ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Fungal Genetics and Biology 38 (2003) 298–309

FUNGAL GENETICS AND BIOLOGY

www.elsevier.com/locate/yfgbi

Congruent evidence from α-tubulin and β-tubulin gene phylogenies for a zygomycete origin of microsporidia

Patrick J. Keeling*

Canadian Institute for Advanced Research, Department of Botany, University of British Columbia, 3529-6270 University Boulevard, Vancouver, BC V6T 1Z4, Canada

Received 18 April 2002; accepted 11 September 2002

Abstract

The origin of microsporidia and the evolutionary relationships among the major lineages of fungi have been examined by molecular phylogeny using α-tubulin and β-tubulin. Chytrids, basidiomycetes, ascomycetes, and microsporidia were all recovered with high support, and the zygomycetes were consistently paraphyletic. The microsporidia were found to branch within zygomycetes, and showed relationships with members of the Entomophthorales and Zoopagales. This provides support for the microsporidia having evolved from within the fungi, however, the tubulin genes are difficult to interpret unambiguously since fungal and microsporidian tubulins are very divergent. Rapid evolutionary rates a characteristic of practically all microsporidian genes studied, so determining their evolutionary history will never be completely free of such difficulties. While the tubulin phylogenies do not provide a decisive conclusion, they do further narrow the probable origin of microsporidia to a zygomycete-like ancestor.

Keywords: Microsporidia; Chytrid; Zygomycete; Tubulin; Phylogeny; Evolution

1. Introduction

Microsporidia are a group of obligate intracellular parasites that infect a wide variety of animals, as well as certain species of ciliates and gregarine apicomplexa (Vivier, 1975). The primary distinguishing features of microsporidia are found in the spore, the only stage of the microsporidian life cycle that is viable outside of a host cell. Microsporidian spores are small, typically unadorned, highly resistant structures protected by a double layered wall consisting largely of glycoprotein and chitin. Microsporidian spores are also highly reduced cytologically, lacking a variety of structures common to most other eukaryotes, including flagella and recognisable mitochondria (Vávra and Larsson, 1999). While relatively simple in these respects, microsporidian spores do contain several highly specialised structures that mediate infection; notably the polaroplast, posterior vacuole, and the polar filament or polar tube. The parasite gains entry to its host by building up osmotic pressure within the spore, and using this pressure to force the polar filament to break through the spore and evert, or turn inside out, at which point the filament becomes a tube. Once the polar tube is discharged, the cytoplasm in the spore is forced through the tube. If the discharging tube hits a potential host cell, it can penetrate the membrane of that cell so that the microsporidian emerges from the tube directly into the host cytoplasm where it will grow and divide, ultimately producing more spores.

This unique infection mechanism, together with the emerging importance of microsporidia as parasites of domestic animals and humans, have generated increasing interest in the group since they were first discovered in the late nineteenth century. In recent years, the group has also attracted considerable attention from an evolutionary perspective [for review see, (Keeling and Fast, 2002)]. Because they are so highly adapted and reduced, it has proved difficult to determine how microsporidia are related to other eukaryotes. They have, at various times, been proposed to be closely related to a number of other spore-forming parasites such as myxosporidia and actinomyxidia (Lom and Vávra, 1962), and have also been

* Fax: +1-604-822-6089.
E-mail address: pkeeling@interchange.ubc.ca.

1087-1845/02/$ - see front matter © 2002 Elsevier Science (USA). All rights reserved.
doi:10.1016/S1087-1845(02)00537-6

proposed to represent the most ancient and primitive lineage of eukaryotes (Cavalier-Smith, 1983). Early molecular data from microsporidia seemed to corroborate the view that microsporidia were ancient, as phylogenies of small subunit ribosomal RNA (SSU rRNA), elongation factors 1α and 2, and valyl-tRNA synthetase all placed microsporidia as an early lineage of eukaryotes (Brown and Doolittle, 1995; Kamaishi et al., 1996; Vossbrinck et al., 1987). In addition, the microsporidian 5.8S and large subunit (LSU) rRNA are fused, as in prokaryotes (Vossbrinck and Woese, 1986). Altogether, these data appeared to show that microsporidia were primitive eukaryotes. However, doubts continued to linger because the known microsporidian genes sequences were all very divergent, which is known to cause difficulties in reconstructing phylogenies (Kuhner and Felsenstein, 1994), and cast suspicion on the deep-branching position of microsporidia (Cavalier-Smith, 1993).

These doubts proved to be well-founded, as additional sampling of microsporidian protein-coding genes has altered this picture considerably. The first molecular phylogenies to challenge the ancient origin of microsporidia were those of α-tubulin and β-tubulin; both of which provided strong support for a relationship between microsporidia and fungi (Edlind et al., 1996; Keeling and Doolittle, 1996). Since then, a number of other gene phylogenies have demonstrated the same relationship, notably, the largest subunit of RNA polymerase II (RPB1), TATA-box binding protein, glutamyl-tRNA synthetase, seryl-tRNA synthetase, vacuolar ATPase, and transcription factor IIB (Fast et al., 1999; Hirt et al., 1999; Katinka et al., 2001). Moreover, a number of molecules that once supported an early origin of microsporidia have been re-examined with methods that consider rate variation between sites in a sequence. Upon re-analysis nearly all of these molecules do not support a deep placement of microsporidia, and in some cases weakly support the fungal relationship [e.g., LSU rRNA and EF-2 (Hirt et al., 1999; Van de Peer et al., 2000)]. Most recently, the complete genome of the microsporidian *Encephalitozoon cuniculi* has been sequenced, providing a wealth of molecular data to test the evolutionary position of microsporidia, and preliminary results indicate that the most robust phylogenies support the fungal affinity of microsporidia (Katinka et al., 2001).

While these data provide compelling support for the notion that microsporidia are somehow related to fungi, exactly how these groups are related to one another remains uncertain. Most gene phylogenies that have been used to address the origin of microsporidia are poorly sampled: typically only one microsporidian sequence is known, and fungi are represented solely by ascomycetes. This lack of sequence diversity renders it impossible to make the critical distinction between microsporidia and fungi originating from a common ancestor (i.e., microsporidia are a sister group to fungi), versus microsporidia evolving from a *bona fide* fungus (i.e., microsporidia are themselves fungi). The only gene that has been sampled sufficiently from both microsporidia and fungi to make any such distinction is β-tubulin. In this case, phylogenies suggest that microsporidia evolved from within the fungi and hint that they are related to either ascomycetes or zygomycetes, but fail to identify a more specific position (Keeling et al., 2000).

Here, the relationship between microsporidia and fungi has been addressed with α-tubulin and β-tubulin independently, and in combination. The tubulin genes from microsporidia and all fungi except chytrids are very divergent (perhaps because these organisms lack $9+2$ microtubule structures, and accordingly their tubulins are under relaxed evolutionary constraints). This accelerated rate of divergence in microsporidia and some fungi means that the position of microsporidia within tubulin phylogenies must be interpreted with a great deal of caution since between taxon rate heterogeneity can lead to erroneous trees with strong support (for instance the original SSU rRNA trees that placed microsporidia at the base of fungi). However, it is difficult to avoid this problem with the microsporidia since the accelerated rate of substitution seen in the first genes characterised from the group appears to be true for most or all of the genome (Katinka et al., 2001). It appears that resolving the exact relationship between microsporidia and fungi will require a careful analysis of many proteins, none of which will be ideal individually. Of the microsporidian genes that have been analysed to date, the tubulins show the greatest consistent support for any relationship between microsporidia and fungi, so they are a credible starting point, although once again, they must be interpreted cautiously. Tubulin genes from several microsporidia, chytrids, and zygomycetes have been sequenced so that the two tubulin genes are roughly equally sampled and represent a broad spectrum of fungal diversity. Phylogenies based on both tubulins individually and in combination all converge on the conclusion that microsporidia are derived from zygomycete fungi, and are most likely related to insect pathogens belonging to the Entomophthorales and Zoopagales. The tubulins are also now among the more widely sampled protein-coding genes from fungi, and lend some interesting insights into deep fungal evolution and fungal gene structure.

2. Materials and methods

2.1. Strains

Nosema locustae (ATCC 30860) spores were a gift from M&R Durango (Bayfield, CO). DNA was prepared from approximately 5×10^9 spores as described (Fast and Keeling, 2001). Microsporidian DNA from *Glugea*

plecoglossi and *Trachipleistophora hominis* were gifts from T. Hashimoto (Rockefeller University) and Sarah Cheney (Imperial College of Science, Technology and Medicine), respectively. DNA from the chytrid fungi *Karlingiomyces* sp. (strain JEL93), *Rhizophydium* sp. (strain JEL138) were gifts from Timothy James (Duke University) from the culture collection of Joyce Longcore (University of Maine). DNA from the chytrid fungi *Nowakowskiella hemisphaeorspora* (strain BK85-6), *Nowakowskiella elegans* (strain BK50-1), *Gaertneromyces semiglobiferus* (strain BK91-10), *Powellomyces variabilis* (strain BK85-1) were gifts from John Taylor (University of California, Berkeley). DNA from the zygomycete *Rhizopus microsporus* var. *oligosporus* (strain NRRL 2710) was a gift from Franz Lang (Université du Montréal). DNA from the zygomycetes *Micromucor ramannianus* (strain NRRL 5844), *Furculomyces boomerangus* (strain AUS-42-7), *Spiromyces minutus* (strain NRRL 3067), *Syncephalis depressa* (strain NRRL 22627), *Basidiobolus ranarum* (strain NRRL 20525), *Conidiobolus coronatus* (strain NRRL 1912), "*Conidiobolus lamprauges*" (strain NRRL 28637), *Entomophaga maimaiga* (strain ARSEF 1400), "*Entomophaga destruens*" (strain NRRL 3727), and *Capniomyces stellatus* (strain MIS-10-108) were gifts from Kerry O'Donnell (United States Department of Agriculture, National Center for Agricultural Utilization Research).

2.2. Amplification and sequencing of tubulin genes

α-Tubulin genes were amplified from *B. ranarum*, *S. depressa*, *C. stellatus*, "*E. destruens*", "*C. lamprauges*," and *C. coronatus* using the primers GGGCCCCAGG TCGGCAAYGCNTGYTGG and GGGCCCCGAGA ACTCSCCYTCYTCCAT. All other α-tubulins were amplified using the primers TCCGAATTCARGTNGG NAAYGCNGGYTGGGA and CGCGCCATNCCYT CNCCNACRTACCA. Both of these primer sets amplify approximately 95% of the α-tubulin gene. So that α-tubulin and β-tubulin data could be compared and combined, β-tubulin genes were amplified from *N. locustae*, "*C. lamprauges*," "*E. destruens*," *N. hemisphaeorspora*, *N. elegans*, *Karlingiomyces* sp., *Rhizophydium* sp., and *P. variabilis* using primers GCCTGCAGGN CARTGYGGNAAYCA and GGCCTCAGTRAAYT CCATYTCRTCCAT. All PCR products were cloned into the vector pCR 2.1 using the TOPO TA cloning kit (Invitrogen), and both strands of multiple clones sequenced using BigDye dideoxy terminator chemistry. Thirty-seven new sequences have been deposited in GenBank under Accession Nos. AY138769–AY138805.

2.3. Phylogenetic analysis

New sequences were added to existing protein alignments and trees inferred using distance and maximum likelihood methods. The alignments used to infer phylogenies consisted of 58 sequences and 374 characters for α-tubulin and 64 sequences and 387 characters for β-tubulin. Tubulins are highly conserved in sequence and length, so no significant gaps exist in the data. All sequences were tested for significant deviations from amino acid composition by a χ^2 test using TREE-PUZZLE version 5.0 (Strimmer and von Haeseler, 1996), and none of the sequences used were found to deviate significantly (two highly divergent ascomycetes α-tubulins from *Neurospora* and *Emericella* were found to deviate in amino acid composition, so these sequences were excluded from all subsequent analyses). Insertions are typically restricted to one sequence, and these were not considered in the analysis. For individual α-tubulin and β-tubulin trees, distance trees including representatives of major eukaryotic groups, as well as trees including only animals, choanoflagellates, fungi, and microsporidia, were inferred using distance methods. Maximum likelihood distances were calculated using the WAG substitution frequency matrix, and corrected for site-to-site rate variation according to a gamma distribution with the eight rate categories plus invariable sites and the shape parameter α estimated from the data (α parameters were estimated at 0.75 and 0.44 for α- and β-tubulins, respectively, and the proportion of invariable sites was estimated at 0.06 and 0.00 for α- and β-tubulins, respectively). Trees were inferred by weighted neighbor-joining [using WEIGHBOR version 1.0.1 α (Bruno et al., 2000)] and Fitch–Margoliash [using FITCH version 3.6 α (Felsenstein, 1993)]. All methods gave very similar results, and all strongly supported nodes were found with all methods, so only weighted neighbor-joining trees are shown. Bootstrap resampling was carried out using PUZZLEBOOT (shell script by A. Roger and M. Holder, http://www.tree-puzzle.de) under the same conditions, except that the alpha parameter and the proportion of invariable sites from the original data set were used with each replicate.

Trees combining α-tubulin and β-tubulin were also inferred. The two individual phylogenies are very similar, and only one branching position was found to conflict between the two trees with bootstrap support above 65% in both trees, which was the criterion for combining tubulins and other protein data recently applied to a question of even deeper eukaryotic phylogeny (Baldauf et al., 2000). The only taxon for which this criterion did not hold was the microsporidian *Sprague lophii*, which branched on either side of the "root" of microsporidia. This taxon was left in the analysis as there were so few microsporidia, and the position of the root of microsporidia is discussed in the text (see below). Representatives were chosen for this analysis with the aim of optimising the taxonomic diversity of each of the four fungal groups and the microsporidia. In general, each fungus for which both genes are known was

included. In cases where two copies of one or the other gene were known, but the two copies were closely related (which was nearly always the case), the more conserved copy was selected. In cases where multiple paralogues existed for both genes (e.g., *Rhizopus*), they were always closely related, so paralogues were combined according to their relative branching position in each individual tree, so that genes in the same relative position were combined with one another. In two cases, one gene was chosen in order to maximise the congruency between the two individual gene phylogenies. In the first instance, one *Spiromyces* β-tubulin branches at the same position as *Spiromyces* α-tubulin (at the base of the harpellales), while a second copy branches elsewhere in the zygomycetes. In this instance the first copy was chosen as its phylogenetic position matched that of the gene with which it was combined. In the second instance, a divergent paralogue of β-tubulin that was described previously in chytrids (Keeling et al., 2000) was found to be widespread among members of this group. The two chytrid paralogues were not observed to branch together in phylogenies, however the extreme difference between the substitution rates of these genes could easily cause the more divergent type to branch with the other more divergent fungal genes. This was tested using AU tests (see below) which suggests that the chytrid tubulins probably all descended from a common ancestral gene. With this in mind, only the short-branched chytrid genes were considered in the combined analysis. A data set of 45 sequences and 761 characters was used to infer distance trees (as described above) and with protein maximum likelihood using ProML version 3.6 α (Felsenstein, 1993), as the reduced number of taxa in the combined data allowed ML analysis. In this case, the JTT substitution frequency matrix was used and site-to-site rate variation was modeled on a gamma distribution using the R-option with six rate categories plus invariable sites, with rates and frequencies estimated by TREE-PUZZLE (the α parameter was estimated to be 0.61 and the proportion of invariable sites 0.09). Protein ML bootstrapping was done in the same way, except with four categories of rates and invariable sites as the computational requirements were too great to consider six categories of sites.

To more directly test whether microsporidia really do branch within the fungi, as opposed to branching as sister group to the fungi, alternative topologies of the individual genes and the combined data were tested using the Approximately Unbiased (AU) test (Shimodaira, 2002). Hypothetical alternative topologies were constructed by moving microsporidia (and in β-tubulin both microsporidia and chytrids) to alternative positions in the trees shown. The likelihood at each character position was calculated for each alternative topology, as well as the topology under scrutiny, using PAML version 3.12 (Yang, 1997). Likelihoods were calculated using the WAG and gamma distributed rates, with α parameters either estimated by PAML or as estimated previously by TREE-PUZZLE (both estimated very similar parameters). Likelihoods were used to calculate *p*-values using CONSEL version 0.1d (Shimodaira and Hasegawa, 2001).

3. Results and discussion

3.1. Characterisation of tubulin genes from microsporidia, chytrids, and zygomycetes

Using primers that span approximately 95% of the gene, 26 new α-tubulins were amplified and sequenced from 6 chytrid species, 11 zygomycete species, and 2 microsporidian species. In addition, to make the α-tubulin and β-tubulin data sets as complementary as possible, 11 new β-tubulin genes were amplified and sequenced from 5 chytrid species, 2 zygomycete species, and one microsporidian (see Section 2 for species and strain designations). With these new data, the representation of chytrid, zygomycete, and microsporidian tubulins encompasses a broad taxonomic range. A recent phylogenetic analysis of chytrid SSU rRNA genes identified eight distinct clades, and a number of independent taxa (James et al., 2000). Of these, six are represented here by either α-tubulin or β-tubulin, with the "*Chytrium*-clade" and the Neocallimastigales being underrepresented in either data set. It was previously noted that some chytrids contained two β-tubulins: one highly conserved and one highly divergent (Keeling et al., 2000). The presence of this second, divergent β-tubulin was confirmed in five genera representing four of the eight chytrid lineages. Traditional classifications and recent molecular phylogenetic analyses recognise eight distinct groups of Zygomycota (including zygomycetes and trichomycetes) (Tanabe et al., 2000), of which five are represented here in both α-tubulin and β-tubulin data sets (including several harpellalean trichomycetes). Lastly, phylogenies of microsporidia based on SSU rRNA and RPB1 are consistent with four distinct microsporidian groups, plus a poorly understood collection of earlier-diverging lineages (Baker et al., 1995, 1997; Bell et al., 2001; Cheney et al., 2001; Keeling and McFadden, 1998; Nilsen and Chen, 2001; Vossbrinck et al., 1996). The characterised microsporidian α-tubulin and β-tubulin genes now span three of the major groups of microsporidia.

Since protein-coding genes have not been widely sampled from fungi in many instances, some characteristics of the diverse chytrid and zygomycete tubulin sequences deserve note. In particular, the distribution and characteristics of spliceosomal introns is unusual and interesting. In general, the zygomycete tubulins are very divergent and have a very low intron density (most

genes lack introns). In contrast, both α- and β-tubulins from chytrids are among the most conserved tubulins known, and the intron density of both is very high: every chytrid gene contained at least one intron, and in total, 22 independent intron positions were identified. In some respects, the distribution of these positions (Fig. 1) is consistent with chytrid taxonomy and the molecular phylogeny of the genes where they are found (see below), however, most aspects of the distribution are consistent with rapid intron gain and loss in chytrid tubulins. For example, β-tubulin intron position 1 is occupied only in *Powellomyces* and *Rhizophydium* (which are classified in two different orders and are unrelated in SSU phylogeny), while position 2 is shared by *Rhizophydium*, *Rhizophlyctis*, and *Spizellomyces*, but not *Powellomyces*, and intron position 11 is found only in *Powellomyces*, *Rhizophlyctis*, and *Spizellomyces* (all classified in one order, but *Rhizophlyctis* is unrelated to the others in SSU phylogeny). Similar incongruence can be seen at several positions, altogether suggesting that intron evolution in chytrids is very active, and probably involves both gain and loss at a higher than typical frequency.

In addition to their distribution, some of the sequence characteristics of the chytrid introns also bear notice. In general, chytrid introns tend to be fairly short, but those of *Rhizophydium* are especially so. To date, the smallest introns known occur in the reduced algal genome of the *Chlorarachnion* nucleomorph, where all introns are 18, 19, or 20 bp in length (Gilson and McFadden, 1996). Other than this unusual genome, the ciliate *Paramecium* has the smallest introns, typically from 25–26 bp, with extremes of 18–35 bp (Sperling et al., 2002). The four introns found in *Rhizophydium* tubulins are only 22, 23, 25, and 37 bp in length, in the same league as those of *Paramecium*. More interesting, however, are the introns of *Karlingiomyces*. Spliceosomal introns are almost universally bounded by GT and AG splice boundaries (referred to as GT–AG introns). Only a few exceptions to this rule are known, the best characterised being the non-canonical AT–AC introns found in animals, plants, and recently zygomycete fungi. In animals, AT–AC introns have been shown to use a spliceosome with some unique components, suggesting that the alteration of splice sites is a non-trivial event. In the *Karlingiomyces* β-tubulin paralogue 2, however, three introns are bounded by a novel, non-canonical sequence, CT–AC. In all three cases, the positions of these introns corresponds to the position of a canonical GT–AG intron in the closely related *N. elegans* β-tubulin paralogue 2, and two of the three positions are also occupied by canonical introns in *Spizellomyces* and *Powellomyces* paralogue 2 genes (Fig. 1). This eliminates any doubt that these are in fact introns, and also indicates that they arose very

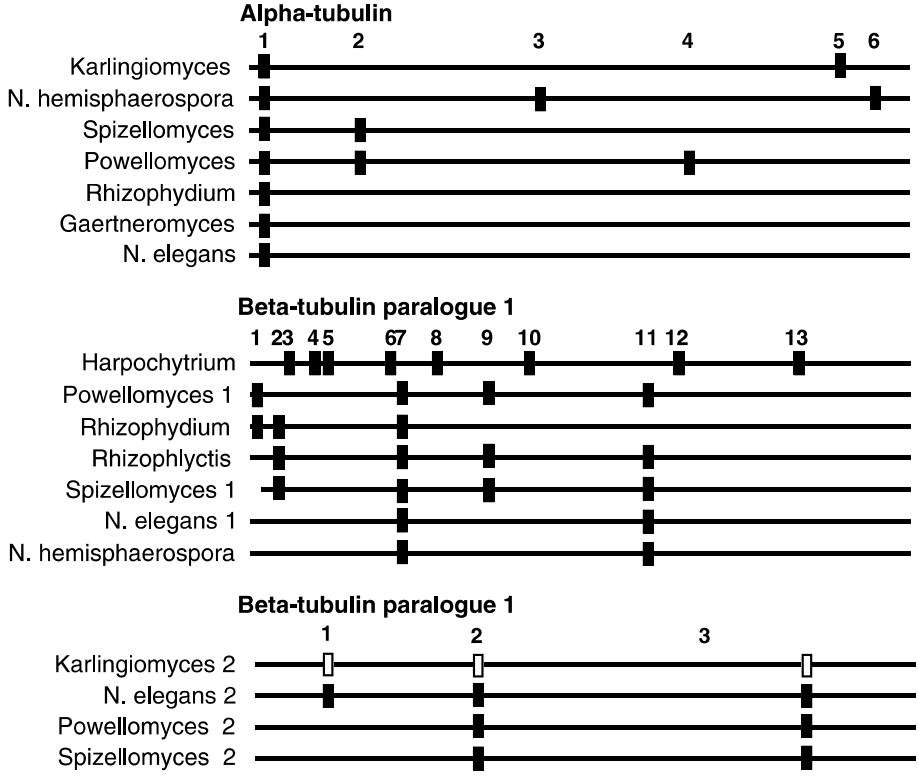

Fig. 1. Intron positions of chytrid α-tubulins and β-tubulins. The position number is indicated at the top of each diagram and introns are marked by boxes in the taxa where they appear beneath the number. Sequences are grouped according to their traditional taxonomic position and their position in SSU phylogenies. The non-canonical CT–AC introns in *Karlingiomyces* are open boxes.

recently from canonical GT–AG introns in this gene by two G-to-C substitutions, presumably though another, intermediate splice site sequence. Whether *Karlingiomyces* has more such introns is unknown, but the two α-tubulin genes from this organism each contain two canonical GT–AG introns, so it is clear that this organism uses two different classes of intron, and non-canonical introns do evolve directly from canonical GT–AG introns.

3.2. Phylogeny of α-tubulin and β-tubulin, independently and combined

New α-tubulins and β-tubulins were added to an existing alignment representing available eukaryotic diversity, and phylogenies inferred using sequences from a wide selection of eukaryotes (not shown), confirming the relationship between fungi and microsporidia in global phylogenies that has been demonstrated in numerous previous analyses (Baldauf et al., 2000; Edlind et al., 1996; Keeling and Doolittle, 1996; Keeling and Fast, 2002; Keeling et al., 2000). In order to allow a comprehensive analysis of a wide diversity of fungal sequences, all subsequent analyses were restricted to animals, a choanoflagellate (a close relative of animals), and fungi, including the microsporidia. The animal/choanoflagellate outgroup was chosen since these are the closest relatives of microsporidia and fungi in tubulin phylogenies (Keeling and Doolittle, 1996; Keeling et al., 2000), and other phylogenies [e.g., see (Baldauf et al., 2000)].

In the α-tubulin phylogeny of fungi (Fig. 2), fungi form a monophyletic group with strong support, as do the chytrids and basidiomycetes. Ascomycetes are monophyletic (though not well supported), and zygomycetes are paraphyletic. Microsporidia form a strongly supported monophyletic group within the zygomycetes, and fall specifically within a lineage composed of the Entomophthorales (*Conidiobolus* and *Entomophaga*) and Zoopagales (*Syncephalis*). This grouping was found in all analyses of α-tubulin, and is relatively well supported by bootstrap (82% in weighted distance and Fitch–Margoliash), although the exact position of microsporidia within this group was not supported. This position was compared with alternate topologies using the AU test. In topologies where the microsporidia were placed as sisters to chytrids, basidiomycetes, or ascomycetes were rejected, while topologies where microsporidia were sisters to all fungi except chytrids or, interestingly, sisters to fungi as a whole, were not rejected at the 5% level.

The microsporidia have been proposed to be related to harpellalean trichomycetes based on superficial similarities between the microsporidian polar tube and the harpellalean apical spore body (Cavalier-Smith, 1998). However, microsporidia were never found to show a specific relationship with the harpellales included here, *Furculomyces* and *Capniomyces*. Instead, the harpellales showed a consistent and well-supported relationship with Spiromyces, a relationship is seen in some SSU analyses (O'Donnell et al., 1998; Tanabe et al., 2000). Several relationships within the Entomophthorales were resolved with apparently strong support, and the genera *Conidiobolus* and *Entomophaga* were intermixed. Although there is a potential for paralogy to interfere with resolving such close relationships in tubulin phylogenies, it does appear that some generic designations should be examined, in particular "*E. destruens*" and "*C. lamprauges*," which are relatively divergent and closely related to one another and *E. maimaiga*.

The relationships within the microsporidia are also largely consistent with existing taxonomy and previous molecular analyses, but the relationships among microsporidia have only been addressed at the molecular level using SSU rRNA, and recently RPB1. rRNA analyses indicate four major microsporidian groups (called Groups I–IV: (Baker et al., 1997; Keeling and McFadden, 1998). The available sampling of RPB1 is not completely overlapping with SSU analyses, but in general RPB1 phylogeny supports the monophyly of Groups I and IV [unfortunately the nomenclature for RPB1 phylogeny is different from SSU, and these two groups are referred to as Groups II and I, respectively (Cheney et al., 2001)]. In α-tubulin phylogenies, *Glugea* and *Trachipleistophora* form a strongly supported clade (corresponding to Group I in SSU phylogeny), while the *Nosema* (Group IV) and *Encephalitozoon* (Group III) are united by a highly supported branch also seen in SSU rRNA trees (Baker et al., 1995, 1997; Keeling and McFadden, 1998; Nilsen and Chen, 2001; Vossbrinck et al., 1996). In contrast, the position of *Spraguea* as sister to Groups III and IV is in conflict with both SSU and RPB1 phylogenies where it falls within Group I and Group IV, respectively (Cheney et al., 2001; Nilsen and Chen, 2001).

Perhaps the most unusual group are the chytrids, where there is practically no molecular diversity at the amino acid level (as shown by the very short branch lengths in Fig. 2). Of the α-tubulins in Fig. 2, four are identical at the amino acid level, and the other five are nearly so (varying at only five positions among all five genes). In contrast, the variation at synonymous sites is high (only 17 synonymous positions have not been substituted in at least one sequence), and there is no detectable sequence conservation between introns that fall at homologous positions (e.g., intron position 1 in Fig. 1, which is found in all chytrids). Altogether, it appears that the chytrid α-tubulins are under strong selection as opposed to representing limited divergence time.

β-Tubulins have already been relatively widely sampled from chytrids, zygomycetes, and microsporidia

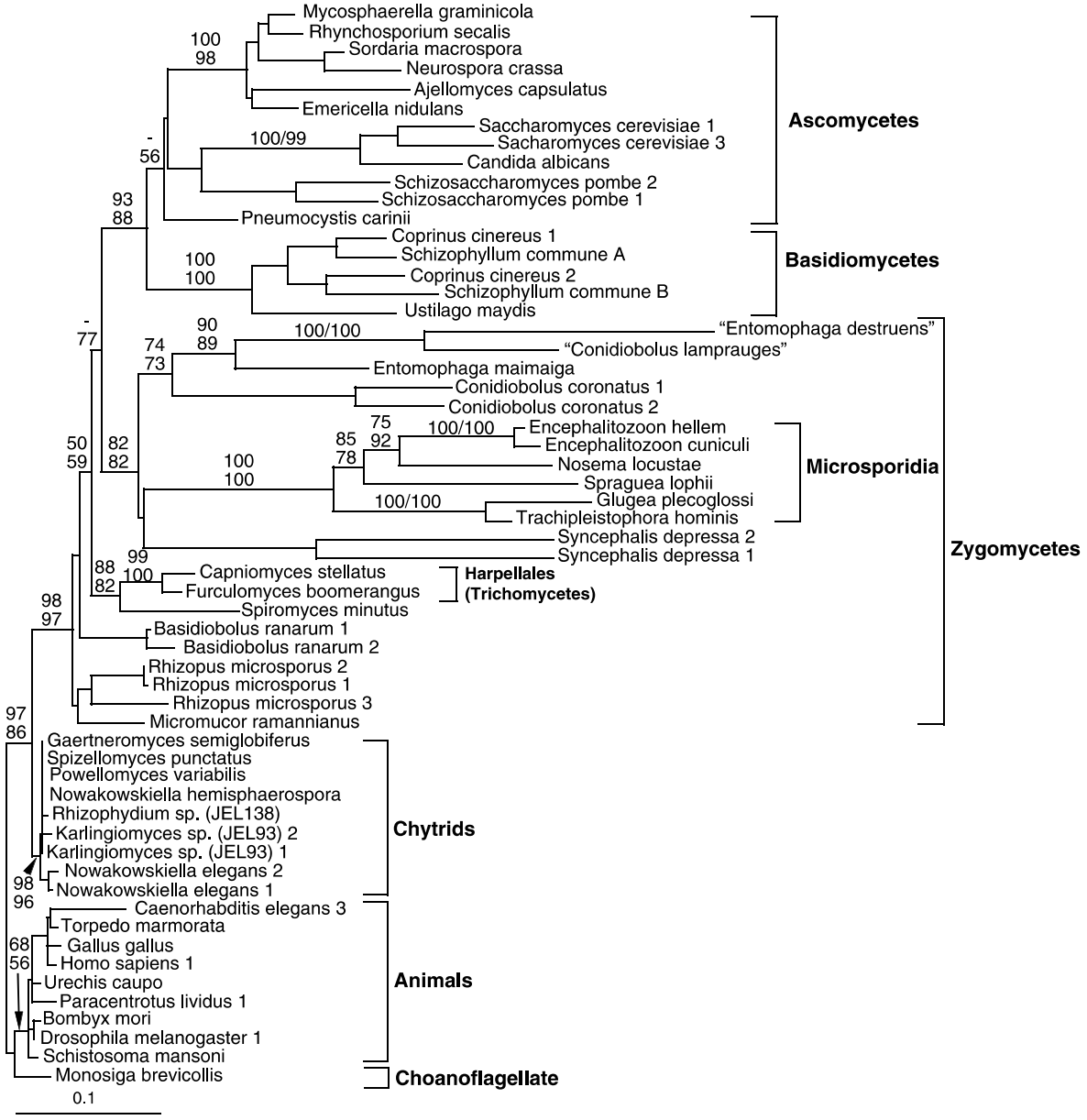

Fig. 2. Weighted neighbor joining tree based on gamma-corrected distances of α-tubulin from diverse fungi and microsporidia, with animals and choanoflagellates as outgroup taxa. Taxonomic groups are bracketed to the right. Note that the harpellalean trichomycetes are nested within the zygomycetes. Taxa are labeled by genus and species, unless no species designation exists, in which case the strain designation is given in brackets. Numbers following taxon names and outside brackets indicate the copy number where multiple genes have been characterised. Numbers at nodes correspond to bootstrap support from weighted neighbor joining (top) and Fitch–Margoliash (bottom). Scale bar represents 0.1 (corrected) substitutions per site.

(Keeling et al., 2000), but the taxon selection in α-tubulin and β-tubulin is non-overlapping in many instances. Accordingly, β-tubulins were sequenced from members of each group and the phylogeny re-analyzed with these new sequences (Fig. 3). In all analyses the ascomycetes, basidiomycetes, microsporidia, and the fungi as a whole formed strongly supported monophyletic lineages. The chytrids formed a strongly supported monophyletic lineage of conserved sequences (paralogue 1), and a second, weak lineage of uncertain position made up of the highly divergent sequences (paralogue 2), previously observed in *Spizellomyces* and a partial gene sequence from *Harpochytrium* (Keeling et al., 2000). Although not well supported by bootstrap, the divergent paralogue 2 genes all contain two common introns (Fig. 1), and almost certainly diverged from a common ancestor. To test the relationship between the two chytrid paralogues, an alternative topology where the chytrid genes were monophyletic was compared with the topology where they are paraphyletic (shown in

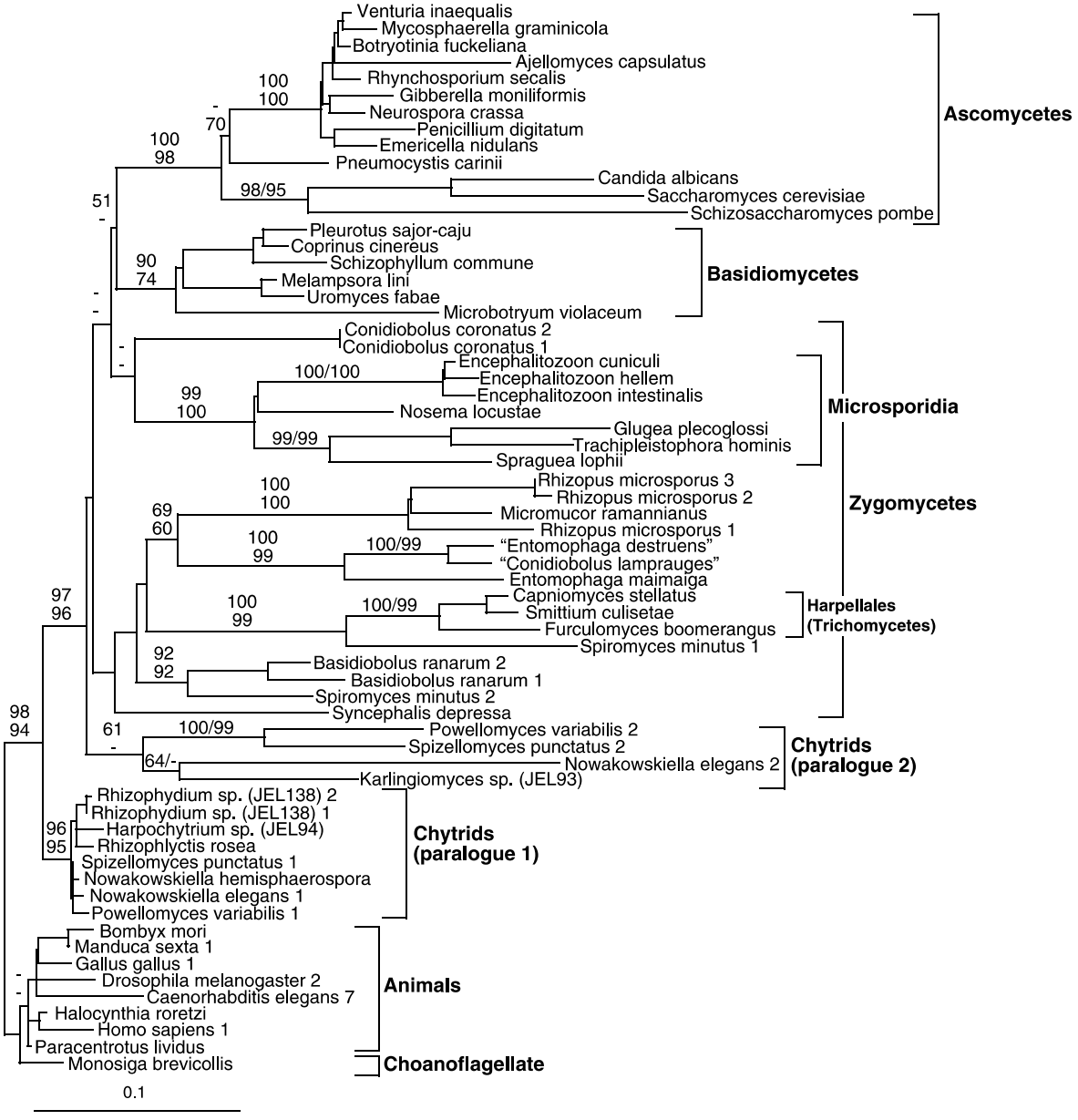

Fig. 3. Weighted neighbor joining tree based on gamma-corrected distances of β-tubulin from diverse fungi and microsporidia, with animals and choanoflagellates as outgroup taxa. Figure notations are as in Fig. 2.

Fig. 3), and interestingly not only was the alternate, monophyletic topology not rejected, but was actually found to be the favoured topology by the AU test. Not surprisingly, the paraphyletic topology was not rejected at the 5% level.

As in α-tubulin phylogenies, microsporidia fall within the zygomycetes, forming a weak but consistent clade with *C. coronatus*, but not the remainder of the Entomophthorales (which branch with the Mucorales, *Micromucor* and *Rhizopus*). Alternative topologies were also tested for β-tubulin using the AU test, and in this case topologies where microsporidia were sisters to ascomycetes, basidiomycetes, or chytrids were all rejected, while topologies where microsporidia sisters to all fungi except chytrids were not rejected. The topology where microsporidia were sisters to fungi as a whole was rejected at the 5% level. Once again, the microsporidia show no relationship to the harpellales, *Capniomyces*, *Smittium*, and *Furculomyces*, which, as in the α-tubulin phylogenies, show a strong relationship with one of two paralogues from *Spiromyces*. Within the Entomophthorales, the topology is congruent with that of α-tubulin, where "*E. destruens*" is closely related to "*C. lamprauges*." This suggests that this relationship is less likely to be an artifact of paralogy, and more likely reflects the actual branching order of these species.

Within the microsporidia, the overall branching order of major groups is similar to that seen in α-tubulin trees. *Nosema* and *Encephalitozoon* form an unsupported clade, while the relationship between *Glugea* and *Trachipleistophora* is well supported. The position of *Spraguea* is inconsistent with α-tubulin, but it is important to note that this reflects uncertainty with respect to the position of the root of the microsporidia rather than the topology within the group, and so either result should be considered with caution.

The overlap in taxon sampling between the two data sets is substantial, and the overall topology of the two trees is also very similar, so a combined analysis was performed with a total of 761 amino acid characters. A maximum likelihood tree (ln likelihood = −16093.21013) inferred from this data is shown in Fig. 4. Here, the overall topology of the fungi is much as expected, and the chytrids, basidiomycetes, ascomycetes, microsporidia, and fungi as a whole are all highly supported by distance and maximum likelihood bootstrap analyses. Not surprisingly, zygomycetes are paraphyletic, and the microsporidia fall within the zygomycetes forming a moderately supported clade with the Entomophthorales and Zoopagales. The branching orders within the various groups are much the same as the well-supported relationships seen in individual analyses. Alternative topologies were also tested using this combined data using the AU test. In this case the topology where mi-

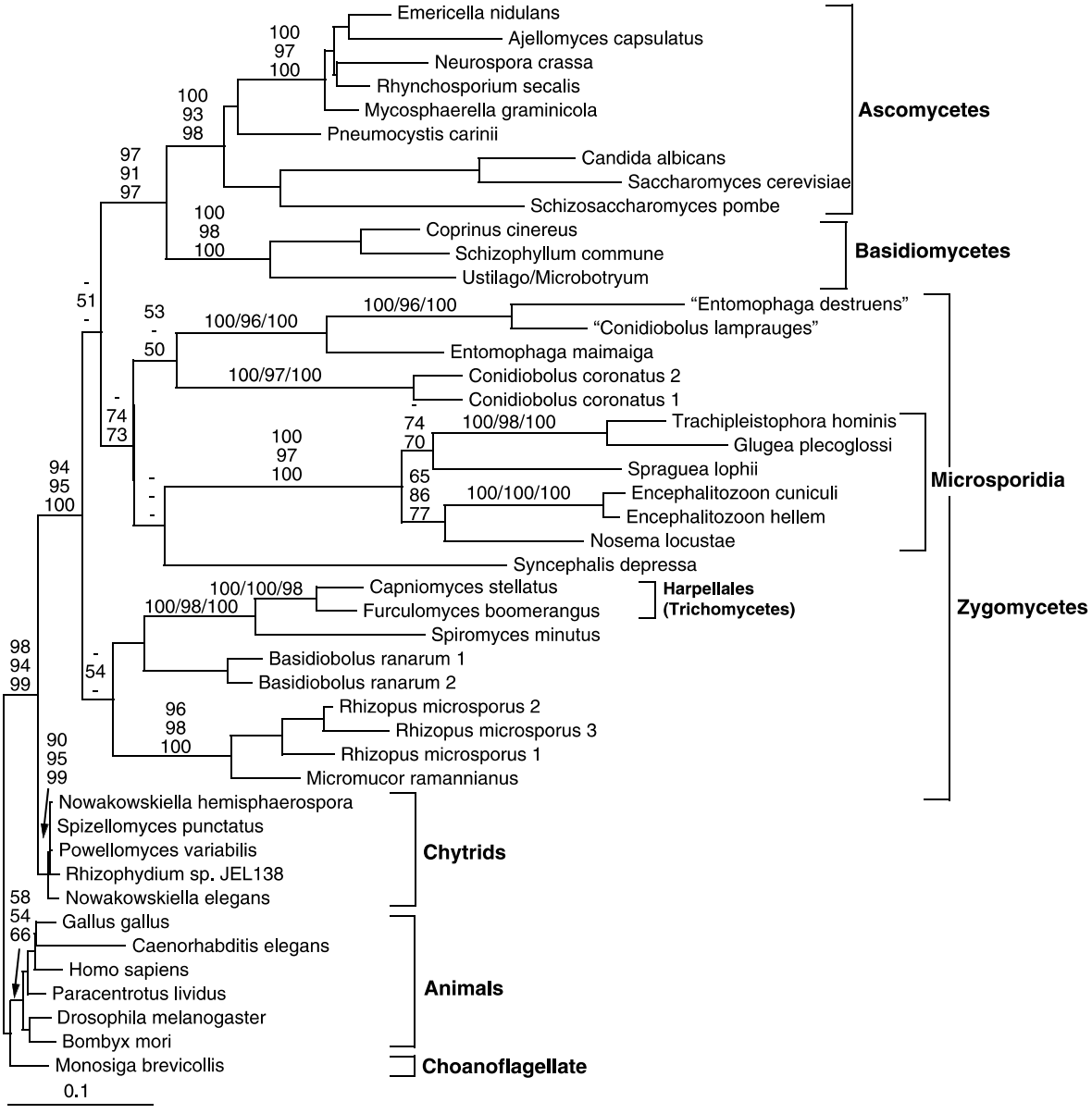

Fig. 4. Gamma-corrected protein maximum likelihood phylogeny based on combined α- and β-tubulin sequences. Figure notations are as in Fig. 2, except that numbers at node correspond to bootstrap support from gamma-corrected protein maximum likelihood (top), weighted neighbor joining (centre), and Fitch–Margoliash (bottom).

crosporidia were sisters to all fungi except chytrids was not rejected, while topologies where microsporidia were sisters to ascomycetes, basidiomycetes, chytrids, or fungi as a whole were all rejected at the 5% level.

3.3. Insights into fungal molecular phylogeny and the origin of microsporidia

The fungi are one of the major lineages of eukaryotes, and yet many of the fundamental aspects of fungal molecular phylogeny remain debated. In part, this is due to the relatively limited number of genes that have been applied to the problem of fungal phylogeny. The vast majority of molecular analyses of fungal phylogeny have focused on rRNA and, not surprisingly, reach similar conclusions and have difficulties with the same regions of the tree (Bruns et al., 1992; James et al., 2000; Tanabe et al., 2000). Only recently have protein-coding genes been sampled from diverse fungi to address these questions; EF-1 α and actin have been sampled from a large number of zygomycetes (Voigt and Wostemeyer, 2001), β-tubulin has been used to test the position of microsporidia within fungi (Keeling et al., 2000), and whole mitochondrial genomes have been analysed for a small number of diverse fungi (Paquin and Lang, 1996). From these analyses it is clear that fungal phylogeny is like other phylogenetic problems: each gene has its strengths and weaknesses, and no believable, comprehensive phylogeny will emerge from the analysis of a single gene.

Tubulins are clearly not ideal markers for fungal phylogeny, due to the obvious rate heterogeneity between different taxa. However, a number of important features of fungal evolution are supported by both tubulin phylogenies, as well as the combined analyses, and these should be noted. First, the major lineages of fungi are all well supported, except the zygomycetes, which are paraphyletic. In other analyses of fungal phylogeny based on rRNA, the ascomycetes and basidiomycetes are both highly supported, but the distinction between zygomycetes and chytrids is not clear, and is at odds with traditional taxonomy based on characteristics such as presence of flagella (Bruns et al., 1992; James et al., 2000; Tanabe et al., 2000). Here, however, the separation of chytrids and zygomycetes is well supported and matches traditional taxonomy, although there are reasons to be cautious in making this claim with genes involved in microtubule structure (see below). In addition, some of the relationships with zygomycetes are of interest, in particular relationships among the genera *Conidiobolus* and *Entomophaga*, the grouping of *Spiromyces* and harpellales, and the association between the Entomophthorales and Zoopagales.

One of the more intriguing single characteristic of these trees is the placement of the microsporidia. Two outstanding questions relating to microsporidian evolution are whether they are truly fungi or sisters to fungi and, if the former is true, from what kind of fungus did they evolve? In all tubulin phylogenies observed here, the microsporidia fall within the zygomycetes, and in α-tubulin and combined-tubulin phylogenies they specifically branch with the Zoopagales and the (or some of the) Entomophthorales with relatively strong bootstrap support. In addition, the branches uniting the basidiomycetes and ascomycetes at the exclusion of microsporidia are highly supported in both α-tubulin and combined-tubulin phylogenies, and trees in which microsporidia are sisters to either of these groups are rejected by AU tests. This suggests that microsporidia are not related to either of these groups, despite earlier data based on β-tubulin alone (Keeling et al., 2000). On the other hand, topology comparisons using the AU tests consistently fail to reject some alternate positions (microsporidia branching after chytrids, but before all other fungi), and a sister relationship between microsporidia and fungi is not significantly rejected by α-tubulin. Taken together, these results lend support to the notion that microsporidia are highly adapted, extremely reduced parasitic fungi, perhaps related to zygomycetes such as Entomophthorales or Zoopagales. Although a sister relationship with fungi is not rejected by α-tubulin data alone, there is currently no evidence whatsoever that supports microsporidia being sisters to fungi. Members of the Entomophthorales and Zoopagales are frequently parasitic, many on other fungi and many on invertebrate animals (Tanabe et al., 2000). The apparently basal lineages of microsporidia are invertebrate parasites, so it is interesting to speculate on an ecological connection between the origin of microsporidia and the insect parasites of these two groups of zygomycetes. In addition, the earlier suggestion that microsporidia might be somehow related to *Conidiobolus* led to the suggestion that the spore dispersal mechanism of *Conidiobolus* might somehow be related to the eversion of the polar tube in microsporidia (Keeling et al., 2000), since both of these processes involved everting invaginated membrane structures using a controlled release of osmotic pressure (Ingold, 1971; Keohane and Weiss, 1999). While these speculations are tempting to consider, they remain somewhat premature since the data are still a very long way from providing a strong consensus for the origin of microsporidia, and even the present suggestion of a relationship with Entomophthorales and Zoopagales only narrows the putative origin of microsporidia to a relatively diverse group of zygomycetes.

It is fitting to end on a cautionary note regarding the tubulin genes and phylogenies. One of the defining features of the ascomycetes, basidiomycetes, and zygomycetes is their lack of flagella and any other 9 + 2 microtubule structures throughout their life cycle. Chytrids also lack flagella most of the time, but do produce them in their motile gametes, so it is thought

that the loss of flagella took place between the divergence of chytrids and zygomycetes. However, this event is confused by the considerable controversy regarding the demarcation between these two groups, mostly based on SSU rRNA phylogenies failing to support the traditional taxonomic divisions (Bruns et al., 1992; James et al., 2000; Tanabe et al., 2000). It has been observed previously that the loss of flagella in various lineages has led to a noticeable acceleration of substitution rates in tubulin genes (Keeling and Doolittle, 1996), including those of fungi other than chytrids. Microsporidia also lack any 9 + 2 microtubule structures and their tubulins are accordingly divergent. In light of this, tubulin is perhaps not an ideal marker with which to examine the relationships among the fungal groups and microsporidia. However, no other gene that has been characterised thus far is a better candidate because no other microsporidian gene retains as strong a phylogenetic signal for a fungal origin as that seen in the tubulins. This will certainly soon change as the complete sequence of the *E. cuniculi* genome contains a number of genes (e.g., TF IIB, V-ATPase-A, Ser-tRNA synthetase, and Glu-tRNA synthetase), previously uncharacterised from microsporidia, that also retain a very strong phylogenetic signal of their fungal origin (Katinka et al., 2001). Each of these genes has a strong potential to be a superior phylogenetic marker than either of the tubulin genes. Developing the fungal diversity of these genes will be the next step in determining the relationships between microsporidia and other fungal groups, but at present these genes are only known from one microsporidian and a handful of fungi.

Acknowledgments

This work was supported by a grant from the Canadian Institutes for Health Research (MOP-42517). PJK is a Burroughs-Wellcome Fund New Investigator in Molecular Pathogenic Mycology, and Scholar of the Canadian Institute for Advanced Research and the Michael Smith Foundation for Health Research. Thanks to Naomi Fast, James Harper, and Celeste Leander for critical comments on the manuscript, Andrew Roger for discussing phylogenetic methods, and T. Hashimoto, Franz Lang, Sarah Cheney, Tim James, Joyce Longcore, John Taylor, and Kerry O'Donnell for generously providing samples to make this work possible.

References

Baker, M.D., Vossbrinck, C.R., Becnel, J.J., Maddox, J.V., 1997. Phylogenetic position of *Amblyospora* Hazard & Oldacre (Microspora: Amblyosporidae) based on small subunit rRNA data and its implication for the evolution of the microsporidia. J. Eukaryot. Microbiol. 44, 220–225.

Baker, M.D., Vossbrinck, C.R., Didier, E.S., Maddox, J.V., Shadduck, J.A., 1995. Small subunit ribosomal DNA phylogeny of various microsporidia with emphasis on AIDS related forms. J. Eukaryot. Microbiol. 42, 564–570.

Baldauf, S.L., Roger, A.J., Wenk-Siefert, I., Doolittle, W.F., 2000. A kingdom-level phylogeny of eukaryotes based on combined protein data. Science 290, 972–977.

Bell, A.S., Aoki, T., Yokoyama, H., 2001. Phylogenetic relationships among microsporidia based on rDNA sequence data, with particular reference to fish-infecting *Microsporidium* Balbiani 1884 species. J. Eukaryot. Microbiol. 48, 258–265.

Brown, J.R., Doolittle, W.F., 1995. Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. Proc. Natl. Acad. Sci. USA 92, 2441–2445.

Bruno, W.J., Socci, N.D., Halpern, A.L., 2000. Weighted neighbor joining: a likelihood-based approach to distance-based phylogeny reconstruction. Mol. Biol. Evol. 17, 189–197.

Bruns, T.D., Vilgalys, R., Barns, S.M., Gonzalez, D., Hibbett, D.S., Lane, D.J., Simon, L., Stickel, S., Szaro, T.M., Weisburg, W.G., Sogin, M.L., 1992. Evolutionary relationships within the fungi: analyses of nuclear small subunit rRNA sequences. Mol. Phylogenet. Evol. 1, 231–241.

Cavalier-Smith, T., 1983. A 6-kingdom classification and a unified phylogeny. In: Schenk, H.E.A., Schwemmler, W.S. (Eds.), Endocytobiology II: Intracellular space as oligogenetic. Walter de Gruyter & Co, Berlin, pp. 1027–1034.

Cavalier-Smith, T., 1993. Kingdom protozoa and its 18 phyla. Microbiol. Rev. 57, 953–994.

Cavalier-Smith, T., 1998. A revised six-kingdom system of life. Biol. Rev. 73, 203–266.

Cheney, S.A., Lafranchi-Tristem, N.J., Bourges, D., Canning, E.U., 2001. Relationships of microsporidian genera, with emphasis on the polysporous genera, revealed by sequences of the largest subunit of RNA polymerase II (RPB1). J. Eukaryot. Microbiol. 48, 111–117.

Edlind, T.D., Li, J., Visvesvara, G.S., Vodkin, M.H., McLaughlin, G.L., Katiyar, S.K., 1996. Phylogenetic analysis of β-tubulin sequences from amitochondrial protozoa. Mol. Phylogenet. Evol. 5, 359–367.

Fast, N.M., Keeling, P.J., 2001. α and β subunits of pyruvate dehydrogenase E1 from the microsporidian *Nosema locustae*: mitochondrion-derived carbon metabolism in microsporidia. Mol. Biochem. Parasitol. 117, 201–209.

Fast, N.M., Logsdon Jr., J.M., Doolittle, W.F., 1999. Phylogenetic analysis of the TATA box binding protein (TBP) gene from *Nosema locustae*: evidence for a microsporidia-fungi relationship and spliceosomal intron loss. Mol. Biol. Evol. 16, 1415–1419.

Felsenstein, J., 1993. PHYLIP (Phylogeny Inference Package). J. Felsenstein, University of Washington, Seattle.

Gilson, P.R., McFadden, G.I., 1996. The miniaturized nuclear genome of eukaryotic endosymbiont contains genes that overlap, genes that are cotranscribed, and the smallest known spliceosomal introns. Proc. Natl. Acad. Sci. USA 93, 7737–7742.

Hirt, R.P., Logsdon Jr., J.M., Healy, B., Dorey, M.W., Doolittle, W.F., Embley, T.M., 1999. Microsporidia are related to fungi: evidence from the largest subunit of RNA polymerase II and other proteins. Proc. Natl. Acad. Sci. USA 96, 580–585.

Ingold, C.T., 1971. Fungal Spores: Their Liberation and Dispersal. Clarendon Press, Oxford.

James, T.Y., Porter, D., Leander, C.A., Vilgalys, R., Longcore, J.E., 2000. Molecular phylogenetics of the Chytridiomycota supports the utility of ultrastructural data in chytrid systematics. Can. J. Bot. 78, 336–350.

Kamaishi, T., Hashimoto, T., Nakamura, Y., Masuda, Y., Nakamura, F., Okamoto, K., Shimizu, M., Hasegawa, M., 1996. Complete nucleotide sequences of the genes encoding translation elongation factors 1α and 2 from a microsporidian parasite, *Glugea plecoglossi*:

implications for the deepest branching of eukaryotes. J. Biochem. (Tokyo) 120, 1095–1103.

Katinka, M.D., Duprat, S., Cornillot, E., Méténier, G., Thomarat, F., Prenier, G., Barbe, V., Peyretaillade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J., Vivarès, C.P., 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature 414, 450–453.

Keeling, P.J., Doolittle, W.F., 1996. α-Tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin family. Mol. Biol. Evol. 13, 1297–1305.

Keeling, P.J., Fast, N.M., 2002. Microsporidia: biology and evolution of highly reduced intracellular parasites. Annu. Rev. Microbiol. 56, 93–116.

Keeling, P.J., Luker, M.A., Palmer, J.D., 2000. Evidence from β-tubulin phylogeny that microsporidia evolved from within the fungi. Mol. Biol. Evol. 17, 23–31.

Keeling, P.J., McFadden, G.I., 1998. Origins of microsporidia. Trends Microbiol. 6, 19–23.

Keohane, E.M., Weiss, L.M., 1999. The structure, function, and composition of the microsporidian polar tube. In: Wittner, M., Weiss, L.M. (Eds.), The Microsporidia and Microsporidiosis. ASM Press, Washington, pp. 196–224.

Kuhner, M.K., Felsenstein, J., 1994. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. Mol. Biol. Evol. 11, 459–468.

Lom, J., Vávra, J., 1962. A proposal to the classification within the subphylum Cnidospora. Syst. Zool. 11, 172–175.

Nilsen, F., Chen, W.-J., 2001. rDNA phylogeny of *Intrapredatorus barri* (Microsporida: Amblyosporidae) parasitic to *Culex fuscanus* Wiedemann (Diptera: Culicidae). Parasitology 122, 617–623.

O'Donnell, K., Cigelnik, E., Benny, G.L., 1998. Phylogenetic relationships among the Harpellales and Kickxellales. Mycologia 90, 624–639.

Paquin, B., Lang, B.F., 1996. The mitochondrial DNA of *Allomyces macrogynus*: the complete genomic sequence from an ancestral fungus. J. Mol. Biol. 255, 688–701.

Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51, 492–508.

Shimodaira, H., Hasegawa, M., 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17, 1246–1247.

Sperling, L., Dessen, P., Zagulski, M., Pearlman, R.E., Migdalski, A., Gromadka, R., Froissard, M., Keller, A.-M., Cohen, J., 2002. Random sequencing of *Paramecium* somatic DNA. Eukaryot. Cell 1, 341–352.

Strimmer, K., von Haeseler, A., 1996. Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. Mol. Biol. Evol. 13, 964–969.

Tanabe, Y., O'Donnell, K., Saikawa, M., Sugiyama, J., 2000. Molecular phylogeny of parasitic zygomycota (Dimargaritales, zoopagales) based on nuclear small subunit ribosomal DNA sequences. Mol. Phylogenet. Evol. 16, 253–262.

Van de Peer, Y., Ben Ali, A., Meyer, A., 2000. Microsporidia: accumulating molecular evidence that a group of amitochondriate and suspectedly primitive eukaryotes are just curious fungi. Gene 246, 1–8.

Vávra, J., Larsson, J.I.R., 1999. Structure of the microsporidia. In: Wittner, M., Weiss, L.M. (Eds.), The Microsporidia and Microsporidiosis. ASM Press, Washington, pp. 7–84.

Vivier, E., 1975. The microsporidia of the protozoa. Protistol. 11, 345–361.

Voigt, K., Wostemeyer, J., 2001. Phylogeny and origin of 82 zygomycetes from all 54 genera of the Mucorales and Mortierellales based on combined analysis of actin and translation elongation factor EF-1α genes. Gene 270, 113–120.

Vossbrinck, C.R., Baker, M.D., Didier, E.S., 1996. Comparative rDNA analysis of microsporidia including AIDS related species. J. Eukaryot. Microbiol. 43, 110S.

Vossbrinck, C.R., Maddox, J.V., Friedman, S., Debrunner-Vossbrinck, B.A., Woese, C.R., 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature 326, 411–414.

Vossbrinck, C.R., Woese, C.R., 1986. Eukaryotic ribosomes that lack a 5.8S RNA. Nature 320, 287–288.

Yang, Z., 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13, 555–556.