Journal of Nematology 30(1):22–39. 1998.

A Phylogenetic Analysis of *Heterorhabditis* (Nemata: Rhabditidae) Based on Internal Transcribed Spacer 1 DNA Sequence Data[1]

B. J. ADAMS,[2] A. M. BURNELL,[3] AND T. O. POWERS[2]

Abstract: Internal transcribed spacer 1 sequences were used to infer phylogenetic relationships among 8 of the 9 described species and one putative species of the entomopathogenic nematode genus *Heterorhabditis*. Sequences were aligned and optimized based on pairwise genetic distance and parsimony criteria and subjected to a variety of sequence alignment parameters. Phylogenetic trees were constructed with maximum parsimony, cladistic, distance, and maximum likelihood algorithms. Our results gave strong support for four pairs of sister species, while relationships between these pairs also were resolved but less well supported. The ITS1 region of the nuclear ribosomal repeat was a reliable source of homologous characters for resolving relationships between closely related taxa but provided more tenuous resolution among more divergent lineages. A high degree of sequence identity and lack of autapomorphic characters suggest that sister species pairs within three distinct lineages may be mutually conspecific. Application of these molecular data and current morphological knowledge to the delimitation of species is hindered by an incomplete understanding of their variability in natural populations.

Key words: entomopathogenic nematode, evolution, *Heterorhabditis*, ITS1, nematode, phylogenetic analysis, ribosomal DNA, species concepts.

The insect-parasitic nematode genus *Heterorhabditis* Poinar (Heterorhabditidae) (Poinar, 1975) and its bacterial symbiont *Photorhabdus* (Enterobacteriaceae) (Boemare et al., 1993) have been shown to be effective agents in the biological control of many insect pests (Smart, 1995). This has led to numerous studies of their biology, ecology, biogeographic distribution, identification, and characterization; yet, the genus *Heterorhabditis* has not undergone a thorough systematic treatment. A phylogenetic framework is a necessary component of the comparative method in evolutionary biology and provides a critical ingredient for studies of gene flow, population structure, biogeography, coevolution, coadaptation, cospeciation, and historical ecology. Systematic information is also critical to rational implementation and monitoring agendas when these nematodes are used as biological control agents.

Cross-breeding, morphometrics, and molecular characters have been used to diagnose phenetic or biological species of heterohabditids (e.g. Akhurst, 1987; Curran and Webster, 1989; Dix et al., 1992; Gardner et al., 1994; Griffin et al., 1994; Joyce et al., 1994a,b; Liu and Berry, 1996; Nasmith et al., 1996; Nguyen and Smart, 1996; Stock and Kaya, 1996; Stock et al., 1996). An evolutionary or phylogenetic species concept has not been considered. Drawbacks to these studies are missing species and isolates, taxonomic methods of varying resolution, and suboptimal analytic methods such as using overall similarity as the basis for grouping taxa. Although the stated intent of these earlier analyses does not include the recovery of evolutionary relationships, the analyses do reveal the need for careful consideration of characters and methodology if phylogenetic relationships are to be accurately recovered. Previous work (Joyce et al., 1994a) suggested that DNA sequences of the internally transcribed spacer region of the rRNA tandem repeat (ITS) could provide the heritable

Received for publication 1 May 1997.

[1] Journal Series Number 12057, Agricultural Research Division, University of Nebraska. Research funded in part by the U.S. Department of Agriculture grant 95-37312-1660 and by the European Community STD program grant CT 940278.

[2] Graduate student and Associate Professor, Department of Plant Pathology, University of Nebraska, Lincoln, NE 68583-0722.

[3] Senior Lecturer, Department of Biology, National University of Ireland Maynooth, Maynooth, County Kildare, Ireland.

E-mail: bjadams@biocomp.unl.edu

The authors thank the researchers who helped identify and supply nematode cultures, T. Harris and D. Sui for their expert help with molecular biology protocols and techniques, D. Casey for help in rearing nematode cultures, J. Trueman for sharing a BASIC program he wrote to calculate PTP and T-PTP values, and D. S. Hibbett for help with Templeton's test. We also thank E. C. Bernard, J. Lyons-Weiler, S. P. Stock, J. Liu, and two anonymous reviewers for their constructive criticisms.

characters requisite for a thorough phylogenetic analysis. The advantages and taxonomic suitability of this marker for addressing phylogenetic relationships among populations, species, and supraspecific taxa have been addressed by Baldwin et al. (1995) and Hillis and Dixon (1991). These include PCR amplification and sequencing by universal primers, forced uniformity of paralogues via rapid concerted evolution, variation due primarily to point mutations, apparent independence of variable sites, and phylogenetic information appropriate for species level investigations.

The goal of this study was to infer phylogenetic relationships among the described taxa of *Heterorhabditis* using DNA sequences of the ITS1 region of the ribosomal tandem repeating unit. We show that this region performs better at resolving relationships among closely related sister taxa than among more inclusive clades. Though it appears as if some of these sister taxa are actually conspecific, a more thorough examination of character variability within these species is required before an evolutionary species delimitation can be accomplished with confidence.

Materials and Methods

Isolates examined: Nine isolates of *Heterorhabditis* representing eight described species and one putative species were examined: *H. bacteriophora, H. hawaiiensis, H. hepialius, H. marelatus, H. indicus, H. zealandica, H. megidis, H. argentinensis,* and Irish. One described heterorhabditid species (*H. brevicaudis* Liu, 1994) could not be examined because of restricted availability. The Irish isolate has not been described as a species, but cross-breeding and PCR-RFLP experiments suggest that this isolate is a distinct species (Dix et al., 1992; Griffin et al., 1994; Joyce et al., 1994a,b). Outgroup taxa consisted of three species representing two other families within the order Rhabditida (*Caenorhabditis elegans, Steinernema carpocapsae,* and *Pellioditis typica*). The taxa used in this study, strain identification, geographic location of isolation, and source of material are listed in Table 1. The DNA sequence of the rDNA ITS1 region of *Caenorhabditis elegans* was obtained from GenBank, accession number X03680. Sequences for all other taxa were deposited in the GenBank database with the following accession numbers: *H. zealandica,* AF029705; *H. argentinensis,* AF029706; *H. hawaiiensis,* AF029707; *H. bacteriophora,* AF029708; *H. hepialius,* AF029709; *H. indicus,* AF029710; *H. megidis,* AF029711; *H.* sp. "Irish K122," AF029712; *H. marelatus,* AF029713; *Pellioditis typica,* AF036946; *Steinernema carpocapsae,* AF036947.

Extraction of DNA: Individuals were removed from White traps (Woodring and Kaya, 1988) in such quantities as to yield a

Table 1. Source and origin of Heterorhabditis taxa included in the analysis.

Species	Strain	Location	Source
Heterohabditis bacteriophora Poinar 1975	Brecon	Australia	R. Akhurst[a]
H. hawaiiensis Gardner, Stock & Kaya 1994	KH3	Hawaii, USA	S. P. Stock[b]/P. Grewal[c]
H. hepialius Stock, Strong & Gardner 1996	Bodega Bay	California, USA	S. P. Stock
H. marelatus Liu & Berry 1996	OH-10	Oregon, USA	J. Liu[d]
H. indicus Poinar, Karunakar & David 1992	LN2	India	S. K. Easwaramoorthy[e]
H. zealandica (Wouts 1979) Poinar 1990	NZH3	New Zealand	R. Bedding[f]
H. megidis Poinar Jackson & Klein 1987	OH-1	Ohio, USA	P. Smits[g]
H. argentinensis Stock 1993	Rafaela	Argentina	S. P. Stock
"Irish"	K122	Ireland	C. T. Griffin[h]
Rhabditis (Caenorhabditis) elegans Maupas 1899	N2	England	T. Stiernagle[i]
Steinernema carpocapsae (Weiser 1955) Poinar 1990	ALL	Georgia, USA	J. Jackson[j]
Rhabditis (Pellioditis) typica Stefanski 1922	CGC#2226	Kenya	W. Sudhaus[k]/T. Stiernagle

[a] CSIRO, Canberra, Australia. [b] University of California at Davis, USA. [c] BIOSYS, Columbia, MD, USA. [d] Oregon State University, USA. [e] Sugar Cane Breeding Institute, Coimbatore, India. [f]CSIRO, Canberra, Australia. [g] Institute for Plant Protection, Wageningen, The Netherlands. [h] National University of Ireland, Maynooth, Co. Kildare, Ireland. [i] Caenorhabditis Genetics Center, University of Minnesota, USA. [j] Northern Grain Insects Research Laboratory, Brookings, SD, USA. [k] Institut für Zoologie, FU Berlin, Germany.

50-µl pellet after a 5-minute microcentrifugation at 12,000*g*. The pellets were incubated in 150 µl of extraction buffer (100 mM Tris, pH 8.0; 100 mM EDTA, pH 8.0; 100 mM NaCl; 0.5% SDS) for 3 hours and then digested by RNAse A (3.0 µl of 10 mg/ml) and Proteinase K (1 µl of 20 mg/ml) at 37°C for 1 hour. The samples were diluted with 100 µl of buffer-saturated phenol (0.1 M Tris, pH 8.0; 0.2% b-mercaptoethanol: 50% total volume) and incubated at 55°C for 15 minutes, during which they were vortexed briefly every 2 minutes. One hundred microliters of a 24:1 chloroform:isoamylalcohol (CHCl3:IAA, 50% w/v) solution was added to each tube, vortexed for 1 minute, and microfuged for 5 minutes at 12,000*g*. Each lysate was transferred to a new tube to which was added 200 µl of 24:1 CHCl3:IAA. This solution was vortexed briefly and microfuged at 12,000*g* for 5 minutes. The lysate was again removed and this time added to a tube containing 400 µl of 100% ethanol. This solution was again vortexed briefly and spun at 12,000*g* for 20 minutes. The supernatant was removed and the remaining pellets were vacuum-dried. The pellets for each sample were resuspended in 50 µl of TE (10.0 mM Tris, 1.0 mM EDTA, pH 8.0). Aliquots of the extracted DNA were run on a 1% agarose gel along with a quantitatively diagnostic ladder (BioMarker Low, Bioventure, Murfreesboro, TN) to provide estimates of relative DNA concentration. The working stock of DNA for PCR amplification was diluted to approximately 10 ng/ml.

PCR amplification: The ITS-1 region was amplified from the diluted DNAs (10 ng/ml) with the rDNA 1.58s primer (5′-ACGAGCCGAGTGATCCACCG-3′) described by Cherry et al. (1997) and rDNA2 primer (5′-TTGATTACGTCCCTGCCCTTT-3′) (Vrain et al., 1992). DNA was amplified according to Joyce et al. (1994a) with the exception of the following cycling parameters: One cycle of 94°C for 2 minutes was followed by 35 cycles of denaturation at 94°C for 1 minute, annealing at 55°C for 45 seconds, and extension at 72°C for 2 minutes. Each 25-µl reaction required 2 µl of extracted DNA (10 ng/ml).

DNA cloning and sequencing: The plasmid pBluescript (Stratagene Cloning Systems, LaJolla, CA) was linearized with Sma I, end-repaired, ligated to the PCR product, and transformed into *E. coli* XL-1 blue (Stratagene) (Sambrook et al., 1989). Between two and six clones of each isolate were entirely and bidirectionally sequenced with forward/reverse or T3/T7 primers by the dideoxy chain-termination method (Sanger et al., 1977) using Sequenase reagent kits (U.S. Biochemical, Cleveland, OH). Sequencing reaction products were separated on a 6% polyacrylamide gel. Cloned products also were sequenced with a LI-COR Model 4000 DNA sequencer (LI-COR, Lincoln, NE) at the University of Nebraska-Lincoln DNA sequencing laboratory. In addition, direct sequencing of PCR products was performed at the Iowa State University sequencing facility, Ames, Iowa.

Sequence fidelity: After completion of DNA sequencing, individual DNA sequences were used to obtain the predicted endonuclease restriction site map of the ITS1 region. The DNA sequences were then digested with enzymes corresponding to the map (Rsa I, Hha I, Dde I, Alu I, Taq I, Hae I, Pst I, Hinf I, Hind II, Sau3a, Cla I, EcoR I) to verify that sequence variation of the cut DNA matched the predicted sites of the sequenced DNA. Enzyme restriction conditions followed Cherry et al. (1997).

Multiple sequence alignment: To infer homology of nucleotide sites, ingroup sequences first were aligned to one another. Then, outgroup sequences were aligned such that homology statements along the ingroup remained internally consistent. To test the effect of different gap-penalty weights on the alignment, the PILEUP command in GCG (Genetics Computer Group, Madison, WI) was used to vary 16 gap initiation and extension penalties that varied from strict to lenient by a factor of 0.9. The effects of these penalties on multiple sequence alignments were observed in the resulting tree topologies. For phylogenetic analysis, sequences were aligned with MALIGN (Wheeler and Gladstein, 1994). Costs for gap initiation, length, and exten-

sion were estimated according to Wheeler (1990). To test whether the presence or absence of different output taxa had an effect on the alignment and resulting phylogenetic trees, each of the outgroup taxa was aligned separately and in combination to the ingroup taxa. To investigate non-independence of sites and the possibility that selection maintains certain conformational similarities, secondary structures were estimated with FOLDRNA, SQUIGGLES, and GCG figure (Genetics Computer Group, Madison, WI) and covarying sites were mapped onto the best estimate of phylogenetic relationships.

Phylogenetic analysis: Parsimony analyses on the alignments were conducted with PAUP 3.1.1 (Swofford, 1993) on phylogenetically informative characters only, with gaps being excluded or treated as a fifth character state. Because multiple gaps may arise due to single insertion or deletion events (indels), including all gaps as independent characters can inflate the actual number of events. To account for this, indel coding was employed. Apparent single indels were treated as a single character, regardless of the actual length of the insertion or deletion. The matrix for the indels was appended to the end of the data set, and the indels in the actual data set that corresponded to those in the matrix were excluded from the analysis. Midpoint rooting (MinF) also was performed to test whether an analysis of the alignment of the ingroup taxa only, when rooted by the midpoint of all ingroup branch lengths, would produce alternative topologies.

In contrast, cladistic analysis was performed by strict outgroup comparison. All variable character states of the outgroup taxa were considered plesiomorphic, and inferred homologous character states of the ingroup were considered apomorphic. Multi-state characters were eliminated from the analysis because of the inherent difficulty of objectively establishing their polarity (example: outgroup = A,G; ingroup = C,T). Autapomorphic and monomorphic characters also were removed from the phylogenetic analysis as they provided no hierarchical information. For each remaining character, a tree of relationships was created, and from these an overall consensus tree was constructed based on combinable components (Nelson and Platnick, 1981).

The DNADIST program of PHYLIP v. 3.57c (Felsenstein, 1993) was used to calculate genetic distances according to the Kimura 2-parameter and maximum likelihood models of sequence evolution. PUZZLE v. 3.1 (Strimmer and von Haeseler, 1996, 1997) was used to generate distances according to the Tamura-Nei (1993) model of substitution-by-way-of-likelihood criterion. Trees were constructed from these distances with the NEIGHBOR and FITCH programs to create Neighbor-Joining (Saitou and Nei, 1987), Fitch-Margoliash (Fitch and Margoliash, 1967), and UPGMA (Sokal and Michener, 1958) trees. BIONJ, an improved neighbor-joining algorithm that more accurately incorporates higher substitution rates among lineages, also was used (Gascuel, 1997). The FITCH, NEIGHBOR, and BIONJ algorithms allow for unequal branch lengths and, hence, heterogeneous rates of sequence evolution. The UPGMA algorithm assumes a molecular clock, such that all branch lengths are constrained to be equidistant from their root. Since the empirical base frequencies of the data set did not appear to be significantly biased, and because of the non-coding nature of the ITS1 region, the Kimura 2-parameter model (Kimura, 1980) with equal rates for variable sites probably was adequate for approximating corrected genetic distances (Gaut and Lewis, 1995). However, the maximum likelihood and Tamura-Nei models (estimated by likelihood), which allow for varying nucleotide frequencies, also were employed. These frequencies were determined empirically from actual nucleotide frequencies as they occur in the multiple sequence alignment. When possible, to minimize the effect of input order on the resulting tree search, taxa were randomized and jumbled three times prior to each search, and the global rearrangements option was employed.

Maximum likelihood analysis was com-

pleted with PUZZLE v. 3.1 (Strimmer and von Haeseler, 1996, 1997) and the DNAML an DNAMLK programs in PHYLIP v. 3.57c (Felsenstein, 1993). PUZZLE uses a quartet puzzling algorithm (Strimmer and von Haeseler, 1996) to search for the optimal tree. DNAML also employs the statistical maximum likelihood search for the optimal tree but uses stepwise addition plus branch swapping to heuristically find the best tree (Felsenstein, 1981, 1993). Both allow for variable nucleotide frequencies and different models of sequence evolution (e.g. Hasegawa et al., 1985; Schöniger and von Haeseler, 1994; Tamura and Nei, 1993). In addition, transition-to-transversion and purine-to-pyrimidine transition parameters also were determined empirically and incorporated into the search algorithm. The DNAMLK program has a search algorithm identical to DNAML, except that all branch lengths are constrained to be equidistant from their root, invoking a molecular clock. A Chi-square test of the difference between clock and non-clock imposed likelihoods was done to test whether the ITS1 region evolves at a constant rate (Felsenstein, 1993). To examine the effect of outgroup taxa on ingroup topology, an alignment of the ingroup only also was analyzed.

The PUZZLE maximum likelihood tree search utilized quartet puzzling of 1,000 steps, with *Steinernema* selected as the single outgroup taxon for rooting purposes only. All models of sequence evolution were tested, even though the assumptions of some models (e.g., Schöniger and von Haeseler, 1994) may be violated by the presumed independence of nucleotides across all sampled sites in the ITS1 region. Transition-to-transversion parameters, purine-to-pyrimidine transition parameters, and nucleotide frequencies were estimated from the data set. DNAML and DNAMLK tree searches utilized a 2:1 transition-to-transversion ratio parameter. Empirical base frequencies were used along with one-category substitution rates, global rearrangements, and randomized input order of the taxa. Because these analyses ultimately resulted in unrooted trees, each tree was arbitrarily rooted with *Steinernema carpocapsae*. All trees were visualized with Treeview v. 1.4 (Page, 1997).

Evaluation of tree topology: Bootstrapping was performed by generating 100 data sets with the DNABOOT program in PHYLIP, in which input sequences were randomized at each replication. Most-parsimonious trees were constructed from the randomized data sets by DNAPARS, and a majority rule and combinable component consensus of these trees were constructed using the CONSENSE program in PHYLIP (trees considered rooted at *S. carpocapsae*). Bremer support (Bremer, 1988, 1994; Källersjö et al., 1992) for the parsimony tree was derived using AutoDecay v. 2.95 (Eriksson, 1996). Alternate topologies were evaluated as to their overall length, T-PTP tests (topology-dependent permutation tail probability tests; Faith and Cranston, 1991), Templeton's non-parametric test (Templeton, 1983), and Kishino-Hasegawa log likelihood tests (Kishino and Hasegawa, 1989). Overall strength of phylogenetic signal was estimated with Relative Apparent Synapomorphy Analysis, RASA v. 2.1 (Lyons-Weiler, 1996; Lyons-Weiler et al., 1996) and PTP tests. These tests also were performed on data sets that were manipulated as to outgroup and ingroup taxa in an effort to determine the origin of the majority of phylogenetic signal.

Results

Multiple sequence alignments: The amount of shared sequence identity among the different heterorhabditid species showed a high degree of variability. *Heterorhabditis marelatus* and *H. hepialius* differed at only one nucleotide position (C-T transition). Similarly, *H. hawaiiensis* and *H. indicus* differed at two sites (one transition and one tranversion), while *H. bacteriophora* and *H. argentinensis* differed by a single transition and deletion. However, the differences between the remaining taxa varied from a low of 35 substitution events between K122 and *Heterorhabditis megidis* to a high of 119 nucleotide positions between *H. zealandica* and *H. hawaiiensis*.

An optimized multiple sequence alignment of the ingroup and outgroup resulted in 730 base pairings, of which 192 were phylogenetically informative for parsimony analysis and 89 could be polarized by strict outgroup comparison for cladistic analysis (Fig. 1). (The multiple sequence alignment of the ingroup only is available upon request.) Alignments that contained any single outgroup taxon, or a combination of any two outgroup taxa, did not generate any variation of tree topology among the ingroup taxa for any of the algorithms that produced unrooted trees. Nonoptimized alignments produced by successively decreasing gap initiation and extension penalties from strict to lenient caused all of the different tree-building algorithms to produce trees with a different topology than those resulting from the optimized alignment.

Parsimony analysis: Parsimony analysis of the aligned sequences in which gaps were treated as missing data or coded as charac-

```
                          10         20         30         40         50         60
H. hawaiiensis    GTACACACCGCCCGTCGCTGTCCGGGACTGAGCTGTTTCGAGAAGAGTGGGGACTG-CTA
H. indicus        ...........................................................-...
H. argentinensis  ..........................................................A.....-..G
H. bacteriophora  ..........................................................A.....-..G
H. megidis        ..........................................................A.....-..G
'Irish (K122)'    ..........................................................A.....-..G
H. hepialius      ..........................................................A.....-..G
H. marelatus      ..........................................................A.....-..G
H. zealandica     ..........................................................A.....-..G
P. typica         ..........................................C..A....C.-..T
S. carpocapsae    ..................C..........T..........A..C..A..T..-.G.
C. elegans        ..................A..........A...A..................T.GC

                          70         80         90         100        110        120
H. hawaiiensis    TATCG-GGGCTTTCGGGC-TCTGGTATGATGGAAACCATTTTAATCGCAATGGCTTGAAC
H. indicus        .....-..........-.....
H. argentinensis  ...T.-..........-.....
H. bacteriophora  ...T.-..........-.....
H. megidis        .....-..........-.....
'Irish (K122)'    .....-..........-.....
H. hepialius      .....-........C...-.....
H. marelatus      .....-........C...-.....
H. zealandica     .....-....C.......-...A.............GC.....
P. typica         .TGT.-A...-...-...-.T..C.C.AG.....G...G.....
S. carpocapsae    .G.T.AAC.T......A.GGTCTT.G.TGC.AG....GCG.......A.TC.....
C. elegans        .-...-A..T..AAC.A.-.TC.T.G.TGC.....T......T...T.....

                          130        140        150        160        170        180
H. hawaiiensis    CGGGCAAAAGTCGTAACAAGGTATCTGTAGGTGAACCTGCAGATGGATCATCGTCG-ATA
H. indicus        ...........................................................-...
H. argentinensis  ..........................................................C..-.A.
H. bacteriophora  ..........................................................C..-.A.
H. megidis        ..........................................................CT.A.A.
'Irish (K122)'    ..........................................................CT.A.A.
H. hepialius      ..........................................................CT.A.A.
H. marelatus      ..........................................................CT.A.A.
H. zealandica     ..........................................................CT.A.A.
P. typica         ......................C.................G..........-.C-..T
S. carpocapsae    ......................T..C.............G..A.......-------.-
C. elegans        ....T.................G.................C.........C..--.G
```

FIG. 1. Multiple sequence alignment of ingroup (*Heterorhabditis*) and outgroup taxa. Periods indicate nucleotide identity, and hyphens indicate gaps.

```
                          190        200        210        220        230        240
H. hawaiiensis    CCT-TATA--GGT-ACAT-GCTG---AT-CAC-GAG-ATGCCGATA----ATCAT-G--G
H. indicus        ...-....--...-....-....---..-...-...-.........----.....-.--.
H. argentinensis  ...-...G--...-.-.-...T-TG..-...-...-.GAT..G..-CC-.--..-.--.
H. bacteriophora  ...-...G--...-.-.-...T-TG..-...-...-.GAT..G..-CC-.--..-.--.
H. megidis        ...-...G--.-A-T-..-...T-TGG.-...-...-.GAT..G.G--CT.--CC-.--.
'Irish (K122)'    ...-...G--.-.-T-..-...T-TGG.-...-...-.GAT..G.G--CT.--C.-.--.
H. hepialius      ...-...G--.-.-T-..-...T-TGG.-...-...-.GAT..G.G--CT.-.-.-.--.
H. marelatus      ...-...G--.-.-T-..-...T-TGG.-...-...-.GAT..G.G--CT.-.-.-.--.
H. zealandica     ...-...G--.-.-T-..-...T-TGG.-...-...-C.AT..G.G--CTC--..-.--.
P. typica         ...-.-..-C..A-GG..-T.C.CT-..A.T.TT..C...G.....GCCGG..G.G.TT.
S. carpocapsae    -A.-.GAG--C-.-.-.--A.T---T.-.-.TT----.TT-A..C----.-----A--.
C. elegans        ...C.G.GCCC..G.T..CCACA---.A-.-.TT.--.GAT-.C.C---G.-.-.-.--.

                          250        260        270        280        290        300
H. hawaiiensis    AAT-CAG-G-C--TT---G-TTC-TT--G----G-TTCCA-GT-C--GGT--GTCT--CA
H. indicus        ...-...-.-.--..---.-...-..--.----.-.....-..-.--...--....--..
H. argentinensis  ...-...-.-.--..---.-...-..--.-----A-..T..-A.-.--...--T...--..
H. bacteriophora  ...-...-.-.--..---.-...-..--.-----A-..T..-A.-.--...--T...--..
H. megidis        ...-...-.-.--..--GCTCC--.C--.-----A-..T.G-A.-.--...--A...--..
'Irish (K122)'    ...-...-.-.--..--GC-.C--..--.-----A-..T..-A.-.--...--A...--..
H. hepialius      ...-...-.-.--..---.-C..-..--.-----A-..T..-A.-.--...--A...--..
H. marelatus      ...-...-.-.--..---.-C..-..--.-----A-..T..-A.-.--...--A...--..
H. zealandica     ...-...-.-.--.C--AC-G.--..--.-----A-..T..-A.-.--...--AG..--..
P. typica         CC.GT.T-.-.A-..GAA.-.C.-..A-.---AC-.GAG.T..A.--CA.--.A..AG..
S. carpocapsae    -T.-TTC-.-.TG..--T.T...-.AA-.----C-..TA.-C.-T--.A.--C..---T.
C. elegans        CT.-..CG.T.AG..G-A.TG..GAAAT.TCAACG.....-..-TGA.A.GCCA.-AA..

                          310        320        330        340        350        360
H. hawaiiensis    -C-CCC---A-T--CTA-AGCTCTCG---GT--GAG-GT-G--TCTATTCTTGATTGGAG
H. indicus        -.-...---.-.--...-.....---..--...-..-.--.....C.....
H. argentinensis  -.-...---.-.--...-.....AT.---.A--...-..-.--....G..CCA.....
H. bacteriophora  -.-...---.-.--...-.....AT.---.A--...-..-.--....G..CCA.....
H. megidis        C.-...---.-.--...-.....---.A--...-..-.--....A..CCA..C....
'Irish (K122)'    -.-...---.-.--...-.....---.A--...-..-.--....A..GCA..C....
H. hepialius      -.-...---.-.--...-.....---.A--...-..-.--....A..CCA.....
H. marelatus      -.-.T.---.-.--...-.....---.A--...-..-.--....A..CCA.....
H. zealandica     -.-...---.-.--...-.....T.---..--...-..-.--.....CCA.....
P. typica         C.-...AA-.-.--T..-.....-..---CA--...-..-.--.TC.....A..-
S. carpocapsae    A.-GG.--------T---T--.---.----A---.A----.--G-.-..-C.AC--A..-
C. elegans        A.TGG.AAG.G.AGTG.CT.-..-..ACCCATG..AC..G.CAGT....-CCA.--AAC-

                          370        380        390        400        410        420
H. hawaiiensis    CCGCTT-T-GAGTGACGGC-A-ATGAT-AGTT--GGGTATG--TTCCCCGTGAG-----G
H. indicus        .....-.-.....-.-.....-....--.....--.....-----.
H. argentinensis  T.....-.-.....-T-....A-.A..--.....--.....-----.
H. bacteriophora  T.....-.-.....-T-....A-.A..--.....--.....-----.
H. megidis        T.....-.-.....-.-....G-GA..--....G..CCA.A...-CATACGTGG.
'Irish (K122)'    T.....-.-.....-.-....A-GA..--CT..G..---.T...-C.TA----A.
H. hepialius      T.....-.-.....-.-....A-G...--....G..--.....A..-----.
H. marelatus      T.....-.-.....-.-....A-G...--....G..--.....A..-----.
H. zealandica     T.....-G-.....-.-G....-GA..--....G..---.TG..-C.TA--T-G.
P. typica         T..T..G.A....TG...--T-T.AC.-...C--AA..CCA--A.GG..-.T..--T--C
S. carpocapsae    T-.T..-G-...-C.--.--T--..-.-.-.G---A.--C.---.GA-.-..T.-----C
C. elegans        TACT..-.-.T..C.G...TTA....G...GAAAA.--..--..TG-.-..TT-----C
```

FIG. 1. Continued.

```
                          430        440        450        460        470        480
H. hawaiiensis     GTAGAGCA-TAGACTTTATGA-ACAG-A-GCT-GGGCTGTCGCCTCA-CC-AAAAACCAT
H. indicus         ........-..........-....-.-...-............-..-.........
H. argentinensis   ..C.....-..........-....T--...G.A..........-..A.....T...
H. bacteriophora   ..C.....-..........-....T--...G.A..........-..A.....T...
H. megidis         ........-.....G...G..-....T--...C..A..........-..---..C-.G.
'Irish (K122)'     ........-.........A..-....T--...G..A.....C.....-..---..C-.G.
H. hepialius       ........-..........-....T--...G..A..........-..A....CT...
H. marelatus       ........-..........-....T--...G..A..........-..A....CT...
H. zealandica      ..G.C...-.......A....-....T.CT..C..A....T......G..--...CT-..
P. typica          ..T.-.-T-.GA...CA.A..TG.T.---C..-.A.-..........--.-----.-.G.
S. carpocapsae     --T--.-.-.GA.--CAT..--T-.-----.--A-T-...-T-A..---T----..--GC
C. elegans         --T--...G....--CGT..-TT..-----.-GA.T-C...T.AA..--.----....GC

                          490        500        510        520        530        540
H. hawaiiensis      --CGA-TAA-CTGGTGGCT-G---A-GT-GA-G-----AAATCAC--T-GG-ATCTGCT-
H. indicus          --...-...-.........-.---.-..-..-.-----.......--.-..-.......-
H. argentinensis    --...-...-.........-A-TGT-..-..-C---ATT.G....A-.A..T.......G
H. bacteriophora    --...-...-.........-A-TGT-..-..-C---ATT.G....A-.A..T.......G
H. megidis          --...-.-GA.........-A-TG.-..-..-C--T-TCGG....--.C.AG.......-
'Irish (K122)'      --...-.-TA.........-A-TG.-..-..-C--T-TCGG....--.C.CG.......-
H. hepialius        --...-.-.A.........-A-TG.T..-..-C--T-TCGG..G.A-.C.AG...C...-
H. marelatus        --...-.-.A.........-A-TG.T..-..-C--T-TCGG..G.A-.C.AG...C...-
H. zealandica       --..T-G-GA.C.......-.-TG.T..-..-C--T-TCGG....A-.C.AG...C...-
P. typica           ---T.-..T-...C...-.-.-T-.-T.-T.-.---CT.GT....--C-TC-.-A..-A-
S. carpocapsae      ---.T-.-T-.--..C.-.-T-T-C--.-T--.------..TG.TTAG.-.A-TGA-.-A-
C. elegans          TAT.TG.CTC.........ATATGC-..CT.G.CTTCTTCT..TTTGC-..-.GAG.-A-

                          550        560        570        580        590        600
H. hawaiiensis      A-TGC----AGGG-AGCCTT-AATG--A-GTTGGTC---TT-------C---A-CCG--A
H. indicus          .-...----....-.....-....--.-.....---..-------.---.-...--.
H. argentinensis    .-...----..A.-.....-....--.-....T..-GTG.-------.---.T.T.ACC
H. bacteriophora    .-...----..A.-.....-....--.-....T..-GTG.-------.---.T.T.ACC
H. megidis          .-...----..A.-..T.-.C....--.-...-..--G-..-------.-----GT--CC
'Irish (K122)'      .-C..----..A.-....-.C....--.-...-.-----..-------.-----G.--CC
H. hepialius        T-...----G.CT-....-.C....--.-...-.----..-------G-----.T---C
H. marelatus        T-...----G.CT-....-.C....--.-...-.----..-------G-----.T---C
H. zealandica       .-...----G.A.-....-.C....--.-...-.----.C-------.-----GT---C
P. typica           .-.TG----.C..--T....CGGG.-CT-.....---GG-------.---.-G.T--C
S. carpocapsae      .-.TA----.A.A-G.T.-.-GC..ACTC.CCAT..--T..-------G---.-TT.--C
C. elegans          .G.TAGTTG.C.AT..T...-G...A-TC..G.T..GGT..CGGCCGAGCCA.GG..-CC

                          610        620        630        640        650        660
H. hawaiiensis      --C-ACA-A-CC--GCCACTATCGGT--AATC--T-A-TT---C-C-C-AA-TT-A----
H. indicus          --.-...-.-..-.....--....--.-.-..---.-.-.-.-..-.-T----
H. argentinensis    --T-...-.-..--....G.....--..-A--.-C-AA---.-.-.-.-..-.-.----
H. bacteriophora    --T-...-.-..--.....--..-A--.-C-AA---.-.-.-.-..-.-.----
H. megidis          --A-..C-G-..--.-GT-G-...A.A-G.-A-A.T-T..---T-.-.T.---.-T----
'Irish (K122)'      --A-..C-G-..--.-GT-G-...A.A-G.-A-A.TCT..---T-.-.T.---.-T----
H. hepialius        --A-..C-G-..--.-GT-G-...A.--..-A---TCT.A---T-.T.-..--.-T---A
H. marelatus        --A-..C-G-..--.-GT-G-...A.--..-A---TCT.A---T-.T.-..--.-T---A
H. zealandica       --A-..C-G-..----TT-G--..A.ATTT-G-A.T-T..GGGTA.-GT.TCG.-T----
P. typica           --T-GT.-.-.A--A...-.-..CA.--T..A--.-G-..---T-A-.-GT-..GTC---
S. carpocapsae      --TA...-.-AA--A.GT-.-.T-..--CTCG----.-.A-----A-TTG---.GTC---
C. elegans          AAAA...C.TAAGCA..T-.-.TCA.A-TT.GAA.C.G.A---.-A-.TG.-..GCCAAA
```

FIG. 1. Continued.

```
                          670        680        690        700        710        720
H. hawaiiensis      A--CTT--GTTTCTAGT----A-AA-AGGCTAA-A-TT-AGTCA-GTGG-AA--A----A
H. indicus          .--...--.........----.-..-.....-.-..-.....-....-..--.----.
H. argentinensis    .--...--.....T..----G-TC-GT.T.-.-.-.A-CA.-.CC...C..--.GTGT.
H. bacteriophora    .--...--.....T..----G-TC-GT.T.-.-.-.A-CA.-.-C...C..--.GTGT.
H. megidis          .-A...--.....CT.-----.TTC-GT.T.-.-.-.A-CA.-T-T...C-.CA.-TGT.
'Irish (K122)'      .-A...--C.....T-C----.TTC-GT.T.-.-.-.A-CA.-T-T...C-.CA.-TGT.
H. hepialius        .--...--.....T.-----.TTC-GT.T.-.-.-.A-CA.-T-T...C-.CA.-TGT.
H. marelatus        .--...--.....T.-----.TTC-GT.T.-.-.-.A-CA.-T-T...C-.CA.-TGT.
H. zealandica       .-T...--..C...T.-G---.T.C-GT.T.-.-.-.A-CA.-T-T...C-.CA.-TGT.
P. typica           .--A..--.G....T----T-.-TT-GA...-G-.-..-.A.-T-C.-.--.--C----.
S. carpocapsae      .--.-.--CG..GAT.C----.-TT-TT-T.-.-.-..--A--T-C--A-.GT-C---T-
C. elegans          .GT...CACGGA.AT.CGGTG.TTTGTT.T.G.G.T..GCA..T-C-.AT..TAC--GT.

                          730
H. hawaiiensis      -TAGCCTTAG
H. indicus          -.........
H. argentinensis    T....T....
H. bacteriophora    T....T....
H. megidis          T....T.C..
'Irish (K122)'      T....T.C..
H. hepialius        T....T.C..
H. marelatus        T....T.C..
H. zealandica       T....T.C..
P. typica           --...T....
S. carpocapsae      -..---.C--
C. elegans          C....T.C..
```

FIG. 1. Continued.

ters resulted in trees that had different overall lengths but were topologically identical. *Heterorhabditis zealandica* was depicted as the sister to *H. megidis* + K122. *Heterorhabditis zealandica* + *H. megidis* + K122 appeared as the sister to *H. hepialius* + *H. marelatus*. *Heterorhabditis argentinensis* + *H. bacteriophora* represented the sister group to *H. hepialius* + *H. marelatus* + *H. zealandica* + *H. megidis* + K122. The lineage of *Heterorhabditis indicus* + *H. hawaiiensis* comprised the sister group to the rest of the genus (MP/ML/GD topology; Fig. 2). There were only 18 instances where multiple states within the ingroup could possibly be due to a single indel event, and there were no topological differences between trees made using all gaps as informative, as missing data, or indel coding (the matrix created by indel coding is available upon request). When the alignment of the ingroup taxa only was examined and midpoint rooting was enforced, the lineage of *H. argentinensis* + *H. bacteriophora* was depicted as the sister to *H. indicus* + *H. hawaiiensis*. Three equally parsimonious solutions for the placement of *H. zealandica* were represented as an unresolved polytomy (MinF topology; Fig. 3A). The treelengths and log likelihoods for all trees are summarized in Table 2.

Cladistic analysis: A cladistic analysis using strict outgroup comparison produced a single tree (SO topology; Fig. 3B) compatible with the maximum parsimony hypothesis (MP/ML/GD topology; Fig. 2). There appeared to be few homoplasies, and the tree showed a good fit with the transformation series (CI = 0.80; RI = 0.83). However, three characters supported the position of *H. zealandica* as sister to *H. hepialius* + *H. marelatus*, while three others supported it as sister to *H. megidis* + K122, resulting in failure to unambiguously resolve this node.

Genetic distance and maximum likelihood analyses: All genetic distance and maximum likelihood analyses produced a topological arrangement congruent with the maximum parsimony tree. The UPGMA algorithm produced a slightly different tree, with *H. zealandica* as sister to the clade comprising *H.*

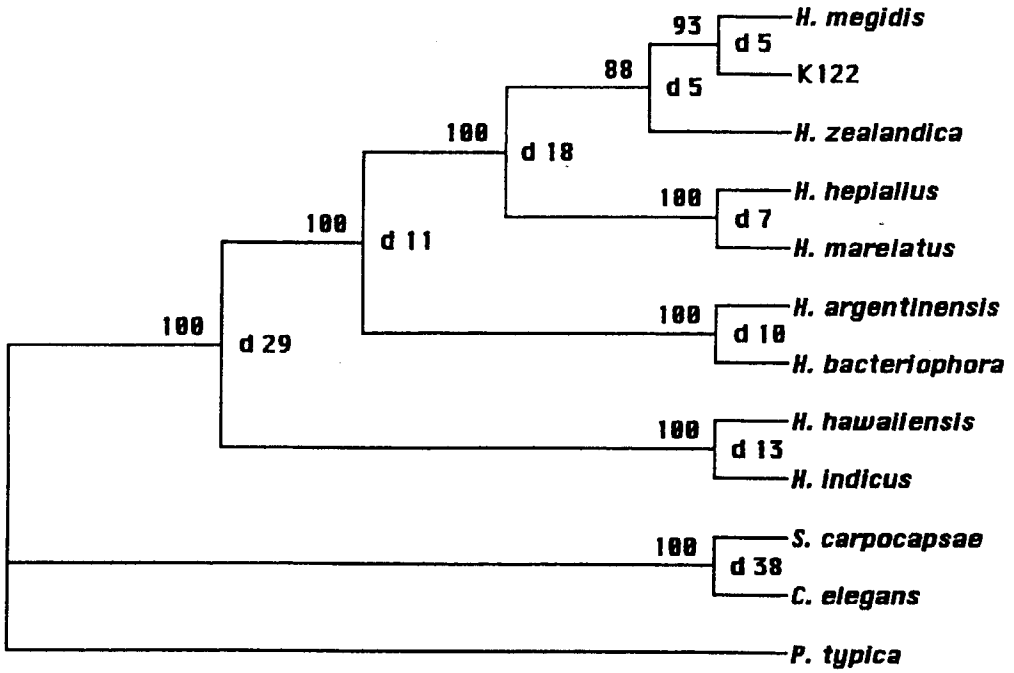

FIG. 2. Best supported hypothesis of phylogenetic relationships for *Heterorhabditis* based on nuclear ribosomal DNA and produced by maximum parsimony (PAUP), maximum likelihood (DNAML, DNAMLK, PUZZLE), and distance (BIONJ, FITCH, NEIGHBOR) tree-building algorithms (MP/ML/GD topology). Bremer support indices (the number of extra steps taken to not find the clade in a longer tree) are mapped at the nodes and preceded by the letter "d." Bootstrap frequencies (100 replicates) appear above and to the left of each decay index.

marelatus, H. hepialius, H. megidis, and K122. This arrangement also favored the lineage of *H. argentinensis* + *H. bacteriophora* as sister to *H. hawaiiensis* + *H. indicus* (UPGMA topology; Fig. 3C). Models of sequence evolution for distance correction or use of an ingroup-only alignment did not affect the topology produced by these algorithms. Although maximum-likelihood trees were identical when a molecular clock was imposed, the hypothesis that the ITS1 region evolves in a clock-like manner was rejected ($P < 0.001$).

Phylogenetic signal: Results of the RASA and PTP analyses are presented in Table 3. The presence of hierarchical signal was high for the data set overall and was independent of any combination of outgroup taxa ($P < 0.001$). However, when taxa representing uncontroversial arrangements at a particular node were removed (*H. argentinensis, H. indicus, H. hepialius*), the amount of phylogenetic signal was not significant ($P > 0.2$). PTP analysis revealed that all randomized data sets produced longer trees than the most economical hypothesis, regardless of which taxa were included in the outgroup, suggesting the presence of ample phylogenetic signal. However, when uncontroversial sister taxa were removed, the standard deviation decreased substantially, from more than 16 steps longer when all taxa were included to less than 3 steps.

Tests of alternative topologies and tree robustness: Table 2 presents a comparative summary of alternative tree topologies. When relationships among outgroup taxa were unresolved, the MP/ML/GD tree (Fig. 2) was the most economical solution with a length of 334. The SO (Fig. 3B) tree was three steps shorter (331), but this was due to an unresolved polytomy. The MinF topology (Fig. 3A) had a treelength of 338, and the UPGMA arrangement (Fig. 3C) was the longest with 346 steps. A Kishino-Hasegawa test of likelihoods and Templeton's non-parametric test of parsimony rejected all alternative topological arrangements. The T-PTP test rejected the MinF and UPGMA topologies but failed to reject the SO arrangement. Bremer support and bootstrap analyses showed strong internal support for the overall MP/ML/GD topology. The weakest support was for the relationship of *H. zealandica* as sister

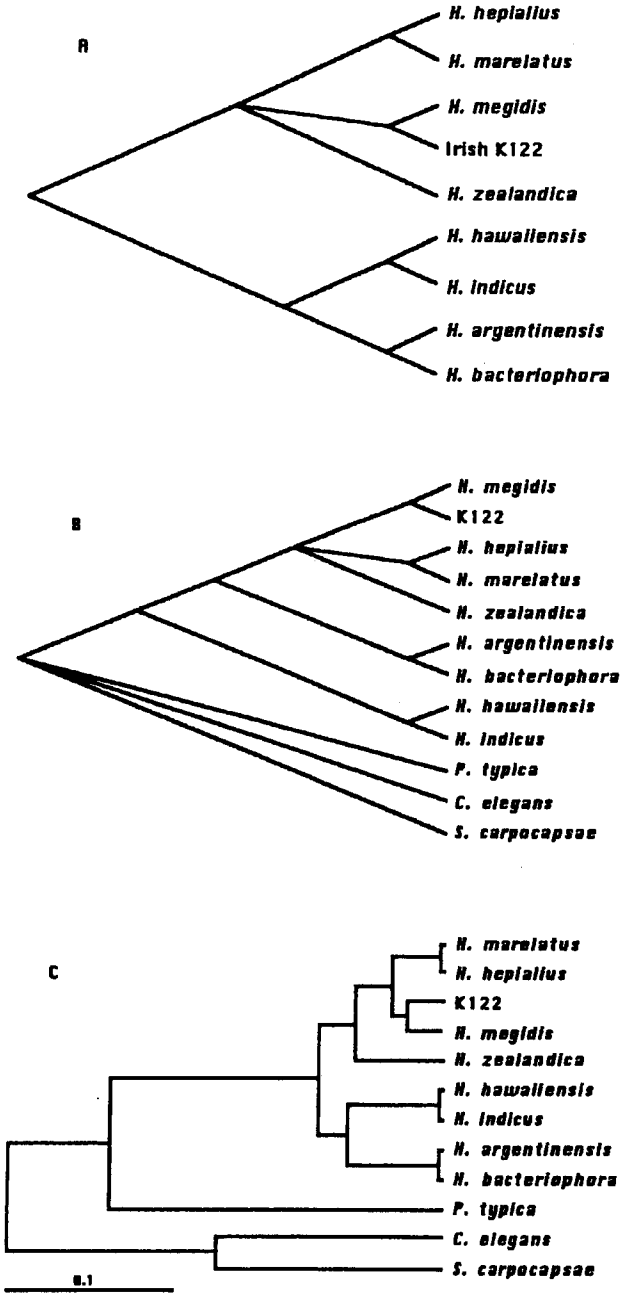

FIG. 3. Alternative hypotheses of *Heterorhabditis* evolution produced by different tree-building methods (scale bars accompanying phenograms represent 10% sequence divergence). A) Parsimony tree produced using an alignment of ingroup taxa only and rooted by the midpoint of all branchlengths (MinF topology). B) Tree produced by strict outgroup comparison (SO topology). C) Tree produced by UPGMA distance algorithm (UPGMA topology).

to K122 + *H. megidis*. Still, this placement required 5 extra steps prior to not finding this arrangement in a longer tree, and was supported by 88% of the bootstrap replicates (Fig. 2).

DISCUSSION

Alignment and homology: We assumed that for every taxon each nucleotide position in the alignment is a homologous, independent estimate of evolutionary history. However, in the absence of more information (i.e., an independent phylogeny) this assumption is only weakly supported as rRNA secondary structure almost certainly leads to interdependence among some sites. In addition, portions of the ITS1 transcript may play a role in the maturation of nuclear RNAs (Musters et al., 1990; van Nues et al., 1994), suggesting that nucleotide changes in one portion of the molecule could affect other nucleotide sites. We checked this hypothesis by mapping potentially non-independent sites onto the best estimate of phylogenetic relationships but could not detect clear patterns of non-independence. While non-independence is likely a confounding factor in this analysis, lacking more specific models and data we are constrained to methods presupposing independence.

Contrary to conventional alignment algorithms that are order-dependent (such as PILEUP) or based on phenetic optimality criterion such as CLUSTAL (Higgins et al., 1996), MALIGN produces and tests homology statements (alignments) by way of parsimony criteria and, for this reason, was deemed superior to other methods (Wheeler, 1996). However, due to the high amount of interspecific variation, some portions of the ingroup multiple sequence alignment could only tenuously infer homologous nucleotide bases. This high variability led to alignment difficulties that were exacerbated by the addition of the more distantly related outgroup taxa. For example, the alignment of closely related species, such as *H. marelatus* and *H. hepialius*, yielded a data set in which the shared base pairs were almost certainly homologous. However, this confidence dropped precipitously as more inclusive clades were added to the alignment. Future studies that include outgroup taxa known to be more closely related to *Heterorhabditis* should increase confidence in homology statements.

Choosing a single outgroup species would have simplified the alignment problem but would have obviated our ability to recover

TABLE 2. Comparison of alternative tree topologies.

Tree topology by method of tree reconstruction	Actual length[a] (length when outgroup resolution is collapsed)	Kishino-Hasegawa test[b]			T-PTP test[b]		Templeton test[b]		
		In likelihood ±	S.D.	P	# trees longer	P	n	Ts	P
MP/ML/GD	334 (334+)	-1499.72	—	optimal	–	optimal	–	–	optimal
MinF[c]	403 (338+)	-1527.88 ±	8.58	<0.01	100	0.01	65	62	<0.01
SO	396 (328+)	-1527.90 ±	8.55	<0.01	74	0.26	51	51	<0.01
UPGMA	400 (346+)	-1515.29 ±	6.69	<0.05	100	0.01	22	18	<0.01

[a] Plus signs following treelengths indicate incomplete resolution of the outgroup taxa. Treelengths are from parsimony informative characters only and exclude gaps and indel coding.

[b] All values are corrected for missing data (gaps) and do not reflect indel coding or treatment of all gaps as informative. P-values ≤ 0.05 reject the hypothesis that the topology is not significantly different from the optimal tree.

[c] The MinF tree was constructed using an alignment of only the ingroup taxa, and its actual length from that data set is 142. To compare this topology with the other trees, the unresolved outgroup taxa were added.

monophyly among the ingroup. In addition, as ingroup characters are polarized based on character states of outgroup taxa, this procedure would have treated the single outgroup taxon as if it were the actual ancestor to *Heterorhabditis,* a hypothesis not supported by any previous research of the Heterorhabditidae, as well as a violation of cladogenic speciation.

Another option to the problem of aligning outgroup sequences to those of the ingroup include tree-building algorithms that do not require the presence of outgroup taxa, such as maximum likelihood and distance methods. To explore whether the alignment of the ingroup taxa was significantly different from the alignment that included the outgroup taxa, we used these

TABLE 3. RASA and PTP tests of phylogenetic strength and signal.

Parameters	RASA test results					PTP test results	
	Observed slope	Null slope	tRASA	df	P-value[a]	# unrooted trees ≥ the most parsimonious solution	P-value[a]
All taxa, all outgroup taxa specified	3.75	2.85	5.38	24	p < .01	100	0.01
All taxa, no outgroup specified	3.54	3.61	-0.29	51	0.10	100	0.01
All ingroup taxa; *C. elegans* specified as sole outgroup taxon	4.13	2.57	7.91	24	p < 0.01	100	0.01
All ingroup taxa; *S.carpocapsae* specified as sole outgroup taxon	3.99	2.47	7.19	24	p < 0.01	100	0.01
All ingroup taxa; *P. typica* specified as sole outgroup taxon	3.38	2.58	4.75	24	p < 0.01	100	0.01
All taxa; *H. argentinensis* specified as sole outgroup taxon	6.40	4.90	5.60	41	p < 0.01	100	0.01
Ingroup only; no outgroup specified	3.52	2.53	5.72	24	p < 0.01	100	0.01
Pruned ingroup only; no outgroup specified	1.63	1.51	0.54	6	0.10	100	0.01
Pruned ingroup, outgroup taxa are included; no outgroup specified	3.07	3.03	0.14	24	0.10	100	0.01
Pruned ingroup, outgroup taxa are included; all outgroup taxa specified	1.75	1.65	1.07	6	0.10	100	0.01

[a] Values of p ≤ 0.05 reject the hypothesis that the data set does not contain significant phylogenetic signal.

programs to search for trees using only the alignment of the ingroup taxa. With the exception of the unrooted parsimony arrangement (MinF), the resulting topologies from these methods did not differ from the arrangements produced when the outgroup taxa were present in the data. These results uphold the validity of the alignment as a well-supported homology statement. Still, as is apparent when gap initiation and extension penalties are manipulated, even small perturbations of the data set can result in different topological arrangements. These findings support those of Morrison and Ellis (1997), who showed that different approaches to aligning sequences can account for more topological discrepancy than do the different types of tree-building algorithms used to construct them.

Presence of phylogenetic signal: The RASA test utilizes a relaxation of Hennig's auxiliary principle (1966) and suggests that all characters in the data set can be assumed to be noise until evidence sufficient to reject this null hypothesis is found (Lyons-Weiler et al., 1996). Unlike bootstrapping and PTP tests, which rely on resampling or *a posteriori* permutations, RASA is independent of tree-building assumptions and algorithms, indicating a statistical test that can be evaluated in terms of sensitivity and power. Results of this test on our data set revealed a significant amount of hierarchical information, but most was informative only for closely related sister species. For example, when uncontroversial relationships among sister species were removed from consideration (i.e. *H. argentinensis, H. indicus, H. hepialius*), the data did not contain a significant amount of hierarchical phylogenetic information. This finding suggests that in the Heterorhabditidae the region has become saturated with change and is most appropriate for inferring relationships among recently divergent lineages.

A PTP test investigates the presence of hierarchical signal based on the possibility that a particular tree could have arisen by chance alone (Archie, 1989; Faith and Cranston, 1991). If randomized data sets of the original can produce equal or more parsimonious trees, the null hypothesis is not rejected and it is likely there is little hierarchic signal in the original data set. Under four different constraints (presence or absence of outgroup taxa and uncontroversial sister species) the trees produced by randomized data were less economical than the most parsimonious solution, an indication of the presence of significant phylogenetic signal. These results were congruent with those of the RASA test, with the exception of the pruned ingroup data set that contained all the outgroup taxa designated as such. In this study, RASA suggested that hierarchic signal was weak, whereas the PTP test showed support for a strong signal. Still, the standard deviation decreased greatly as uncontroversial taxa were pruned. When designated outgroups are included in the test, PTP may be biased toward rejection of the null hypothesis, or suggest more support for strong signal than actually exists (Trueman, 1996). This bias does not exist in the data sets that do not contain outgroup taxa but may explain the discrepancy between the PTP and RASA results for the pruned data set, which does contain designated outgroup taxa. Alternatively, RASA may underestimate phylogenetic signal in the presence of long branch lengths and multiple outgroup taxa (J. Lyons-Weiler, pers. comm.). These caveats may explain some of the disparity between the PTP and RASA test results. For example, the outgroup taxa have exceptionally long branch lengths relative to the ingroup taxa, and the branch lengths of three pairs of sister taxa (*H. marelatus* + *H. hepialius, H. indicus* + *H. hawaiiensis, H. argentinensis* + *H. bacteriophora*) are short relative to the branches representing their common lineage. Long branch attraction may be a problem for accurate reconstructions of evolutionary history by parsimony methods (Felsenstein, 1978; Kuhner and Felsenstein, 1994). However, since maximum parsimony and maximum likelihood trees did not differ topologically, it appears that long branch attraction did not affect the ability of parsimony methods to recover an optimal tree by maximum likelihood standards.

Tree robustness: Bootstrapping, usually con-

sidered a statistical indication of confidence in tree topology, can be a misleading indication of support (Carpenter, 1992; Hillis and Bull, 1993; Kluge and Wolf, 1993). For example, in our study all nodes appeared to have strong bootstrap and Bremer support indices. However, these indices were equally high for very different trees produced by slight perturbations to the multiple sequence alignment (B. J. Adams, unpubl.).

Alternative topologies: Parsimony analysis without outgroup comparison and the UPGMA method of tree building have been shown to be incompatible with the goals of recovering phylogenetic relationships congruent with evolutionary history (Farris, 1980, 1981, 1982). For this reason, coupled by their rejection by all three tree comparison tests, the UPGMA and MinF arrangements are rejected as not reflecting phylogenetic relationships in *Heterorhabditis.* Among common methods of phylogenetic analysis, maximum parsimony and maximum likelihood have been shown to result in reliable estimates under simulated variations of DNA sequence evolution (Huelsenbeck, 1995; Huelsenbeck and Hillis, 1993) and distance algorithms appear to be improving (Gascuel, 1997). In this analysis these three algorithms produced the arrangement favored by evaluation of tree-length and all three of the tree comparison tests. Therefore, the MP/ML/GD tree is our best estimate of phylogenetic relationships within the genus, with reservations only as to the correct position of *H. zealandica* relative to *H. hepialius* + *H. marelatus* and *H. megidis* + K122.

The ITS1 region appears to perform well at establishing relationships among sister species and some populations of *Heterorhabditis* (B. J. Adams and T. O. Powers, unpubl.) but, as taxa become more phylogenetically divergent, change accumulates such that relationships with more distantly related taxa must be inferred with less confidence. This is especially true with methods requiring outgroup comparison because increased homoplasy can corrupt the validity of character homology and polarity, a condition exacerbated by the properties of DNA. For instance, strict outgroup character polarization of the nucleotide bases is compromised by the fact that the bases are restricted to one of only four character states, from which we must deduce evolutionary history. Limiting the number of evolutionary possibilities of each character increases the number of homoplasious characters masquerading as synapomorphies, especially as the region becomes saturated with change. For large data sets such as the one in this study, it is often assumed that synapomorphies will eventually drown out the false (homoplasious) signals, which may well be the case for the maximum parsimony and maximum likelihood algorithms. However, it is unknown whether the trend of synapomorphy dominance occurs under the stringent character selection of the strict outgroup method, where most noise is removed prior to analysis. Although the SO topology is a compatible subset of the MP/ML/GD solution, we question the compatibility of highly divergent DNA sequences with the methodology employed by the cladistic strict outgroup method.

Species delimitation and taxonomy: Despite rigorous morphometric analyses and keys for distinguishing species (Nguyen and Smart, 1996; Stock and Kaya, 1996), some authors have maintained that morphological characters are unreliable for the identification of heterorhabditid species (Liu and Berry, 1996). However, the source of much of this presumed unreliability may simply be a function of the limitations of traditional light microscopy and underrepresentation of morphological variation within and among populations. For example, since measurements often are taken from the progeny of a few soil-baited insect hosts, it is unlikely that they represent the range of variation present in the population. Consequently, it is impossible to discriminate between characters that are variable within or among different species. This consideration is important, especially since some morphometric studies of natural populations have shown such high levels of variability (Roman and Figueroa, 1995). This caveat also applies to this study, since the DNA sequences from

a single isolate are assumed to be representative of its species. It remains likely that some portion of the sequence dissimilarity between these species is attributable to variation within the species each population is purported to represent.

Heterorhabditidae have been delimited on the basis of Linnean species concepts and phenetic morphometrics, and tested for adherence to the biological species concept through cross-breeding analysis. These concepts can lead to species delimitations that are incompatible with recovered evolutionary history (Frost and Kluge, 1994). A species concept more compatible with this goal has been proposed for nematology (Adams, 1998). According to this species concept, the discovery operations of the Phylogenetic Species Concept (Cracraft, 1983; Nixon and Wheeler, 1990) are modified and used to recover the Evolutionary Species of Wiley (1978). Adams' method utilizes character polarization by way of outgroup comparison to identify autapomorphies (unique, derived characters) among comparable individuals. Lineages that possess an autapomorphy exhibit sufficient evidence that they are on independent evolutionary trajectories and can be delimited as species (Adams, 1998).

The ITS region can reveal diagnostic differences at the species level among most nematode taxa studied thus far (Powers et al., 1997). However, as proposed by Adams (1998), we distinguish species diagnosis from species delimitation as the latter having to do with characters and operations that are meaningful in the context of recovering relationships among historical entities. Although the focus of this paper is primarily the reconstruction of phylogenetic history and not species delimitation, a careful analysis of ITS1 ribosomal DNA characters reveals that each lineage is well supported by autapomorphies with the exception of the closely related sister taxa *H. marelatus* + *H. hepialius, H. indicus* + *H. hawaiiensis,* and *H. argentinensis* + *H. bacteriophora,* suggesting that these pairs of sister taxa may actually be conspecific (Table 4). Restricted to our study, conspecificity (or synonymy) is not entirely conclusive, since each of these taxa has at least one character state not shared by its sister appearing among other members of the genus. If further analyses reveal that these character states are uniquely derived, then there may be sufficient evidence that each of these sister taxa represents separate species. However, in support of conspecificity, a recent morphological and morphometric reexamination of *H. marelatus* and *H. hepialius* resulted in a taxonomic proposal for their synonymization (Stock, 1997). While confident delimitation of these species requires careful analyses of all potentially informative characters (morphological, genetic, behavioral, etc.), adopting a more satisfying concept of species and gaining a better understanding of character variability will further empower taxonomic statements.

TABLE 4. List of autapomorphies possessed by each *Heterorhabditis* taxon or lineage. Character numbers correspond to nucleotide position in the multiple sequence alignment.

Taxon	Autapomorphic character
H. megidis	235, 266, 280, 415, 418, 435, 439, 674
H. zealandica	71, 83, 218, 231, 255, 294, 383, 423, 425, 449, 450, 461, 468, 487, 492, 581, 615, 617, 622, 628, 641, 642, 643, 645, 648, 652, 653, 671
Irish K122	351, 395, 396, 462, 543, 575
H. argentinensis + *H. bacteriophora*	324, 381, 423, 505, 517, 540, 594, 598, 639, 682, 716
H. indicus + *H. hawaiiensis*	174, 179, 196, 206, 221, 361, 447, 521, 600, 603, 632, 675, 685, 687, 688, 692, 701, 706, 726
H. hepialius + *H. marelatus*	524, 541, 552, 553, 647

While relationships among the three families of the order Rhabditida represented in this study are beyond the scope of this paper, unrooted trees depict the outgroup taxon *Pellioditis* as being more closely related to *Heterorhabditis* than to *Caenorhabditis* and *Steinernema.* This relationship is also supported by other analyses, including an 18S rRNA gene phylogeny (T. O. Powers, unpubl.; P. De Ley, pers. comm.), a cladistic

analysis (Sudhaus, 1993), and anecdotal evidence (Poinar, 1993).

Future considerations: This work provides an introductory framework for studies that depend on evolutionary history as the basis for investigation (i.e., Brooks and McLennan, 1991; Harvey and Pagel, 1991; Slatkin and Maddison, 1989). For example, coevolutionary patterns between these nematodes and their symbiotic bacteria can be investigated. As a biological control agent, the origin and maintenance of traits such as environmental tolerance and host finding behavior can be studied within a historical context. Ecology, biogeography, gene flow, and population structure can be explored by indirect phylogenetic methods. This is a critical consideration because, after many years of field application and pest control performance research, little is known about the dispersal and fate of exotic strains of *Heterorhabditis* and its bacterial symbiont.

Literature Cited

Adams, B. J. 1998. Species concepts and the evolutionary paradigm in modern nematology. Journal of Nematology 30:1–21.

Akhurst, R. J. 1987. Use of starch gel electrophoresis in the taxonomy of the genus *Heterorhabditis* (Nematoda: Heterorhabditidae). Nematologica 33:1–9.

Archie, J. W. 1989. A randomization test for phylogenetic information in systematic data. Systematic Zoology 38:239–252.

Baldwin, B. G., M. J. Sanderson, J. M. Porter, M. F. Wojciechowski, C. S. Campbell, and M. J. Donoghue. 1995. The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. Annals of the Missouri Botanical Garden 82:247–277.

Boemare, N. E., R. J. Akhurst, and R. G. Mourant. 1993. DNA relatedness between *Xenorhabdus spp.* (Enterobacteriaceae), symbiotic bacteria of entomopathogenic nematodes, and a proposal to transfer *Xenorhabdus luminescens* to a new genus, *Photorhabdus gen. nov.* International Journal of Systematic Bacteriology 43: 249–255.

Bremer, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42:795–803.

Bremer, K. 1994. Branch support and tree stability. Cladistics 10:295–304.

Brooks, D. R., and D. A. McLennan. 1991. Phylogeny, ecology and behavior: A research program in comparative biology. Chicago: University of Chicago Press.

Carpenter, J. M. 1992. Random cladistics. Cladistics 8:147–153.

Cherry, T., A. L. Szalanski, T. C. Todd, and T. O. Powers. 1997. The internal transcribed spacer region of *Belonolaimus* (Nemata: Belonolaimidae). Journal of Nematology 29:23–29.

Cracraft, J. 1983. Species concepts and speciation analysis. Pp. 159–187 *in* R. F. Johnston, ed. Current ornithology, vol. 1. New York: Plenum Press.

Curran, J., and J. M. Webster. 1989. Genotypic analysis of *Heterorhabditis* isolates from North Carolina. Journal of Nematology 21:140–145.

Dix, I., A. M. Burnell, C. T. Griffin, S. A. Joyce, M. J. Nugent, and M. J. Downes. 1992. The identification of biological species in the genus *Heterorhabditis* (Nematoda: Heterorhabditidae) by crossbreeding second-generation amphimictic adults. Parasitology 104:509–518.

Eriksson, T. 1996. AutoDecay version 2.9.5 <http://www.botan.su.se/Systematik/Folk/Torsten.html> (29 October 1996).

Faith, D. P., and P. S. Cranston. 1991. Could a cladogram this short have arisen by chance alone?: On permutation tests for cladistic structure. Cladistics 7:1–28.

Farris, J. S. 1980. Naturalness, information, invariance, and the consequences of phenetic criteria. Systematic Zoology 29:360–381.

Farris, J. S. 1981. Distance data in phylogenetic analysis. Pp. 3–23 *in* V. A. Funk and D. R. Brooks, eds. Advances in cladistics I. New York: New York Botanical Garden.

Farris, S. J. 1982. Outgroups and parsimony. Systematic Zoology 31:328–334.

Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Systematic Zoology 27:401–410.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. Journal of Molecular Evolution 17:368–376.

Felsenstein, J. 1993. PHYLIP v. 3.572 <http://evolution.genetics.washington.edu/phylip/html> (15 January 1996).

Fitch, W M., and E. Margoliash. 1967. Construction of phylogenetic trees. Science 155:279–284.

Frost, D. R., and A. G. Kluge. 1994. A consideration of epistemology in systematic biology, with special reference to species. Cladistics 10:259–294.

Gardner, S. L., S. P. Stock, and H. K. Kaya. 1994. A new species of *Heterorhabditis* from the Hawaiian islands. Journal of Parasitology 80:100–106.

Gascuel, O. 1997. BIONJ: An improved version of the NJ algorithm based on a simple model of sequence data. Molecular Biology and Evolution 14:685–695.

Gaut, B. S., and P. O. Lewis. 1995. Success of maximum likelihood phylogeny in the four-taxon case. Molecular Biology and Evolution 12:152–162.

Griffin, C. T., S. A. Joyce, I. Dix, A. Burnell, and M. J. Downes. 1994. Characterisation of the entomopathogenic nematode *Heterorhabditis* (Nematoda: Heterorhabditidae) from Ireland and Britain by molecular and cross-breeding techniques, and the occurrence of the genus in these islands. Fundamental and Applied Nematology 17:245–253.

Harvey, P. H., and M. D. Pagel. 1991. The comparative method in evolutionary biology. Oxford: Oxford University Press.

Hasegawa, M., H. Kishino, and K. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mi-

tochondrial DNA. Journal of Molecular Evolution 22: 160–174.

Hennig, W. 1966. Phylogenetic systematics. D. D. Davis and R. Zangerl, trans. Urbana, IL: University of Illinois Press.

Higgins, D. G., J. D. Thompson, and T. J. Gibson. 1996. Using CLUSTAL for multiple sequence alignments. Methods in Enzymology 266:383–402.

Hillis, D. M., and J. J. Bull. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Systematic Biology 42:182–192.

Hillis, D. M., and M. T. Dixon. 1991. Ribosomal DNA: Molecular evolution and phylogenetic inference. Quarterly Review of Biology 66:411–453.

Huelsenbeck, J. P. 1995. Performance of phylogenetic methods in simulation. Systematic Biology 44:17–48.

Huelsenbeck, J. P., and D. M. Hillis. 1993. Success of phylogenetic methods in the four-taxon case. Systematic Biology 42:247–264.

Joyce, S. A., A. M. Burnell, and T. O. Powers. 1994a. Characterization of *Heterorhabditis* isolates by PCR amplification of segments of mtDNA and rDNA genes. Journal of Nematology 26:260–270.

Joyce, S. A., C. T., Griffin, and A. M. Burnell. 1994b. The use of isoelectric focusing and polyacrylamide gel electrophoresis of soluble proteins in the taxonomy of the genus *Heterorhabditis* (Nematoda: Heterorhabditidae). Nematologica 40:601–612.

Källersjö, M., J. S., Farris, A. G., Kluge, and C. Bult. 1992. Skewness and permutation. Cladistics 8:275–287.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16:111–120.

Kishino, H., and M. Hasegawa. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. Journal of Molecular Evolution 29:170–179.

Kluge, A. G., and A. J. Wolf. 1993. Cladistics: What's in a word? Cladistics 9:183–199.

Kuhner, M. K., and J. Felsenstein. 1994. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. Molecular Biology and Evolution 11:459–468.

Liu, J., and R. E. Berry. 1996. *Heterorhabditis marelatus* n. sp. (Rhabditida: Heterorhabditidae) from Oregon. Journal of Invertebrate Pathology 67:48–54.

Lyons-Weiler, J. 1996. RASA 2.0. <http://loco.biology.unr.edu/archives/rasa/rasa.html> (4 November 1996).

Lyons-Weiler, J., G. A. Hoelzer, and R. J. Tausch. 1996. Relative apparent synapomorphy analysis (RASA) I: The statistical measurement of phylogenetic signal. Molecular Biology and Evolution 13:749–757.

Morrison, D. A., and J. T. Ellis. 1997. Effects of nucleotide sequence alignment on phylogeny estimation: A case study of 18s rDNAs of Apicomplexa. Molecular Biology and Evolution 14:428–441.

Musters, W., K. Boon, C. A. F. M. van der Sande, H. van Heerikhuizan, and R. J. Planta. 1990. Functional analysis of transcribed spacers of yeast ribosomal DNA. EMBO 9:3989–3996.

Nasmith, C. G., D. Speranzini, R. Jeng, and M. Hubbes. 1996. RFLP analysis of PCR amplified ITS and 26S ribosomal RNA genes of selected entomopathogenic nematodes (Steinernematidae, Heterorhabditidae). Journal of Nematology 28:15–25.

Nelson, G. J., and N. I. Platnick. 1981. Systematics and biogeography: Cladistics and vicariance. New York: Columbia University Press.

Nguyen, K. G., and G. C. Smart, Jr. 1996. Identification of entomopathogenic nematodes in the Steinernematidae and Heterorhabditidae (Nemata: Rhabditida). Journal of Nematology 28:286–300.

Nixon, K. T., and Q. D. Wheeler. 1990. An amplification of the phylogenetic species concept. Cladistics 6: 211–223.

Page, R. D. M. 1997. TreeView ver. 1.4. <http://taxonomy.zoology.gla.ac.uk/rod/treeview.html> (15 March 1997).

Poinar, G. O., Jr. 1975. Description and biology of a new insect parasitic rhabditoid *Heterorhabditis bacteriophora* n. gen., n. sp. (Rhabditida; Heterorhabditidae n. fam.). Nematologica 21:463–470.

Poinar, G. O., Jr. 1993. Origins and phylogenetic relationships of the entomophilic rhabditids, *Heterorhabditis* and *Steinernema*. Fundamental and Applied Nematology 16:333–338.

Powers, T. O., T. C. Todd, A. M. Burnell, P. C. B. Murray, C. C. Fleming, A. L. Szalanski, B. J. Adams, and T. S. Harris. 1997. The rDNA internal transcribed spacer region as a taxonomic marker for nematodes. Journal of Nematology 29:441–450.

Roman, J., and W. Figueroa. 1995. Morphometric evaluation of 20 *Heterorhabditis* isolates from Puerto Rico. Journal of Agriculture of the University of Puerto Rico 79:51–64.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406–425.

Sambrook, J., E. F., Fritsch, and T. Maniatis. 1989. Molecular cloning: A laboratory manual (2nd ed.). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sanger, F., S. Nicklen, and M. Coulson. 1977. DNA sequencing with chain terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America 162:5463–5467.

Schöniger, M., and A. von Haeseler. 1994. A stochastic model for the evolution of autocorrelated DNA sequences. Molecular Phylogenetics and Evolution 3: 240–247.

Slatkin, M., and W. P. Maddison. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics 123:603–613.

Smart, G. C. 1995. Entomopathogenic nematodes for the biological control of insects. Supplement to the Journal of Nematology 27:529–534.

Sokal, R. R., and C. D. Michener. 1958. A statistical method for evaluating systematic relationships. University of Kansas Science Bulletin 38:1409–1438.

Stock, S. P. 1997. *Heterorhabditis hepialius* Stock, Strong & Gardner, 1996 a junior synonym of *H. marelatus* Liu & Berry, 1996 (Rhabditida: Rhabditidae) with a

redescription of the species. Nematologica 43:455–463.

Stock, S. P., and H. K. Kaya. 1996. A multivariate analysis of morphometric characters of *Heterorhabditis* species (Nemata: Heterorhabditidae) and the role of morphometrics in the taxonomy of species of the genus. Parasitology 82:806–813.

Stock, S., D. Strong, and S. Gardner. 1996. Identification of *Heterorhabditis* (Nematoda: Heterorhabditidae) from California with a new species isolated from the larvae of the ghost moth *Hepialius* [sic] *californicus* (Lepidoptera: Hepialidae) from the Bodega Bay Natural Reserve. Fundamental and Applied Nematology 19: 585–592.

Strimmer, K., and A. von Haesler. 1996. Quartet puzzling: A quartet maximum likelihood method for reconstructing tree topologies. Molecular Biology and Evolution 13:964–969.

Strimmer, K., and A. von Haeseler. 1997. PUZZLE ver. 3.1. <http://www.zi.biologie.uni-muenchen.de/~strimmer/puzzle.html> (12 March 1997).

Sudhaus, W. 1993. Die mittels symbiontischer Bakterien entomopathogenen Nematoden Gattungen *Heterorhabditis* und *Steinernema* sind keine Schwestertaxa. Verhandlungen der Deutschen Zoologischen Gesellschaft 86:146.

Swofford, D. L. 1993. PAUP: Phylogenetic analysis using parsimony. Champaign, IL: Illinois Natural History Survey.

Tamura, K., and M. Nei. 1993.Estimation of the number of nucleotide substititions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10:512–526.

Templeton, A. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37:221–224.

Trueman, J. 1996. Permutation tests and outgroups. Cladistics 12:253–261.

van Nues, R. W., J. M. J. Rientjes, C. A. F. M. van der Sande, S. F. Zerp, C. Sluiter, J. Venema, J. Planta, and H. A. Raue. 1994. Separate structural elements within internal transcribed spacer 1 of *Saccharomyces cerevisiae* precursor ribosomal RNA direct the formation of 17S and 26S rRNA. Nucleic Acids Research 22:912–919.

Vrain, T. C., D. A. Wakarchuk, A. C. Levesque, and R. I. Hamilton. 1992. Intraspecific rDNA restriction fragment length polymorphism in the *Xiphinema americanum* group. Fundamental and Applied Nematology 15:563–573.

Wheeler, W. 1990. Combinatorial weights in phylogenetic analysis: A statistical parsimony procedure. Cladistics 6:269–278.

Wheeler W. 1996. Optimization alignment: The end of multiple sequence alignment in phylogenetics? Cladistics 12:1–9.

Wheeler, W., and D. Gladstein. 1994. MALIGN. New York: American Museum of Natural History. <ftp://ftp.amnh.org/pub/molecular> (25 January 1996).

Wiley, E. O. 1978. The evolutionary species concept reconsidered. Systematic Zoology 27:17–26.

Woodring, J. L., and H. K. Kaya. 1988. Steinernematid and heterorhabditid nematodes: A handbook of biology and techniques. Fayetteville, AR: Arkansas Agricultural Experiment Station.