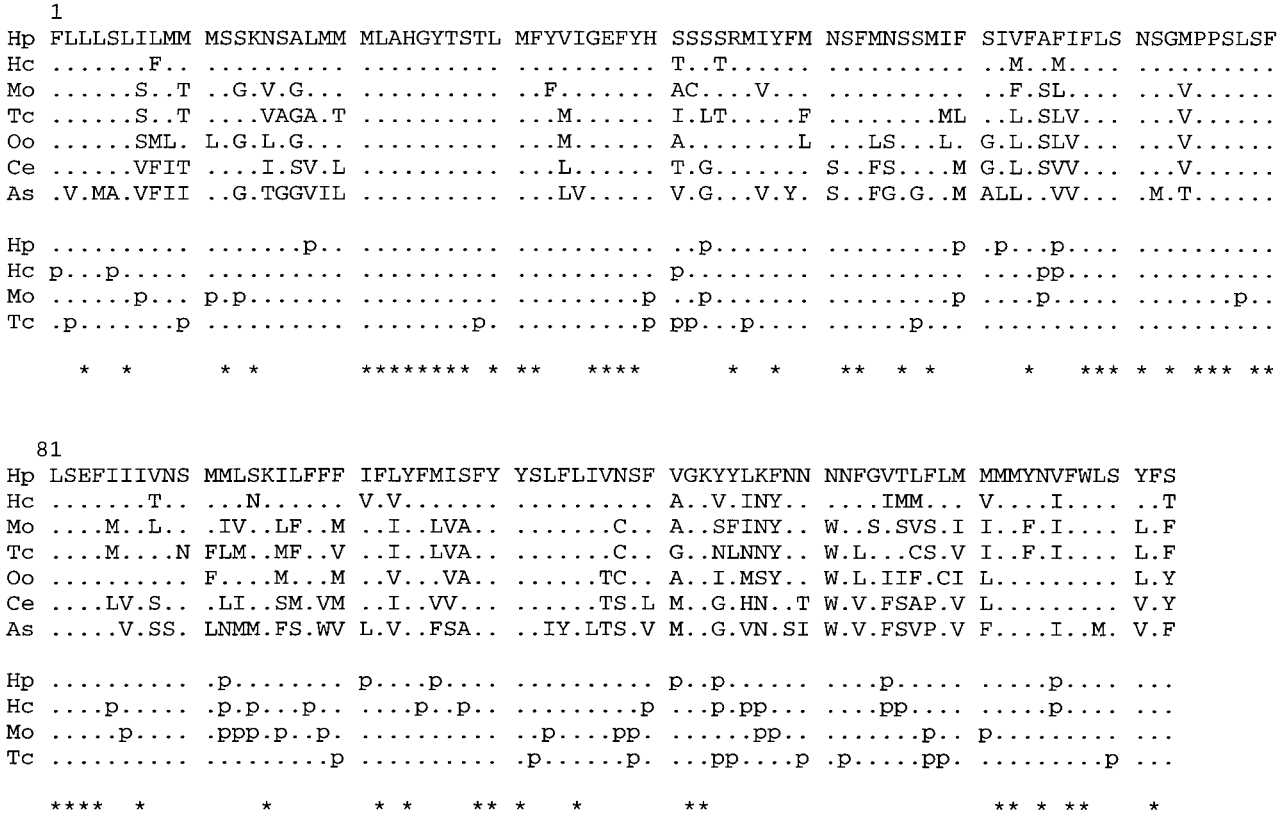

FIG. 1.—Amino acid sequences for the ND4 region sequenced in this study (the last 153 codons on the 3′ end of the gene, not including the stop codon). Designators to the left of each sequence are the initials of that species. For each of the four trichostrongylids for which we have multiple sequences, the amino acid sequence of an arbitrarily chosen individual is shown. For each of these species, sites that are polymorphic within the species are indicated by a ''p'' in the lines below the seven-species layout. Asterisks indicate amino acid sites that are invariant among all sequences for all of the species used in this study. Note the two conserved blocks in positions 21–32 and 75–84.

Results and Discussion

General

All ND4 sequences aligned in the same reading frame and share the *C. elegans*/*A. suum* gene arrangement in which the 3′ end of the ND4 gene is separated from the cytochrome oxidase I gene by an ~100-bp noncoding region of unknown function. Rates of substitution within trichostrongylid species appear to be high, with average pairwise sequence differences among conspecific individuals ranging from 1.9% to 2.6% (see also Blouin et al. 1995), and with 11%–19% of sites being variable within species. Most of the variable sites are in third codon positions, and, as is typical for protein coding regions, overall substitution rates are highest in third-position sites, and lowest in second-position sites (table 2). Overall levels of amino acid substitution are also high, with pairwise aa distances between conspecifics averaging 1% (appendix). The 3′ end of the fragment appears to be less conserved than the 5′ end, and there are two stretches of sequence, each approximately 10 amino acids long, whose compositions are almost invariant across all species and individuals in this study (fig. 1).

Nucleotide Composition and Substitution Bias

The trichostrongylids all have highly A+T-rich mtDNA (table 3), as is typical of other nematodes. It has been suggested, based on patterns of nucleotide and amino acid usage, that there is an A+T mutational bias in nematodes (Thomas and Wilson 1991; Okimoto et al. 1992; Hyman and Azevedo 1996; Hugall, Stanton, and Moritz 1997). Substitution matrices generated for the trichostrongylids here show that relative rates of substitution from a C or G to an A or T are always higher than rates of substitution in the opposite direction (e.g., the rate of C→T is higher than the rate of T→C, and so on, for all types of substitution in all four species; table 4). These asymmetric rates suggest that the high A+T composition is indeed caused by an underlying mutational bias (Sueoka 1992). Mutational pressure is also consistent with the fact that A+T composition is lower in first- and second-position sites than in third-position sites (e.g., 72% A+T in second positions vs. 87% in third positions; see table 3; Jukes and Bushan 1986), and with the low transition/transversion (ts/tv) ratios for these species (Jermiin and Crozier 1994). Here, the ratio ranges from 6:1 to 10:1 among the four species (table 4). For comparison, typical ratios for silent DNA in vertebrates are at least 15:1 (Aquadro and Greenberg 1983; Tamura and Nei 1993; Yang and Kumar 1996). Note also that an excess of A↔T substitutions alone does not explain the low ts/tv ratio. Although A↔T is by far the most common transversion (appendix), when substitu-

Table 3
Nucleotide Composition of the ND4 Gene of each Species, Broken Down by Codon Position

Species	First Position				Second Position				Third Position				All Positions			
	A	T	C	G	A	T	C	G	A	T	C	G	A	T	C	G
Haemonchus contortus	44.3	40.5	4.0	11.2	20.3	53.5	15.7	10.5	34.5	51.8	5.7	8.0	33.0	48.6	8.5	9.9
Haemonchus placei	39.2	44.5	5.2	11.2	19.6	53.7	15.6	11.1	33.2	52.8	4.2	9.9	30.7	50.3	8.3	10.7
Teladorsagia circumcincta	36.8	43.1	5.7	14.4	17.7	53.5	14.4	14.3	36.8	46.9	4.7	11.6	30.4	47.8	8.3	13.4
Mazamastrongylus odocoilei	36.5	44.6	5.6	13.3	15.1	54.8	15.7	14.4	32.6	59.1	3.7	4.6	28.0	52.8	8.3	10.8
Average	39.2	43.2	5.1	12.5	18.2	53.9	15.4	12.6	34.4	52.7	4.6	8.5	30.5	49.9	8.4	11.2
Percentage of A+T		82.4				72.1				87.1				80.4		

Note.—A+T content increases going from second to first to third position.

tions are expressed as an absolute rate, A→T and T→A are not the most rapidly occurring transversions (table 4).

Saturation Rates of Different Types of Substitutions

A striking feature of this data set is the rapid rate at which transitions, and then A↔T transversions, saturate as more distantly related sequences are compared (appendix). For example, in intraspecific comparisons, each transition is five to six times more common than the most abundant transversion, A↔T. But even by the time one compares species within a genus, the number of A↔T's has outpaced either transition (and note that, by nematode standards, *H. placei* and *H. contortus* are closely related, being difficult to tell apart morphologically and able to produce hybrids; LeJambre 1979). Thomas and Wilson (1991) also noted a very rapid drop in the ts/tv ratio as one goes from intra- to interspecific comparisons for *Caenorhabditis.* These results illustrate the importance of using only the most closely related sequences in order to infer patterns of substitution accurately for nematodes or for other taxa having strongly composition-biased DNA. The rapid saturation at silent sites is illustrated even more dramatically in plots of third-position substitutions against first- plus second-position substitutions and against 18S distance (figs. 2 and 3*C*). Third-position transitions appear to be saturated at the level of congeneric species, and transversions begin to plateau at the level of confamilial genera. Figure 3 suggests that even replacement substitutions in ND4 may be saturating at the higher taxonomic levels.

Use of mtDNA in Nematode Molecular Systematics

Figures 2 and 3 show that for comparisons much beyond the intraspecific level, third-position sites should probably be excluded. Not surprisingly, other authors have also found third-position sites in nematode mtDNA to be saturated and uninformative in higher-level phylogenetic reconstructions (among genera in the superfamily Ascaridoidea, using cytochrome oxidase (CO) II: Nadler 1995; among *Meloidogyne, Romanomermis, Ascaris,* and *Caenorhabditis* using ND3 and cytochrome *b*: Powers, Harris, and Hyman 1993; among trichostrongylid genera using COI: D. Zarlenga, personal communication). At least for less conserved mitochondrial genes, such as those in the NADH dehydrogenase complex (e.g., table 1 in Simon et al. 1994), one might be cautious about using even amino acid sequence for phylogeny reconstructions much beyond the level of confamilial genera (figs. 1 and 3). At this level, we recommend using the more slowly evolving mitochondrial genes, such as cytochrome oxidases, and using amino acid or only first and second position sites. For example, of the 34 fixed differences between *H. placei* and *H. contortus* at ND4, 10 result in an amino acid change (Blouin et al. 1997). In contrast, all 51 substitutions separating the two species at COI were in silent sites (one individual sequenced from each species; D. Zarlenga, personal communication).

Highly A+T-rich mtDNA poses several problems for phylogeny reconstruction (Simon et al. 1994). When taxa differ in composition bias, they can be grouped based more on shared nucleotide composition than on shared history (Hasegawa and Hashimoto 1993; Steel, Lockhart, and Penny 1993). There are also effectively fewer possible character states, so sites saturate more quickly, reducing resolution in the deeper parts of trees when distance or parsimony methods are used (e.g., Wolfe, Sharp, and Li 1989; Wolfe and Sharp 1993; Brower and DeSalle 1998). These two problems can be minimized for nematodes by using mtDNA for phylogenetics on only closely related species. Nevertheless, even for closely related species, it will be important to use phylogenetic methods that account for the rapid rate of saturation, the biased mode of substitution, and the rate variation among codon positions. For example, in a parsimony analysis, one might account for substitution bias by using the step matrix approach implemented in PAUP, in which the elements of the step matrix are based on the data in table 4. For highly composition-biased and rapidly saturating sequences such as these, however, maximum-likelihood approaches may be most appropriate. A variety of multiparameter models are available that adjust for substitution bias and rate variation to different extents, although one must consider the trade-off in increased variance of the estimate that occurs as the number of parameters that must be estimated from the data increases (Kumar, Tamura, and Nei 1993; Yang 1994; Yang and Kumar 1996). One simple approach for es-

Table 4
Substitution Patterns Inferred for Each Species (Q_{ij}*dt = prob[$i \rightarrow j$; dt])

	Haemonchus contortus				*Haemonchus placei*				*Teladorsagia circumcincta*				*Mazamastrongylus odocoilei*			
	T	C	A	G		T	C	A	G		T	C	A	G		T	C	A	G
T	−0.45	0.40	0.03	0.01	T	−0.47	0.37	0.10	0.00	T	−0.79	0.67	0.10	0.01	T	−0.48	0.43	0.04	0.00
C	2.09	−2.42	0.22	0.10	C	3.54	−3.92	0.40	0.00	C	3.33	−3.49	0.16	0.00	C	3.88	−4.08	0.09	0.10
A	0.03	0.04	−0.77	0.69	A	0.13	0.05	−0.90	0.70	A	0.09	0.03	−0.70	0.58	A	0.13	0.03	−0.95	0.78
G	0.05	0.08	3.02	−3.16	G	0.00	0.00	2.17	−2.15	G	0.02	0.00	1.02	−1.05	G	0.00	0.14	3.11	−3.25
Average ts/tv = 10.1					Average ts/tv = 6.4					Average ts/tv = 8.7					Average ts/tv = 10.2				

NOTE.—The Q matrix for each species was estimated using fourfold degenerate sites only. In these matrices, the diagonals are set to the negative of the substitution rate of nucleotide i so that rows sum to zero (slight deviations owing to roundoff errors). Note that rates of substitution from C+G to A+T are always greater than rates of substitution in the opposite direction.

FIG. 2.—Observed numbers of transitions and transversions at third-position sites plotted against total substitutions at first- plus second-position sites for pairwise comparisons within and between species. Small letters roughly denote taxonomic levels of comparisons: a = intraspecific; b = intrageneric (*H. placei* vs. *H. contortus*); c = intrafamilial (between trichostrongylid genera); d = interordinal (*C. elegans* vs. *A. suum* vs. the trichostrongylids). Note how rapidly the ts/tv ratio changes in going from intraspecific to interspecific comparisons.

timating distances between closely related sequences would be to use the maximum-likelihood distance in PHYLIP. This model at least accounts for a rate difference between transitions and transversions and for biased nucleotide composition and allows the user to input the expected ts/tv ratio and the expected rates of substitution at different sites. For example, from these data on trichostrongylids, one might choose relative rates of 3:1:18 for first-, second-, and third-position sites (table 2) and a ts/tv ratio of around 7 (table 4).

So what is the most appropriate use of mtDNA in nematode evolutionary genetics? While the high rate of substitution in genes such as ND4 make them excellent markers for population genetics (e.g., Blouin et al. 1995), this study shows that the biased substitution pattern and rapid rate of saturation make them undesirable markers for higher-level phylogenetics. Nevertheless, genes such as ND4 should be very useful in two types of applications, particularly if one corrects for multiple hits and the substitution bias. First, mtDNA should be excellent for determining relationships among closely related species when other markers show insufficient variation. For example, nuclear ribosomal RNA genes are very useful for resolving genus- and family-level relationships in nematodes, but often fail to adequately resolve relationships among closely related congeners (e.g., Fitch, Bugaj-Gaweda, and Emmons 1995; Liu, Berry, and Moldenke 1997). As a case in point, *H. placei* and *H. contortus* have identical 18s rDNA sequences

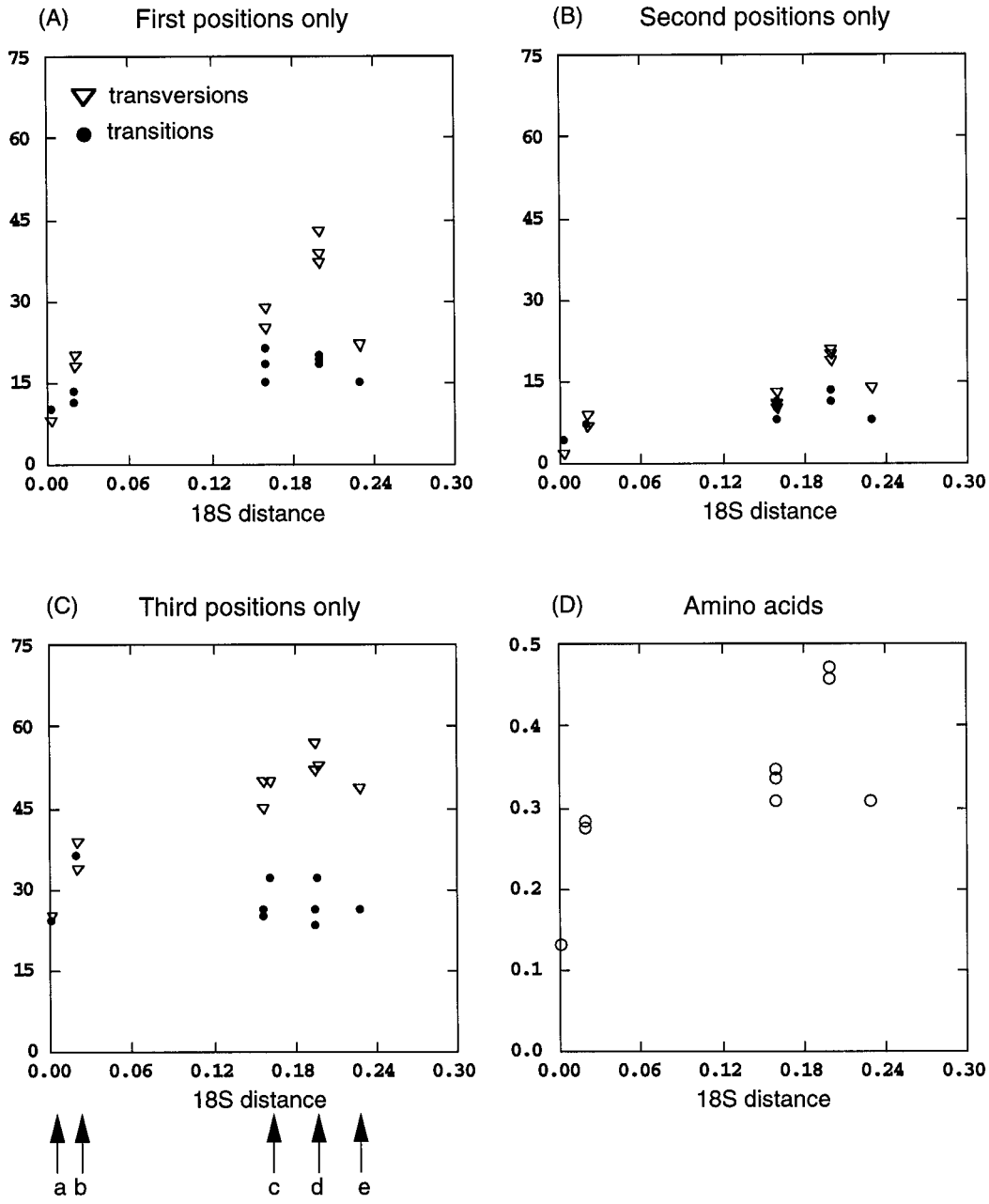

FIG. 3.—Observed numbers of transitions and transversions for each codon position class (*A–C*), and amino acid distance (*D*) for the ND4 fragment plotted against 18S distance for pairwise comparisons within and between species. The 18S distance is the percent sequence difference over the entire 18S rRNA subunit. a = *H. contortus* versus *H. placei*; b = *O. ostertagi* versus *Haemonchus* sp.; c = *O. ostertagi* and *Haemonchus* sp. versus *C. elegans*; d = *O. ostertagi* and *Haemonchus* sp. versus *A. suum*; e = *C. elegans* versus *A. suum.*

(Zarlenga et al. 1994) and differ by only 1% in the ribosomal ITS-2 sequence (Stevenson, Chilton, and Gasser 1995), but they are well differentiated by a 13%–16% difference in mtDNA coding genes (table 5). A second very useful application of nematode mtDNA may be to aid in decisions on whether isolates deserve species status. A surprising number of presumed monospecific parasite species have recently been shown to actually consist of several cryptic species (reviewed in Anderson, Blouin, and Beech 1998). Fast-evolving genes such as ND4 may quickly reach reciprocal monophyly between reproductively isolated populations (barring complications of the breeding system, which can cause the opposite effect via enhanced organelle effective size; Hoelzer 1997), and thus may be useful for identifying cryptic species. For example, for nematode species that were originally defined on the basis of morphology and for which mtDNA data are now available, we see that most interspecific differences are in the 10%–20% range (table 5), while the maximum pairwise difference observed within any species so far is 7% (table 6), and this value is between individual *Mazamas-*

Table 5
mtDNA Percent Sequence Difference Between Congeneric Nematode Species that are Recognized on the Basis of Morphology

mtDNA Region	Species Compared	Percent Sequence Difference[a]	Source of Data
ND4	*Haemonchus contortus* vs. *H. placei*	16	This study; Blouin et al. (1997)
COI	*H. contortus* vs. *H. placei*	13	D. Zarlenga (personal communication)
	Ostertagia ostertagi vs. *O. mossi*	12	D. Zarlenga (personal communication)
ND4	*Heterorhabditis bacteriophora, H. megidis, H. marelatus,* and *H. indicus*	8–19[b]	Unpublished data
COII	*Caenorhabditis elegans, C. briggsae,* and *C. remanei*	9.5 to 12	Thomas and Wilson (1991)
1.8-kb fragment (3′ end of COII to the 5′ end of Cyt b)	*Meloidogyne hapla* vs. *M. javanica*	20[c]	Hugall, Stanton, and Moritz (1997)

[a] Range is given for all pairwise comparisons among three or more species. Average difference between species is shown when more than one individual was sequenced per species.

[b] The distance between *H. bacteriophora* and *H. megidis* is 8%; other distances between *Heterorhabditis* species range from 13% to 19%.

[c] Estimated from figure 1 in Hugall, Stanton, and Moritz (1997).

trongylus from the A and B groups, which may be reproductively isolated. Thus, if sympatric populations form distinct clades that differ by more than, say, 10%, one might question whether they really are conspecifics. Obviously, an arbitrary genetic distance cutoff is not a sound way to define species, and mtDNA should be used only in conjunction with other evidence to make such decisions. For example, in a recent study on worldwide variation in *T. circumcincta*, samples from the Arctic were found to differ from all other populations by an average of 13% at ND4 (unpublished data). Host associations, biogeography, and subsequent morphological analyses all indicate that the Arctic samples represent a distinct species that had been overlooked until now. Thus, in our opinion, mtDNA's most appropriate place in the molecular tool bag for nematode molecular systematics should be for population genetics, for resolving relationships at the tips of trees whose deeper branches are resolved by other methods, and to aid in identifying species.

Acknowledgments

Thanks to Z. Yang for providing his pamp software and for advice on its use and interpretation. Thanks also to D. Zarlenga for permission to cite unpublished data, and to B. Hyman and D. Lunt for comments on an earlier draft. Sequences used in this analysis are available under GenBank accession numbers AF070736–AF070916.

Table 6
Maximum Percent Sequence Difference Observed Between Individuals of the Same Species

mtDNA Region	Species	Number of Individuals Sampled	Maximum Percent Sequence Difference	Source of Data
ND4	*Haemonchus contortus*	50	5.6	This study
	Haemonchus placei	40	4.1	This study
	Teladorsagia circumcincta	40	4.5	This study
	Mazamastrongylus odocoilei			
	Group A[a]	13	4.7	This study
	Group B	38	5.4	This study
	All 51 sequences		7.0	This study
	Heterorhabditis marelatus	40	2.5	Unpublished data
	Heterorhabditis megidis	4	3.3	Unpublished data
	Heterorhabditis bacteriophora	8	0.4	Unpublished data
COII	*Caenorhabditis elegans*	11 strains	1.9	Thomas and Wilson (1991)
	Caenorhabditis briggsae	2 strains	1.5	
16S rRNA and ND3	*Meloidogyne hapla*	4 strains	2.0	Hugall, Stanton, and Moritz (1997)
RFLP of entire mtDNA	*Ostertagia ostertagi*	50	6.0	Blouin et al. (1992)
RFLP of 55% of the entire mtDNA	*Ascaris suum*/*lumbricoides*[b]	50	5.8	Anderson and Jaenike (1997)

[a] *Mazamastrongylus odocoilei* haplotypes cluster into two distinct lineages (groups A and B) that are also geographically separated (Blouin et al. 1995). It is not known if these two groups represent cryptic species or just highly differentiated populations.

[b] mtDNA phylogeny does not completely support the species distinction between *A. suum* and *A. lumbricoides.* The maximum distance observed among all haplotypes is reported here.

APPENDIX

Counts of Each of the Six Types of Substitutions in Pairwise Comparisons at Different Taxonomic Levels (all nucleotide positions included)

Level of Comparison	Type of Substitution						Amino Acid Distance
	AG	TC	AT	AC	TG	CG	
Between families							
Trichostrongylids vs. *A. suum*							
Hp–As	34	21	62	5	40	12	0.47
Hc–As	40	19	60	8	32	10	0.46
Mo-As	36	25	51	4	34	11	0.46
Tc-As	40	21	55	7	36	8	0.46
Oo-As	32	29	56	8	40	8	0.46
Average	36.4	23.0	56.8	6.4	36.4	9.8	0.46
Trichostrongylids vs. *C. elegans*							
Hp-Ce	32	21	54	8	21	7	0.33
Hc-Ce	38	20	55	8	13	7	0.35
Mo-Ce	32	14	56	10	15	6	0.34
Tc-Ce	35	16	45	9	18	7	0.33
Oo-Ce	31	24	53	11	21	4	0.31
Average	33.6	19.0	52.6	9.2	17.6	6.2	0.33
A. suum vs. *C. elegans*							
Ce-As	31	18	49	5	26	5	0.31
Between genera within family							
Hp-Mo	21	19	48	3	18	4	0.27
Hp-Tc	27	21	50	2	20	5	0.31
Hp-Oo	34	20	38	7	14	4	0.27
Hc-Mo	24	21	45	2	13	4	0.27
Hc-Tc	29	22	51	4	18	5	0.33
Hc-Oo	38	18	40	6	14	4	0.28
Mo-Tc	20	22	35	2	13	4	0.21
Mo-Oo	24	29	36	3	20	3	0.27
Tc-Oo	28	27	40	2	12	6	0.27
Average	27.2	22.1	42.6	3.4	15.8	4.3	0.28
Between species within genus							
Hp-Hc	20	18	27	2	5	1	0.13
Within species (average of all pairwise comparisons)							
Hc	6.3	4.2	0.77	0.22	0.15	0.12	0.007
Hp	5.0	3.2	0.53	0.14	0.05	0.00	0.007
Tc	4.6	4.9	1.26	0.11	0.06	0.00	0.012
Mo	6.0	5.5	0.77	0.07	0.07	0.04	0.013
Average	5.5	4.5	0.83	0.14	0.08	0.04	0.010

NOTE.—Amino acid distance is the fraction of sites that are different. Values for the within-species comparisons are the averages for all pairwise comparisons within each listed species. As = *Ascaris suum*; Ce = *Caenorhabditis elegans*; Hc = *Haemonchus contortus*; Hp = *Haemonchus placei*; Mo = Mazamastrongylus odocoilei; Oo = *Ostertagia ostertagi*; Tc = *Teladorsagia circumcincta.*

LITERATURE CITED

ANDERSON, T. J. C., M. S. BLOUIN, and R. N. BEECH. 1998. Population biology of parasitic nematodes: applications of genetic markers. Adv. Parasitol. **41**:219–283.

ANDERSON, T. J. C., and J. JAENIKE. 1997. Host specificity, evolutionary relationships and macrogeographic differentiation among *Ascaris* populations from humans and pigs. Parasitology **115**:325–342.

ANDERSON, T. J. C., M. E. ROMERO, and J. JAENIKE. 1993. Genetic structure and epidemiology of *Ascaris* populations: patterns of host affiliation in Guatemala. Parasitology **107**: 319–334.

ANDERSON, T. J. C., M. E. ROMERO-ABAL, and J. JAENIKE. 1995. Mitochondrial DNA and *Ascaris* microepidemiology: the composition of parasite populations from individual hosts, families and villages. Parasitology **110**:221–229.

AQUADRO, C. F., and B. D. GREENBERG. 1983. Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. Genetics **103**:287–312.

BLAXTER, M. L., P. DE LEY, J. R. GAREY et al. (9 co-authors). 1998. A molecular evolutionary framework for the phylum nematoda. Nature **392**:71–75.

BLOUIN, M. S., J. B. DAME, C. A. TARRANT, and C. H. COURTNEY. 1992. Unusual population genetics of a parasitic nematode: mtDNA variation within and among populations. Evolution **46**:470–476.

BLOUIN, M. S., C. A. YOWELL, C. H. COURTNEY, and J. B. DAME. 1995. Host movement and the genetic structure of populations of parasitic nematodes. Genetics **141**:1007–1014.

BLOUIN, M. S., C. A. YOWELL, C. H. COURTNEY, and J. B. DAME. 1997. *Haemonchus placei* and *Haemonchus contortus* are distinct species based on mtDNA evidence. Int. J. Parasitol. **27**:1383–1387.

BROWER, A. V. Z., and R. DESALLE. 1998. Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies; the utility of *wingless* as a source of characters for phylogenetic inference. Insect Mol. Biol. **7**: 73–82.

CLARY, D. O., and D. R. WOLSTENHOLME. 1985. The mitochondrial DNA molecular of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. J. Mol. Evol. **22**:252–271.

FELSENSTEIN, J. 1993. PHYLIP (phylogeny inference package). Version 3.5. Distributed by the author, Department of Genetics, University of Washington, Seattle.

FITCH, D. H., B. BUGAJ-GAWEDA, and S. W. EMMONS. 1995. 18S ribosomal RNA gene phylogeny for some rhabditidae related to *Caenorhabditis.* Mol. Biol. Evol. **12**:346–358.

HASEGAWA, M., and T. HASHIMOTO. 1993. Ribosomal RNA trees misleading? Nature **361**:23.

HOELZER, G. A. 1997. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees revisited. Evolution **51**:622–626.

HUGALL, A., C. MORITZ, J. STANTON, and D. R. WOLSTENHOLME. 1994. Low but strongly structured mitochondrial DNA diversity in root knot nematodes (Meloidogyne). Genetics **136**:903–912.

HUGALL, A., J. STANTON, and C. MORITZ. 1997. Evolution of the AT-rich mitochondrial DNA of the root knot nematode, *Meloidogyne hapla.* Mol. Biol. Evol. **14**:40–48.

HYMAN, B. C., and J. L. B. AZEVEDO. 1996. Similar evolutionary patterning among repeated and single copy nematode mitochondrial genes. Mol. Biol. Evol. **13**:221–232.

HYMAN, B. C., and T. M. SLATER. 1990. Recent appearance and molecular characterization of mitochondrial DNA deletions within a defined nematode pedigree. Genetics **124**: 845–853.

HYMAN, B. C., and L. E. WHIPPLE. 1996. Application of mitochondrial DNA polymorphism to Meloidogyne molecular population biology. J. Nematol. **28**:268–276.

JERMIIN, L. S., and R. H. CROZIER. 1994. The cytochrome b region in the mitochondrial DNA of the ant *Tetraponera rufoniger*: sequence divergence in hymenoptera may be associated with nucleotide content. J. Mol. Evol. **38**:282–294.

JERMIIN, L. S., D. GRAUR, and R. H. CROZIER. 1995. Evidence from analyses of intergenic regions for strand-specific directional mutation pressure in metazoan mitochondrial DNA. Mol. Biol. Evol. **12**:558–563.

JOYCE, S. A., A. M. BURNELL, and T. O. POWERS. 1994. Characterization of *Heterorhabditis* isolates by PCR amplification of segments of mtDNA and rDNA genes. J. Nematol. **26**:260–270.

JUKES, T. H., and V. BHUSHAN. 1986. Silent nucleotide substitutions and G+C content of some mitochondrial and bacterial genes. J. Mol. Evol. **24**:39–44.

KUMAR, S., K. TAMURA, and M. NEI. 1993. MEGA: molecular evolutionary genetics analysis. Version 1.0. Pennsylvania State University, University Park.

LEJAMBRE, L. F. 1979. Hybridization studies of *Haemonchus contortus* (Rudolphi, 1803) and *H. placei* (Place, 1893) (Nematoda: Trichostrongylidae). Int. J. Parasitol. **9**:455–463.

LIU, J., R. E. BERRY, and A. F. MOLDENKE. 1997. Phylogenetic relationships of entomopathogenic nematodes (Heterorhabditidae and Steinernematidae) inferred from partial 18S rRNA gene sequences. J. Invertebr. Pathol. **69**:246–252.

LOCKHART, P. T., M. A. STEEL, M. D. HENDY, and D. PENNY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. **11**:605–612.

LUNT, D. H., and B. C. HYMAN. 1997. Animal mitochondrial DNA recombination. Nature **387**:247.

MADDISON, W. P., and D. R. MADDISON. 1992. MacClade Version 3. Sinauer, Sunderland, Mass.

MAGGENTI, A. 1981. General nematology. Springer, New York.

NADLER, S. A. 1995. Advantages and disadvantages of molecular phylogenetics: a case study of Ascaridoid nematodes. J. Nematol. **27**:423–432.

OKIMOTO, R., J. L. MACFARLANE, D. O. CLARY, and D. R. WOLSTENHOLME. 1992. The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum.* Genetics **130**:471–498.

OKIMOTO, R., J. L. MACFARLANE, and D. R. WOLSTENHOLME. 1994. The mitochondrial ribosomal RNA genes of the nematodes *Caenorhabditis elegans* and *Ascaris suum*: consensus secondary-structure models and conserved nucleotide sets for phylogenetic analysis. J. Mol. Evol. **39**:598–613.

POWERS, T. O., T. S. HARRIS, and B. C. HYMAN. 1993. Mitochondrial DNA sequence divergence among *Meloidogyne incognita, Romanomermis culicivorax, Ascaris suum* and *Caenorhabditis elegans.* J. Nematol. **25**:564–572.

SIMON, C., F. FRATI, A. BECKENBACH, B. CRESPI, H. LIU, and P. FLOOK. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. **87**:651–701.

STEEL, M. A., P. J. LOCKHART, and D. PENNY. 1993. Confidence in evolutionary trees from biological sequence data. Nature **364**:440–442.

STEVENSON, L. A., N. B. CHILTON, and R. B. GASSER. 1995. Differentiation of *Haemonchus placei* from *H. contortus* (Nematoda: Trichostrongylidae) by the ribosomal DNA second internal transcribed spacer. Int. J. Parasitol. **25**:483–488.

SUEOKA, N. 1992. Directional mutation pressure, selective constraints, and genetic equilibria. J. Mol. Evol. **34**:95–114.

TAMURA, K., and M. NEI. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. **10**:512–526.

TARRANT, C. A., M. S. BLOUIN, C. A. YOWELL, and J. B. DAME. 1992. Suitability of mitochondrial DNA for assaying interindividual genetic variation in small helminths. J. Parasitol. **78**:374–378.

THOMAS, W. K., and A. C. WILSON. 1991. Mode and tempo of molecular evolution in the nematode *Caenorhabditis*: cytochrome oxidase II and calmodulin sequences. Genetics **128**:269–279.

WOLFE, K. H., and P. M. SHARP. 1993. Mammalian gene evolution: nucleotide sequence divergence between mouse and rat. J. Mol. Evol. **37**:441–456.

WOLFE, K. H., P. M. SHARP, and W.-H. LI. 1989. Rates of synonymous substitution in plant nuclear genes. J. Mol. Evol. **29**:298–211.

YANG, Z. 1994. Estimating the pattern of nucleotide substitution. J. Mol. Evol. **39**:105–111.

———. 1995. Phylogenetic analysis by maximum likelihood (PAML). Version 1.1. Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park.

YANG, Z., and S. KUMAR. 1996. Approximate methods for estimating the pattern of nucleotide substitution and the variation of substitution rates among sites. Mol. Biol. Evol. **13**:650–659.

ZARLENGA, D. S., F. STRINGFELLOW, M. NOBARY, and J. R. LICHTENFELS. 1994. Cloning and characterization of ribosomal RNA genes from three species of *Haemonchus* (Nematoda: Trichostrongyloidea) and identification of PCR primers for rapid differentiation. Exp. Parasitol. **78**:28–36.

DAVID M. RAND, reviewing editor

Accepted September 4, 1998