srRNA Evolution and Phylogenetic Relationships of the Genus *Naegleria* (Protista: Rhizopoda)

P. R. Baverstock, S. Illana, P. E. Christy, B. S. Robinson, and A. M. Johnson

*Evolutionary Biology Unit, South Australian Museum; †Unit of Clinical Microbiology, School of Medicine, Flinders University of South Australia; and ‡Protozoology Section, State Water Laboratory, South Australian Engineering and Water Supply

A rapid RNA sequencing technique was used to partially sequence the small-subunit ribosomal RNA (srRNA) of four species of the amoeboid genus *Naegleria*. The extent of nucleotide sequence divergence between the two most divergent species was roughly similar to that found between mammals and frogs. However, the pattern of variation among the *Naegleria* species was quite different from that found for those species of tetrapods characterized to date. A phylogenetic analysis of the consensus *Naegleria* sequence showed that *Naegleria* was not monophyletic with either *Acanthamoeba castellanii* or *Dictyostelium discoideum*, two other amoebas for which sequences were available. It was shown that the semiconserved regions of the srRNA molecule evolve in a clocklike fashion and that the clock is time dependent rather than generation dependent.

Introduction

The nucleotide sequence of the small-subunit ribosomal RNA (srRNA) (16S–18S) is proving extremely valuable for probing phylogenetic relationships among distant relatives because many regions remain conserved or semiconserved over large periods of time. Indeed, homologous sequences are evident between eukaryotes and prokaryotes, with homology extending even to mitochondrial and chloroplast genomes (Huysmans and De Wachter 1985).

In contrast, some regions show relatively high levels of variation. The distribution and extent of this latter class of variation are, however, poorly understood, being based exclusively on sequence data for tetrapods (Gonzalez and Schmickel 1986). All other sequences so far available are from organisms that are distantly related (Huysmans and De Wachter 1985; Johnson et al. 1987, 1988). Among the five tetrapods (human, mouse, rat, rabbit, and frog) for which sequences are available, mouse and rat show 9–16 nucleotide differences, (two slightly different rat sequences are given in Huysmans and De Wachter 1985), while human and rodent differ by 14–17 bases. Mouse and rat are believed to have diverged ∼20 Mya (Sarich 1985), while rodents and primates diverged ∼80 Mya (Goodman et al. 1982). This gives an average rate of divergence of ∼1 nucleotide substitution/4 Myr for these regions. The rabbit sequence differs from those of the other three mammals at 55–70 bases. Most of the variation among the mammals is restricted to two short regions, one within the region of bases 194–
Table 1
Origin of \textit{Naegleria} Strains Used in the Present Study

<table>
<thead>
<tr>
<th>Species</th>
<th>State Water Laboratory Serial No.</th>
<th>Cross Reference</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{N. fowleri}</td>
<td>NG-052</td>
<td>.</td>
<td>Water supply, Taplan, South Australia</td>
</tr>
<tr>
<td>\textit{N. lovaniensis}</td>
<td>NG-051</td>
<td>.</td>
<td>Water supply, Wudinna, South Australia</td>
</tr>
<tr>
<td>\textit{N. australiensis}</td>
<td>NG-035</td>
<td>ATCC3098</td>
<td>Port Pirie, South Australia</td>
</tr>
<tr>
<td>\textit{N. gruberi}</td>
<td>NG-004</td>
<td>CCAP1518/1e</td>
<td>Pond, Wisconsin</td>
</tr>
</tbody>
</table>

334 and one within the region of bases 1754–1782 (human numbering system), which are referred to by Gonzalez and Schmickel (1986) as the V3 region and the V8 region, respectively. The frog, \textit{Xenopus laevis}, differs from the mammals at about 100 sites (Gonzalez and Schmickel 1986), many of which occur in the V3 and V8 regions.

Woese (1987) has suggested that the ribosomal RNAs may be the ultimate molecular clocks. There is a large body of evidence that suggests that the rate of molecular evolution in general is relatively constant with time (Wilson et al. 1977; Ayala 1986), although much of this evidence is based on analyses of molecules evolving within taxa at levels lower than those within the phylum. Because the srRNA molecule is common to virtually all life forms, it becomes possible to test the clock for a single molecule across many phyla.

The present study addresses srRNA evolution in the genus \textit{Naegleria}. \textit{Naegleria} is a genus of free-living amoeboflagellates in the family Vahlkampfiidae which is placed at present in the protistan phylum Rhizopoda. Its phylogeny is of interest because the transformation from amoeba to flagellate form has been studied extensively in one species, \textit{N. gruberi}, as a model of cell differentiation, and because another species, \textit{N. fowleri}, is an important opportunistic pathogen of man. Two other members of the phylum Rhizopoda for which srRNA sequences are available are the soil amoeba \textit{Acanthamoeba castellanii} and the slime mold \textit{Dictyostelium discoideum}.

This study was undertaken with the following three questions in mind: 1. What is the pattern of base substitution in a group of closely related protozoa? 2. What are the phylogenetic relationships of \textit{Naegleria}, especially with respect to other species of amoeba for which srRNA sequences are available? 3. Does the srRNA gene evolve in a clocklike manner?

Material and Methods

The four strains of \textit{Naegleria} used are listed in table 1. NG-051 and NG-052 were isolates from recent field studies and were used because they adapted readily to axenic culture with high yields. NG-052 was identified as \textit{N. fowleri} on the basis of zymogram comparison with clinical isolates. NG-051 was identified as \textit{N. lovaniensis} by electrophoretic comparison with the reference strain Aq/q/1/45D from de Jonckheere (Institute of Hygiene and Epidemiology, Brussels) but differed at some variable allozyme loci. NG-035 is the type strain of \textit{N. australiensis}. NG-004 has been characterized as a “smooth” strain of \textit{N. gruberi}, although some doubt exists about its provenance (Page 1975). Strains were cultured axenically in 50 ml PYNFH medium (SCGYEM for \textit{N. fowleri}) in 1-liter glass tissue-culture bottles (media recipes as in American Type Culture Collection 1984). Trophozoites were harvested every 48–72
h, depending on the strain; were pelleted by centrifugation at 5,000 g; were resuspended in 10 vol 0.6 M guanidine hydrochloride, 0.2 M sodium acetate pH 5.4, 1 mM 2-mercaptoethanol; and were stored at -20°C. The harvested cells were pooled progressively until (1.5–2.5) × 10^8 cells of each strain had been accumulated.

The sequencing of srRNA in the *Naegleria* species followed the method of Lane et al. (1985) with the modifications given in Johnson et al. (1987). In brief, cellular RNA was extracted and purified by a modification of the method of Brooker et al. (1980). The srRNA sequencing method of Lane et al. (1985) relies on the observation that three regions of the srRNA (referred to, respectively, as the A, B, and C regions) are highly conserved in evolution, even across kingdoms. DNA sequences complementary to these conserved regions were used as primers in three separate dideoxynucleotide sequencing reactions using reverse transcriptase.

To facilitate direct comparison among the *Naegleria* species, sequences for the four species were run on the same gels, so that postulated base substitutions could be compared directly. For each reverse transcription, gels were run for both 2.5 h and 6.0 h to extend the readable sequence. A minimum of three sequencing reactions were performed for each species for each primer.

For the purposes of phylogenetic analysis, the *Naegleria* sequence was aligned by eye with srRNA sequences from five vertebrates (*Homo*, *Mus*, *Rattus*, *Oryctolagus*, and *Xenopus*), an invertebrate (*Artemia*), two fungi (*Saccharomyces* and *Neurospora*), five ciliates (*Oxytricha*, *Stylonchus*, *Tetrahymena*, *Paramecium*, and *Euplotes*), three plants (*Oryza*, *Zea*, and *Glycine*), two flagellates (*Trypanosoma* and *Euglena*), a dinoflagellate (*Prorocentrum*), a microsporidium (*Vairimorpha*), and two other genera of amoeba (*Acanthamoeba* and *Dictyostelium*) (see Johnson et al. 1987 for sources; also see Vossbrinck et al. 1987, for *Vairimorpha*; Gunderson and Sogin 1986, for *Acanthamoeba*).

Phylogenetic analysis was restricted to the semiconserved regions, since the more variable regions provide no useful phylogenetic information at this level of analysis.

The aligned data were analyzed phylogenetically by two general methods—character-state analysis and distance analysis. The former treats each nucleotide site as a character, with four possible character states (A, G, C, or U), while the latter uses genetic distances between taxa, calculated as \( K_{\text{Nuc}} \) values (Jukes and Cantor 1969). The character-state data were analyzed by the parsimony method, using DNAPARS in J. Felsenstein’s PHYLIP package, version 3.0. Parsimony (Felsenstein 1982) aims to find a tree with the minimum number of total postulated changes. The robustness of the parsimony analysis was tested by bootstrapping (Felsenstein 1985), using DNA-BOOT in PHYLIP version 3.0. Bootstrapping randomly samples subsets of the data, with the idea that nodes on the tree strongly supported by the data will also be supported by subsets of the data. The distance data were analyzed by two methods, neither of which assumes constant rates of evolution. These were the Fitch and Margoliash (1967) method, using FITCH in PHYLIP version 3.0, and the distance Wagner procedure (Farris 1972). The former aims for a tree with topology and branch lengths that minimize the discrepancy between the input and output data, while the latter aims for a tree of minimum total length. Both methods yield trees with branch lengths that reflect the actual amount of change along each branch. PHYLIP version 3.0 performs branch swapping in an effort to find better trees but does not output any but the best or equally best trees found.
FIG. 1—srRNA sequences for the A, B, and C sequences of *Naegleria fowleri* (NFOWL), *N. lovaniensis* (NLOV), *N. australiensis* (NAUS), and *N. gruberi* (NGRUB). A dash (−) = gap; an asterisk (*) = no data; a dot (·) = same as *N. lovaniensis* for the A and B sequences, same as *N. gruberi* for the C sequence; N = nucleotide of undetermined type. The numbering system follows that of Gonzalez and Schmickel (1986) for the proposed homologous regions for the human sequence.
Table 2
Number of Base Differences among the Four *Naegleria* Species at 278 Sites
Common to All Species

<table>
<thead>
<tr>
<th>SPECIES</th>
<th>N. fowleri</th>
<th>N. lovaniensis</th>
<th>N. australiensis</th>
<th>N. gruberi</th>
</tr>
</thead>
<tbody>
<tr>
<td>N. fowleri</td>
<td>. . . . . . .</td>
<td>2</td>
<td>. . . . . . .</td>
<td>. . . . . . .</td>
</tr>
<tr>
<td>N. lovaniensis</td>
<td>. . . . . . .</td>
<td>. . . . . . .</td>
<td>7</td>
<td>. . . . . . .</td>
</tr>
<tr>
<td>N. australiensis</td>
<td>10</td>
<td>7</td>
<td>. . . . . . .</td>
<td>. . . . . . .</td>
</tr>
<tr>
<td>N. gruberi</td>
<td>17</td>
<td>15</td>
<td>9</td>
<td>. . . . . . .</td>
</tr>
</tbody>
</table>

Results

A. Comparisons among *Naegleria* Species

The reverse transcriptase method that was used yields base sequences for regions 5' to the three primers—the A primer, the B primer, and the C primer. The sequences are referred to here as the A sequence, B sequence and C sequence, respectively. Figure 1 shows the A, B, and C sequences obtained in this study for the four *Naegleria* species.

The reverse transcriptase method of srRNA sequencing has an error rate of <1% (Lane et al. 1985). The major component of the error is due to "false" insertions and deletions. Therefore, in aligning the sequences for the four species, gaps were postulated except where it was quite clear that base substitutions were involved.

When the sequences for the four *Naegleria* species were aligned in this way, a striking pattern emerged—only two base-substitution differences were evident in the C sequence, five in the B sequence, and 47 in the A sequence (fig. 1). Moreover, the base substitutions in the A sequence were concentrated in the 60 bases immediately 5' to the A primer.

To quantitatively assess the extent of divergence among the four species, it was necessary to restrict the analysis to sequences common to all four. There were 116 sites in common for the A sequence, 82 for the B sequence, and 80 for the C sequence. Table 2 shows the total number of base differences detected among the four species at these 278 sites. Here, *N. fowleri* and *N. lovaniensis* differ at only two sites, *N. australiensis* differs from them at 7–10 sites, while *N. gruberi* differs from the remaining three species at 9–17 sites.

The only comparable data set available is that for the tetrapods (four mammalian species and a frog). The numbers of base differences in these same 278 sites are shown in table 3. The extent of base divergence between the most divergent *Naegleria* species is higher than that among the mammals but is similar to that among the tetrapods (table 4).

B. Comparison of *Naegleria* Sequences with Other Sequences

For the purposes of comparing the sequence of the genus *Naegleria* with other available sequences, the A and B sequences of *N. lovaniensis* and the C sequence of *N. gruberi* were used. This was done because more of the C sequence of *N. gruberi* than of *N. lovaniensis* was available (fig. 1). In terms of placing *Naegleria* in a total phylogenetic framework, such a procedure is justified so long as *N. lovaniensis* and *N. gruberi* are monophyletic to the exclusion of all other taxa for which sequences are
Table 3
Number of Base Differences among Five Tetrapod Species (four mammals and a frog) for the Same 278 Sites (from Huysmans and De Wachter 1985) as in Table 2

<table>
<thead>
<tr>
<th>SPECIES</th>
<th>Homo sapiens</th>
<th>Mus musculus</th>
<th>Rattus norvegicus</th>
<th>Oryctolagus cuniculus</th>
<th>Xenopus laevis</th>
</tr>
</thead>
<tbody>
<tr>
<td>H. sapiens</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. musculus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. norvegicus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O. cuniculus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X. laevis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H. sapiens</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. musculus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. norvegicus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O. cuniculus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X. laevis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4
Number of Sites at Which Base Substitutions Occur in the A, B, and C Regions among the Four Naegleria Species, Four Mammal Species, and Five Tetrapod Species at the Same 278 Sites as in Tables 2 and 3

<table>
<thead>
<tr>
<th>SPECIES</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naegleria</td>
<td>19</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Mammal</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Tetrapod</td>
<td>3</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

NOTE.—Values in parentheses are % substitutions.
srRNA Evolution in *Naegleria* 249

**A-SEQUENCE**

\[\text{Hs} \quad \text{GAGAGGGAGC} \quad \text{CUUGAAGAC} \quad \text{GCUAACCAAU} \quad \text{CAGGAAAGG} \quad \text{CAGGCAAC-} \quad \text{GGGCAAAAU} \\
\text{Ns} \quad \text{GAGAGGUAGC} \quad \text{CUUGAAGAC} \quad \text{GCUAACCAAU} \quad \text{GAGGAAAGG} \quad \text{CAGGCAAC-} \quad \text{GGGCAAAAU} \\
\text{Ac} \quad \text{ACCCACUCCC} \quad \text{GACCGGGG-} \quad \text{AGGUAGUGAC} \quad \text{GAAAGAAUAAAC} \quad \text{AUAACAGGGAC} \quad \text{UCUUUAGG-G} \\
\text{ACCCAAUCCC} \quad \text{AACGAGGGG-} \quad \text{AGGUAGUGAC} \quad \text{GAAAGAAUAAAC} \quad \text{AUAACAGGGAC} \quad \text{UCUUUAGG-G} \\
\text{GCCCUUGUAAU} \quad \text{UGGAAAGAC} \quad \text{CUCACUAAA} \quad \text{--CCUUAAGAAGAC} \quad \text{CUGAAAGAC} \quad \text{CUGAAAGAC} \quad \text{UCA}^{591} \\
\text{GUCUUGUAAU} \quad \text{UGGAAAGUGA} \quad \text{AACAUAUAAAC} \quad \text{--CCCUUAAAGAC} \quad \text{CUGAAAGAC} \quad \text{CUGAAAGAC} \\
\]

**B-SEQUENCE**

\[\text{Hs} \quad \text{---_------} \quad \text{ACGAAUCGA} \quad \text{UGGCAAUAC} \quad \text{AUGGAAGAC} \quad \text{UGGAAUGAGU} \\
\text{Ns} \quad \text{AGCUGCUACG} \quad \text{UCUAGGGAAG} \quad \text{GAUAGUACG} \quad \text{GAUAGUACG} \quad \text{UGGAAUGAGU} \\
\text{Ac} \quad \text{CGCUGGUGCC} \quad \text{CUCUGGUGCC} \quad \text{GUGUGCCUGG} \quad \text{GUGUGCCUGG} \quad \text{GUGUGCCUGG} \\
\text{CCUCGUGGAC} \quad \text{CUUCUGGUGG} \quad \text{GUGUGCCUGG} \quad \text{GUGUGCCUGG} \quad \text{GUGUGCCUGG} \\
\text{GCGGUGGAC} \quad \text{CUUCUGGUGG} \quad \text{GUGUGCCUGG} \quad \text{GUGUGCCUGG} \quad \text{GUGUGCCUGG} \\
\text{GCGGUGGAC} \quad \text{CUUCUGGUGG} \quad \text{GUGUGCCUGG} \quad \text{GUGUGCCUGG} \quad \text{GUGUGCCUGG} \\
\text{GCGGUGGAC} \quad \text{CUUCUGGUGG} \quad \text{GUGUGCCUGG} \quad \text{GUGUGCCUGG} \quad \text{GUGUGCCUGG} \\
\]

**C-SEQUENCE**

\[\text{Hs} \quad \text{CCUGGGCCCAA} \quad \text{CUUCUUAGAG} \quad \text{---_------} \quad \text{GCGUUCAGC} \quad \text{CA-CC-CGAG} \quad \text{AUUGA-GCA-} \\
\text{Ns} \quad \text{CUUUGUCCAG} \quad \text{CUUCUUAGAN} \quad \text{ACUUC-AUUC} \quad \text{GUAAACUAG} \quad \text{GAUGAGGAAG} \quad \text{AUUUGAGCC-} \\
\text{Ac} \quad \text{CGCGGCGUCG} \quad \text{CCUGGCGUCG} \quad \text{GUGUGCCUGG} \quad \text{GUGUGCCUGG} \quad \text{GUGUGCCUGG} \\
\]

**Fig. 2.—** A and B sequences for *Naegleria lovaniensis* and the C sequence for *N. gruberi* aligned with those of *Homo sapiens* (data from Huysmans and De Wachter 1985) and *Acanthamoeba castellanii* (data from Gunderson and Sogin 1986). The blocks used for phylogenetic analysis are bracketed. The numbering system is for *H. sapiens* (Gonzalez and Schmickel 1986). \(N\) = nucleotide of undetermined type; a dash (\(-\)) = a gap.

(see fig. 2) to leave a total of 398 sites in the semiconserved regions for analysis. The results of the phylogenetic analyses are shown in figures 3 and 4.

The parsimony analysis yielded a single tree (fig. 3A), whose topology is in general agreement with previous analyses using srRNA sequences, except for the placement
Fig. 3.—A, Most parsimonious trees found, using DNAPARS, for the semiconserved sequences involving 398 base positions. The tree had a consistency index of 0.64. Escherichia coli and Halobacterium cutirubrum were used as outgroups to root the trees. B, Results of bootstrapping (from DNABOOT) the data used to construct fig. 3A for 20 bootstraps. The values at the nodes indicate the number of bootstrap trees (of 20) that supported that node. C, The best-fit Fitch-Margoliash tree found for the $K_{nuc}$ values in table 5. Branch lengths shown are proportional to the proposed amount of evolutionary change. The standard deviation for goodness of fit for this tree (Fitch and Margoliash 1967) was 5.04%. D, The distance Wagner tree derived by the method of Farris (1972) for the $K_{nuc}$ values in table 5. Branch lengths shown are proportional to the proposed amount of evolutionary change.
Fig. 4.—Strict consensus tree for the trees shown in figs. 3A, 3C, and 3D
Table 5

<table>
<thead>
<tr>
<th>Species</th>
<th>Hs</th>
<th>As</th>
<th>Zm</th>
<th>Sc</th>
<th>Ac</th>
<th>Pm</th>
<th>Pb</th>
<th>Dd</th>
<th>Tt</th>
<th>Ns</th>
<th>Eg</th>
<th>Tb</th>
<th>Vn</th>
<th>Hc</th>
<th>Ec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hs</td>
<td></td>
<td>114</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>As</td>
<td></td>
<td></td>
<td>161</td>
<td>187</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zm</td>
<td>181</td>
<td>185</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sc</td>
<td>187</td>
<td>214</td>
<td>122</td>
<td>137</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ac</td>
<td>201</td>
<td>213</td>
<td>119</td>
<td>117</td>
<td>123</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pm</td>
<td>359</td>
<td>383</td>
<td>278</td>
<td>253</td>
<td>275</td>
<td>211</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pb</td>
<td>298</td>
<td>292</td>
<td>220</td>
<td>194</td>
<td>210</td>
<td>235</td>
<td>335</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dd</td>
<td>310</td>
<td>311</td>
<td>240</td>
<td>207</td>
<td>258</td>
<td>235</td>
<td>368</td>
<td>287</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tt</td>
<td>390</td>
<td>430</td>
<td>398</td>
<td>394</td>
<td>390</td>
<td>374</td>
<td>459</td>
<td>447</td>
<td>451</td>
<td>480</td>
<td>504</td>
<td>544</td>
<td>464</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ns</td>
<td>469</td>
<td>475</td>
<td>371</td>
<td>410</td>
<td>414</td>
<td>407</td>
<td>491</td>
<td>445</td>
<td>423</td>
<td>507</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eg</td>
<td>499</td>
<td>531</td>
<td>447</td>
<td>465</td>
<td>442</td>
<td>477</td>
<td>569</td>
<td>480</td>
<td>504</td>
<td>544</td>
<td>464</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb</td>
<td>730</td>
<td>758</td>
<td>678</td>
<td>661</td>
<td>655</td>
<td>671</td>
<td>666</td>
<td>661</td>
<td>648</td>
<td>702</td>
<td>693</td>
<td>759</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vn</td>
<td>795</td>
<td>748</td>
<td>712</td>
<td>731</td>
<td>704</td>
<td>741</td>
<td>832</td>
<td>724</td>
<td>766</td>
<td>842</td>
<td>722</td>
<td>805</td>
<td>912</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hc</td>
<td>799</td>
<td>834</td>
<td>824</td>
<td>799</td>
<td>785</td>
<td>774</td>
<td>851</td>
<td>845</td>
<td>822</td>
<td>886</td>
<td>828</td>
<td>761</td>
<td>875</td>
<td>541</td>
<td></td>
</tr>
<tr>
<td>Ec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.—Hs = Homo sapiens; As = Artemia salina; Zm = Zea mays; Sc = Saccharomyces cerevisiae; Ac = Acanthamoeba castellanii; Pm = Prorocentrum micans; Pb = Plasmodium berghei; Dd = Dictyostelium discoideum; Tt = Tetrahymena thermophila; Ns = Naegleria species; Eg = Euglena gracilis; Tb = Trypanosoma brucei; Vn = Vairimorpha necator; Hc = Halobacterium cutirubrum; Ec = Escherichia coli.

of the ciliates (represented here by *Tetrahymena*), which on previous analyses have fallen below the node leading to *Dictyostelium* (Elwood et al. 1985; Gunderson et al. 1986; Sogin et al. 1986; Johnson et al. 1987). On this tree, *Naegleria* is the sister taxon to a clade consisting of all taxa except *Vairimorpha*, *Trypanosoma*, and *Euglena*. Of special interest are the relative positions of *Naegleria* and *Acanthamoeba*, the latter being allied more closely to metazoa than to *Naegleria*.

Bootstrapping of these data (fig. 3B) provides reasonably strong support for such a placement of *Naegleria* with respect to *Acanthamoeba*, with 20 of 20 bootstrap trees showing monophyly of *Acanthamoeba* with *Homo* and *Artemia* to the exclusion of *Naegleria*.

The $K_{nu}$ values among the 15 taxa are shown in table 5. These data were used to construct trees by both the Fitch and Margoliash (1987) method and the distance Wagner procedure (Farris 1972). The results are shown in figures 3C and 3D, respectively. These trees are topologically similar to each other and to the character-state analyses in showing *Naegleria* to be outside the *Dictyostelium*/*Tetrahymena*/*Homo* clade. Moreover, both analyses agree in placing *Naegleria* well apart from both *Acanthamoeba* and *Dictyostelium*.

A strict consensus tree of the parsimony tree, the Fitch and Margoliash tree, and the distance Wagner tree is shown in figure 4. All trees agree in placing *Vairimorpha* as the sister taxon to all other eukaryotes, in agreement with the analysis of Vossbrinck et al. (1987). The remaining eukaryotes form an unresolved tetrachotomy, with branches leading to *Trypanosoma*, *Euglena*, *Naegleria*, and all other eukaryotes (including *Acanthamoeba* and *Dictyostelium*) for which sequences are available.
Discussion

Patterns of Variation among the *Naegleria* Species

Among the *Naegleria* species, by far the majority of base-substitution differences occur in the A sequence. For the 278 sites common to all four species, 19 base substitutions occur in the A sequence (16% of bases), with one difference in the B sequence (1%) and no differences in the C sequence. By contrast, for the mammals, the base substitutions are evenly distributed among the three regions. Indeed, for the mammals, all but one of the eight substitutions occurring in these regions are seen in the rabbit. When the data for the one species of frog for which a sequence is available are added to those for the mammals, the pattern is even more divergent from the *Naegleria* pattern, with more substitutions in the B and C sequences than in the A sequence.

Extent of Variation among the *Naegleria* Species

For the 278 sites common to all four *Naegleria* species, *N. fowleri* and *N. lovaniensis* are most similar, showing only two base-substitution differences. Such a level of difference is within the technical error of the method. *Naegleria australiensis* differs from them at 7–10 bases, while *N. gruberi* is most divergent, differing from the remaining three at 9–17 bases. Such differences are well beyond the error of the method and represent real differences. Thus, the divergence between *N. fowleri*/*N. lovaniensis* and *N. australiensis* is about equivalent to that among the orders of mammals, while *N. gruberi* is about as divergent from the other three *Naegleria* species as mammals are from frogs. However, such a comparison must be tempered by the lack of an appropriate data base for the tetrapods, the current analysis relying on only four mammal sequences and a single frog sequence. Moreover, it has been shown recently that *Plasmodium berghei* actually has two distinct gene families for srRNA, each of which is transcribed at different life-cycle stages (Gunderson et al. 1987). Thus, the possibility that we have sequenced paralogous families of srRNA in *Naegleria* must be considered.

From the srRNA differences among the four *Naegleria* species, it would seem that *N. fowleri* and *N. lovaniensis* are more closely related to each other than either is to *N. australiensis* or *N. gruberi*. Such a result is in accord with both the isozyme study of Pemin et al. (1985) and the serological work of Stevens et al. (1980).

Phylogenetic Relationships of *Naegleria*

This study bears on the relationship between *Naegleria* and *Acanthamoeba*, both of which include pathogens responsible for superficially similar infections of man (Martinez 1985, pp. 63–71).

The amoebas have long been considered the most difficult protozoa to place satisfactorily in a taxonomic scheme. Until recently, they have been grouped loosely with all flagellate protists in the phylum Sarcomastigophora. Organisms with both amoeboid and flagellate stages, such as *Naegleria* and certain slime molds, have had a pivotal if ill-defined position. A number of former groupings of eukaryotic microorganisms are now recognized as taxa of convenience—consider, for example, the disappearance of Protozoa as a formal taxon—and Page (1976) has referred to "the truism that amoebae are polyphyletic."

Differences between *Naegleria* and *Acanthamoeba* that suggest widely separate origins include the mode of nuclear division and biochemical features such as the occurrence of cellulose in the *Acanthamoeba* cyst wall. Recent studies of amoebas have emphasized the significance of ultrastructural features, particularly details of
mitochondrial and Golgi structure, which are believed (from studies of other protists) to be highly conservative (Page and Blanton 1985). These studies have led to the separation of groups that include Naegleria, Acanthamoeba, and Dictyostelium at progressively higher rank (class) but that still retain the Rhizopoda, also at higher rank (phyllum) (Page 1987).

We have analyzed the partial srRNA sequence data of Naegleria in relation to proposed homologous sequence data for a range of other taxa representing a wide range of phyla, including Dictyostelium and Acanthamoeba (figs. 3, 4). The analysis used a suite of phylogenetic methods, the results of which largely concur with each other. Moreover, the tree topology that we have found by using partial sequence analysis is in general agreement with all other analyses using more-complete sequence data for the srRNA gene (Elwood et al. 1985; Gunderson et al. 1986; Sogin et al. 1986). There are a number of topological discrepancies vis-à-vis the trees based on 5SRNA (e.g., see Hori and Osawa 1987), but this molecule is probably not suitable for resolving the deep branchings for eukaryotes (Baroin et al. 1988).

For Naegleria, all analyses agree in placing it as the sister taxon to all other taxa except Vairimorpha, Trypanosoma, and Euglena. The precise relationships of Acanthamoeba were unresolved by the analysis, but all agree in placing Acanthamoeba monophyletic with Homo, Artemia, Zea, Saccharomyces, and Plasmodium to the exclusion of Dictyostelium (fig. 4). Thus, on the basis of the phylogenetic analysis of partial srRNA sequence data, there is no monophyletic group that includes any two, let alone three, of the trio Naegleria, Acanthamoeba, and Dictyostelium to the exclusion of nonamoeboid organisms. Most significantly for those interested in protist phylogeny, the branching point for ciliates intervenes between Naegleria and Dictyostelium.

Amoeboid locomotion, the only unique and universal criterion for inclusion in the Rhizopoda, is not a single mechanism. It is a convenient collective term for an array of locomotive processes that depend on flexibility of cell form and that involve contraction. This diversity is reflected in the array of cell processes listed in the diagnosis of the phylum (Page 1987). The molecular mechanisms may well be homologous, but they are common to the process of mobility in many metazoan cell types. Thus, the links between many rhizopod groups are tenuous. Furthermore, the “simple” amoeobas, represented here by Acanthamoeba, may well not be monophyletic even when those organisms with more complex life cycles involving flagellate stages (e.g., Naegleria) and/or fruiting structures (e.g., Dictyostelium) are separated at a higher rank.

Does srRNA Evolve at a Constant Rate?

A striking feature of the phylogenetic analyses of the distance data (figs. 3C, 3D) is the relative uniformity of rates of change of the srRNA sequence in the various eukaryote lineages. The algorithms used to generate these trees do not assume constant rates of evolution and, indeed, provide estimates of the actual amount of change in each branch. This view is reinforced if we apply the “relative rates” test (Wilson et al. 1977). In this test, the variation in the distance of the outgroup(s) (here Escherichia and Halobacterium) to each member of the ingroup (here the eukaryotes) is measured to estimate variation in the rate of evolution of the sequence in question. Here, distances of the various ingroup members to Halobacterium vary from 704 to 912, while distances to Escherichia vary from 761 to 886. It is data of this kind that lead one to postulate a molecular clock. Since a molecular clock is expected to be, at best, stochastic rather than metronomic (Wilson et al. 1977; Sarich 1985), such minor variation is well
within the precepts of clocklike behavior and supports the argument that the srRNA molecule can be used as a molecular clock (Woese 1987).

Whether the molecular clock is time dependent or generation dependent has been debated in the literature (e.g., see Laird et al. 1969; Kohne et al. 1972; Lovejoy et al. 1972; Sarich 1972; Wilson et al. 1977; Sarich and Cronin 1980). The taxa included in our analysis undoubtedly represent groups whose generation lengths differ by several orders of magnitude (consider, e.g., metazoa vs. various rapidly dividing protists). Yet there is uniformity in the rates of change of the srRNA sequence, at least for the sequences we included. Our data are therefore consistent with the hypothesis that, at least for the srRNA molecule, the rate is time dependent and not generation dependent. Hasewaga et al. (1985) reached the same conclusion for the srRNA molecule from a more restricted set of taxa.

Acknowledgments

We thank J. Felsenstein, Department of Genetics, University of Washington, Seattle, for his PHYLIP package and P. Hakendorf, A. Gunjko, and M. Adams for computing assistance. We are grateful to S. Blackler, P. Dobson, and A. Thanou for technical assistance and to P. Kidd for typing the manuscript. For critical comments on the manuscript we thank M. Adams, R. Andrews, S. Donnellan, M. Mahony, M. Nei, and W. Fitch. This work was funded by a grant from the National Health and Medical Research Council of the Commonwealth of Australia.

LITERATURE CITED


GUNDERSON, J. H., T. F. McCUTCHEON, and M. L. SOGIN. 1986. Sequence of the small subunit


MASATOSHI NEI, reviewing editor

Received February 23, 1988; revision received December 16, 1988