

ms# 1130670, Schaap et al.,

SUPPORTING ONLINE MATERIAL

MATERIALS AND METHODS

Cell culture and genomic DNA extraction

All species were cultured on LP agar (*S1*). When necessary, fruiting body formation of the more delicate species was facilitated by placing about 20 pellets of activated charcoal in the petri dish lid. To isolate genomic DNA, cells were harvested and washed three times with 10 mM K-phosphate buffer pH 6.5. Around 3×10^7 cells were lysed in 1 ml HMN (10 mM NaCl, 30 mM Mg-acetate, 10% (w/v) sucrose and 0.5% (v/v) Nonidet P-40 in 30 mM HEPES, pH 7.5), and nuclei were pelleted by centrifugation for 5 minutes at 6500 rpm. Pellets were resuspended in 50 μ l HMN and 150 μ l of resuspension solution from the GenElute Mammalian Genomic DNA Kit (Sigma, St.Louis, USA). Further DNA isolation was performed according to the kit's protocol.

PCR amplification, cloning and DNA sequencing

PCR and sequencing of SSU rDNA. A ~2000 base pair (bp) fragment of SSU rDNA was amplified by PCR (*S2*) using primers A (forward) and B (reverse) as previously described (*S3*) with 50°C annealing. Following amplification, PCR products were separated on 1% agarose gels, and appropriate sized bands excised and extracted from the gel, using the QIAquick gel extraction kit (Qiagen, Crawley, UK). Purified DNA was then cloned into the pGEM-T easy T-tailed vector (Promega, Southampton, UK) and transformed into DH5 α competent cells (Invitrogen, Paisley, UK). A minimum of eight positive colonies were screened by PCR using flanking primer sequences (Sp6 and T7) to confirm the presence of inserts and to screen for possible multiple products (*S4*).

For sequencing, DNA was first amplified by PCR using the T7 and SP6 primers, precipitated with polyethylene glycol and processed further as described (S3). Two clones for each PCR product were sequenced completely on complementary strands using the T7 or SP6 primers and one walking primer (forward primer D542F: 5'-ACAATTGGAGGGCAAGTCTG-3'; reverse primer D1340R: 5'-TCGAGGTCTCGTCCGTTATC-3'). To control for a possible multiple of SSU rDNA genes in *P. violaceum*, 50-150 PCR product clones for each of three independent isolates were screened by restriction fragment polymorphism (RFLP) analysis (S3). Sequencing was performed using ABI PRISM BigDye Terminator v3.0 (Applied Biosystems) on a 3730 DNA Analyzer at the Oxford Sequencing Facility (<http://polaris.bioch.ox.ac.uk/dnaseq/>). Sequences were initially analyzed using Chromas version 2.23 (Technelysium Pty Ltd) and then assembled into contigs using BioEdit version 5.0.9 (S5).

PCR and sequencing of α -tubulin (tubA). Amplification, cloning and sequencing of *tubA* followed the same protocol as above except for using the primers atF3 and atR3 (S3). Some *tubA* PCR reactions also required annealing at 40°C, and others required a second round of amplification using 1 µl of the original reaction as substrate in order to obtain sufficient material for cloning. A minimum of 12 positive clones from each PCR product was screened by reamplification and size fractionation, which always yielded only identical sized clones from any given DNA. Two clones were then sequenced completely on complementary strands. Since *tubA* reactions utilized degenerate PCR primers, some reactions yielded multiple products of the correct or larger size. In all cases, these additional products were also cloned, screened and completely sequenced. However, in no case were multiple *tubA* genes detected.

Multiple sequence alignment

Alignment of SSU rDNA sequences consisted of five steps. 1) Based on a rough preliminary phylogeny, sequences were assigned to one of four major groups. 2) Sequences for each group and for the amoebozoan outgroup were each aligned separately in ClustalX (S6) using default parameters. 3) Each alignment was imported separately into the alignment editor Bioedit (S5) where minor modifications were made by eye to minimize hypothetical insertion/deletion events, and 70% consensus sequences were calculated (also 50% and 60% sequences for amoebas). 4) The resulting consensus sequences were imported into PAUP* (S7) and aligned to each other by eye along with individual sequences not obviously belonging to any group (*D. polycarpum*, *D. polycephalum*, *P. violaceum*, *D. laterosorum*). 5) Finally, the five separate alignments were imported into the same file. These were aligned to each other following their respective consensus sequences in the master consensus alignment, with minor final adjustments made by eye.

Only unambiguously aligned regions of the SSU rDNA alignment were used for phylogenetic analyses. These were defined as regions bordered by a consensus sequence entry for all Groups and with no internal insertions or deletions greater than 1 nucleotide in length (1674 sites). In order to resolve the highly similar Group 4 sequences, separate analyses were conducted including additional sites only alignable within the group (1861 total sites). The position of the Dictyostelia root was tested using only sites alignable among all 4 groups plus the amoebas (1374 sites). Very little length variation is seen in α -tubulin across all eukaryotes (S8). Therefore, *tubA* nucleotide sequences were aligned manually in Bioedit (S5) and then automatically translated into amino acids following intron removal. All analyses were conducted at the amino acid level, and all sites were included.

Phylogenetic analysis

Bayesian inference. Bayesian inference was used to calculate posterior probability of clades (biPP) utilizing the program MrBayes (version 3.1 or 3.1.1) (S9). Final analyses consisted of two sets of four chains each (one cold and three heated) run for 1-10 million generations with trees saved every 10 generations and parameters sampled every 100. All analyses were run at least until a split frequency of ≤ 0.01 between the two run sets was reached. Posterior probabilities were averaged over the final 75% of trees (25% burn in), which was well passed convergence (log probability plateau, usually reached within $\leq 10,000$ generations) for all analyses. Bayesian analyses of SSU rDNA sequences utilized the general time reversible model with a proportion of sites designated invariant, and rate variation among sites modelled after a gamma distribution divided into six categories (GTR+I+G) (S10), with all variable parameters estimated by the program based on BioNJ starting trees. Amino acid sequences of α -tubulin were analyzed using a mixed amino acid model with sites again weighted according to a six-category gamma distribution and relevant parameters estimated by the program as above.

Maximum likelihood. Maximum likelihood was used to calculate bootstrap support values (mlBP) utilizing the program PHYML (S11) on a PC or over the web (<http://atgc.lirmm.fr/phyml/>). A total of 500 (web-based analyses) or 1000-10,000 bootstrap replicates were conducted using the GTR+I+G model for SSU rDNA and the WAG+I+G model (S12) for amino acids. All variable parameters were estimated by the program from BioNJ starting trees.

Phylogenetic controls. All sets of analyses were also conducted with sequential and combined exclusion of relevant disproportionately long-branched sequences (overall tree and root analyses – *D. multistipes*, *A. ellipticum*, *D. polycarpum*, *Planoprotostelium*, *Thecamoeba*; Group 4 analyses - *D. septentrionalis*). From these results it was determined that neither topology nor support values were substantially affected by long-branches (S13). That is, there

were no changes in the significantly resolved portions of the topology (clades with mlBP>70% and/or biPP>0.90), and no substantial change in support values (mlBP +/- 5%, biPP +/- 0.01).

Combined sequence analyses. Combined analyses utilized concatenated alignments of SSU rDNA with either α -tubulin deduced amino acid sequences (tubA-AA) or first and second codon position nucleotides (tubA-NT). Differential weighting of the α -tubulin matrix was also performed in an attempt to compensate for the roughly 2 or 4 fold difference in the number of phylogenetically informative sites between SSU rDNA versus tubA-NT and tubA-AA datasets (564, 273 and 161 parsimony informative sites, respectively). Weighting was accomplished by double, triple or quadruple entry of the tubA (NT or AA) matrix. Bayesian analyses utilized separate models for each data type (partition), with separate GTR+I+G models for SSU rDNA and tubA-NT, and a mixed amino acid model tubA-AA.

Maximum likelihood (PHYML) and paralinear distance (LogDet) analyses were performed at the nucleotide level only (SSU rDNA + tubA-NT), as mixed nucleotide and amino acid model implementations are not currently available for either of these methods. PHYML analyses were performed as described above. LogDet analyses were performed under the minimum evolution model with PAUP 4b10 (S7) and a fraction of 0.34 sites designated as invariant as estimated by the program based on an initial neighborjoining tree calculated under the Kimura 2-parameter model (S14). Analyses consisted of 1000 bootstrap replicates with trees derived by neighborjoining.

Phylogenetic depth. A set of structure-based pre-aligned SSU rDNA sequences from animals, fungi and *D. discoideum* were downloaded from the European Ribosomal DNA Database (ERD, <http://www.psb.ugent.be/rDNA/>), with the new dictyostelid SSU rDNA sequences added following the ERD-aligned *D. discoideum* sequence. Distances are based on a conservative set of 1567 universally aligned positions and were derived separately from trees

constructed by neighborjoining using uncorrected pairwise (“p”) distances and maximum likelihood distances calculated under the GTR+I+G model, in both cases using PAUP4b10 (S7). GTR+I+G model parameters were estimated by the program based on a neighborjoining tree constructed under the Kimura 2-parameter model (S14). Trees were derived separately for animals, fungi and Dictyostelia, to allow for more accurate estimation of model parameters and branch lengths. The animal tree was rooted with hydrozoans and the fungal tree with chytrids. Distances were then taken directly from the Table of Linkages (S7) and calculated as the sum of branch lengths from the common root of the respective tree to the end of its longest terminal branch. The latter correspond to *Caenorhabditis elegans* and *Dugesia japonica* in the animal tree (with and without nematodes, respectively), and *Rhodotorula* and *Rhizomucor* in the fungal tree.

Animal sequences used were from Porifera (*Ephydatia muelleri* - AF121110), Cnidaria (*Hydra littoralis* - U32392), Ctenophora (*Beroe cucumis* - D15068), Placozoa (*Trichoplax adhaerens* - L10828), Nematoda (*Ascaris suum* - AF036587, *Brugia malayi* - AF036588, *Caenorhabditis elegans* - X03680, *Gnathostoma turgidum* - Z96948, *Haemonchus contortus* - L04153, *Trichuris muris* - AF036637, and *Trichinella spiralis* - U60231), Annelida (*Hirudo medicinalis* - AF116011, *Lumbricus terrestris* - AJ272183), Platyhelminthes (*Dugesia japonica* - AF013153), Mollusca (*Aplysia punctata* - AJ224919, *Mytilus edulis* - L24489), Arthropoda (*Artemia salina* - X01723, *Daphnia galeata* - Z23111, *Anopheles clowi* - AF178683, *Drosophila melanogaster* - M21017), Echinodermata (*Strongylocentrotus purpuratus* - L28056), and Chordata (*Styela plicata* - M97577, *Latimeria chalumnae* - L11288, *Salmo trutta* - X98839, *Xenopus laevis* - X04025), *Mus musculus* - X00686, *Rattus norvegicus* - K01593, *Homo sapiens* - K03432). Fungal sequences used were from Chytridiomycota (*Chytridiomycetes sp.* - AF051932, *Chytridium confervae* - M59758, *Neocallimastix frontalis* - X80341, *Smittium culisetae* - AF007540, *Spizellomyces acuminatus*

- M59759), Zygomycota (*Absidia glauca* - AF113409, *Conidiobolus coronatus* - D29947, *Mucor racemosus* - AF113430, *Rhizomucor miehei* - AF192506), Glomeromycota (*Glomus versiforme* - X86687), Basidiomycota (*Agaricus bisporus* - U23724, *Boletus satanas* - M94337, *Coprinus cinereus* M92991, *Filobasidiella neofor* - X60183, *Phanerochaete chrysosporium* - AF026593, *Rhodotorula glutinis* - AB016292, *Schizophyllum commune* - X54865, *Ustilago maydis* - X62396), and Ascomycota (*Arxula adenivorans* - AB018123, *Emericella nidulans* - AB008403, *Neurospora crassa* - X04971, *Penicillium allii* - AF218787, *Pichia mexicana* - AB013570, *Pneumocystis carinii* - L27658, *Candida albicans* - AB013586, *Kluyveromyces aestuarii* - X89520, *Eremothecium gossypii* - AF113137, *Taphrina californica* - D14166, *Trichoderma viride* - AF218788, *Schizosaccharomyces pombe* - X58056, *Saccharomyces cerevisiae* - Z75578).

LEGENDS TO FIGURES

Figure S1. Statistical support for the SSU rDNA phylogeny of Dictyostelia. The statistical support values for the SSU rDNA phylogeny depicted in Figure 1 of the main text are shown above and below the lines for Bayesian inference posterior probabilities (biPP) and maximum likelihood bootstrap (mlBP), respectively. Values of biPP<0.50 and mlBP<50% are indicated by “-“. Values are shown separately for separate analyses to determine A) the overall structure of the tree (72 sequences, 1674 sites), B) the detailed phylogeny of Group 4 (35 sequences, 1861 sites), and C) the position of the root (49 sequences, 1374 sites). For the latter tree, support values are only shown for deep nodes (heavy lines), as there is inadequate information from this highly conserved subset of sites to clearly differentiate many of the more terminal nodes (fine lines). The branch lengths in this tree (C) are also drawn to scale, as indicated by the scale bar, to indicate their actual lengths as opposed to those in figure 1, which have been scaled to compensate for the additional 300 sites used to construct the

ingroup-only tree (A). Topological rearrangements in the maximum likelihood (ML) tree relative to the Bayesian inference (BI) results are indicated by broken lines, with an “x” below the BI branches not found in the ML tree.

Figure S2. Phylogeny of the Dictyostelia based on α -tubulin amino acid sequences. The tree shown was derived from analyses of 322 universally aligned α -tubulin amino acid positions using Bayesian inference and including a broad representation of all four major groups identified by SSU rDNA phylogeny (Fig. 1). Major groups are indicated in colors and with brackets as in Figure 1. Branch lengths are drawn to scale as indicated by the scale bar (0.1 substitutions per site). Values for biPP over 0.89 and mlBP over 60% are shown above and below the lines, respectively.

Figure S3. Phylogeny of Dictyostelia based on combined SSU rDNA and α -tubulin sequences. The tree shown was derived by Bayesian Inference using combined SSU rDNA and double-weighted α -tubulin first and second codon position nucleotide (NTx2) sequences to give nearly equal numbers of informative sites for both genes. Additional analyses were performed with single weighted tubA nucleotide (NTx1), unweighted (AAx1) and four-fold weighted (AAx4) deduced amino acid sequences. Heavy branches indicate nodes receiving 0.98-1.00 biPP for all data types and weighting schemes. Otherwise, values are shown for major deep branches only and are displayed separated by parenthesis in the order NTx2/NTx1/AAx4/AAx1 with a “-” to indicate biPP values less than 0.70. The single strongly supported alternative branching pattern is indicated with dotted lines.

Figure S4. Relative molecular depth of the dictyostelid SSU rDNA tree. The chart shows the maximum diameter of the SSU rDNA phylogeny of Dictyostelia versus a broad sampling

of animals and fungi. Depth is measured as total distance from the taxon root to the tip of its longest terminal branch. Solid boxes indicate values with highly divergent sequences removed (nematodes in the case of animals, *Rhizomucor* in the case of fungi, *D. multi-stipes* for Dictyostelia). Raw pairwise distances (PD) and maximum likelihood (ML) GTR+I+G distances were calculated using PAUP 4b10 based on taxon-specific neighborjoining trees constructed under the same model.

Figure S5. Mapping of species characters. A total of 20 characters were mapped onto the SSU rDNA phylogeny of all Dictyostelid species. The first set (columns 1-7) represents taxonomic traits. Cell size shows no strong group-specific trend (column 1), while spores are consistently smaller in Group 1. The acellular stalk is found only in a subset of Group 2 species, the acytostelids (column 4). The tips of acellular stalks are thinly pointed (piliform) (column 5). Pointy (acuminate) or blunt (obtuse) tips also mark the cellular stalks of other Group 2 species and of some Group 3 species. Group 1 and Group 4 species usually have extended (capitate or clavate) stalk tips. The color of the spores and stalk is represented as an approximation of the observed color. Species with similarly colored stalks and spore heads are often related, but no color is specific for any of the four major groups (columns 6 and 7).

The second set of characters (columns 8-12) represents alternative modes of behaviour: Many species in Groups 1-3 form microcysts, an alternative survival strategy that is also used by solitary amoebas. However, microcysts have not been observed for Group 4 species (column 8). Species that can form sexual macrocysts are present in all four groups (column 9). The chemoattractant (acrasin) that is used for aggregation is known for only a few species (column 10). It is cAMP for all investigated Group 4 species, while in the other groups at least three other compounds are used. To aggregate, cells can either come together as individuals, creating mound-shaped aggregates, or line up to form streams. Stream formation

is the most common mode of aggregation, with only a few species in Groups 1 and 3 aggregating as individuals (column 11). Species in all groups form migrating slugs, but stalkless or free migration is only shown by a small cluster of Group 4 species and a single non-Group 4 species, *D. polycephalum* (column 12). Phototropism, the tendency of fruiting structures to veer towards the light (*I*), is also be common to all four groups, but not to all species within the groups (column 13). The third set contains characters relate to fruiting body size and shape. These are discussed in the main text (Fig. 2).

LEGENDS TO TABLES

Table S1. Consistently documented characters were retrieved from the original species diagnoses and from secondary publications, such as the *Dictyostelium* monographs by Raper and Hagiwara (*S1, S15*). The varieties (states) described for qualitative characters, such as the shapes of spores or the branching patterns of fruiting bodies were numerically coded. The alphabetical code key for each character and the numerical code key for each character state are listed in table S2.

For quantitative characters (usually size) a frequency distribution of data for all species was prepared first, and the data range was subdivided into 3-4 intervals in such a manner that each interval contained about the same number of species. Each interval then represents a character state, which was again numerically coded as shown in table S2. Quantitative characters have usually been reported as a range of observed values. For trait mapping this range had to be reduced to a single parameter. For characters such as the size of aggregates, stalks and sori, the upper and intermediate value of the range were first analysed separately. However, this revealed virtually identical trends in character evolution (data not shown). We chose to show the analysis of the upper values here, because this is a better indication of the maximal number of cells that can be incorporated into a single structure (all species will form

small structures when few cells are available, but only few will form large structures at high cell densities). For spore and cell sizes, intermediate values were used or average values, when reported. Spore sizes were usually reported as length (l) and width (w) of oblong spores and diameter (d) of globose spores. These values were further reduced to a single parameter by calculating the volume of oblong spores as a cylinder ($lB(\frac{1}{2}w)^2$) and that of globose spores as a sphere ($\frac{4}{3}B(\frac{1}{2}d)^3$).

Table S2.

This table lists the letter code for character names and the numerical code for character states as are used in table S1. The color code for character states are as used in figure 4 (main text), and figure S4 (supporting material) is indicated in the last column.

REFERENCES

- S1. K. B. Raper, *The Dictyostelids* (Princeton University Press, Princeton, New Jersey, 1984)
- S2. L. Medlin, H. J. Elwood, S. Stickel, M. L. Sogin, *Gene* **71**, 491 (1988).
- S3. E. S. Steenkamp, J. Wright, S. L. Baldauf, *Mol. Biol. Evol.* **in press** (2005).
- S4. L. Nitschke, M. Kopf, M. C. Lamers, *Biotechniques* **14**, 914 (1993).
- S5. T. A. Hall, *Nucl. Acids. Symp. Ser.* **41**, 95 (1999).
- S6. J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, D. G. Higgins, *Nucl. Acids Res.* **25**, 4876 (1997).
- S7. D. L. Swofford, *PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4* (Sinauer Associates, Sunderland, Massachusetts, 1998).
- S8. S. L. Baldauf, A. J. Roger, I. Wenk-Siefert, W. F. Doolittle, *Science* **290**, 972 (2000).
- S9. F. Ronquist, J. P. Huelsenbeck, *Bioinformatics* **19**, 1572 (2003).

- S10. J. Felsenstein, G. A. Churchill, *Mol. Biol. Evol.* **13**, 93 (1996).
- S11. S. Guindon, O. Gascuel, *Syst. Biol.* **52**, 696 (2003).
- S12. S. Whelan, N. Goldman, *Mol. Biol. Evol.* **18**, 691 (2001).
- S13. S. Gribaldo, H. Philippe, *Theor. Popul. Biol.* **61**, 391 (2002).
- S14. M. Kimura, T. Ohta, *J. Mol. Evol.* **2**, 87 (1972).
- S15. H. Hagiwara, in *The taxonomic study of Japanese Dictyostelid cellular slime molds.* (Nat. Science Museum, Tokyo, 1989).
- S16. H. Hagiwara, *Bull. Natn. Sc. Mus.* **10**, 27 (1984).
- S17. O. Brefeld, *Abhandlungen der Senckenbergischen Naturforschenden Gesellschaft* **7**, 85 (1869).
- S18. H. Hagiwara, *Bull. Natn. Sc. Mus.* **9**, 45 (1983).
- S19. H. Hagiwara, *Bull. Natn. Sci. Mus.* **22**, 47 (1996).
- S20. H. Hagiwara, *Bull. Natn. Sci. Mus.* **17**, 103 (1991).
- S21. E. W. Olive, *Proc. Amer. Acad. Arts Sci.* **37**, 333 (1901).
- S22. J. C. Cavender, A. C. Worley, K. B. Raper, *Amer. J. Bot.* **68**, 373 (1981).
- S23. J. C. Cavender, *Can. J. Bot.* **57**, 1326 (1978).
- S24. H. Hagiwara, *Bull. Natn. Sci. Mus. Tokyo* **10**, 63 (1984).
- S25. H. Hagiwara, *Bull. Natn. Sci. Mus.* **18**, 1 (1992).
- S26. C. A. J. A. Oudemans, *Aanwinsten voor de Flora Mycologia van Nederland* **9-10**, 39 (1885).
- S27. H. Hagiwara, *Bull. Natn. Sci. Mus.* **14**, 351 (1971).
- S28. K. B. Raper, J. C. Cavender, *J. Elisha Mitchell. Sci. Soc.* **84**, 31 (1968).
- S29. H. Hagiwara, *Bull. Natn. Sci. Mus.* **9**, 149 (1983).
- S30. E. G. Olive, *Proc. Boston Soc. Nat. His.* **30**, 451 (1902).
- S31. H. Hagiwara, Z.-Y. Yeh, C.-Y. Chien, *Bull. Natn. Sci. Mus. Tokyo* **11**, 103 (1985).

- S32. K. B. Raper, *J. Agric. Res.* **50**, 135 (1935).
- S33. E. M. Vadell, M. T. Holmes, J. C. Cavender, *Mycologia* **87**, 551 (1995).
- S34. J. C. Cavender, *J. Gen. Microbiol.* **62**, 113 (1970).
- S35. J. C. Cavender, *Amer. J. Bot.* **63**, 60 (1976).
- S36. K. Kawabe, *Trans. Mycol. Soc. Japan* **23**, 91 (1982).
- S37. B. N. Singh, *J. Gen. Microbiol.* **1**, 11 (1947).
- S38. H. Hagiwara, *Bull. Natn. Sci. Mus.* **18**, 101 (1992).
- S39. H. Hagiwara, *Bull. Natn. Sci. Mus.* **24**, 81 (1998).
- S40. O. Brefeld, *Unters. Gesammtgeb. Mykol.* **6**, 1 (1884).
- S41. J. C. Cavender, S. L. Stephenson, J. C. Landolt, E. M. Vadell, *New Zealand J. Bot.* **40**, 235 (2002).
- S42. H. Hagiwara, *Bull. Natn. Sci. Mus.* **16**, 493 (1973).
- S43. S. Kawakami, H. Hagiwara, *Mycoscience* **40**, 357 (1999).
- S44. J. C. Cavender, E. Vadell, J. C. Landolt, S. L. Stephenson, *Mycologia* **97**, 493 (2005).
- S45. K. B. Raper, *Mycologia* **33**, 633 (1941).
- S46. J. C. Cavender, K. B. Raper, A. M. Norberg, *Amer. J. Bot.* **66**, 207 (1979).
- S47. K. B. Raper, D. I. Fennell, *Amer. J. Bot.* **54**, 515 (1967).
- S48. J. T. Bonner, *Amer. Nat.* **119**, 530 (1982).
- S49. M. P. Van Tieghem, *Bull. de la Soc. Botan. de France* **27**, 317 (1880).
- S50. J. C. Cavender, (unpublished results).
- S51. D. R. Waddell, *Nature* **298**, 464 (1982).
- S52. K. B. Raper, *J. Gen. Microbiol.* **14**, 716 (1956).
- S53. F. Traub, H. R. Hohl, J. C. Cavender, *Amer. J. Bot.* **68**, 162 (1981).
- S54. G. Kauffman, J. Cavender, H. R. Hohl, *Botanica Helvetica* **98**, 123 (1988).
- S55. J. C. Cavender, (unpublished results).

- S56. E. M. Vadell, J. C. Cavender, (in prep.).
- S57. E. M. Vadell, J. C. Cavender, *Mycologia* **90**, 715 (1998).
- S58. H. Hagiwara, *Bull. Natn. Sci. Mus.* **5**, 67 (1979).
- S59. H. Hagiwara, *Bull. Natn. Sci. Mus.* **29**, 127 (2003).
- S60. K. B. Raper, M. S. Quinlan, *J. Gen. Microbiol.* **18**, 16 (1958).
- S61. J. C. Cavender, E. M. Vadell, *Mycologia* **92**, 992 (2000).
- S62. H. Hagiwara, *Bull. Natn. Sci. Mus.* **4**, 27 (1978).
- S63. H. Hagiwara, *Bull. Natn. Sci. Mus.* **12**, 99 (1986).
- S64. J. S. Anderson, D. I. Fennell, K. B. Raper, *Mycologia* **60**, 49 (1968).

Table S1. The data matrix for inference of character evolution

Species	Character	Character states																	Citation	
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q		R
<i>D.brefeldianum</i>		1	0	1	0	?	2	1	1	4	0	3	2	3	2	?	0	0	2	(S15,S16)
<i>D.mucoroides</i>		1	0	1	0	0	2	2	1	4	0	2	3	2	2	2	0	2	2	(S1, S15,S 17)
<i>D.capitatum</i>		0	0	1	1	?	2	1	0	0	0	2	1	2	0	?	0	0	0	(S15,S18)
<i>D.pseudo-brefeldia.</i>		1	0	1	1	?	2	2	2	4	0	3	2	3	3	?	0	1	2	(S19)
<i>D.aureocephalum</i>		1	0	1	1	?	2	2	1	4	0	2	3	2	3	?	0	1	2	(S20)
<i>D.aureum</i>		0	0	1	1	0	2	2	3	4	0	3	2	2	0	?	0	?	2	(S1, S21, S22)
<i>D.septentrionalis</i>		2	0	1	1	?	2	1	3	4	0	3	3	3	3	2	0	0	1	(S1, S15, S23)
<i>D.implicitum</i>		2	0	1	1	?	2	2	2	4	0	3	2	3	3	?	0	3	2	(S15, S24)
<i>D.medium</i>		1	0	1	1	?	2	1	1	4	0	2	2	2	3	?	0	?	2	(S25)
<i>D.crassicaule</i>		1	0	1	1	?	2	1	2	4	0	1	3	3	3	?	0	0	0	(S15, S24)
<i>D.sphaerocephalum</i>		1	0	1	0	?	2	3	1	4	1	2	3	3	0	2	0	?	1	(S1, S26, S27)
<i>D.rosarium</i>		2	5	1	0	0	2	1	3	2	4	3	1	2	0	2	0	?	0	(S1, S28)
<i>D.clavatum</i>		0	0	1	1	?	2	2	1	4	0	2	3	2	2	?	0	1	2	(S25)
<i>D.longosporum</i>		1	0	1	1	?	2	1	3	4	0	3	2	3	0	?	0	0	2	(S29)
<i>D.purpureum</i>		0	0	1	0	0	2	2	3	4	0	3	3	3	2	2	0	0	2	(S1, S15, S30)
<i>D.macrocephalum</i>		2	0	1	1	?	2	1	2	4	0	1	2	3	3	?	0	2	2	(S15, S31)
<i>D.discoideum</i>		1	0	1	0	0	2	4	3	4	0	2	3	3	3	2	0	1	2	(S1, S15, S32)
<i>D.citrinum</i>		1	0	1	1	0	2	4	?	4	0	2	1	3	3	?	0	0	?	(S33)
<i>D.dimigraformum</i>		2	0	1	1	?	2	3	?	4	0	3	3	3	0	1	0	?	2	(S1, S34)
<i>D.intermedium</i>		1	0	1	1	?	2	3	1	4	0	2	3	2	0	2	0	?	2	(S1, S35)
<i>D.firmibasis</i>		1	0	1	1	?	2	1	2	4	0	3	3	3	3	?	0	0	2	(S1, S15, S27)
<i>D.brunneum</i>		1	0	1	1	?	2	2	1	4	0	3	2	3	2	?	0	0	2	(S1, S36)
<i>D.giganteum</i>		1	0	1	0	0	2	2	3	4	0	3	3	3	3	2	0	0	2	(S1, S37-39)
<i>D.robustum</i>		2	0	1	1	?	2	3	3	4	0	3	3	3	4	?	0	0	2	(S19)
<i>D.laterosorum</i>		2	2	1	1	?	2	1	?	0	4	3	2	2	1	1	0	?	1	(S1, S34)
<i>P.violaceum</i>		0	2	1	0	3	2	2	3	4	6	3	2	2	0	2	0	1	2	(S1, S15, S40)
<i>D.australe</i>		1	2	1	1	?	1	0	?	4	1	0	2	2	0	2	0	?	?	(S41)
<i>D.monochasioides</i>		1	2	1	0	?	1	2	1	0	2	2	1	2	0	1	0	1	1	(S1, S15, S42, S43)
<i>D.potamoides</i>		0	2	0	1	?	1	0	0	2	1	1	0	0	0	0	?	?	?	(S44)
<i>D.minutum</i>		0	0	0	0	1	0	0	?	2	1	1	1	1	0	1	0	3	0	(S1, S15, S45)
<i>D.tenue</i>		1	3	0	1	?	0	0	1	1	5	2	1	1	0	2	0	?	2	(S1, S46)
<i>D.gracile</i>		1	2	1	1	?	1	2	1	0	0	2	1	2	0	?	0	2	0	(S29)
<i>D.lavandulum</i>		1	2	0	1	?	2	2	3	2	0	2	1	2	1	2	0	?	2	(S1, S15, S47)
<i>D.vinaceo-fuscum</i>		1	0	0	1	4	2	2	?	0	0	3	0	2	1	2	0	?	2	(S1, S47, S48)
<i>D.rhizopodium</i>		1	2	0	1	?	2	2	1	2	0	2	2	2	1	2	0	0	2	(S1, S15, S47)
<i>D.coeruleo-stipes</i>		0	2	1	1	?	2	2	?	4	0	2	2	2	1	?	0	?	2	(S1, S47)
<i>D.lacteam</i>		0	4	0	1	2	0	0	?	2	0	1	1	1	0	2	0	3	?	(S1, S49)
<i>D.menorah</i>		1	2	1	1	?	0	1	?	4	1	0	2	0	0	0	0	?	?	(S50)
<i>D.caveatum</i>		0	2	1	1	3	0	?	0	2	0	1	0	1	0	2	0	?	1	(S1, S51)
<i>D.polycephalum</i>		1	1	0	1	?	2	4	3	3	0	0	1	1	0	0	0	3	0	(S1, S15, S52)
<i>D.polycarpum</i>		2	1	1	1	?	2	0	2	3	0	2	1	2	0	2	0	3	0	(S1, S53)
<i>P.filamentosum</i>		1	1	1	1	?	2	?	2	2	6	3	1	1	0	2	0	?	0	(S1, S53)
<i>P.luridum</i>		2	1	0	1	4	2	?	2	4	6	3	3	2	0	2	0	2	0	(S54)
<i>P.equisetoides</i>		2	4	0	1	?	2	2	1	2	6	3	2	0	0	2	0	2	2	(S55)
<i>P.arachnoideum</i>		1	1	0	1	?	2	1	0	4	6	2	2	1	0	0	0	3	1	(S56)

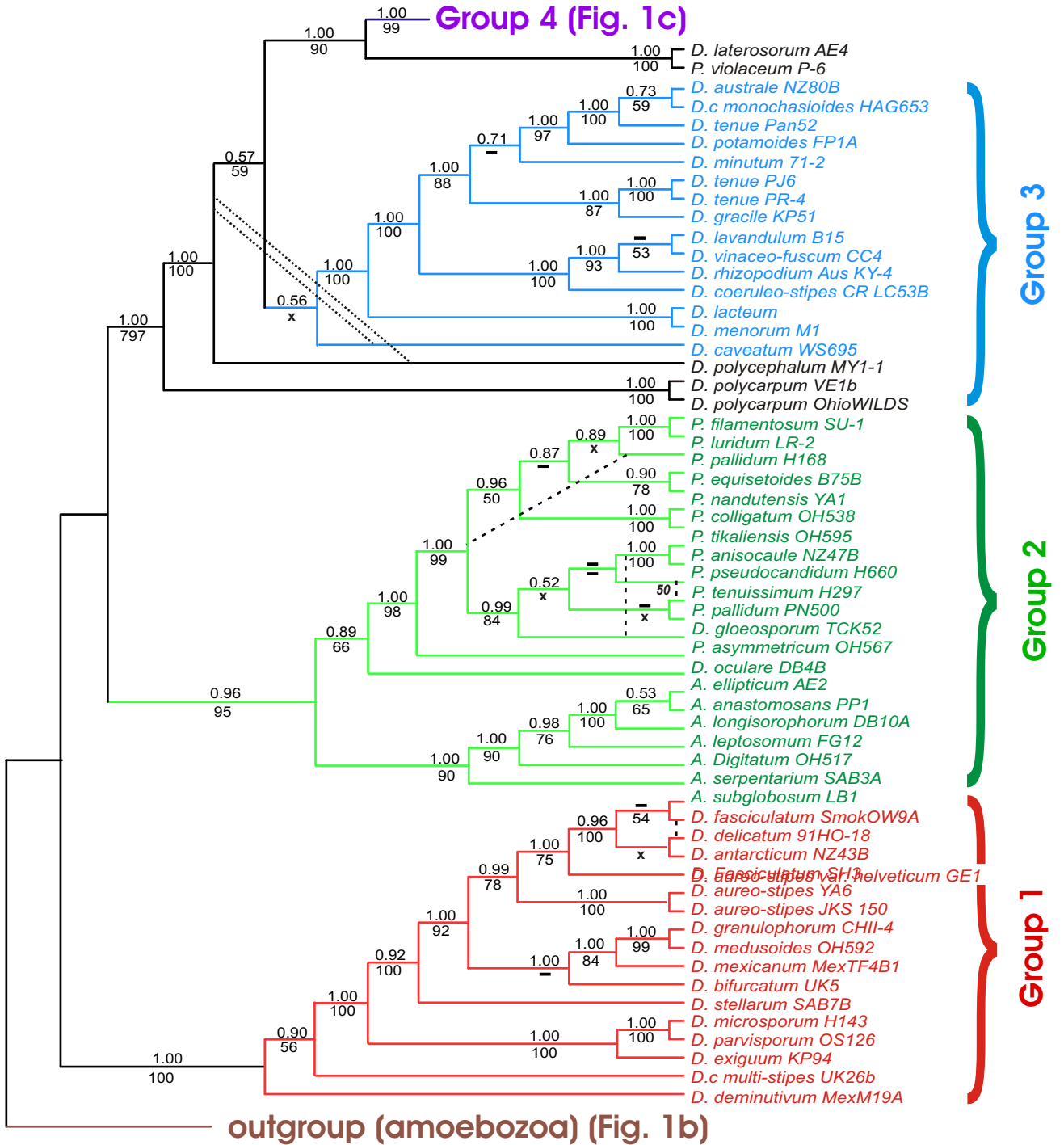
<i>P.colligatum</i>	1	1	1	1	?	2	?	?	3	6	3	3	1	0	?	0	?	2	(S57)
<i>P.tikaliensis</i>	0	1	1	1	?	2	2	3	2	6	3	1	2	0	?	0	?	2	(S57)
<i>P.anisocaulis</i>	1	1	0	1	?	2	?	0	4	6	3	1	2	0	?	0	3	2	(S41)
<i>P.pseudo-candidum</i>	1	1	0	0	?	2	2	?	4	6	3	1	1	0	?	0	3	0	(S1, S15, S43, S58)
<i>P.tenuissimum</i>	0	1	0	0	?	2	1	3	4	6	3	2	1	0	?	0	3	?	(S1, S15, S58)
<i>D.gloeosporum</i>	1	0	1	0	?	1	1	?	4	0	1	1	2	0	?	0	1	2	(S59)
<i>P.pallidum</i>	0	1	0	0	3	2	2	1	2	6	3	1	2	0	2	0	3	1	(S1, S15, S21)
<i>P.asymmetricum</i>	2	1	0	1	?	2	2	?	2	6	2	1	1	0	?	0	?	2	(S57)
<i>D.oculare</i>	0	2	0	1	?	1	0	0	4	0	0	0	0	0	?	0	?	?	(S44)
<i>A.ellipticum</i>	0	1	1	1	?	1	0	0	1	0	1	0	0	0	1	1	4	0	(S1, S34)
<i>A.anastomosans</i>	1	4	1	1	?	1	0	0	0	0	0	0	0	0	1	?	?	?	(S44)
<i>A.longisorophorum</i>	1	4	0	1	?	2	1	0	2	0	1	0	0	0	1	1	?	?	(S44)
<i>A.leptosomum</i>	2	4	0	1	4	2	0	0	0	0	1	0	0	0	1	1	4	0	(S1, S48, S60)
<i>A.digitatum</i>	2	4	0	1	?	2	?	0	0	0	1	0	1	0	?	1	4	?	(S61)
<i>A.serpentarium</i>	0	4	0	0	?	2	0	0	2	0	0	0	0	0	0	1	?	?	(S44)
<i>A.subglobosum</i>	2	4	1	1	?	1	0	0	0	0	0	0	0	0	1	1	?	?	(S1, S35, S61)
<i>D.antarcticum</i>	1	2	1	1	?	2	0	0	0	1	1	3	2	0	?	0	?	2	(S41)
<i>D.fasciculatum</i>	0	2	1	1	?	2	2	?	2	0	3	1	3	0	2	0	?	2	(S1, S53)
<i>D.delicatatum</i>	0	2	1	1	?	2	1	3	0	0	3	2	2	0	2	0	1	1	(S1, S15, S27)
<i>D.aureo-stipes</i>	0	2	0	1	4	2	1	3	4	5	3	1	2	0	1	0	1	2	(S1, S15, S46)
<i>D.granulophorum</i>	0	2	0	1	4	2	0	1	2	0	2	2	3	0	?	0	0	?	(S33)
<i>D.medusoides</i>	0	2	0	1	4	2	2	0	0	0	2	1	3	0	?	0	0	2	(S33)
<i>D.mexicanum</i>	1	2	0	0	4	2	0	1	2	0	2	3	3	3	2	0	0	?	(S1, S22)
<i>D.bifurcatum</i>	0	2	1	1	?	2	0	1	2	3	1	2	2	0	1	0	0	2	(S1, S35)
<i>D.stellatum</i>	0	2	1	1	?	2	1	0	4	2	1	0	0	0	1	0	2	?	(S44)
<i>D.microsporium</i>	0	2	1	1	?	2	1	1	0	0	1	1	1	0	?	0	2	0	(S1, S15, S62)
<i>D.parvisporum</i>	0	2	1	1	?	2	1	1	1	0	3	1	2	0	?	0	1	2	(S15, S63)
<i>D.exiguum</i>	0	2	1	1	?	2	2	1	0	1	2	2	2	0	?	0	1	2	(S29)
<i>D.multi-stipes</i>	0	3	1	1	?	0	0	0	2	0	1	0	0	0	?	0	?	0	(S1, S35)
<i>D.deminutivum</i>	0	2	0	1	?	0	0	?	2	0	0	0	0	0	1	0	?	0	(S1, S64)

Table S2. Code key for character names and states

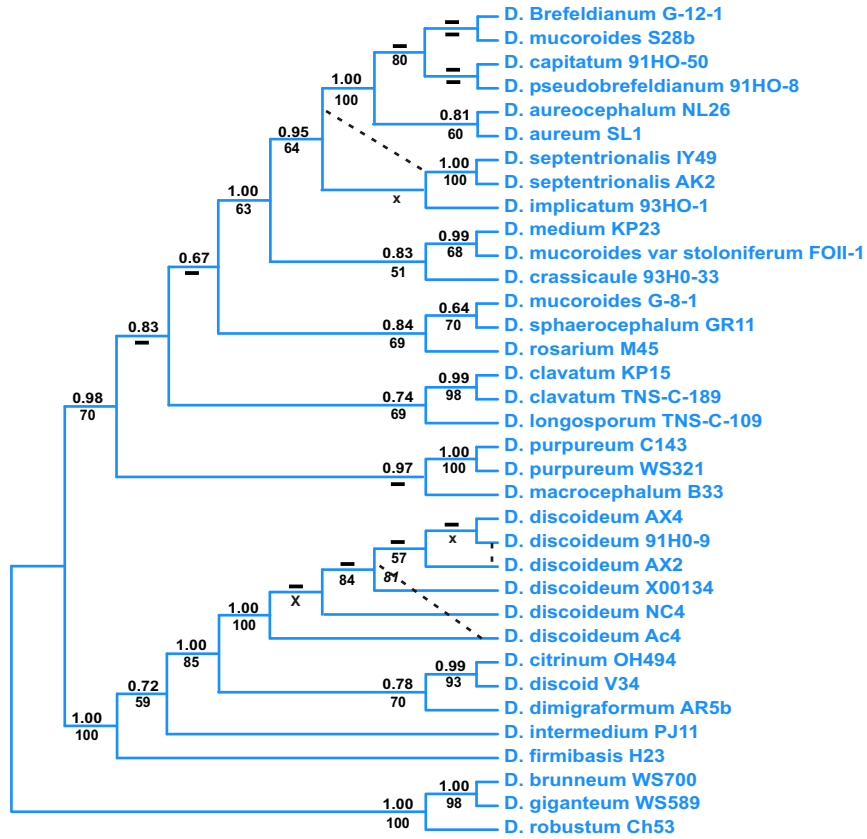
Character name	Code	Character state	Code	Color code
Spore volume	A	0-50 μm^3	0	white
		50-80 μm^3	1	gray
		>80 μm^3	2	black
		not measured or unknown	?	absent
Spore shape and granule position	B	oblong, no polar granules	0	blue
		oblong, unconsolidated polar granules	1	yellow
		oblong, consolidated polar granules	2	red
		oblong, granule position equivocal	3	orange
		globose or subglobose	4	grey
Microcyst	C	observed	0	red
		not observed	1	green
Macrocyt	D	observed	0	red
		not observed	1	green
Acrasin	E	cAMP	0	blue
		glorin	1	green
		folate	2	red
		pterin	3	yellow
		not cAMP	4	orange
Mode of aggregation	F	mounds	0	red
		minor streaming	1	yellow
		streaming	2	green
Slug migration	G	none	0	red
		briefly	1	yellow
		with stalk being formed	2	green
		with and without stalk	3	green/blue
		without stalk formation	4	blue
Aggregate size	H	0-2 mm	0	white
		2-4 mm	1	light gray
		4-7 mm	2	dark grey
		>7 mm	3	black
Sorocarp habit	I	gregarious (loosely grouped)	0	yellow
		gregarious and clustered	1	yellow/red
		clustered (closely grouped)	2	red
		coremiform (bunched)	3	purple
		solitary	4	black
Branching pattern	J	unbranched	0	black
		lateral branches	1	orange
		repeated lateral	2	red
		bifurcation	3	yellow
		abstriction from posterior - sessile	4	green
		abstriction - irregular whorls	5	light blue
		abstriction - regular whorls	6	dark blue
Stalk height	K	0-1 mm	0	white
		1-3 mm	1	light gray
		3-7 mm	2	dark gray

		>7 mm	3	black
Stalk diameter	L	0-10 μm	0	white
		10-20 μm	1	light gray
		20-30 μm	2	dark gray
		>30 μm	3	black
Sori diameter	M	0-75 μm	0	white
		75-150 μm	1	light gray
		150-250 μm	2	dark gray
		> 250 μm	3	black
Cellular support	N	none	0	red
		crampon	1	yellow
		supporter	2	green
		disk	3	blue
Cell diameter	O	0-9 μm	0	white
		9-12 μm	1	gray
		>12 μm	2	black
Stalk shape	P	cellular	0	green
		acellular	1	red
Stalk tip shape	Q	capitate (head-shaped)	0	red
		clavate (club-shaped)	1	orange
		obtuse (blunt)	2	green
		acuminate (pointed)	3	light blue
		piliform (thread-shaped)	4	dark blue
Phototropism	R	none	0	red
		weak	1	yellow
		strong	2	green

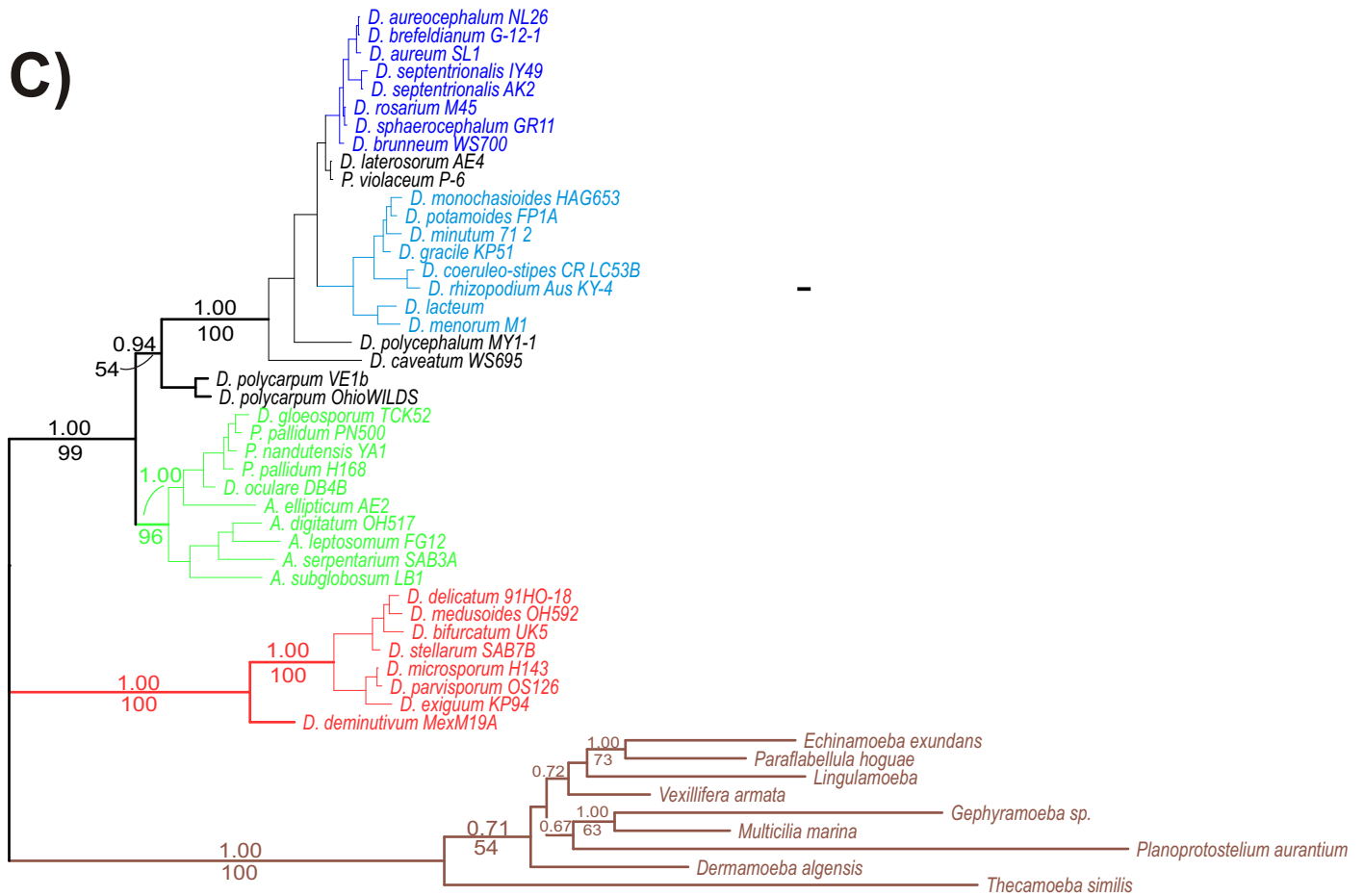
A)

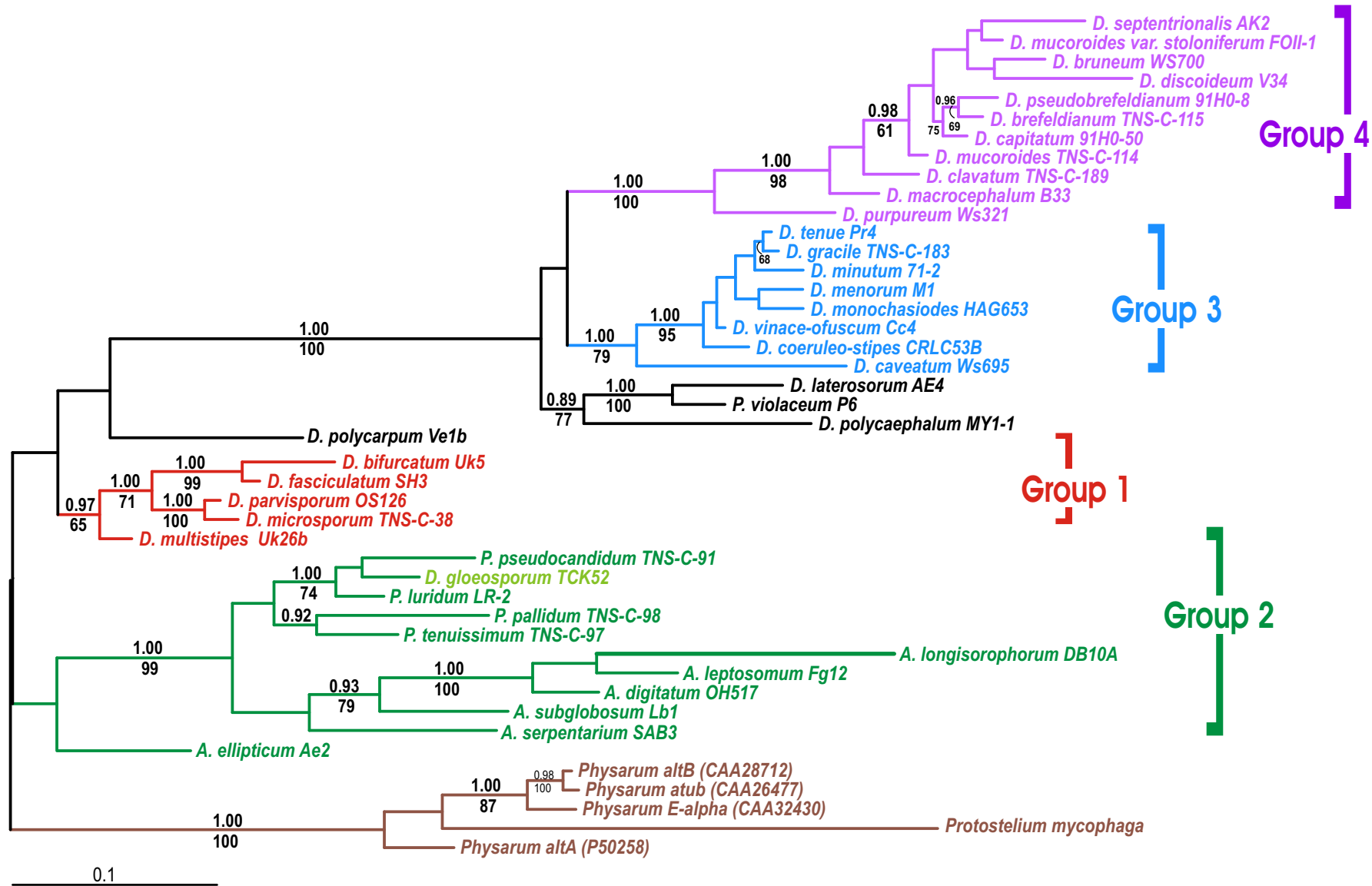



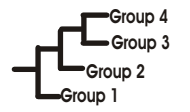
B)



C)





	 root1	 root2
- Ntx1 bayes	0.85	
logdet	---	100%
ML	---	68%
- Ntx2 bayes	---	0.98
ML	---	69%
- Aax1	0.97	---
- Aax2	0.92	---
- Aax3	---	---
- Aax4	0.70	---

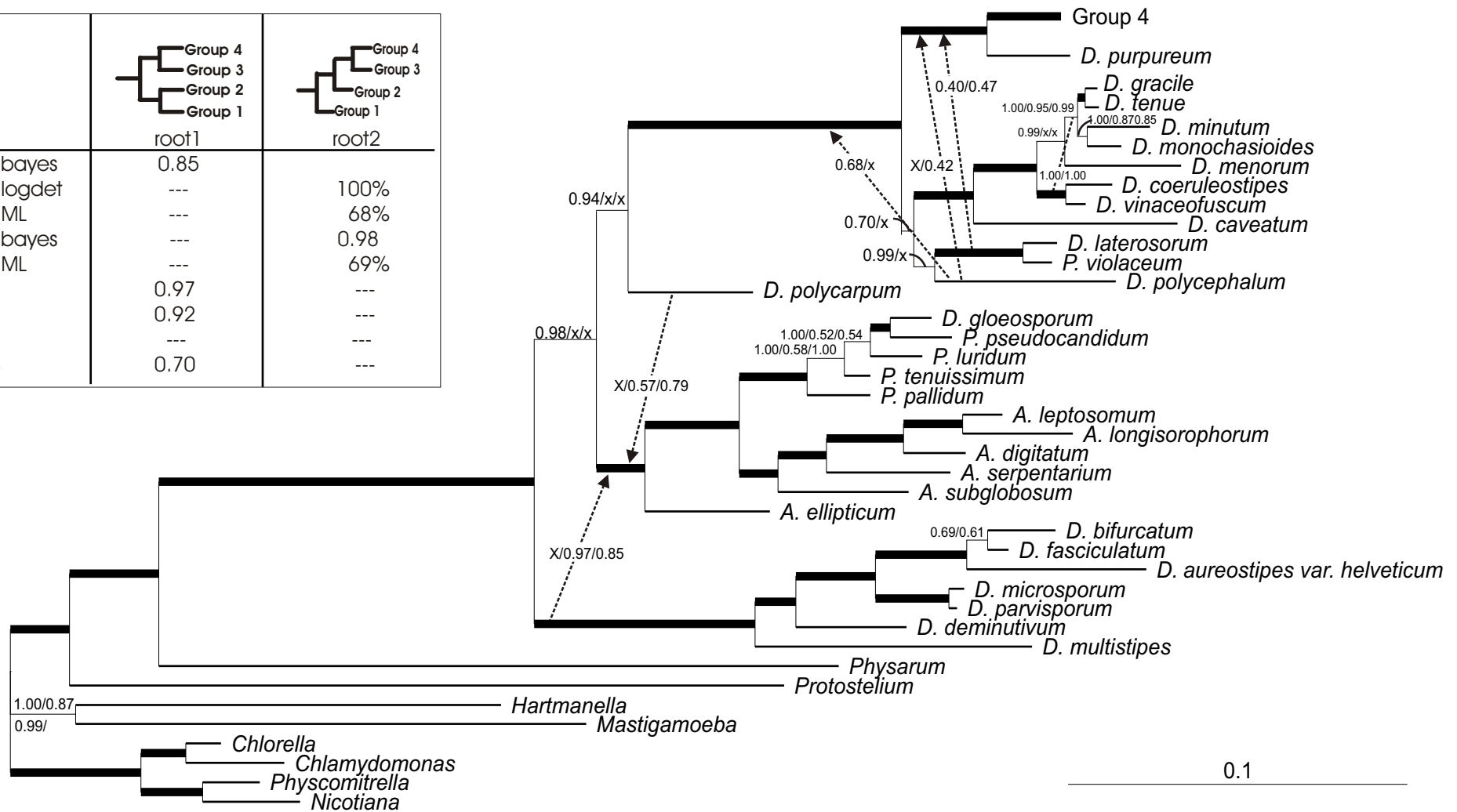


Figure S3. Phylogeny of Dictyostelia based on combined SSU rRNA and tubA sequences. The tree shown was derived by Bayesian Inference using combined SSU rRNA and tubA first and second codon nucleotide sequences, with the latter given twice the weight of the former to give nearly equal numbers of informative sites for both partitions. Model parameters for the GRT+I+G model were determined separately for the two partitions. Additional analyses were performed with single weight tubA sequences (Ntx1), and tubA deduced amino acid sequences unweighted or double weighted. Values for all analyses for all nodes are displayed separated by parenthesis in the order NTx2/NTx1/Aax1/Aax2. Branches supported by 0.98-1.00 biPP for all weighting schemes are indicated by heavy lines. Support values for the alternative two rootings are also given in the table to the upper right, along with values obtained for treble (Aax3) and quadruple (Aax4) weighted amino acid sequences and LogDet analyses of nucleotide sequences, which should compensate for any effects due to nucleotide compositional bias.



