Mitochondrial DNA Sequence Divergence among *Meloidogyne incognita, Romanomermis culicivorax, Ascaris suum,* and *Caenorhabditis elegans*[1]

T. O. Powers,[2] T. S. Harris,[2] and B. C. Hyman[3]

Abstract: Mitochondrial DNA sequences were obtained from the NADH dehydrogenase subunit 3 (ND3), large rRNA, and cytochrome *b* genes from *Meloidogyne incognita* and *Romanomermis culicivorax*. Both species show considerable genetic distance within these same genes when compared with *Caenorhabditis elegans* or *Ascaris suum*, two species previously analyzed. *Caenorhabditis, Ascaris,* and *Meloidogyne* were selected as representatives of three subclasses in the nematode class Secernentea: Rhabditia, Spiruria, and Diplogasteria, respectively. *Romanomermis* served as a representative outgroup of the class Adenophorea. The divergence between the phytoparasitic lineage (represented by *Meloidogyne*) and the three other species is so great that virtually every variable position in these genes appears to have accumulated multiple mutations, obscuring the phylogenetic information obtainable from these comparisons. The 39 and 42% amino acid similarity between the *M. incognita* and *C. elegans* ND3 and cytochrome *b* coding sequences, respectively, are approximately the same as those of *C. elegans*–mouse comparisons for the same genes (26 and 44%). This discovery calls into question the feasibility of employing cloned *C. elegans* probes as reagents to isolate phytoparasitic nematode genes. The genetic distance between the phytoparasitic nematode lineage and *C. elegans* markedly contrasts with the 79% amino acid similarity between *C. elegans* and *A. suum* for the same sequences. The molecular data suggest that *Caenorhabditis* and *Ascaris* belong to the same subclass.

Key words: *Ascaris suum, Caenorhabditis elegans,* DNA, *Meloidogyne incognita,* molecular evolution, mtDNA, nematode, nucleotide sequence, *Romanomermis culicivorax.*

The alignment of amino acid and nucleotide sequence data permits an estimation of relative levels of genetic divergence. These comparisons can be conducted at any taxonomic level, provided sufficient evidence for homology exists among the comparisons. Theoretically, because the degree of genetic divergence is directly related to time of separation of two lineages, the longer two nematode taxa have been separated evolutionarily, the greater the degree of divergence (37). This linear relationship between time and nucleotide substitution proceeds at variable rates in different genes (34,37) until a plateau is reached, beyond which little additional divergence is possible because of the saturation of particular loci with mutations

(27,34). At this point, mutations upon mutations confound the measurement of divergence and often lead to convergent nucleotide changes. Therefore, molecular comparison of distantly related taxa may require more slowly evolving genes.

In an evolutionarily old phylum like the Nematoda, expected divergences of basic lineages would likely extend hundreds of millions of years into the past. Classification of higher taxa in the Nematoda has reflected the perceived fundamental differences among these lineages (22–24,29). Secernentea, one of the two classes in the phylum, has been divided on the basis of morphological and behavioral criteria into three subclasses, Spiruria, Rhabditia, and Diplogasteria (22–24).

Given the previously estimated age of subclass separation within the Secernentea (29), a molecular study of mitochondrial DNA (mtDNA) surprisingly indicated a relatively recent divergence of ca. 65 million years ago for the separation of *Ascaris suum* (28), a representative of the Spiruria (23), and *Caenorhabditis elegans,* a member of the Rhabditia. Although other classification schemes have recognized a some-

what closer relationship between the ascarid and rhabditid nematodes (22,24), alternative explanations, such as a slowdown in the rate of mtDNA evolution in nematodes, could influence the interpretation of the molecular data. In this study, we expand molecular comparisons of mitochondrial genes to include *Meloidogyne incognita,* a representative of the Diplogasteria, and *Romanomermis culicivorax,* a member of the remaining class, the Adenophorea.

Materials and Methods

A 1.8-kilobase (kb) HindIII fragment from *M. incognita* race 2, which contained genes for the large (16S) ribosomal RNA, NADH dehydrogenase subunit 3 (ND3), and 360 base pairs (bp) of the cytochrome *b* mitochondrial gene, was cloned into the vector pUC18. Sequence analysis was conducted by dideoxynucleotide chain termination (31,32) from single-stranded and double-stranded DNA templates (26). A series of 11 primers (12) was used to obtain the complete sequence from both DNA strands using the reaction components of Sequenase II Kit (United States Biochemical, Cleveland, OH).

Romanomermis culicivorax strain 3B4 (2) mitochondrial genes were isolated and identified as follows. A cDNA library was constructed from total nematode polyadenylated (poly A+) RNA. The poly (A+) RNA (2) population was converted into cDNA molecules and introduced into the λ ZAP cloning vector (31,33). The lambda library was packaged in vitro into phage particles with the Gigapack II Gold packaging extract (18; Stratagene, La Jolla, CA) and amplified in the *Escherichia coli* strain PLK-F'. MtDNA-specific cDNA clones were identified by plaque hybridization with highly purified *R. culicivorax* mtDNA molecules that had been ^{32}P-labeled in vitro by random priming (31).

The DNA sequences of sequential exonuclease III-generated deletions (14,31) were obtained as described above. Nucleotide sequences were analyzed with a Genetics Computer Group (GCG) software package (11) operating on a VAX 8700 and by CLUSTAL (15), RDF (20), and BLASTP (1) programs. Results in this study with *M. incognita* and *R. culicivorax* were compared with those obtained with *A. suum* and *C. elegans* (28). Modifications of the universal genetic code used in computer-assisted translation of the mt-nucleotide sequences were as follows: AGA and AGG = serine; TGA = tryptophan; ATA = methionine (28). Nematode mitochondrial protein genes were identified by alignment of predicted translations with amino acid sequences of known mitochondrial protein genes and by hydropathic profile comparisons (8,19, 20).

Results and Discussion

Mitochondrial DNA can be used to examine a wide range of taxonomic questions (34). Although primarily known for its rapid evolution relative to nuclear genes (6), highly conserved genes in the mitochondrial genome have been used to address ancient divergences. For example, cytochrome oxidase subunit II (COII) sequences have been applied to the relationships among insect orders (21), the small ribosomal RNA gene (12S rRNA) for the monophyly of arthropods (3), and the cytochrome *b* gene for the evolution of major vertebrate groups (16,17). The striking conservation in the 12S ribosomal gene in the mitochondrial genome led to the concept of "universal primers" widely applied to both vertebrate and invertebrate taxa (17).

Figures 1 and 2 present the nucleotide sequence for the complete ND3 and the large rRNA mitochondrial genes, and for the 5′ 320 nucleotides of the cytochrome *b* coding region derived from the four nematode species. Gene length varies among these nematodes. For ND3, *C. elegans* and *A. suum* coding sequences are 333 bp long, whereas the *M. incognita* and *R. culicivorax* genes are 24 and 15 bp shorter.

One large difference exists in the length of the 16S rRNA genes. In *C. elegans* and *Ascaris,* these genes are respectively 953

```
ASU    ---------- ---------- ---------- ---------- ---------- ----------  -TAGTATTTT TTTTTTCATT
CEL    ---------- ---------- ---------- ---------- ---------- ---------- -..AAT...A C..G..T.AA
MIN    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
RCU    ---------- ---------- ------GATA ..TA..TGT. A.AT.TCCTT ...TTT.ACA AA.A...A.G CAACAA..C.

ASU    TTTTAGTATT GTGTACTATT TAGGGAAAAT AACATTTGTT TATGTTAAAT AGAAAGAAAA GGTAGAAGAT TAATAAGTTT
CEL    .A..G..... .CA....... ACAAT..... TT..--.... A...AA.... ...C...G ......GT.A AT..T.....
MIN    ---------- ---------- ---------- --------A. .T.T.ATTG. GATT.A.... .T.TTTG.C. A...T.T...
RCU    ....TTAC.. --A..T...A AGCTT...T. TTA.GA.TA. G.GTC.T.TG .CTTTTTT.G TC..A.CTTA A...T.TAAA

ASU    ATTTGTTTCA CAAATAAAAT TATTTGTTTT GGTGTTGATA TGTCGACTTT TTTGTTATCT GTTTATTTTT TTTTATTAGA
CEL    TA........ T.CTA..... ...A.T.A.. A.A....... .......C.. .G..A..A.. ....T.A... ..A.......
MIN    T..A.A...T TTTT..TTG. .GAAAAAA.. AAAAACA.AT ...TTTT.AC AA.AA.TAAA A.....AA.A ..TC.A.TTT
RCU    .A..A...AG TTTT.CTGG. GTA..TGAA. TT.A..TT.. .AAAACT..A ..AAAATAGA AA...AAA.. ..A....CT.

ASU    AAGTTATATG GTTTGTTTGT ATGTTAGGAA GTTAAGATTT TGTAGTGTTC TTTTTTTGTG TTTTATAACG TTT-T-CTTT
CEL    ..A......A T.A.A.AAT. ..T....... A..T.A.A.. ..A......T .AAA...A.. .....C...A ...-.C..AA
MIN    T.A..T.TA. T...AAA.AA .AA...CA.. TA.G.A.AA. AAA.AAA..T .G..A.AAAT .AAAT...TT .A.-.TA.A.
RCU    ..T..TAT.A TAAAAAA.AA ..A...TTT. AA..TTCA.. ATA.AAAA.A AAA....AAT .A.ACCT.TT .A.A.T.C.C

ASU    TTTTTTAATT GGATGGTTTT GTTTTTATTA ATTTTTTAAT AAATTTATAA TTTTGAATTT AGTAGT--TT TGTTTAAATT
CEL    ....A.TTAA .TT.AA.... TAA.....A. .AG.....T. ....AA.... ...GT..A.. ....AA--.. .TA.A..--.
MIN    .AAAAAT... TTT.ATAA.. T..A..T..T T.A....TT. ....AAA.A. .AA.TT.A.. .TA.T.--A. .TA..T..A.
RCU    ..A.AAGT.A AAT.ATAAG. AT.A..CA.. .A.A..C... ...AAATA.. ..A.AT.AAA .T....AAAA AAA..T..A.

ASU    TTTGTTTTTT ATTTTAATTT TTTTGTGTGA ACTTGTCTTT TGGTCAAATG TTTTTTAAAG ACTTAGG-TC T---TTTTTG
CEL    ..AA...A.. ...AA...A. AA...AAGA. CT.GAAG.C. ..A....... .......... .......C.T .---A.A..A
MIN    ..AA.AA..A .A.A...AA. ....TAT.A. .TAAA.T.AA .AA.A..TGT .....A..TT CT..GA.G.T .TGA......
RCU    A.AT...AAA ...A.TT.AA AAC.AAAGA. CTAA..A..A .AACTG.C.. ....C....A .A.TTTT.T ATTAA...AA
```

```
                                          ┌──────a──────┐      ┌───b───┐
ASU    AGACTGGCCT CTGCTCTATG TTTTTATAAA TG│GCAGTCTT AGCGTGA│GGA CATT│AAGGTA GCAAAAT│AAT TTGTGCTTTT
CEL    .AG.....T. ....C..... A.A..-.... ..│........ .......│... ....│...... ......│... ..........
MIN    .TTT.TTGT. TCTGCTC..T G..AA-AG.. AA│....C.T.. .......│TTG GTCA│...... ..G.G..│... ....TT....
RCU    .ATA.---T. ......AC.. .......... .A│.....T.. .......│TT. .T.A│..A... .......│CCA .A.AT....A
```

```
ASU    AATGGGTTCT AGTATGAATG GGGTTAGTGG CTAATT-TTT TACTTTTATT TTTATGAATT AGTTTTGTGT TTAAGAAATA
CEL    .T..A....C .......... AA....T... T..G..C.A. .TA.G..T.A .G.T...... TAA...T.A. ........A.
MIN    .T...A.... .......... .A-..TT.T. T..T.AAA.A ..T.GA.T.. AAA.A..... TTAAA.T.A. ....A.....
RCU    ..A.A..... T.AC....A. AAA.A.C..T TAT.A.T... .TTA...T.A A....TG.AA TTA.A.AAA. ..TGA.....

ASU    ATGGTTGATA TTACACAAAG ATAAGTCTTC GGAAATTTTT TTATTAAATA TTTTTTTATT TTTTTAATAT GTTTTCTAGG
CEL    ..AAA.ATAT ...T...... .......... .......C.G ......C.C. A..AAA..A. .G.G.....A A.........
MIN    ..AAA.A.A. .A.A..T... .A.......T A........A. .AT.------ ---------- -------..A T..AAT....
RCU    .AAT..TT.. A.TT.TT.C. .G...A.CCT AAG.T..AA- -----..... AA..A....A AC..A...TG .G...T..AA
```

```
                                                                  ┌──────c──────┐
ASU    GTAGAATGTT TTGTAA-CTT TTTTTACTAA TTTTAATTTA AAAAATTACT TCG│GAGTTAA CAG│AAAGTCA TATCTGATTT
CEL    .C.....A.. A.A...TAG. A...C....T A.........G...... C..│....... ...│...A... ..C..A..C.
MIN    .A.AT.AAA. AAT...TA.. .....CT... --..C..AAA. .......T.. CTA│....... ...│G..TAA. ...TAAT..-
RCU    T..AT.AA.. ..T....---A ....C.A..T A.......... ..----...C .TA│.G.A... ...│CTTAAA. .C..CTT.--
```

```
                                             ┌──────d──────┐
ASU    AGTTCTTATA ATAATGATAA GTTTT│ACATC GATGTTG│TAT TTTAGTTCTT AAAGGGAGGA GAGGATTTTA GGTTTTAGAC
CEL    ...A...... G...-.G... .....│..... .......│... .CAGA.AA.C T...A..... ..A.GC..AG TAG.......
MIN    ---------- -------..T. T....│..... .......│-T. ..ATT.AA.A TT.AAA.T.. ..AAC....G A..GAG....
RCU    --------..- TA..G.GGT. A.A.G│..C.. .......│AC. ------.AAAA ...TAA.TTT A.AA.A...T TAA....A..
```

```
                                        ┌──────e──────┐
ASU    TGTTCTTCTA TTATTAAACT AAAC│GTGATA TTAGTTTAA│T TCATCGTGAG ATAGAATTGT TTATCTTGGT AATGTCTTAA
CEL    .......... ...A.T..TC TG..│...... .........│. ....T..... .......... .......A. ..-A.A..T.
MIN    .........-T ..T..T..T- --..│....A. .........│. ..G.T...A .....G.... ......A.T. ---T-A....
RCU    .....----- G.T..T..T. .TTT│....GC .G.......│A A..CT.AA.A G...TT.G.. ..C.A.AAAA .TA..A..G.
```

```
              ┌──────f──────┐
ASU    TTTTAATTAC TGAA│AGTACG AAAGGA│AATT TGTTGGAGTT TGATAACTTT TTGAAAGTGT TAGTACTTTT T
CEL    .A..T.A... ATTT│...... ......│.CA. ...AAA.... .T.A.CT..A AA..TTT..A A.TCTT.--- -
MIN    A..AG.AA.T AA.T│...... ......│..A. .A..TT.AA. .A.ATTT... AA-------- ---------- -
RCU    ..G.....-- .TTT│...... ......│CT.A AAA.ACTT.G ..TT..TAA.C ..A..TT.AA A.T.TA---- -
```

FIG. 1. DNA sequence of the LrRNA gene from *Ascaris suum* (ASU) *Caenorhabditis elegans* (CEL), *Meloidogyne incognita* (MIN), and *Romanomermis culicivorax* (RCU). Dots indicate nucleotide identity with ASU; dashes are inferred deletions. Boxes identify regions of relatively high conservation among all species. Data for *A. suum* and *C. elegans* are from another report (28). Alignments were conducted by the multiple alignment program Pileup from the Genetics Computer Group Software package (11) using a gap weight of 5.0 and gap length weight of 0.3.

ND3

```
ASU     ---------- ---------- ----TTGTTG GTTTTGGTTA TGGTTGTTTT ATTTACGTTA GTTTTGTTAT TTGTGTTTTA
CEL     ---------- ---------- ----A.T..A ..GC.AT.A. .A...T.AG. G......C.. .....A.... ...CT.....
MIN     ---------- ---------- ---------- -------A. .....TTG.T A....T.CT. .A..A....T
RCU     ---------- ---------- ----A..ACA ...A.AA..T .AT.AA.AG. .A...TT..G A.CA.A..TA .AAAT..AAT

ASU     TATTGGTAAT TTTGTTTTGA GATGTAAGGA TTTTTATAAG AATAAGATTT CTTCTTTTGA GTGTGGTTTT GTTAGGATTG
CEL     .T.GAT.... ...T.A..A. ..AT...... .A.AGGA..A .....A...A GAG.G..... A......... ..A..AG...
MIN     .T...TA.C. ...T....T. ATAAA..AA. AAA.A.GTTA .TG..A.A.. ...A...C.. AA.A...... AA.TATT.G.
RCU     .T..AAAG.A ..CTA...T- ---------- -..AA....A .CC.AT... T...TA..... A......... .AGTCA...A

ASU     GTAAAATTCA GAATTCTTTT AGTATTCATT TTTTTATTAT GATGCTTATG TTTGTTATTT TTGATTTGGA GGTGGTGATG
CEL     .A........ A......... ..A....... .......... ...AT.G..A .......... ......A.AA AA.T..T...
MIN     .A...T.A.T ATT.AG..A. ..A....... ......A.. T..TT.A..T ...A..T.A. .....C.... AT.AT.TT.A
RCU     .ATGGTA.TT T...AA.A.. ....A..... ......AA.. TGGCA....T ...A...C.. ....C..A.. .T.AA..TAC

ASU     TTTCTTGGGA TTTTGGTTTC TGATTTAAAT TCTTTGATTA GATTTTTTAT GTTGTTAATA TTTATTTTTG GTGGTTTTTA
CEL     ...T.A..T. ....A..A.. A........G. ..G.AT..C. .G.....A.. AA.A..C..C ..C..C..G. .A..A.....
MIN     ...T.ATTT. ...ATT..AA .TTAAATTTG AT..ATTGA. T.A.....T. A..AA.T..T ......A.GA TAAC...AGT
RCU     ...T.AT--- ..C..TA...T .A......T. AA.A.A.AA. T.A...GGT. AA.T..T.AC .........T T.ACA..A.T

ASU     TATGGAGTGG TGGTATGGTA AGTTGGTTTG GTTGATT---
CEL     C..A.....A ..A........ .A..A..... AG.A...TA-
MIN     .T....A..A AAA...AT.. .A..AA.... A..TT.GTAG
RCU     .C.A..A.TT .TT.TAAA.. CAC..TC... AAAT..GTAG
```

CYTOCHROME B

```
ASU     ---------- -----ATTAA GTTGGATTTT GTTAATTCTA TGGTTGTTAG GTTACCGTCT AGTAAGGTTT TAACTTATGG
CEL     TTGAAAATTA ATAAT.GATT A..AA..... ......GGG. ..T.G..G.C A..G..A... ..A..AAC.. ....A.TAA.
MIN     ---ATTAATT TGATTT..GT AGAAA.A... T....AAAA. .AT.AAAGTT AAATGGAG.A GA.TTTA... .G.A....TT
RCU     ---------- -----..... T..TTTAA.A AA...G..AT .AAAAA...C A...TTT... CCC.TAAA.. ...AC...T.

ASU     TTGGAATTTT GGTAGTATGT TGGGCATGGT TTTGGGTTTT CAGATTTTGA CTGGTACTTT TTTGGCTTTT TATTATTCTA
CEL     A..A...... ........A. ....T..A.. ...AATC... .........A. .A.....A.. ...A..A... .......A.GC
MIN     ..TT...... ......T.T. .A..A..TA. ...TTT.A.G ..AG.AA.TT .....TTA.. AA..AAG...A .T...GAAT
RCU     A..A....A.. ...A..AT.A. .AATT..AT. AA.ATT.A.. ..A.....AT ....ATT... A...A.A..A .T....G---

ASU     ATGATGGTGC TTTGGCCTTT TTGAGTGTTC AATACATTAT ATATGAAGTT AATTTTGGTT GGATTTTTCG TGTTTTACAT
CEL     CC...A.GTT AA.A..A... .CA.CA..G. .G..T..... G.....G..A .........A. .AG.A..... AA....T...
MIN     C.AGAA.AAA ...TT.T... GATTCAA... ....TT.A.. .AT...G... ...AA..... .T.AA..... .T.AA.T...
RCU     ..A.CA.A.A AACCT.A... AAA..AT.AT G..TA..CCA TAT...... TT......C. CA..A..A.A .TA.G.T...

ASU     TTTAATGGTG CTAGTTTGTT TTTTATTTTT TTGTATTTGC ATTTATTTAA GGGATTGTTT TTTATGAGTT AT---CGTTT
CEL     .........G. .C..G..A.. .......... ........A. ..A.T..... A..G..A... .....A.A. ..---.....
MIN     ......ATA. T.TC...A.. ......A..A G.TATA..A. ..A.....A.. ..CT..A... .A.T.T..A. T.---.....
RCU     ..A...TT.T ..TC...TA. ....T.A... A.T...C.T. .....A.... A.....TG.. .A..ATTCA. ....AAA...

ASU     GAAGAAGGTT TGGGTATCTG GTATTGTAAT TCTTCTTTTG GTTATAATGG AGGCT
CEL     A..A..A..A ..AA.G.... ..T.AAC... .TA.T.A..A ..A.....A. .A...
MIN     A..A..A... ...AAATAT. ..T.AT.... .T..T.GA.A T.A...T... .A...
RCU     A..ACTT... ...AA...TA. ..TGAT.... .T.AA....A C.A.....AA TCT.A
```

FIG. 2. Nucleotide sequence of the NADH dehydrogenase subunit 3 gene and the 5′ end of the cytochrome *b* gene from *Ascaris suum* (ASU) *Caenorhabditis elegans* (CEL), *Meloidogyne incognita* (MIN), and *Romanomermis culicivorax* (RCU). Dots indicate nucleotide identity with ASU; dashes are inferred deletions. Data for *A. suum* and *C. elegans* are from another report (28). Alignments were conducted by the multiple alignment program Pileup from the Genetics Computer Group Software package (11) using a gap weight of 5.0 and gap length weight of 0.3.

and 960 bp long and were previously the shortest metazoan 16S genes reported (28). The *R. culicivorax* 16S rRNA gene is also 960 bp long. In contrast, the *M. incognita* 16S rRNA gene appears to be 838 bp in size. In this species, the exact start of the gene was determined with difficulty because noncoding intergenic sequences of varying length immediately precede this gene in many obligatory mitotic parthenogenetic species in the genus (30). However,

by sequencing across the junction of the upstream COII gene into the 5′ portion of the 16S rRNA in *M. hapla* and other *Meloidogyne* species without this intergenic mtDNA sequence, we tentatively identified and assigned the beginning of the 16S rRNA gene. In *M. hapla*, *M. chitwoodi*, and *M. naasi*, the TAG termination codon of the COII gene is immediately followed by a conserved region of 21–23 identical nucleotides among the four *Meloidogyne* spe-

cies (Powers, unpubl.), a region that we interpret to be the beginning of the 16S rRNA gene.

Identification of the *R. culicivorax* 16S rRNA gene was achieved by complete sequencing of an 1,123-bp region between the genes for cytochrome oxidase subunit I and NADH dehydrogenase subunit 5 (Azevedo and Hyman, unpubl.). Large open reading frames were not found in either strand. Computer-assisted alignments revealed this region to contain several highly conserved domains (Fig. 1, boxes A–F) characteristic for 16S RNA genes from phylogenetically distant metazoans (5,9).

In the 5' region of the 16S rRNA gene, there is little obvious sequence homology among *R. culicivorax, M. incognita, C. elegans,* and *A. suum.* Only the 3' half of the gene contains blocks of sequence conserved among all nematodes. At nucleotide 10,844 (by the numbering system of Okimoto et al. (28) for the *C. elegans* mitochondrial genome), 13 of 15 nucleotides are shared among the four taxa (Fig. 1, box a), and seven nucleotides downstream exists another shared region of 10 to 13 nucleotides. Several similarly conserved blocks are located further into the 3' region of the gene (Fig. 1, boxes c–f). The extremely A–T rich composition (87%) of the gene is especially notable in *M. incognita.* At 176 residues from the start of the

gene, there is a string of 154 uninterrupted A–Ts. However, each of the blocks conserved among the four nematode mtDNAs are high in G–C composition relative to the remainder of the gene.

The uncorrected (i.e., observed and without considering gaps) nucleotide similarity for the 16S rRNA gene is between 51 and 74% for the four nematode taxa (Table 1). Reliability of these values as an estimator of evolutionary divergence is diminished because of the considerable A–T nucleotide bias plus the uncertainty of the alignments. Nonetheless, these values do indicate that the relative similarity (74%) between the 16S rRNA genes of *A. suum* and *C. elegans* is far greater than that observed among the other species.

Inferred amino acid alignments for ND3 and the first 120 amino acids of cytochrome *b* are presented in Figure 3. These regions can be aligned with greater certainty than the 16S rRNA gene because of the presence of blocks of conserved amino acids. In ND3, residues 38–42 relative to the *C. elegans* sequence are conserved among all nematode taxa except for a cysteine replaced by serine at site 40 in *M. incognita.* Similarly, residues 55–57 and 66–69 are conserved among the four species. These same regions are conserved among ND3 polypeptides from other widely separated taxa, including *Piaster,* an echinoderm thought to have diverged in

TABLE 1. Uncorrected (observed) amino acid and nucleotide sequence divergence for pairwise comparisons of *Caenorhabditis elegans* (Ce), *Ascaris suum* (As), *Meloidogyne incognita* (Mi), and *Romanomermis culicivorax* (Rc).

	ND3†			Cyt B			LrRNA
	aa	nuc	Ti/Tv	aa	nuc	Ti/Tv	nuc
Ce vs As	79	74	40/46	79	72	52/48	74
Ce vs Mi	39	57	35/99	42	59	88/113	(60)
Ce vs Rc	27	54	44/100	39	55	37/116	(55)
Mi vs As	41	56	36/100	42	57	42/110	(57)
Mi vs Rc	30	58	30/96	38	55	36/100	(53)
As vs Rc	27	55	49/94	41	56	30/96	(51)
Ce vs mouse‡	26	44		44	53	49/94	

† Percentage amino acid (aa) and nucleotide (nuc) similarity are presented for the entire NADH dehydrogenase subunit 3 (ND3), the large ribosomal RNA gene (LrRNA), and the initial 5' 120 amino acids (360 nucleotides) of the cytochrome *b* (CytB) gene. Ti/Tv is the number of transitions (Ti) and transversions (Tv) in each comparison. Parentheses around the LrRNA comparisons indicate uncertainty in alignment.
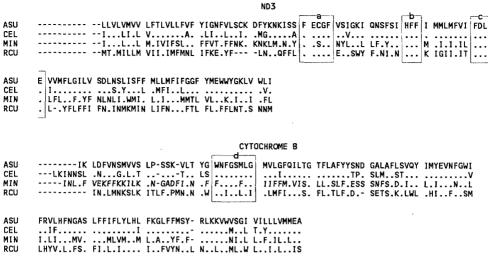‡ Mouse data are from Bibb et al. (5)

ND3

```
                                                          ┌───a───┐        ┌─b─┐      ┌─c─┐
ASU   ---------- --LLVLVMVV LFTLVLLFVF YIGNFVLSCK DFYKNKISS F ECGF VSIGKI QNSFSI HFF I MMLMFVI FDL
CEL   ---------- --I...LI.L V........A. .LI..L..I. .MG.....A . ....V V...... ...  . ....... ...
MIN   ---------- --I...L..L M.IVIFSL.. FFVT.FFNK. KNKLM.N.Y . .S..NYL..L LF.Y.. ...M .I.I.IL ...
RCU   ---------- --MT.MILLM VII.IMFMNL IFKE.YF--- -LN..QFFL . ....E..SWY F.NI.N ...K IGII.IT ...


ASU   E VVMFLGILV SDLNSLISFF MLLMFIFGGF YMEWWYGKLV WLI
CEL   . I........ ...S.Y...L .MFI..L... .......... .V.
MIN   . LFL..F.YF NLNLI.WMI. L.I...MMTL VL..K.I..I .FL
RCU   . L-.YFLFFI FN.INMKMIN LIFN...FTL FL.FFLNT.S NNM
```

CYTOCHROME B

```
                                           ┌──d──┐
ASU   --------IK LDFVNSMVVS LP-SSK-VLT YG WNFGSMLG MVLGFQILTG TFLAFYYSND GALAFLSVQY IMYEVNFGWI
CEL   ---LKINNSL .N...G.L.T ..-...-T.. LS ........ ...I...... .......TP. SLM..ST... .........V
MIN   -----INL.F VEKFFKKILK .N-GADFI.N .F F....F.. IIFFM.VIS. LL.SLF.ESS SNFS.D.I.. L.I...N..L
RCU   ---------- IN.LMNKSLK ITLF.PMN.N .W ..I..L.I .LMFI...S. FL.TLF.D.- SETS.K.LWL .HI..F..SM


ASU   FRVLHFNGAS LFFIFLYLHL FKGLFFMSY- RLKKVWVSGI VILLLVMMEA
CEL   ..IF...... .........I .........- ......M..L T.Y.......
MIN   I.LI...MV. ...MLVM..M L.A..YF.F- ......NI.L L.F.IL.L..
RCU   LHYV.L.FS. FI.L.I.... I..FVYN..L N..L..ML.W L..I.L..IS
```

FIG. 3. Amino acid alignment of the NADH dehydrogenase subunit 3 gene and the 5' portion of the cytochrome *b* gene from *Ascaris suum* (ASU) *Caenorhabditis elegans* (CEL), *Meloidogyne incognita* (MIN), and *Romanomermis culicivorax* (RCU). Dots indicate nucleotide identity with ASU; dashes are inferred deletions. Boxes identify regions of relatively high conservation among all species. Data for *A. suum* and *C. elegans* are from another report (28). Alignments were conducted by the multiple alignment program Pileup from the Genetics Computer Group Software package (11) using a gap weight of 5.0 and gap length weight of 0.3.

the same evolutionary time frame as Secernentean nematodes (35).

The first 120 amino acids of cytochrome *b* display a slightly higher degree of conservation than the complete ND3 proteins. Particularly notable is the region from residues 28–35, which appears homologous to the Q1 redox center functionally involved in electron transport (Fig. 3, box d) (16).

Uncorrected divergence estimates for the ND3 and cytochrome *b* genes and their encoded polypeptides from the four nematode taxa are presented in Table 1. In both genes, *C. elegans* and *A. suum* share nearly twice as many amino acids than any other pairwise comparison. The ND3 amino acid similarity of 79% between these two species is dramatically higher than the next highest value, 41% for *M. incognita* vs *A. suum*. Cytochrome *b*, which evolves at a slower rate than ND3 (6), also displays exceptionally low amino acid similarity for all comparisons involving *M. incognita* and *R. culicivorax*.

The transition/tranversion ratio (Ti/Tv) is another means for evaluating the extent of mutational saturation between two taxa (34). Typically for mitochondrial DNA sequences, transitions initially accumulate at an approximately 10-fold higher rate than transversions (6,34). As the age of divergence increases, the percentage of tranversions gradually increases until they predominate as the major class of substitution. For the sequence comparisons presented here, only the *C. elegans–A. suum* comparisons display transition/transversion ratios within a range indicating that divergence times can be calculated realistically.

For the two protein coding genes, there were 45 invariant and 172 polymorphic amino acid sites, excluding positions where there were missing data due to gaps and unequal gene lengths. Only 21 of the polymorphic positions were "phylogenetically informative" (34), i.e., the amino acids were shared by two taxa but differed from a common amino acid shared by the other two taxa. Each of the phylogenetically informative sites supported a *C. elegans–A. suum* clade. Not a single site supported a *C. elegans–M. incognita* or a *C. elegans–R. culicivorax* grouping, consistent with affinities perviously generated using mitochondrial tRNA and codon usage analysis (38).

Comparative sequence data from the two protein coding genes and the 16S

rRNA genes indicate extremely ancient roots for the separation of basic nematode lineages. Conversely, a relatively recent divergence is supported for the ascarid–rhabditid lineages compared to the lineage leading to the tylenchid plant parasites. The ascarid–rhabditid split is estimated to have occurred approximately 65 million years ago, assuming the rate of nematode mitochondrial evolution is equivalent to mammals (28). However, a slow rate of evolution could also explain the genetic similarity between *C. elegans* and *A. suum*. We consider this possibility unlikely in light of the extreme level of divergence observed when mitochondrial genes from these two nematodes are compared with those derived from *M. incognita* and *R. culicivorax*. These alignments argue against rate deceleration among the Nematoda as the sole explanation for the "close" phylogenetic relationship of *C. elegans* and *A. suum* as assessed by mtDNA analysis. Furthermore, an examination of the mitochondrial COII gene in *C. elegans* and other *Caenorhabditis* species has demonstrated high rates of mitochondrial evolution relative to nuclear rates, as observed in other animal systems (6,7,36,37).

The mitochondrial genomes of *A. suum* and *C. elegans* exhibit A + T compositions of 70.4 and 75.5%, respectively (28). Convergent evolution of nucleotide bias can artifactually support certain phylogenetic affinities, a situation that has confounded the apparent relationship between birds and mammals (13,25). However, several independent lines of evidence argue that *A. suum* and *C. elegans* are the most closely related among the nematode genera characterized in this study. Amino acid alignment of two different mitochondrial genes reveals levels of similarity significantly elevated over any other pairwise comparison (Table 1); nucleotide frequencies in codon positions one and two are nearly identical only among *A. suum* and *C. elegans* mitochondrial protein coding genes (28; Azevedo and Hyman, unpubl.); Ti/Tv ratios are informative only in *A. suum–C. elegans* comparisons (Table 1); and *R. culicivorax*

and *M. incognita* both exhibit mtDNA A + T contents of over 80%, but otherwise are dissimilar in amino acid similarity and codon composition.

The three Secernentean subclasses considered here have previously been classified based upon behavorial and anatomical definitions. Spiruria is characterized, in part, as a wholly parasitic group possessing an esophagus without valves. The most distinctive character of Rhabditia is the uniquely valved esophagus or its derivities indicative of an ancestry involving bacterial feeding. Members of Diplogasteria are generally distinguished by a three-part esophagus, which is often associated with movable stomatal armature. The Secernentean plant parasites are included in this group. The distinction of ancient lineages is also emphasized in a hypothetical evolutionary tree, which depicts separation of the basic Secernentean lineages occurring in the Cambrian period over 550 million years ago (29). From a systematic perspective, the relatively close relationship between *C. elegans* and *A. suum* indicates that either *Ascaris* is misplaced in the Spiruria or that the subclass divergences are highly unequal. In order to address this question, additional members of each subclass require examination.

Dating the diplogasterid split from the other secernenteans will not be a simple matter using the mitochondrial sequence data from the genes analyzed here. For *Meloidogyne* and *Romanomermis*, homologies among the genes are barely recognizable. Even by the most conservative estimates, Ti/Tv ratios, amino acid similarity, and 16S rRNA gene comparisons, the divergence of the basic lineages in Secernentea are too old for resolution. In contrast, *Piaster* and *Asterina*, two sea stars from different Echinoderm orders that diverged over 500 million years ago, have an amino acid sequence identity of 70.9% for ND3, and the *Piaster–Drosophila yakuba* comparison for the same gene is 55.6%. All the nematode ND3 amino acid comparisons (except *C. elegans–A. suum*) were ≤41%. Importantly, there are several highly con-

served amino acid residues in the ND3 gene that are shared among virtually all invertebrates. These include residues ECGF at position 39–42 relative to the echinoderm alignment (35), FF at 56–57, and FDLE at 66–69 (Fig. 3, boxes a–c).

This extreme divergence between the lineages represented by *C. elegans* and *Meloidogyne* has implications beyond that of systematics. Many techniques in molecular biology operate under an assumption of nucleotide similarity in order to isolate genes from organisms with a limited history of genetic investigation. Given the wealth of DNA sequence data generated by research with *C. elegans*, a natural starting point for isolation of a gene from a plant-parasitic nematode might be the *C. elegans* database. Indeed, cloned DNA segments from *C. elegans* have been successfully applied in several cases as hybridization probes for the selection of genes from plant parasites (4,10). Yet, as is evident from these comparisons, successful application of a *C. elegans* gene in a search for homologous coding sequences from a plant parasite should not be expected just because the source is another nematode. Sequence comparisons across a broad taxonomic range could improve strategies for identification of genes of interest from plant-parasitic nematodes.

LITERATURE CITED

1. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local sequence alignment tool. Journal of Molecular Biology 215:403–410.

2. Azevedo, J. L. B., and B. C. Hyman. 1993. Molecular characterization of lengthy mitochondrial DNA duplications from the parasitic nematode *Romanomermis culicivorax*. Genetics 133:933–942.

3. Ballard, J. W. O., G. J. Olsen, D. P. Faith, W. A. Odgers, D. M. Rowell, and P. W. Atkinson. 1992. Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods. Science 258:1345–1348.

4. Beckenbach, K., M. J. Smith, and J. M. Webster. 1992. Taxonomic affinities and intra- and interspecific variation in *Bursaphelenchus* spp. as determined by polymerase chain reaction. Journal of Nematology 24:140–147.

5. Bibb, M. J., R. A. Van Etten, C. T. Wright, M. W. Walberg, and D. A. Clayton. 1981. Sequence and gene organization of mouse mitochondrial DNA. Cell 26:167–180.

6. Brown, W. M. 1985. The mitochondrial genome of animals. Pp. 62–68 *in* M. Nei and R. K. Koehn, eds. Evolution of genes and proteins. Sutherland, MA: Sinauer.

7. Brown, W. M., E. M. Prager, A. Wang, and A. C. Wilson. 1982. Mitochondrial DNA sequence of primates: Tempo and mode of evolution. Journal of Molecular Evolution 18:255–239.

8. Chou, P. V., and G. D. Fasman. 1978. Prediction of the secondary structure of proteins from their amino acid structure. Advances in Enzymology 47:45–147.

9. Clary, D. O., and D. R. Wolstenholme. 1985. The ribosomal RNA genes of *Drosophila* mitochondrial DNA. Nucleic Acids Research 13:4029–4045.

10. Curran, J., and J. M. Webster. 1987. Identification of nematodes using restriction fragment length differences and species-specific DNA probes. Canadian Journal of Plant Pathology 9:162–166.

11. Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Research 12:387–395.

12. Harris, T. S. 1990. Identification of root-knot juveniles by PCR. M.S. thesis, University of Nebraska, Lincoln.

13. Hedges, S. B., K. D. Moberg, and L. R. Maxsom. 1990. Tetrapod phylogeny inferred from 18S and 28S ribosomal RNA sequences and a review of the evidence for amniote relationships. Molecular Biology and Evolution 7:607–633.

14. Henikoff, S. 1987. Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods in Enzymology 155:156–165.

15. Higgans, D. G., and P. M. Sharp. 1988. CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. Gene 73:237–244.

16. Irwin, D. M., T. D. Kocher, and A. C. Wilson. 1991. Evolution of the cytochrome B gene in mammals. Journal of Molecular Evolution 32:128–144.

17. Kocher, T. D., W. K. Thomas, A. Meyer, S. V. Edwards, S. Paabo, F. X. Villablanca, and A. C. Wilson. 1989. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. Proceedings of the National Academy of Sciences USA 86:6196–6200.

18. Kreiz, P. L., and D. M. Short. 1989. Gigapack II: Restriction free (hsd⁻, mcrA⁻, mcrB⁻, mrr⁻) lambda packaging extracts. Strategies 2:25–26.

19. Kyte, D., and R. F. Doolittle. 1982. A single method for displaying the hydropathic character of a protein. Journal of Molecular Biology 157:332–348.

20. Lipman, D. J., and W. R. Pearson. 1985. Rapid and sensitive protein similarity searches. Science 227:1435–1441.

21. Liu, H., and A. T. Beckenbach. 1992. Evolution of the mitochondrial cytochrome oxidase II gene among 10 orders of insects. Molecular Phylogenetics and Evolution 1:41–52.

22. Maggenti, A. R. 1982. Pp. 879–929 *in* S. P. Parker, ed. Synopsis and classification of living organisms, vol. 1. New York: McGraw-Hill.

23. Maggenti, A. R. 1983. Nematode higher classi-

fication as influenced by species and family concepts. Pp. 25–40 *in* A. R. Stone, H. M. Platt, and L. K. Khalil, eds. Concepts in nematode systematics. London: Academic Press.

24. Maggenti, A. R. 1991. Nemata: Higher classification. Pp. 147–187 *in* W. R. Nickle, ed. Manual of agricultural nematology. New York: Marcel Dekker.

25. Marshall, C. R. 1992. Substitution bias, weighted parsimony and amniote phylogeny as inferred from 18S rRNA sequences. Molecular Biology and Evolution 9:370–373.

26. Mierendorf, R. C., and D. Pfeffer. 1987. Direct sequencing of denatured plasmid DNA. Methods in Enzymology 152:556–552.

27. Mindell, D. P., and R. L. Honeycutt. 1990. Ribosomal RNA in vertebrates: Evolution and phylogenetic applications. Annual Review of Ecology and Systematics 21:541–566.

28. 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.

29. Poinar, G. O., Jr. 1983. The natural history of nematodes. Englewood Cliffs, NJ: Prentice-Hall.

30. Powers, T. O., and T. S. Harris. 1993. A polymerase chain reaction method for identification of five major *Meloidogyne* species. Journal of Nematology 25:1–6.

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

32. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences, USA 74:5463–5467.

33. Short, J. M., J. M. Fernansby, D. A. Sorge, and W. D. Huse. 1988. Lambda Zap: A bacteriophage lambda expression vector with in vivo excision properties. Nucleic Acids Research 16:7583–7600.

34. Simons, C. 1991. Molecular systematics at the species boundary: Exploiting conserved and variable regions of the mitochondrial genome of animals via direct sequencing from amplified DNA. Pp. 33–71 *in* G. M. Hewitt, A. W. B. Johnson, and J. P. W. Young, eds. Molecular techniques in taxonomy. Berlin: Springer-Verlag.

35. Smith, M. J., D. K. Banfield, K. Doteval, S. Gorski, and D. J. Kowbel. 1990. Nucleotide sequence of nine protein-coding genes and 22 tRNAs in the mitochondrial DNA of the sea star *Piaster ochraceus.* Journal of Molecular Evolution 31:195–204.

36. 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.

37. Wilson, A. C., H. Ochman, and E. M. Prager. 1987. Molecular time scale for evolution. Trends in Genetics 3:241–247.

38. Wolstenholme, D. R. 1992. Animal mitochondrial DNA: Structure and evolution. International Review of Cytology 141:173–216.