Alpha-Tubulin from Early-Diverging Eukaryotic Lineages and the Evolution of the Tubulin Family

Patrick J. Keeling and W. Ford Doolittle
Department of Biochemistry, Dalhousie University

The tubulin gene family, which includes alpha-, beta-, and gamma-tubulin subfamilies, is composed of highly conserved proteins which are the principle structural and functional components of eukaryotic microtubules. We are interested in (1) establishing when in eukaryotic evolution the duplications leading to paralogous alpha, beta, and gamma subfamilies occurred and (2) the possible utility of tubulin sequences in reconstructing organismal phylogeny. To broaden the taxonomic representation of alpha-tubulins so that it roughly equals that of beta-tubulins, alpha-tubulin genes from three Microsporidia (*Encephalitozoon hellem, Nosema locustae,* and *Spraguea lophii*), two Parabasalia (*Monocercomonas* sp. and *Trichomitus batrachorum*), and one Heterolobosean (*Acrasis rosea*) were sequenced. With these new genes, phylogenetic trees of alpha- and beta-tubulins were constructed and compared. Trees were congruent with each other, but incongruent with other molecular phylogenies. The agreement between alpha- and beta-tubulin trees could arise by the co-adaptation of one molecule to variants of the other as a result of their intimate steric association in microtubules. Thus, these trees may not be providing independent support for the phylogenetic results. However, one of these unexpected results, that microsporidia cluster with fungi, is supported by other circumstantial evidence, and may therefore reflect a real relationship despite the basal position usually assigned to microsporidia. Relationships between the three tubulins were also examined by constructing trees of all three types. These trees were found to be of limited value for determining the position of the root within each subfamily because of the great interfamily distances, but they do confirm the classification of all known genes into three monophyletic subfamilies. Divergent genes from *Caenorhabditis elegans* and *Saccharomyces cerevisiae* that have been proposed to represent the novel classes delta- and epsilon-tubulin were found to be specifically related to gamma-tubulins from animals and fungi respectively, and therefore are best seen as rapidly evolving orthologues of gamma-tubulin.

Introduction

The tubulin gene family consists of three distinct but highly conserved subfamilies, alpha-, beta-, and gamma-tubulin, each defined by sequence conservation, a wide distribution among eukaryotes, and, where studied, a conservation of function. Of the three varieties, alpha- and beta-tubulins are the most abundant in the eukaryotic cell and have been studied most extensively. Heterodimers of these two proteins are the primary constituents of microtubules, which in turn are central to the composition of eukaryotic flagella, cilia, mitotic spindles, and the cytoskeleton. Gamma-tubulin was discovered much later (Oakley and Oakley 1989) and its function is less clear, although it is known to be important in microtubule organizing centers, or MTOCs (Oakley et al. 1990; Zheng, Jung, and Oakley 1991), and has been implicated in several other processes (Gard 1994; Lajoie-Mazenc et al. 1994). Recently, two additional tubulin families have been proposed based on the identification of two unusual and highly divergent sequences, the so-called delta-tubulin found in *Caenorhabditis elegans,* and the epsilon-tubulin from *Saccharomyces cerevisiae* (Burns 1995). While it is true that these sequences are very distant from other known tubulins, their apparent restriction to a single taxon each implies that they may not represent novel gene families but rather unique genes, specific to the lineages in which they have been described.

Key words: alpha-tubulin, evolution, phylogeny, Microsporidia, Parabasalia, Heterolobosea.

Address for correspondence and reprints: Patrick J. Keeling, Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia, B3H 4H7 Canada. E-mail: pkeeling@is.dal.ca.

Mol. Biol. Evol. 13(10):1297–1305. 1996
 ISSN: 0737-4038

Each tubulin orthologue is unique to eukaryotes, but the tubulin family as a whole does have a prokaryotic antecedent in the FtsZ protein, a component of the eubacterial cytokinesis system (see Lutkenhaus 1993; Erickson et al. 1996; Margolin, Wang, and Kumar 1996). Presumably the three tubulins diverged from a single ancestral FtsZ, but it is not known when this triplication took place, or in which order the paralogues arose. Of alpha, beta, and gamma, beta-tubulin currently enjoys the widest taxonomic representation. Beta-tubulin genes have been found even in the earliest-diverging eukaryotes (Kirk-Mason, Turner, and Chakraborty 1988; Katiyar and Edlind 1994; Edlind et al. 1996), demonstrating that at least one gene duplication event took place prior to the divergence of extant eukaryotes. However, the data on archezoal tubulins is restricted to beta: alpha- and gamma-tubulins have been identified in a few protist lineages, but none that diverged so early in eukaryotic evolution (Lai, Remillard, and Fulton 1988; Sanchez et al. 1995). This leaves some uncertainty as to when alpha- and gamma-tubulins diverged, before or after the appearance of extant eukaryotes.

Materials and Methods

Characterization of Tubulin Genes

DNA from a variety of organisms was generously donated by individuals as follows: *Monocercomonas* sp. Ns-1PRR and *Trichomitus batrachorum* G11 DNA were gifts from M. Müller, *Encephalitozoon hellem* CDC: 0291:V213 DNA was a gift from C. G. Clark, *Nosema locustae* ATCC 30860 and *Acrasis rosea* T-235 DNA were gifts from A. J. Roger, and *Spraguea lophii* DNA was a gift from G. Hinkel.

These samples were used as templates in PCR reactions with primers specific for alpha-tubulin, provided by A. J. Roger, as described (Keeling and Doolittle 1996). Clones of the expected size were sequenced in duplicate on both strands and in all cases both duplicates were found to be identical except for that of *Monocercomonas,* where two clones were discovered that differ sufficiently to suggest that they are separate genes. The seven genes were all subcloned into fragments ranging in size from 200 bp to 800 bp in pBluescript, using restriction enzymes appropriate for each individual gene. The actual sequencing was carried out either by manually sequencing subclones and gap-filling using primers, or with ABI 373A or LiCor automated sequencing machines.

Phylogenetic Analysis of the Tubulin Family

An alignment of representatives of all three tubulin types was made by PileUp (Deveraux, Haeberli, and Smithies 1984) and edited by eye. From this template alignment, a larger collection of tubulins consisting of 24 gamma-, 42 beta-, and 81 alpha-tubulins was aligned according to the template. The divergent *S. cerevisiae* and *C. elegans* sequences were also added by eye, aligning them with the globally conserved regions.

Trees based on this alignment were then inferred by neighbor-joining and Fitch-Margoliash analyses on distance measurements calculated according to the Dayhoff PAM250 substitution matrix using the PROTDIST, NEIGHBOR, and FITCH programs from PHYLIP version 3.57 (Felsenstein 1993). Significance of individual nodes was assessed by conducting 100 bootstrap resampling replicates and constructing trees by neighbor joining. Trees inferred by Fitch-Margoliash were found to be nearly identical to those inferred by neighbor joining, and those differences which were observed are discussed in the figure legends. Unweighted parsimony trees were also inferred using PAUP version 3.1.1 (Swofford 1993) for a set of alpha-tubulin sequences, resulting in over 700 equally parsimonious trees, the consensus of which was congruent with the neighbor-joining topology.

```
                                                    *
A.ro  LYCLEHGIQPDGQMPSDKTIGVEDDAFNTFFSETGAGKHVPRAVFLDLEPTVIDE
T.ba  LYCLEHGIQPDGQMPSDKTIGICDDAFNTFFSETGAGKHVPRAVMVDLEPTVVDE
M. 2  LYCLEHGIQPDGQMPSDKTIGVCDDAFNTFFSETGAGKHIPRAVFVDLEPTVVDE
M. 3  LYCLEHGIQPDGQMPSDKTIGVCDDAFNTFFSETGAGKHVPRAVFVDLEPTVVDE
E.he  LYCKEHGILPDGRLDQNRM...DDESAESFFSQTSVGTYVPRTLMVDLEPGVLES
S.lo  LYCKEHGILPDGTPDPNFN..DKESYSSTFFSETSGGNFVPRALMIDLEPGVIDS
N.lo  LYCKEHNIRPDGTTGGV......DDSCSSFFIETSAGTYVPRTLMVDLEPGVIES

A.ro  VRTGTYRQLFHPEQLISGKEDAANNYARGHYTIGKEIVDLCLDRIRKLADNCTGL
T.ba  VRTGTYRQLWHPEQLINGKEDAANNYARGHYTVGKEIIDLTLDRIRKLADQCTGL
M. 2  VRTGTYRQLFHPEQIINGKEDAANNYARGHYTVGKEIIDLTLDRIRKLADQCTGL
M. 3  VRTGTYRQLFHPEQIINGKEDAANNYARGHYTVGKEIIDLTLDRIRKLADQCTGL
E.he  IKTGKYRELYHPGQLISGKEDAANNYARGHYTVGKEIIEPVMEQIRRMADNCDGL
S.lo  IKTSEYKNLYHPSQLIAGQEDAANNYARGHYTAGKEIIEKVTDQIKRIAENCSGL
N.lo  IKNSEYRALYHPSSLINGKEDAANNYARGHYTVGKEIIEPVMEQIRRMADCCDGL

A.ro  QGFLVFNSVGGGTGSGLGALLLERLSVDYGKKSKLGFTVYPSPQVATAVVEPYNS
T.ba  QGFLIFHSFGGGTGAGFGSLLLERLSVDYGKKSKLEFTVYPAPQVSTAVVEPYNS
M. 2  QGFLIFHSFGGGTGAGLGSLLLERLSVDYGKKSKLELTVYPAPQVSTAVVEPYNS
M. 3  QGFLIFHSFGGGTGAGFGSLLLERLSVDYGKKSKLEFTVYPAPQVSTAVVEPYNS
E.he  QGFLIYHSFGGGTGSGFASLMMDRLAAEFGKKSKLEFSVYPAPKIATAVVEPYNS
S.lo  QGFLVFHSFGGGTGSGFGALLMDRLSVEFGKKSKLEFAIYPSPRISTAVVEPYNS
N.lo  QGFLIFHSFGGGTGSGFGGLLMDRLSQEFGKKSKLEFSVYPAPRIATAVVEPYNS

A.ro  VLSTHALLEHTDVAVMLDNEAIYDICRRSLDIQRPTYTNLNRLVAQVISSLTCSL
T.ba  ILATHAMIDHSDCAFMVDNEALYDLCRRALDIERPTYTNLNRLIGQVVSSLTASL
M. 2  ILATHAMIDHSDCAFMVDNEALYDLCRRALDIERPTYTNLNRLMGQVVSSLTASL
M. 3  ILATHAMIDHSDCAFMVDNEALYDLCRRALDIERPTYTNLNRLIGQVVSSLTASL
E.he  ILTTHTTLDYSDCSFLVDNEAIYDMC.RNLGIQRPYYTDINRIIAQVVSSITASL
S.lo  ILTTHTTLNHFDCSFLVDNEAIYDIC.KNLGIAMPHANDLNKCITQVVSSITASL
N.lo  ILTTHTTLDHSDCSFLVDNEAIYDMC.RNLGIERPKYKEINRVLAQVVSSITASL

A.ro  RFDGALDVDVTEFQTNLVPYPRIHFMLCSIAPVISAEKAYHEQLSVAEITNSAFE
T.ba  RFDGALNVDFTEFQTNLVPYARIHFPICSYAPVISAEKAYHEQLTVAEVTNTLFE
M. 2  RFDGALNVDFTEFQTNLVPYPRIHFPICSYAPVISAEKAYHEQLSVAEITNSLFE
M. 3  RFDGALNVDFTEFQTNLVPYSRIHFPICSYAPVISAEKAYHEQLSVAEITNSLFE
E.he  RFPGSLNVDLTEFQTNLVPYPRIHFPLVAYSPMLSKEKAAHEKLSVQEITNACFE
S.lo  RFPGSLNVDLTEFQTNLVPYPRIHFPLVAYFPMLSRERASHEQLSVQEITSACFD
N.lo  RFPGSLNVDLTEFQTNLVPYPRIHFPLVAYAPMLSRNKASHEQLSVSEITSACFN

A.ro  PASMMAKCDPRHGKYMACCLMYRGDVVPKDVNAAVATIKTKRTIQFVDWSPTGFK
T.ba  PANMMVKCDPRHGKYMACTLLYRGDVVPKDVSAAIATIKTKRAIQFVDWCPTGFK
M. 2  PANMMVKCDPRHGKYMACTLLYRGDVVPKDVGAAVATIKTKRTIQFVDWCPTGFK
M. 3  PANMMVKCDPRHGKYMACTLLYRGDVVPKDVGAAVATIKTKRTIQFVDWCPTGFK
E.he  PQSQMVRCDTRKGKYMACCLLFRGDVNPKDANTATANVKAKRTNQFVEWCPTGFK
S.lo  PENQMVKCDPRNGKYMACCLLFRGNVNPKDVNQATSLVKSKRANQFVEWCPTGFK
N.lo  PESQMVKCDPKKGKYMACCLLFRGDVQPKDVNQAMAFVKAKRAAQFVEWCPTGFK

A.ro  CGINYQPPIVVPGGDLAKIQRAVCMISNSTAIAEVFSRIDHKFDLMYAKRAFVH
T.ba  IGINYQPPTVVPGGDLAKVQRAVCMLANTTAVAEAWSRLDHKFDLMYAKRAFVH
M. 2  IGINYQPPTVVPGGDLAKVQRAVCMLANTTAIAEAWSRLDHKFDLMYAKRAFVH
M. 3  IGINYQPPTVVPGGDLAKVQRAVCMLANTTAIAEAWSRLDHKFDLMYAKRAFVH
E.he  VGINSRKPTVLDGEAMAEVSRAVCALSNTTAISEAWKRLNNKFDLMFSKRAFVH
S.lo  IGINDRKPYVFEDGAMAPVDRAVCMLSNTTAISEAWKRLNRKFDLMFSKRAFVH
N.lo  IGMNSRKPTILDDDAMAPVSRAVCLLSNTTAIAEAWQRLNQKFDLMFSKRAFVH
```

FIG. 1.—Inferred amino acid alignment of alpha-tubulins from *Acrasis rosea* (*A.ro*), *Trichomitus batrachorum* (*T.ba*), *Monocercomonas* sp. (*M.*1 and *M.*2), *Encephalitozoon hellem* (*E.he*), *Spraguea lophii* (*S.lo*), and *Nosema lucustae* (*N.lo*). The acetylated lysine at position 40 of the full-length protein is marked with an asterisk.

Results

Identification of Alpha-Tubulin Genes in Ancient Eukaryotic Lineages

Products of the expected size were isolated from the parabasalia *Trichomitus batrachorum* and *Monocercomonas* sp., the heterolobosean *Acrasis rosea,* and the microsporidia *Nosema locustae, Encephalitozoon hellem,* and *Spraguea lophii.* These were cloned and the ends were sequenced, revealing that each encoded a gene with a high resemblance to alpha-tubulin. Two variants from *Monocercomonas* were found which differed at 17 positions (15 transitions and 2 transversions) resulting in two conservative amino acid substitutions (both due to transitions).

An alignment of the inferred amino acid sequences of these genes is shown in figure 1. These genes are from taxa that are among the deepest-branching eukaryotes known according to molecular and ultrastructural data (see Cavalier-Smith 1993), but are nevertheless extremely similar to known alpha-tubulin homologues. There are a number of conserved motifs which have a defined function in tubulin proteins (for review see Burns 1991) which are also maintained in all these sequences except those from the microsporidia, where there are two noteworthy exceptions. The GTP-binding motif at positions 70 to 73 (numbered according to the full length human sequence) is generally LEPT in alpha-tubulins and LEPG in beta-tubulins, but in microsporidia, both alpha- and beta-tubulin sequences contain LEPG. Also, the acetylatable lysine at position 40 of alpha-tubulins (marked in fig. 1) and the highly conserved region around it are both missing in microsporidia as they are in fungi, *Entamoeba histolytica* and *Dic-*

tyostelium discoideum. It is not obvious why constraints on this otherwise highly conserved region have relaxed in these unrelated taxa, but one interesting correlation is that these organisms all lack flagella and cilia in all stages of their life cycle (although the same is true of some other organisms that have maintained the acetylation domain). In any case, the role and importance of acetylation in tubulin function remains unclear, especially since it may be abolished without apparent consequence in *Chlamydomonas* and *Tetrahymena* (Kozminski, Diener, and Rosenbaum 1993; Gaertig et al. 1995), but is always observed when acetylatable alpha-tubulin is present in the cell.

Phylogeny Based on Alpha- and Beta-Tubulins

Tubulin genes, for the most part beta-tubulin, have been used in the past to infer organismal relationships (Baldauf and Palmer 1993; Edlind et al.1996; Li et al. 1996), but the extreme conservation leaves few informative characters. Nevertheless, the utility of three alignable gene-families is attractive, and the substantial diversity of taxa previously known for beta-tubulin has now been roughly matched in the alpha-tubulin branch. From an amino acid alignment composed of 24 gamma-, 42 beta-, and 81 alpha-tubulins, phylogenetic trees were inferred for each tubulin independently, and combined sets were used to reciprocally root one anther.

To make the data more manageable, pairwise distance calculations were used to identify and eliminate closely related sequences. In this way the numbers of sequences were reduced to 58 and 40 for alpha and beta, respectively. Unfortunately, these numbers are still prohibitively large for protein maximum-likelihood analysis, and maximum-parsimony analysis was also hampered by the impractically large number of equally parsimonious trees. However, it should be noted that the strict consensus of over 700 maximum-parsimony trees of alpha-tubulin yielded a topology consistent with neighbor-joining trees (including the fungi-microsporidia clade discussed below). Trees were therefore constructed by neighbor-joining and Fitch-Margoliash analyses of corrected distance measurements calculated according to the Dayhoff PAM250 substitution matrix. Significance of individual nodes on these trees was assessed by conducting 100 bootstrap resampling replicates, the results of which are also shown on each tree.

An alpha-tubulin tree is depicted in figure 2. This tree is based on 406 positions, includes 58 sequences, and has been oriented with a diplomonad outgroup (diplomonads were chosen because they are consistently deep-branching eukaryotes in trees based on ribosomal RNA and EF-1 alpha; Leipe et al. 1993; Hashimoto et al. 1994). Figure 3 is a beta-tubulin tree consisting of 431 positions from 40 sequences, and once again has a diplomonad outgroup. These trees share a number of features with other molecular phylogenies, including the presence of several monophyletic groupings such as animals, plants, fungi, and alveolates. It is also noteworthy that the alpha-tubulins of *Acrasis rosea* and *Naegleria gruberi* branch together, since these taxa are thought to belong to the phylum Heterolobosea (Page and Blanton 1985), for which supporting molecular data have just been introduced (Roger et al. 1996).

While these groups may be consistent with other data, alpha- and beta-tubulin trees also mirror one another in several ways that are not generally supported by other data. Such anomalies may well be artifacts or the results of inappropriate data for the question, but being shared by both trees, they do require some auxiliary explanation.

The first such characteristics is the position of the animals and fungi relative to parabasalia and diplomonads. When the diplomonads are used as an outgroup in these unrooted trees, the result is a deep split in the eukaryotes where animals and fungi fall on one side, and plants, euglenozoa, alveolates, and heterolobosea on the other (slime molds and parabasalia cannot readily be classified into either category as they both branch close to diplomonads and their exact position is inconsistent). This topology (also found by Edlind et al. 1996 and Li et al. 1996) is not a feature of ribosomal RNA or EF-1 alpha phylogenies, in which diplomonads fall at or near the base of a comblike distribution of taxa (Cavalier-Smith 1993; Leipe et al. 1993; Hashimoto et al. 1994).

A second noteworthy characteristic of both trees is the position of microsporidia within the fungi. Microsporidia are generally thought to be archezoa, partly because they lack several cytological features also missing in other archezoa (see Cavalier-Smith 1993), and partly because they normally branch very deeply in eukaryotic trees of ribosomal RNA or translation elongation factors (Vossbrinck et al. 1987; Kamaishi et al. 1996). Considering that fungi and microsporidia share a highly divergent acetylation domain in alpha-tubulin, these residues were excluded and the analysis was repeated. The resulting topology was no different than that of figure 2 (data not shown), suggesting that microsporidian and fungal alpha-tubulins do generally resemble one another outside the acetylation domain.

One last concern with these tree topologies is the position of *Entamoeba histolytica* and its alarmingly long branch. *Entamoeba* tubulins, although easily classifiable by family, are extremely divergent from other orthologues, resulting in a very long branch. The position of *Entamoeba* in these trees is therefore suspect, since it branches with the next longest branch in trees based on both alpha- and beta-tubulins. This conclusion seems to be borne out by removing *Dictyostelium* from alpha-tubulin trees, which results in no change to the topology except that *Entamoeba* moves to the next longest branch on the tree, at the base of the fungi (data not shown). In contrast, the removal of *Entamoeba* resulted in no change at all to the rest of the tree (data not shown). Moreover, although the trees inferred by Fitch-Margoliash differ very little from the neighbor-joining trees shown, the position of *Entamoeba* is variable in both alpha- and beta-tubulin trees, suggesting that this long branch is very unstable.

Gamma-, Delta-, and Epsilon-Tubulins

In addition to the well-represented alpha- and beta-tubulins, there are the more poorly represented gamma-

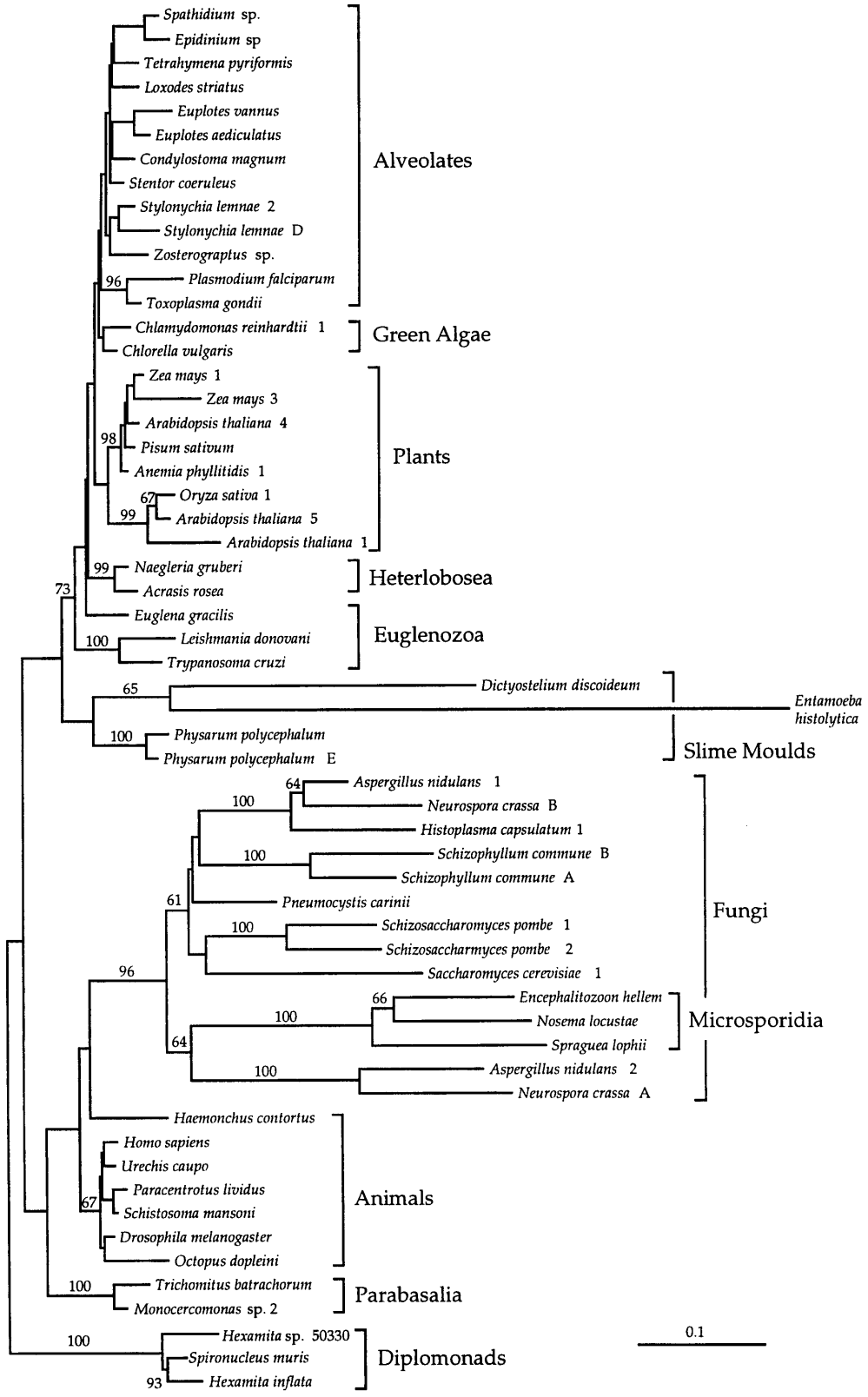

FIG. 2.—Neighbor-joining tree of 58 alpha-tubulins with bootstrap support for nodes over 50%. Fitch-Margoliash analysis on the sam distance matrix resulted in a tree with the same topology except that the slime molds are paraphyletic, and *Entamoeba* branches at the base o the fungi, further reinforcing the tenuity of *Entamoeba*'s position in the tree.

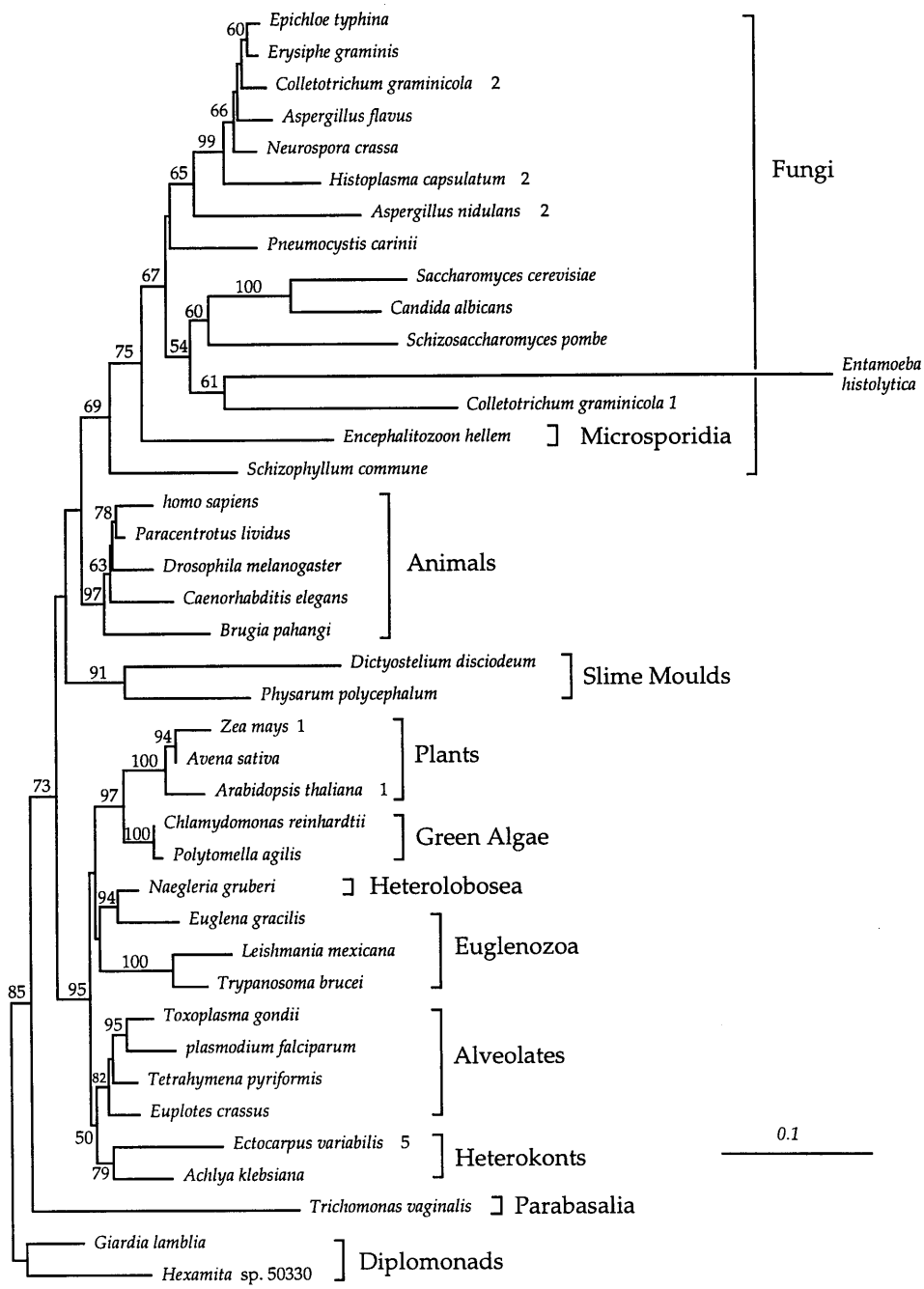

FIG. 3.—Neighbor-joining tree of 40 beta-tubulins with bootstrap support for nodes over 50%. Once again, Fitch-Margoliash analysis on the same distance matrix resulted in an identical tree except that slime molds and animals switch positions, and the position of *Entamoeba* within the fungi changes.

tubulins, and two highly divergent tubulin-like sequences from *Caenorhabditis elegans* and *Saccharomyces cerevisiae*. Based on their extreme distance from other tubulins, these two genes have prompted a proposal to expand the number of tubulin families from three to five, classifying *C. elegans* and *S. cerevisiae* sequences as delta- and epsilon-tubulins, respectively (Burns 1995).

Distance notwithstanding, there are a number of facts that support the contrasting notion that these delta- and epsilon-tubulins are not really novel families, but rather highly divergent orthologues of an existing family that are unique to the lineages where they have been observed. First, the now completed *S. cerevisiae* genome does not contain either a conventional gamma-tubulin gene or a so-called delta-tubulin gene, but only highly conserved alpha- and beta-tubulins and the so-called epsilon-tubulin. Similarly, searching the expressed sequence tag (EST) database for gamma-, delta-, and epsilon-tubulins in *C. elegans* and *C. briggsae* yielded only alpha-, beta-, and the so-called delta-tubulin. The implication from these observations is that neither *S. cerevisiae* nor *C. elegans* contain either a con-

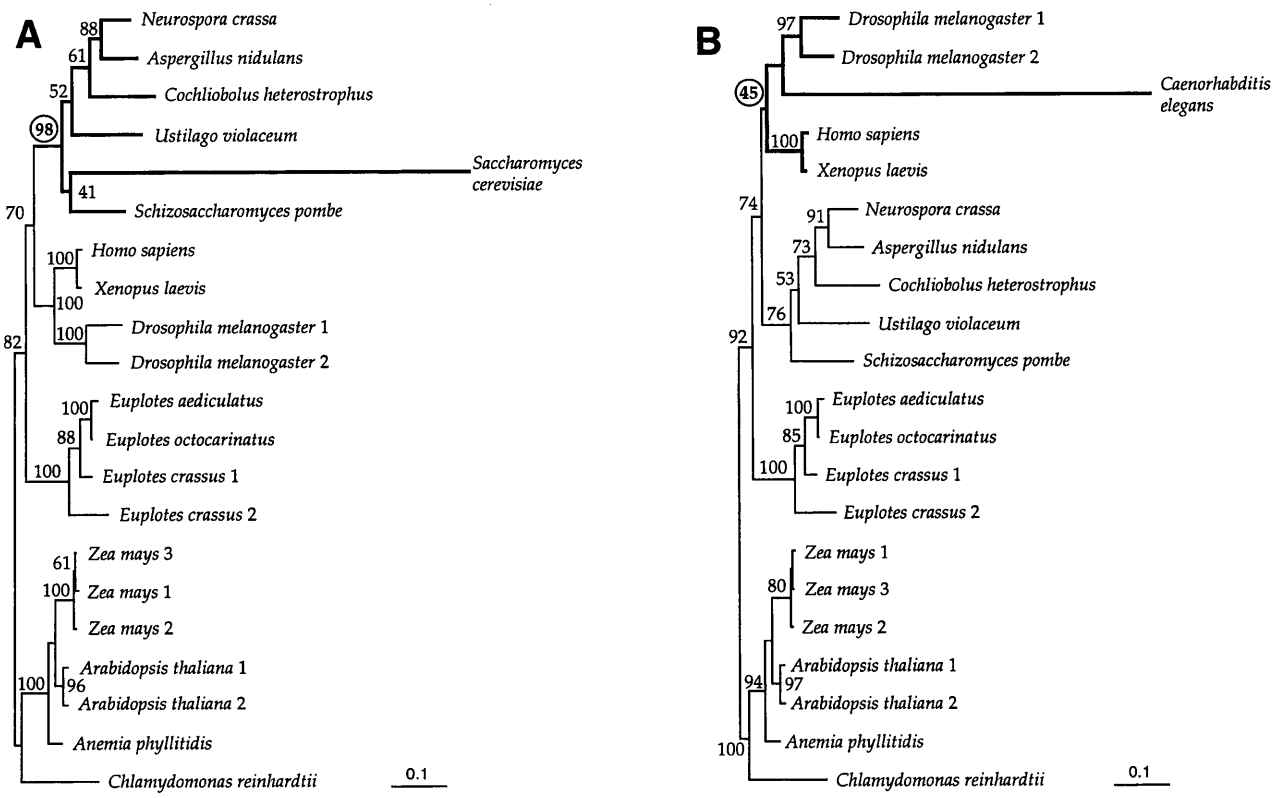

FIG. 4.—*C. elegans* (*A*) and *S. cerevisiae* (*B*) tubulin genes show a specific relationship to the gamma-tubulins of other animals and fungi respectively (support for nodes uniting these groups is circled). In Fitch-Margoliash analyses both distance matricies yield trees nearly identica to those shown, and in both cases the *C. elegans* and *S. cerevisiae* sequences branch within the animals and fungi, respectively.

ventional gamma-tubulin or the supposed novel tubulin found in the other. Indeed, no other organism has ever been found to contain either of these genes except *C. briggsae,* which contains an EST almost identical to the *C. elegans* delta-tubulin.

Greater support for the gamma-tubulin provenance of these unusual sequences comes from phylogenetic reconstruction of all gamma-tubulins with the delta and epsilon gene and outgroups chosen from the alpha- and beta-tubulins. In trees of all three tubulins, the unusual *Caenorhabditis* and *Saccharomyces* sequences branch with the gamma-tubulin subtree with 100% bootstrap support (not shown), but they fall very deep in the gamma subtree and not within the animals and fungi (see also Sobel and Snyder 1995). The possibility that the basal position of these sequences is the result of an attraction to the other long branches on the tree was addressed by including the *C. elegans* and *S. cerevisiae* sequences independently in trees where other long branches were excluded. In analyses excluding *Entamoeba, Reticulomyxa,* and *Plasmodium* (fig. 4), or simply excluding *Entamoeba* (not shown), *S. cerevisiae* branches specifically with the fungi with high statistical support, and *C. elegans* branches specifically with the animals, although with much weaker support.

Lastly, and perhaps most conclusively, several recent functional characterizations of the *Saccharomyces* gene product provided excellent evidence that it is located at the spindle pole body (a MTOC), and that its disruption results in a phenotype similar to gamma-tubulin disruptions in other ascomycetes (Sobel and Snyder 1995; Marschall et al. 1996; Spang et al. 1996) Taken together, these observations leave little room tc doubt the conclusion that both *Saccharomyces* and *Caenorhabditis* tubulin-like genes are lineage-specific, highly divergent orthologues of gamma-tubulin.

Rooting Tubulin Trees

Figure 5 shows the result of combining subsets of the alpha-, beta-, and gamma-tubulin alignments (chosen for representative diversity, but excluding the extremely diverse gamma-tubulins from *Saccharomyces* and *Caenorhabditis* discussed above). This tree is based on 31C positions and 70 sequences. Clearly the three families are each independent, monophyletic groups, and it is not obvious if any two are more similar than the third, reflecting the great interfamily distances relative to intrafamily distances. The actual topologies within the alpha and beta subtrees differ only very subtly from the topologies yielded by individual analyses, but in both cases trees appear different on account of the root falling within the branch leading to *Entamoeba histolytica.* As discussed above, *Entamoeba* tubulins are quite divergent, so it is likely that this relationship is the product of a long branch attraction between *E. histolytica* and the branch leading to the other tubulin families. This conclusion is supported by the three-way rooted Fitch-Margoliash trees, which also retain the topology of the

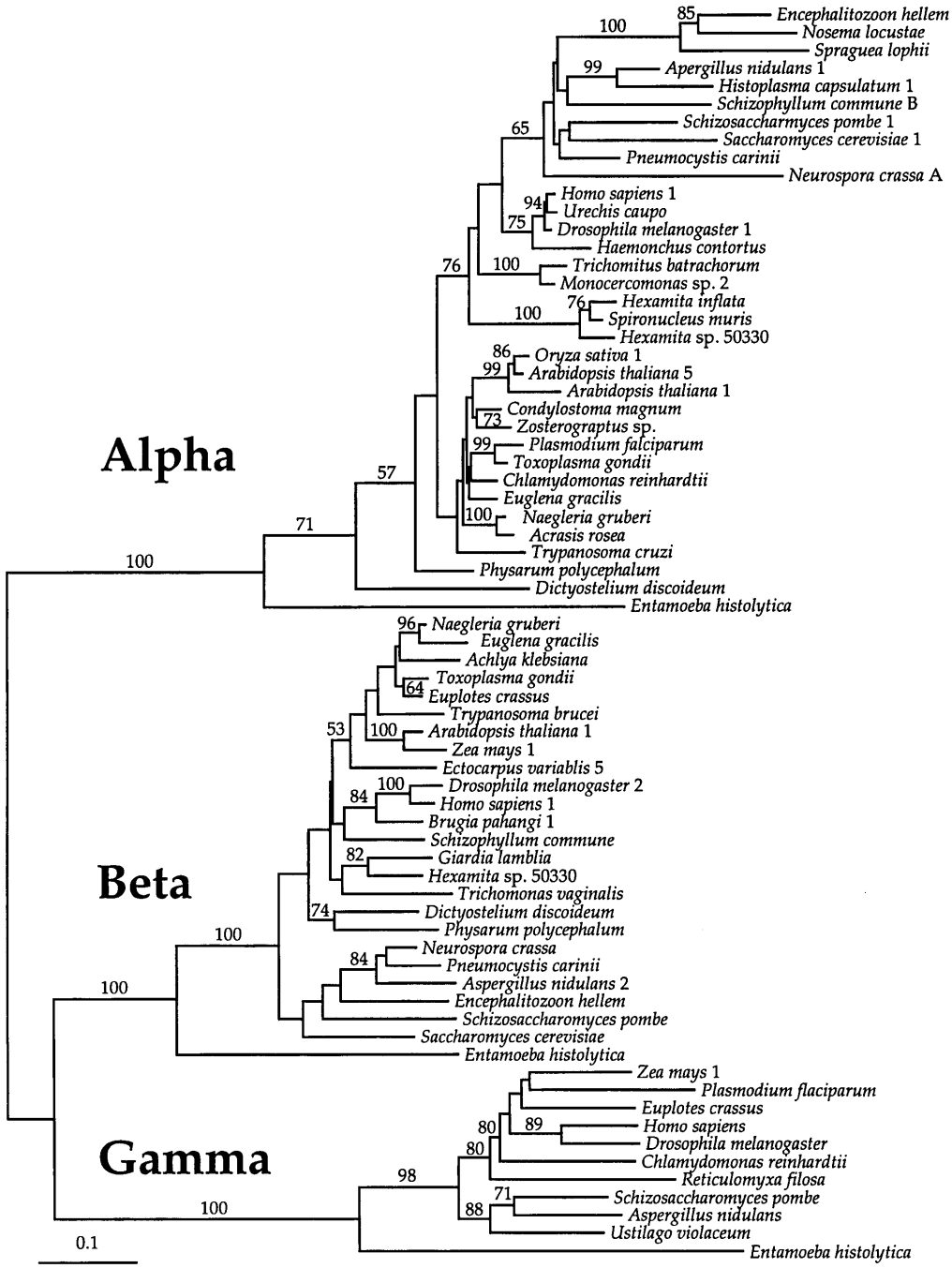

FIG. 5.—Reciprocally rooted neighbor-joining tree of alpha-, beta-, and gamma-tubulins with bootstrap support for nodes over 50%. In Fitch-Margoliash trees of this distance matrix the topology differed very little in the gamma and beta subtrees, but in the alpha subtree the topology was affected by the position of *Entamoeba*. In Fitch-Margoliash trees of alpha-tubulin alone, *Entamoeba* branched with the fungi rather than with the slime molds as in neighbor joining (see fig. 2); this topology was preserved in the rooted Fitch-Margoliash trees with *Entamoeba* falling at the base, resulting in a topology which differs from that shown for neighbor joining.

individual Fitch-Margoliash analysis in each subtree, but root each with *Entamoeba,* resulting in an overall topology much different from that of neighbor joining (not shown). Moreover, when *E. histolytica* alpha- and beta-tubulins were removed and the three-way rooting repeated with neighbor-joining and Fitch-Margoliash, the topology does not change within each subtree, but the position of the root does change, moving in both cases to the next longest branch in the fungi (data not shown). Therefore, the position of the root within each subtree appears to be largely dependent on branch length and should therefore be considered highly suspect.

Discussion

Phylogenetic analysis of alpha-tubulin genes from diplomonads, microsporidia, parabasalia, and heterolobosea provides convincing evidence that alpha- and

beta-tubulins diverged prior to the divergence of extant eukaryotes. Moreover, since each of the three subtrees is holophyletic, it appears that gamma-tubulin was also present by this time. If this were not the case, then, barring any paralogue-specific rate acceleration, one would expect that the gamma subtree would branch from within one of the other two, and not from the basal position seen in figure 5.

These new sequences also provide the opportunity to compare alpha- and beta-tubulin phylogenies with an almost equal representation of major lineages. The alpha- and beta-tubulin trees are nearly identical in topology, but in several major respects this topology is curiously inconsistent with the phylogeny inferred by other molecular markers. This is most evident in the relative branching order of the unrooted trees of figures 2 and 3, where the animals and fungi branch closer to supposedly deep-branching protists such as diplomonads and parabasalia than they do conventionally (Cavalier-Smith 1993; Leipe et al. 1993; Gunderson et al. 1995; Kamaishi et al. 1996). Another unexpected relationship mirrored by both molecules is the firm position of the microsporidia within the fungi. Conventionally, Microsporidia is seen as an ancient phylum, in part because its members display a very degenerate cytology, lacking numerous features also missing in other Archezoa, and in part from the consistently deep position of microsporidia in the few molecular trees where microsporidian data are available (Vossbrinck et al. 1987; Kamaishi et al. 1996).

The congruent alpha- and beta-tubulin trees might be taken as independent support for these unusual results. However, since microtubules are composed of alternating units of closely packed alpha- and beta-tubulin, these trees may instead reflect a strong tendency to covariation between the two tubulin molecules. If there were a finite number of "solutions" to the problem of satisfactory interactions between the proteins (which is supported by the extreme degree of both inter- and intrafamily conservation in tubulins), then the appearance of a certain variant of one tubulin could strongly favor the covariation of the other along predictable lines. Paralogy and loss could also be involved in such a process, but in such a case once again covariance would have to be evoked to explain the congruent loss of paralogues in several supposedly unrelated lineages.

Even if the congruence between tubulin trees is not independent support for this topology, there are truly independent reasons to carefully consider the phylogenetic position of the Microsporidia within the fungi. The extremely derived, obligately parasitic lifestyle of these organisms has raised doubts as to whether their "primitive" cytology is ancestral or a relatively recent adaptation (Cavalier-Smith 1993). Similarly, since microsporidian gene sequences are typically very divergent, the deep phylogenetic position of these sequences may be a consequence of attraction to other long branches. Alpha- and beta-tubulins are not immune to this possibility either, and their position within the fungi may simply be due to the fact that both microsporidian and fungal tubulins are diverging faster than other lineages, but there is other circumstantial evidence for a relationship between microsporidia and fungi. First, the ridged endospore wall of microsporidia is composed of chitin (see Canning 1990), the same material that fungi utilize as a cell wall polymer (chitin is also found in numerous other unrelated lineages: Mulisch 1993). Secondly, unlike other archezoa, which appear to be clonal (see Tibayrenc et al. 1991 for review), microsporidia undergo a form of meiosis. This process is in itself a source of debate, however, as there have been alternative arguments that it is either radically different from meiosis in other eukaryotes (Canning 1988) or fundamentally the same process (Flegel and Pasharawipas 1995). Curiously, the argument that microsporidian meiosis is typically eukaryotic in form is based on similarities observed by the authors between the cell cycles of microsporidia and fungi (Flegel and Pasharawipas 1995). Lastly, although the molecular phylogeny of EF-1 α supports the deep divergence of microsporidia (Kamaishi et al. 1996), the only microsporidian EF-1α gene known to date, that of *Glugea plecoglossi,* contains an 11-codon insertion at exactly the same position as a 12-codon insertion that has been argued to be a determinative feature of the animal–fungal clade (Baldauf and Palmer 1993). An insertion in this general region of the protein may be a common event, but the *Glugea* insertion is at exactly the same location and also bears a weak resemblance to that of animals and fungi (see fig. 2 of Kamaishi et al. 1996).

Individually each of these characters may be inadequate to argue strongly for any relationship between microsporidia and fungi, as each is also shared with other taxa. However, taken together, and considering the relatively strong support for the microsporidia–fungi clade in both alpha- and beta-tubulin trees, the possibility that microsporidia are highly derived fungi certainly should be considered.

Acknowledgments

This work was supported by an operating grant from the Medical Research Council of Canada (MRC). P.J.K. is the recipient of an MRC studentship and W.F.D. is a fellow of the Canadian Institute for Advanced Research. We would like to thank A. J. Roger for primers, C. G. Clark, G. Hinkel, A. J. Roger, and M. Müller for genomic DNA, and J. Logsdon Jr. and A. J. Roger for helpful discussions. New sequences have been deposited in GenBank under the accession numbers U66902–U66908.

LITERATURE CITED

BALDAUF, S. L., and J. D. PALMER. 1993. Animals and fungi are each other's closest relatives: congruent evidence from multiple proteins. Proc. Natl. Acad. Sci. USA **90**:11558–11562.

BURNS, R. G. 1991. Alpha-, beta-, and gamma-tubulins: sequence comparisons and structural constraints. Cell Motil. Cytoskeleton **20**:181–189.

———. 1995. Identification of two new members of the tubulin family. Cell Motil. Cytoskeleton **31**:255–258.

CANNING, E. U. 1988. Nuclear division and chromosome cycle in microsporidia. Biosystems **21**:333–240.

———. 1990. Phylum Microsporidia. Pp. 53–72 *in* L. MARGULIS, J. O. CORLISS, M. MELKONIAN, and D. J. CHAPMAN, eds. Handbook of Protoctista. Jones and Bartlett Publishers, Boston.

CAVALIER-SMITH, T. 1993. The kingdom Protozoa and its 18 phyla. Microbiol. Rev. **57**:953–994.

DEVEREUX, J., P. HAEBERLI, and O. SMITHIES. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.

EDLIND, T. D., J. LI, G. S. VISVESVARA, M. H. VODKIN, G. L. MCLAUGHLIN, and S. K. KATIYAR. 1996. Phylogenetic analysis of b-tubulin sequences from amitochondrial protozoa. Mol. Phylogent. Evol. **5**:359–367.

ERICKSON, H. P., D. W. TAYLOR, K. A. TAYLOR, and D. BRAMHILL. 1996. Bacterial cell division protein FtsZ assembles into protofilament sheets and minirings, structural homologs of tubulin polymers. Proc. Natl. Acad. Sci. USA **93**:519–523.

FELSENTEIN, J. 1993. PHYLIP (phylogeny inference package). Distributed by the author. University of Washington, Seattle.

FLEGEL, T. W., and T. PASHARAWIPAS. 1995. A proposal for typical eukaryotic meiosis in microsporidians. Can. J. Microbiol. **41**:1–11.

GAERTIG, J., M. A. CRUZ, J. BOWEN, L. GU, D. G. PENNOCK, and M. A. GOROVSKY. 1995. Acetylation of lysine 40 in alpha-tubulin is not essential in *Tetrahymena thermophila*. J. Cell Biol. **129**:1301–1310.

GARD, D. L. 1994. Gamma-tubulin is asymmetrically distributed in the cortex of *Xenopus* oocytes. Dev. Biol. **161**:131–140.

GUNDERSON, J., G. HINKLE, D. LEIPE, H. G. MORRISON, S. K. STICKEL, D. A. ODELSON, J. A. BREANAK, T. A. NERAD, M. MÜLLER, and M. L. SOGIN. 1995. Phylogeny of Trichomonads inferred from small-subunit rRNA sequences. J. Eukaryot. Microbiol. **24**:411–415.

HASHIMOTO, T., Y. NAKAMURA, F. NAKAMURA, T. SHIRAKURA, J. ADACHI, N. GOTO, K. OKAMOTO, and M. HASEGAWA. 1994. Protein phylogeny gives a robust estimation for early divergences of eukaryotes: phylogenetic place of a mitochondria-lacking protozoan, *Giardia lamblia*. Mol. Biol. Evol. **11**:65–71.

KAMAISHI, T., T. HASHIMOTO, Y. NAKAMURA, F. NAKAMURA, S. MURATA, N. OKADA, K.-I. OKAMOTO, M. SHIMZU, and M. HASEGAWA. 1996. Protein phylogeny of translation elongation factor EF-1α suggests microsporidians are extremely ancient eukaryotes. J. Mol. Evol. **42**:257–263.

KATIYAR, S. K., and T. D. EDLIND. 1994. Beta-tubulin genes of *Trichomonas vaginalis*. Mol. Biochem. Parasitol. **64**:33–42.

KEELING, P. J., and W. F. DOOLITTLE. 1996. A non-canonical genetic code in an early diverging protist lineage. EMBO J. **15**:2285–2290.

KIRK-MASON, K. E., M. J. TURNER, and P. R. CHAKRABORTY. 1988. Cloning and sequence of beta-tubulin cDNA from *Giardia lamblia*. Nucleic Acids Res. **16**:2733.

KOZMINSKI, K. G., D. R. DIENER, and J. L. ROSENBAUM. 1993. High level expression of nonacetylatable alpha-tubulin in *Chlamydomonas reinhardtii*. Cell Motil. Cytoskeleton **25**:158–170.

LAI, E. Y., S. P. REMILLARD, and C. FULTON. 1988. The alpha-tubulin gene family expressed during cell differentiation in *Naegleria gruberi*. J. Cell Biol. **106**:2035–2046.

LAJOIE-MAZENC, I., Y. TOLLON, C. DETRAVES et al. (13 co-authors). 1994. Recruitment of antigenic gamma-tubulin during mitosis in animal cells: presence of gamma-tubulin in the mitotic spindle. J. Cell Sci. **107**:2825–2837.

LEIPE, D. D., J. H. GUNDERSON, T. A. NERAD, and M. L. SOGIN. 1993. Small subunit ribosomal RNA of *Hexamita inflata* and the quest for the first branch of the eukaryotic tree. Mol. Biochem. Parasitol. **59**:41–48.

LI, J., K. KATIYAR, A. HAMELIN, G. S. VISVESVARA, and T. D. EDLIND. 1996. Tubulin genes from AIDS-associated microsporidia and implications for phylogeny and benzimidazole sensitivity. Mol. Biochem. Parasitol. **78**:289–295.

LUTKENHAUS, J. 1993. FtzZ ring in bacterial cytokinesis. Mol. Microbiol. **9**:403–409.

MARGOLIN, W., R. WANG, and M. KUMAR. 1996. Isolation of an *ftsZ* homologue from the archaebacterium *Halobacterium salinarium*: implications for the evolution of FtsZ and tubulin. J. Bacteriol. **178**:1320–1327.

MARSCHALL, L. G., R. L. JENG, J. MULHOLLAND, and T. SEARNS. 1996. Analysis of Tub4p, a yeast γ-tubulin-like protein: implications for microtubule-organizing center function. J. Cell Biol. **134**:443–454.

MULISCH, M. 1993. Chitin in protistan organisms. Eur. J. Protistol. **29**:1–18.

OAKLEY, B. R., C. E. OAKLEY, Y. YOON, and M. K. JUNG. 1990. Gamma-tubulin is a component of the spindle pole body that is essential for microtubule function in *Aspergillus nidulans*. Cell **61**:1289–1301.

OAKLEY, C. E., and B. R. OAKLEY. 1989. Identification of gamma-tubulin, a new member of the tubulin superfamily encoded by *mip*A gene of *Aspergillus nidulans*. Nature **338**:662–664.

PAGE, F. C., and R. L. BLANTON. 1985. The Heterolobosea (Sarcodina: Rhizopoda), a new class uniting the Schizopyrenida and the Acrasidae (Acrasida). Protisologica **21**:121–132.

ROGER, A. J., M. W. SMITH, R. F. DOOLITTLE, and W. F. DOOLITTLE. 1996. Evidence for the Heterolobosea from phylogenetic analysis of genes encoding glyceraldehyde-3-phosphate dehydrogenase. J. Eukaryot. Microbiol. (in press).

SANCHEZ, M. A., D. A. PEATTIE, D. WIRTH, and E. OROZCO. 1994. Cloning, genomic organization and transcription of the *Entamoeba histolytica* alpha-tubulin-encoding gene. Gene **146**:239–244.

SOBEL, S. G., and M. SYNDER. 1995. A highly divergent gamma-tubulin gene is essential for cell growth and proper microtubule organization in *Saccharomyces cerevisiae*. J. Cell Biol. **131**:1775–1788.

SPANG, A., S. GEISSLER, K. GREIN, and E. SCHIEBEL. 1996. γ-tubulin-like Tub4p of *Saccharomyces cerevisiae* is associated with the spindle pole body substructures that organize microtubules and is required for mitotic spindle formation. J. Cell Biol. **134**:429–441.

SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony. Version 3.1.1. Illinois Natural History Survey, Champaign, Ill.

TIBAYRENC, M., F. KJELLBERG, J. ARNAUD, B. OURY, S. F. BRENIERE, M. L. DARDE, and F. J. AYALA. 1991. Are eukaryotic microorganisms clonal or sexual? A population genetics vantage. Proc. Natl. Acad. Sci. USA **88**:5129–5133.

VOSSBRINCK, C. R., J. V. MADDOX, S. FRIEDMAN, B. A. DEBRUNNER-VOSSBRINCK, and C. R. WOESE. 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature **326**:411–414.

ZHENG, Y., M. K. JUNG, and B. R. OAKLEY. 1991. Gamma-tubulin is present in *Drosophila melanogaster* and *Homo sapiens* and is associated with the centrosome. Cell **65**:817–823.

MITCHELL SOGIN, reviewing editor

Accepted September 5, 1996