LETTER

doi:10.1038/nature09984

Discovery of novel intermediate forms redefines the fungal tree of life

Meredith D. M. Jones[1,2], Irene Forn[3], Catarina Gadelha[4], Martin J. Egan[1,5], David Bass[2], Ramon Massana[3] & Thomas A. Richards[1,2]

Fungi are the principal degraders of biomass in terrestrial ecosystems and establish important interactions with plants and animals[1–3]. However, our current understanding of fungal evolutionary diversity is incomplete[4] and is based upon species amenable to growth in culture[1]. These culturable fungi are typically yeast or filamentous forms, bound by a rigid cell wall rich in chitin. Evolution of this body plan was thought critical for the success of the Fungi, enabling them to adapt to heterogeneous habitats and live by osmotrophy: extracellular digestion followed by nutrient uptake[5]. Here we investigate the ecology and cell biology of a previously undescribed and highly diverse form of eukaryotic life that branches with the Fungi, using environmental DNA analyses combined with fluorescent detection via DNA probes. This clade is present in numerous ecosystems including soil, freshwater and aquatic sediments. Phylogenetic analyses using multiple ribosomal RNA genes place this clade with *Rozella*, the putative primary branch of the fungal kingdom[1]. Tyramide signal amplification coupled with group-specific fluorescence *in situ* hybridization reveals that the target cells are small eukaryotes of 3–5 μm in length, capable of forming a microtubule-based flagellum. Co-staining with cell wall markers demonstrates that representatives from the clade do not produce a chitin-rich cell wall during any of the life cycle stages observed and therefore do not conform to the standard fungal body plan[5]. We name this highly diverse clade the cryptomycota in anticipation of formal classification.

To investigate phylogenetic diversity among the deepest branches of the fungal tree of life, we aligned a broad selection of fungal small subunit ribosomal DNA (SSU rDNA) sequences using published phylogenies[1] as a guide, to which we added a comprehensive sampling of environmental DNA sequences available in GenBank[6–8] (Supplementary Table 1). Our trees are consistent with previous analyses[1] demonstrating that nucleariids, a group of opisthokont amoebae[9], branch as sisters to the fungi, indicating a transition from a phagotrophic to an osmotrophic form very early within the fungal radiation[1]. Our analyses also recovered a highly diverse clade of environmental sequences branching with the fungi and demonstrated that current models of fungal evolution and biodiversity, which are largely based on cultured microbes, have missed a huge fraction of the kingdom (perhaps even approaching half). The analysis demonstrated preliminary bootstrap support (66/33%) for this large clade of environmental sequences branching with *Rozella*[10], the putative primary branch in the fungal phylogeny[1]. We name this clade cryptomycota (Fig. 1a).

The cryptomycota clade contains sequences recovered from diverse habitats and geographical locations, including soils, marine and freshwater sediments, freshwater planktonic samples and oxygen-depleted environments but it seems to be absent from samples of the upper marine water column (Fig. 1a, b). To investigate the ecology and cell biology of these deep-branching fungi, we designed several DNA probes for the detection of subgroups in the cryptomycota clade. Using the ARB program, with comprehensive GenBank and SILVA database sampling, we identified ten probes of about 18 base pairs (bp) that are specific to different sequences in the cryptomycota clade; probes and their target sequences are listed in Supplementary Table 2. Two probes were used successfully as forward PCR primers in combination with a general eukaryotic SSU rDNA reverse primer, 1520r (ref. 8; see Supplementary Table 2 and Fig. 1c). We then used PCR to test for the presence of the cryptomycota sequences termed CM1 and CM2 in multiple samples from a local freshwater pond, three freshwater reservoirs (Dartmoor National Park) and four coastal marine surface water samples (Devon, UK). Of the primer sequences tested, CM1 and CM2 consistently amplified cryptomycota rDNA from the Washington Singer pond (Exeter University, Devon, UK, 50.7339 °N, 3.5375 °W). We constructed clone libraries from both sets of amplicons and sequenced 12 clones from each, recovering only sequences that were 99% similar to Washington Singer CM1 in the first library and to the Lily Stem CM2 sequence previously sampled from Priest Pot pond (Cumbria, UK, 54.372 °N, 2.990 °W) in the second. This process demonstrated that both probes, when used as forward PCR primers, are specific to the two target groups in the Washington Singer pond samples. We did not detect either subgroup in the marine waters tested; however, only 0.8% of the thousands of eukaryotic environmental sequences retrieved from oceanic surface waters are classified as belonging to the Fungi[11], indicating a low density of fungi cells in the upper marine water column.

We then aimed to increase gene sampling so that we could perform additional multi-gene phylogenetic analyses to test the branching position of the environmental sequences. We constructed environmental gene libraries to sample the wider SSU–ITS1–5.8S–ITS2–large subunit (LSU) gene array, using the group-specific forward PCR primers in combination with a general eukaryotic LSU reverse primer, 28Sr1 (ref. 12), see Fig. 1c and Supplementary Table 2). This allowed a second round of phylogenetic analysis with increased character sampling and resulted in improved topology support values throughout the phylogeny (Supplementary Fig. 2). Importantly, CM1 and CM2 sequences clustered as a sister clade to *Rozella* with 96% support using two distinct bootstrap methods (shown in red on Fig. 1a and in full on Supplementary Fig. 2). This confirms that the environmental sequences branch with the genus *Rozella*.

To visualize the cells of the novel clade, we used the ten cryptomycota probe sequences detailed in Supplementary Table 2 as HRP-conjugated probes for tyramide signal amplification fluorescence *in situ* hybridization (TSA-FISH). Of these probes, CM1.1 and CM1.2 were consistently successful when we applied a high-stringency method combined with multiple helper probes[13,14]. We applied the TSA-FISH methods to the same water samples in which we had confirmed the presence of cryptomycota sequence using PCR (the Washington Singer pond and the Trenchford and Tottiford reservoirs sampled from Dartmoor National Park, UK). For both groups, our FISH analyses consistently recovered an ovoid cell type of 3–5 μm (Fig. 1d) among a dense 4′, 6-diamidino-2-phenylindole (DAPI)-stained microbial community ($n = 98$, CM1.1; $n = 58$, CM1.2). In contrast, the other eight

[1]School of Biosciences, University of Exeter, Exeter EX4 4QD, UK. [2]Department of Zoology, Natural History Museum, Cromwell Road, London SW7 5BD, UK. [3]Department of Marine Biology and Oceanography, Institut de Ciències del Mar, CSIC, Passeig Marítim de la Barceloneta 37-49, 08003 Barcelona, Catalonia, Spain. [4]Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QP, UK. [5]Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue, Boston Massachusetts 02115, USA.

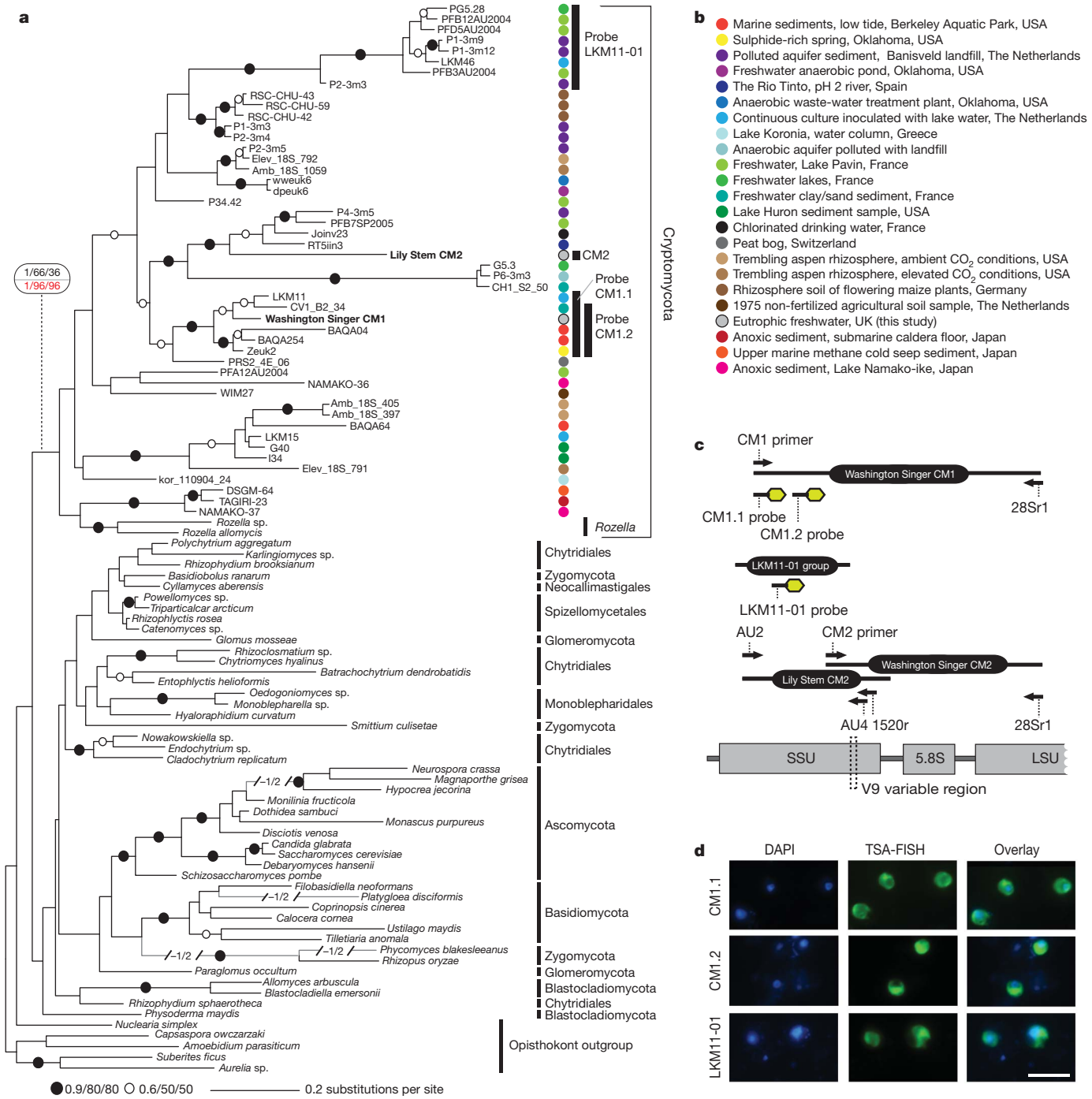

Figure 1 | Identification of the cryptomycota. **a**, Phylogeny demonstrates a diverse clade of environmental sequences (coloured dots) branching at the base of the Fungi. MrBayes tree topology was calculated from an alignment of 100 sequences and 1,012 DNA characters. Support values are summarized by black dots (indicating at least 0.9 Bayesian posterior probability and 80% bootstrap support by ML and Log-Det distance methods) or by black rings (Bayesian posterior probability above 0.6 and bootstrap support above 50%). The key node showing monophyly of cryptomycota encompassing *Rozella* is marked with actual values, with the SSU–5.8S–LSU support value (Supplementary Fig. 2a) shown in red. Shortened branches are in grey. Sequences targeted by the probes/primers are labelled on the tree. **b**, Provenance list of environmental DNA sequences. **c**, Schematic representation of the positions of primers (arrows) and probes (green hexagons) on a partial rDNA gene array (grey). **d**, TSA-FISH identification of cryptomycota cells (green) and DAPI staining of microbial community (blue). Scale bar, 10 μm.

cryptomycota group-specific probes and a panel of negative controls (detailed in the Methods) were consistently negative.

To complement this approach, we also used two LKM11 FISH probes previously used in association with additional eukaryotic probes to determine the eukaryotic community structure in freshwater environments[15]. LKM11 denotes a subsection of the environmental sequences branching within the cryptomycota clade detected in earlier studies[6–8]. We used the probe LKM11-01 (Fig. 1a, c) to identify a second subset of the cryptomycota clade and consistently recovered a cell type similar to that recovered for CM1 (Fig. 1d). The LKM11-02 probe was negative. We note that the abundance of cells identified by LKM11-01 was at least tenfold higher than observed for the CM1 cells, suggesting that the LKM11-01 subclade was more abundant than the CM1 subclade in the environments sampled.

To examine further the life cycle and morphology of the target groups in the freshwater samples, we tested whether the cells identified by

TSA-FISH possessed a flagellum. Flagella are a characteristic of chytrid fungi which, according to our phylogenetic analyses, branch closer to the cryptomycota than other fungal forms (Fig. 1a and Supplementary Figs 1 and 2). Co-localization of our TSA-FISH probes with the monoclonal antibody TAT1[16] against α-tubulin (a major structural component of eukaryotic flagella) was used to test for flagellate forms in the cryptomycota. Evidence of zoospore (flagellate spore) formation was observed for both the CM1 and LKM11-01 subgroups (Fig. 2a and Supplementary Fig. 3). In a comparative FISH experiment using 12 hybridizations from four samples (two from the Washington Singer eutrophic pond and one each from the oligotrophic Trenchford and Tottiford reservoirs, all with three replicates and using the LKM11-01 probe) we found that 47–85% of cells possessed a single flagellum ($n = 455$ of a total of 696 cells). These data demonstrate that the cells are present in a zoospore form, actively growing, reproducing and seeking food. During a typical chytrid cell cycle, a flagellum is usually lacking during the cyst phase when a cell wall is formed[5,17] and the cell enters dormancy; the majority of the remaining, unflagellate cryptomycota cells are therefore hypothesized to be cyst forms or cells in transition to form cysts (Fig. 2c).

We found preliminary evidence of a third phase of the cryptomycota life cycle. In both the LKM11-01 and CM1 groups, we observed non-flagellate cells attached to second-party cells (Fig. 2d, e) in what appeared to be an epibiotic (possibly parasitic or saprotrophic) association ($n = 39$ for LKM11-01, $n = 6$ for the CM1 probes). This represents a low total observation number ($n = 45$), an unavoidable consequence of sampling interacting microbes directly from the environment. However, in support of the interaction hypothesis, the cryptomycota cells were frequently found on diatoms. Furthermore, we saw examples of diatom cells with several attached cryptomycota cells (Supplementary Fig. 4): this would be unlikely if cell attachment were a sampling artefact. In addition, we observed non-illuminated cells attached to diatoms, indicating that the probes target cryptomycota specifically and that this was not an auto-fluorescence artefact affecting attached cells generally. Because cryptomycota are currently uncultured, we have to rely on environmental observations alone to understand their biology; it is therefore likely that several aspects of their life cycle have not yet been observed. Consequently, key cellular apparatuses may remain undetected, such as germ tubes, filter-feeding structures, rhizoids, hyphae, sporangia and cell division characteristics. However, our TSA-FISH observations of three life stages enable us to propose a skeleton life cycle for the cryptomycota lineages sampled, although we note that because of the evolutionary diversity of cryptomycota, this model is unlikely to represent accurately the numerous unobserved forms that branch within the wider clade (Fig. 2f).

The evolution of a chitin-rich cell wall was one of the most important acquisitions which drove the success and diversification of the Fungi[5], enabling these organisms to resist high osmotic pressure and feed by osmotrophy. *Rozella*, the only known genus branching within the cryptomycota clade (Fig. 1a), does not synthesize its own cell wall during many phases of its life cycle; instead, it appears to acquire cell

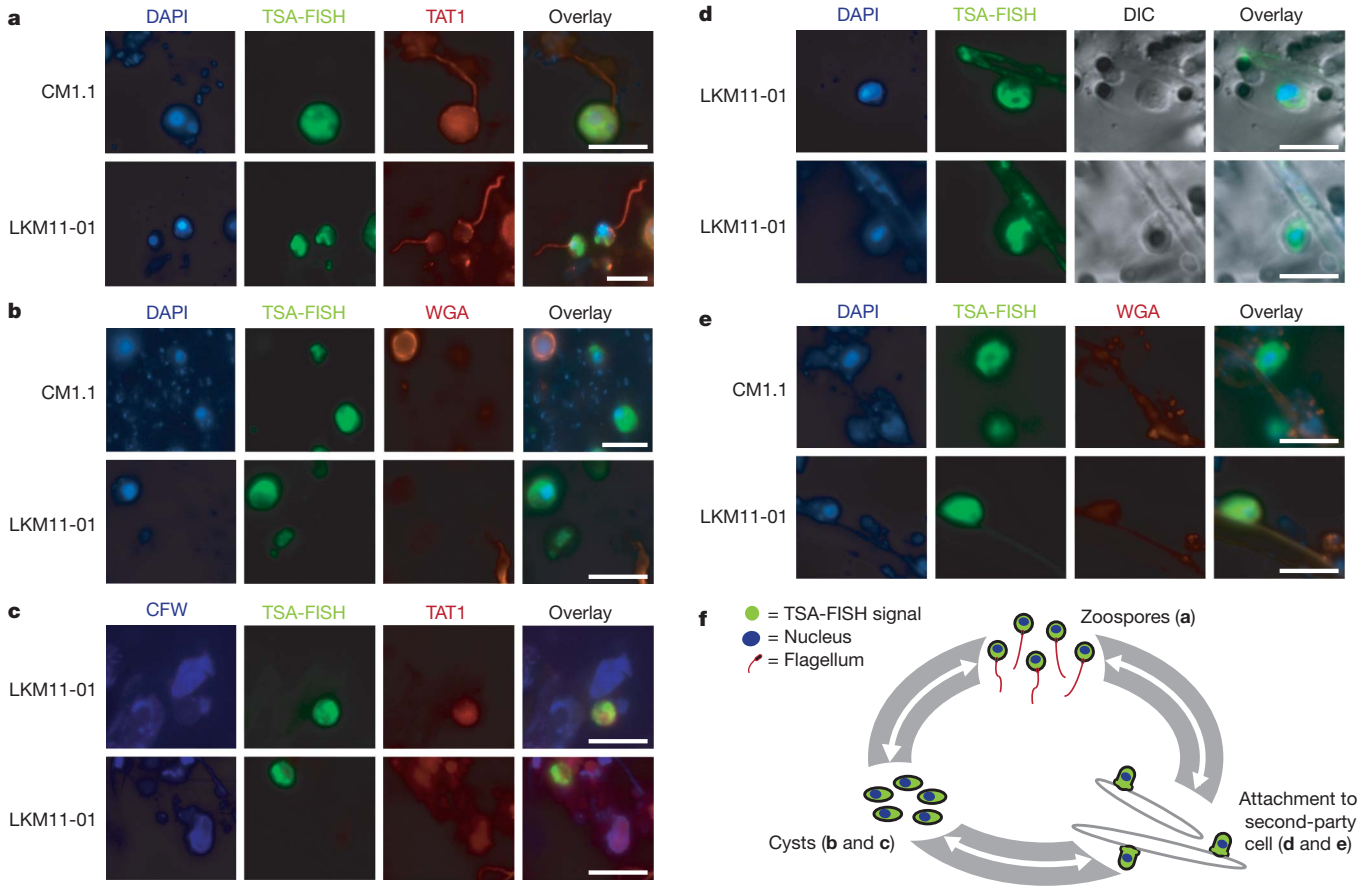

Figure 2 | Structural properties of cryptomycota cells and evidence for different life cycle stages. **a**, Micrographs showing flagella on cryptomycota cells, as detected by TAT1 tubulin antibody. **b**, Single cells without a chitin cell wall, as inferred by non-binding of wheat germ agglutinin (WGA). **c**, Non-flagellate putative cysts (lack of TAT1 signal indicates absence of flagellum), without a chitin/cellulose cell wall, as inferred by non-binding of calcofluor white. **d**, Cryptomycota cells attached to a second-party cell. Bright field differential interference contrast (DIC) shows filamentous structure of the second-party cell. **e**, Cryptomycota cells attached to a second-party cell; residual staining identifies the boundary of the second-party cell and WGA identifies absence of chitin during attachment. Scale bars, 10 μm. **f**, Putative cryptomycota skeleton life cycle (letters in brackets refer to micrographs **a**–**e**). This life cycle is limited to stages identified using TSA-FISH so additional stages are likely to remain unobserved (for example, sporangia stages and cell division) and the order of transition remains hypothetical. Furthermore, the diversity of the cryptomycota group strongly indicates that there are likely to be numerous life-cycle variations within the group, so this life cycle is unlikely to represent the wider diversity of cryptomycota.

wall material from its hosts during infection[18,19]. To investigate the cell biology of the cryptomycota further, it was therefore important to determine whether they produce a chitin- and cellulose-rich cell wall like other fungi (Supplementary Fig. 5), or whether this character is missing, consistent with the biology of *Rozella*[18,19].

We combined a TSA-FISH-based strategy with fluorescent stains specific to key compounds of the fungal cell wall. Calcofluor white binds and illuminates chitin and cellulose, the rigidifying components of fungal cell walls[5], and lectin wheat germ agglutinin binds and illuminates chitin [20]. Using these stains, we readily observed non-cryptomycota cells surrounded by chitin and chitin/cellulose in our samples (Fig. 2b, c) and in additional control experiments (Supplementary Fig. 5), confirming that these methodologies readily detect fungal cell walls. However, we did not observe a chitin/cellulose cell wall in any of the three cryptomycota life-cycle stages identified by FISH (Fig. 2b, c and e; Supplementary Figs 6 and 7). The TSA-FISH-based strategy and our samples may miss some specific stages of the life cycle with a chitin/cellulose-rich cell wall, yet our analyses show that the cryptomycota form non-flagellate putative cysts and attach to other cells (with the flagellum absent), all without a detectable chitin/cellulose cell wall. These results indicate that the key characteristic defining the kingdom Fungi — a growth and development strategy supported by a rigid chitin-rich cell wall[5] — is not present in key phases of the cryptomycota life cycle.

Using a combination of environmental DNA sequencing and fluorescence microscopy we have identified a new component of the fungal tree of life. We have tentatively named this wider group cryptomycota (crypto, hidden; mycota, of the kingdom Fungi) because a chitin-rich cell wall (one of the important fungal-defining characteristics) is so far undetected and because, with the exception of *Rozella*[18,19], the biology of this group is still largely cryptic. Significantly, the biodiversity within this clade is extensive, representing a breadth of rDNA molecular diversity similar to that of the currently sampled fungal kingdom.

METHODS SUMMARY

Phylogenetic analysis. The rDNA sequences from environmental samples and characterized species were obtained from GenBank, aligned using MUSCLE[21] and then manually refined and masked using SEAVIEW[22]. Phylogenies were calculated using 1,000 fast-ML[23] bootstraps, 1,000 Log-Det[24] distance bootstraps and Bayesian[25] methods.

FISH identification of cryptomycota cells. Surface water samples (3–4 litres) were collected from freshwater and marine coastal sites around Devon (Supplementary Table 3) and prefiltered with Miracloth (pore size 22–25 μm, Merck Chemicals). Half of the volume was immediately filtered under vacuum through a 2.0-μm polycarbonate membrane filter (Millipore) for DNA extraction. The remaining half was fixed with formaldehyde (3.7% final volume) overnight before filtration and storage at −80 °C. Oligonucleotide probes were designed against several cryptomycota subclades (Fig. 1c and Supplementary Table 2), conjugated with the HRP enzyme and used for TSA-FISH according to ref. 14. FISH optimization was performed using a range of stringency conditions (see online Methods).

To investigate whether a flagellum was present, filter pieces subjected to TSA-FISH hybridization were re-permeabilized with 0.1% v/v nonidet P-40 in PBS (10 mM Na_2HPO_4, 2 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.2), blocked with 1% w/v bovine serum albumin in PBS and incubated for 1 h with the monoclonal antibody TAT1[16] against α-tubulin, followed by fluorescein isothiocyanate (FITC)-conjugated goat anti-mouse immunoglobulins (Jackson ImmunoResearch/Stratech).

To determine the presence and composition of a cell wall, filter pieces prehybridized with TSA-FISH probes were counter-stained with cell wall markers: calcofluor white (Sigma, 1 mg ml^{-1}, fluorochrome for chitin and/or cellulose) and wheat germ agglutinin (Invitrogen, Alexa 555–580 tetramethylrhodamine conjugate, 5 μg ml^{-1}, for chitin). These were then viewed under epifluorescent ultraviolet radiation and red light respectively.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 July 2009; accepted 7 March 2011.
Published online 11 May 2011.

1. James, T. Y. *et al.* Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* **443,** 818–822 (2006).
2. Pirozynski, K. A. & Malloch, D. W. The origin of land plants: a matter of mycotrophism. *Biosystems* **6,** 153–164 (1975).
3. Wang, B. & Qiu, Y. L. Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza* **16,** 299–363 (2006).
4. Hawksworth, D. L. The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycol. Res.* **105,** 1422–1432 (2001).
5. Bartnicki-Garcia, S. *Evolutionary Biology of the Fungi* (eds Rayner, A. D. M., Brasier, C. M. & Moore, D.) 389–403 (Cambridge University Press, 1987).
6. van Hannen, E. J., Mooij, W., van Agterveld, M. P., Gons, H. J. & Laanbroek, H. J. Detritus-dependent development of the microbial community in an experimental system: qualitative analysis by denaturing gradient gel electrophoresis. *Appl. Environ. Microbiol.* **65,** 2478–2484 (1999).
7. Lepère, C., Domaizon, I. & Debroas, D. Unexpected importance of potential parasites in the composition of freshwater small-eukaryote community. *Appl. Environ. Microbiol.* **74,** 2940–2949 (2008).
8. Lefèvre, E. *et al.* Unveiling fungal zooflagellates as members of freshwater picoeukaryotes: evidence from a molecular diversity study in a deep meromictic lake. *Environ. Microbiol.* **9,** 61–71 (2007).
9. Amaral Zettler, L. A., Nerad, T., O'Kelly, C. J. & Sogin, M. L. The nucleariid amoebae: more protists at the animal-fungal boundary. *J. Eukaryot. Microbiol.* **48,** 293–297 (2001).
10. Lara, E., Moreira, D. & Lopez-Garcia, P. Environmental clade LKM11 and *Rozella* form the deepest branching clade of Fungi. *Protist* **161,** 116–121 (2010).
11. Massana, R. & Pedrós-Alió, C. Unveiling new microbial eukaryotes in the surface ocean. *Curr. Opin. Microbiol.* **11,** 213–218 (2008).
12. Bass, D., Richards, T. A., Matthai, L., Marsh, V. & Cavalier-Smith, T. DNA evidence for global dispersal and probable endemicity of protozoa. *BMC Evol. Biol.* **7,** 162 (2007).
13. Fuchs, B. M., Glöckner, F. O., Wulf, J. & Amann, R. Unlabeled helper oligonucleotides increase the in situ accessibility to 16S rRNA of fluorescently labeled oligonucleotide probes. *Appl. Environ. Microbiol.* **66,** 3603–3607 (2000).
14. Kim, E. *et al.* Newly identified and diverse plastid-bearing branch on the eukaryotic tree of life. *Proc. Natl Acad. Sci. USA* **108,** 1496–1500 (2011).
15. Mangot, J. F., Lepère, C., Bouvier, C., Debroas, D. & Domaizon, I. Community structure and dynamics of small eukaryotes targeted by new oligonucleotide probes: new insight into the lacustrine microbial food web. *Appl. Environ. Microbiol.* **75,** 6373–6381 (2009).
16. Woods, A. *et al.* Definition of individual components within the cytoskeleton of *Trypanosoma brucei* by a library of monoclonal antibodies. *J. Cell Sci.* **93,** 491–500 (1989).
17. Webster, J. & Weber, W. S. *Introduction to Fungi* 3rd edn (Cambridge University Press, 2007).
18. Held, A. A. The zoospore of *Rozella allomycis*: ultrastructure. *Can. J. Bot.* **53,** 2212–2232 (1975).
19. Held, A. A. *Rozella* and *Rozellopsis*: naked endoparasitic fungi which dress up as their hosts. *Bot. Rev.* **47,** 451–515 (1981).
20. Bulawa, C. E. Genetics and molecular biology of chitin synthesis in fungi. *Annu. Rev. Microbiol.* **47,** 505–534 (1993).
21. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5,** 113 (2004).
22. Galtier, N., Gouy, M. & Gautier, C. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.* **12,** 543–548 (1996).
23. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52,** 696–704 (2003).
24. Lockhart, P. J., Steel, M. A., Hendy, M. D. & Penny, D. Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol. Biol. Evol.* **11,** 605–612 (1994).
25. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19,** 1572–1574 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank: N. J. Talbot for advice, K. Gull for the TAT1 antibody, L. Guillou for access to curated SSU database and the Broad Institute of the Massachusetts Institute of Technology and Harvard for making their *Rhizopus* and *Batrachochytrium* genome sequence data publicly available. T.A.R. thanks the Leverhulme Trust for fellowship support. This work was primarily supported by a Natural Environment Research Council grant UK (NE/F011709/1). Additional support came from the Systematic Research Fund (awarded by the Systematics Association and the Linnean Society) to T.A.R., project FLAME (CGL2010-16304, MICINN) to R. M. and the BioMarKs project (European Funding Agencies from the ERA-net program BiodivERsA) to T.A.R. and R.M.

Author Contributions This study was conceived by T.A.R. and M.D.M.J. with assistance from D.B. and R.M.. M.D.M.J. performed the molecular biology experiments with assistance from I.F. (FISH), C.G. (immunolocalization) and M.J.E. (microscopy). T.A.R. performed the bioinformatics and phylogenetic analysis. T.A.R. and M.D.M.J. wrote the paper with assistance from D.B. and R.M.

Author Information Novel sequence data have been deposited in GenBank under accession numbers FJ687265, FJ687267 and FJ687268. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to T.A.R. (thomr@nhm.ac.uk).

METHODS

DNA extraction, PCR and environmental gene library construction. DNA was extracted from an epiphytic microbial community scraped from a submerged lily stem (Priest Pot pond, 54.372 °N, 2.990 °W) using the maximum yield protocol of the MoBio UltraClean Soil DNA Extraction Kit (Cambio), after filtration of a 50-ml stem-scraping suspended in Priest Pot water onto a GF-F filter (Whatman). Amplified PCR products of lily stem-scrapings (using the AU2 and AU4[26,27] primers, see Supplementary Table 2) were excised from the gel, purified (Wizard SV Gel and PCR Clean-Up System, Promega) and cloned (StrataClone, Stratagene) as suggested by the manufacturer. Plasmid purification (Wizard plus Miniprep Kit, Promega) was performed on 24 colonies before sequencing in the forward direction (Cogenics). Chromatograms were checked manually using Sequencher (Genecodes) and used for BLASTn searches of the National Center for Biotechnology Information, after which one positive cryptomycota clone was chosen for double-strand sequencing (GenBank number FJ687267).

Ten novel cryptomycota probes/primers (Supplementary Table 2), designed using the ARB[28] program (using database sampling from Silva[29] and GenBank[30]), were tested for specificity on community DNA extracted from the sites detailed in Supplementary Table 3 by using PCR with a forward-priming version of the probe in combination with the universal eukaryotic SSU reverse primer 1520r[8] (Supplementary Table 2). Two probes/primers were successful, yielding PCR products of the correct estimated amplicon length from Washington Singer environmental DNA using PCR amplification protocol 1 (Supplementary Table 4). The CM1–1520r PCR amplicon size was ~1,600 bp and the CM2–1520r PCR amplicon size was ~420 bp. Both amplicons crossed the variable V9 region (Fig. 1c). Two clone libraries were constructed and 12 clones from each library were sequenced.

Targeting longer rDNA sequences. We used the CM1 forward primer and an extended version of the CM2 forward primer in combination with the reverse LSU primer 28Sr1[12] (primer sequences given in Supplementary Table 2) to enable rRNA gene array PCR amplification. The CM1 and CM2 SSU–ITS1–5.8S–ITS2–LSU PCR amplifications were conducted using PCR protocols 2 and 3 respectively (Supplementary Table 4). Amplicons were purified, cloned as before and representatives of the CM1 (2,600 bp) and CM2 (1,449 bp) groups were double-strand sequenced. The Washington Singer CM2 group SSU-to-LSU sequence overlapped the Lily Stem CM2 SSU sequence by 316 identical bp, which encompassed the variable V9 region (Fig. 1c). For phylogenetic analyses, these sequences were concatenated, generating a sequence of 2,711 bp.

Sequence analysis, alignment and phylogenetic analyses. An SSU multiple-sequence DNA alignment was constructed using the work of James *et al.*[1,31] as a guide, with taxon sampling focusing on 'deep' branching fungal groups and on taxa with five rDNA sequence regions (SSU, ITS1, 5.8S, ITS2 and LSU) available (Supplementary Table 5). For the first alignment, we used the Washington Singer CM1 and Lily Stem CM2 SSU sequences as seeds for BLASTn searches of the GenBank non-redundant database to identify additional SSU sequences not sampled among currently published fungal phylogenies (for example, refs 1, 31). The alignment was calculated using the automatic alignment program MUSCLE[21], then manually refined and masked using SEAVIEW[22]. To check that we had included all available SSU sequences, we searched the GenBank non-redundant database again using all the members of the cryptomycota clade and the *Rozella* sequences as BLASTn search seeds (searches conducted May 2010). Many of the environmental sequences recovered were too short to be included in our phylogenetic analyses (Supplementary Table 1) but they indicated that the diversity of the cryptomycota clade is very extensive and even more ecologically varied than shown in Fig. 1a, b. The final phylogenetic analysis was based on environmental SSU sequences that overlapped both the CM1 and CM2 probe-binding regions; that is, only environmental SSU rDNA sequences of more than ~1,350 bp (BLAST searches conducted in May 2010). We removed all sequences with evidence of chimaeras identified by comparison to the curated SSU rDNA database (gift of L. Guillou).

A second multiple-sequence alignment was generated on the basis of the SSU–ITS1–5.8S–ITS2–LSU DNA sampling often used in the fungal tree of life project[1,31] and including our two cryptomycota sequences. This was processed, as described above, to generate a 54-sequence and 1,877-character data matrix with the ITS1 and ITS2 regions masked out because they were too variable for use. The alignment was checked for patterns of sequence variation that could be the products of chimaeras, both by visual analysis of signature sequences across the alignment and by constructing trees from one-sixth divisions of the masked alignment.

Both the SSU and SSU–5.8S–LSU sequence alignments were run through the program MODELGENERATOR[32] to identify the most appropriate model parameters for phylogenetic analysis (SSU: model of substitution GTR[33] + Γ (0.28, eight rate categories); SSU–5.8S–LSU: model of substitution GTR[33] + Γ (0.35, eight rate categories + I = 0.22)). Bayesian tree topologies were inferred for both DNA alignments using a Metropolis-coupled Markov chain Monte Carlo (MCMCMC) method. For this, we used MrBayes[25] with two runs, each with four Markov chains for 5,000,000 generations with default 'temperature' settings, a sampling frequency of 250 generations, six substitution categories (nst = 6) and an eight-category gamma model with spatial autocorrelation between rates at adjacent sites (rates = adgamma) and a covarion-like model (covarion = yes). The MCMCMC log likelihood results were compared and a 'burn-in' of 1,000 generations (SSU–5.8S–LSU analysis) and 2,500 generations (SSU analysis) was removed to sample only trees from the stationary phase of the MCMCMC searches. The resulting samples of trees were then used to construct the majority-rule consensus trees. Tree topologies were evaluated using two bootstrap methods. For each phylogeny, 1,000 bootstraps were calculated using the fast-ML method implemented in the PHYML program[23] using the model parameters from the MODELGENERATOR analyses. As compositional bias can cause phylogenetic reconstruction artefacts[34], we ran an additional bootstrap analysis using Log-Det methods[24] with stepwise addition (ten random starting trees per replicate) and a tree-bisection-reconnection branch-swapping algorithm. In both Log-Det bootstrap analyses, the 'proportion of invariant sites' parameter was estimated using PAUP[35]: SSU = 0.457 and SSU–5.8S–LSU = 0.509.

FISH identification of cryptomycota cells. Surface water samples (3–4 litres) were collected from freshwater and marine coastal sites around Devon (Supplementary Table 3) and prefiltered (pore size 22–25 μm) with Miracloth (Merck Chemicals). Half of the volume was immediately filtered under vacuum through a 2.0-μm polycarbonate membrane filter (Millipore) for DNA extraction. The remaining half was fixed with formaldehyde (3.7% final volume) overnight before filtration and storage at −80 °C. Oligonucleotide probes were designed against several cryptomycota subclades (Fig. 1a, c and Supplementary Table 2), conjugated with the HRP enzyme and used for TSA-FISH[13,14]. FISH optimization was performed using a range of stringency conditions (see below).

FISH stringency, helper probe, chitin wash and control experiments. The thirteen FISH probes, including a negative control probe (Supplementary Table 2), conjugated with the HRP enzyme, were used for TSA-FISH according to the method detailed in ref. 14. Initial hybridizations were at 35 °C with 30% formamide hybridization buffer. These were unsuccessful in revealing our target group. Formamide concentrations (30%, 40%, 45% and 50%) were varied in combination with a range of hybridization temperatures (35, 42 and 46 °C). Probes CM1.1 and CM1.2, both with a minimum of two mismatches to all non-target database sequences, were successful in illuminating candidate cryptomycota cells. These cells were not seen in the negative controls (detailed below).

To improve the rate of detection further, we designed additional helper probes (without the HRP conjugate) specific to the CM1.1 and CM1.2 target groups, which bind either side of the TSA-FISH probe-binding site (Supplementary Table 2). This approach should theoretically open up the tertiary structure of the ribosome and increase accessibility for the TSA-FISH probe to the target binding site[13,14]. Group-specific helper probes (Supplementary Table 2) were used in conjunction with the addition of 10% dextran sulphate to the hybridization reaction. For both probes, the greatest number of target cells was viewed using a 30% formamide hybridization buffer at 42 °C, with the addition of 10% dextran sulphate and helper probes.

We then combined our TSA-FISH protocol with a chitinase incubation step before hybridization, to test whether a chitin cell wall was inhibiting detection of the target cell[36]. Filter pieces were incubated in a 1 mg ml^{-1} solution of chitinase (Sigma) in modified PBS (10 mM Na_2HPO_4, 2 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 5.5 rather than 7.2) with 1% SDS. Filters were incubated for 10 min at 30 °C, rinsed with distilled water and subjected to the TSA-FISH protocol as already described. The chitinase wash step did not appear to increase detection frequency and therefore a chitin cell wall was not considered an inhibitor of detection. The chitinase wash step was not used during further TSA-FISH analyses.

For the LKM11-01 probe, we used the methodology described in ref. 15 with filter pieces hybridized at a range of formamide concentrations (30%, 40%, 45% and 50%). The optimum concentration was 40% and this was used for all subsequent hybridizations. Again, for LKM11-01 TSA-FISH, six replicate preparations were completed using the chitinase wash step. These extra analyses again failed to increase detection frequency, indeed it often reduced the frequency of detection in replicate filters indicating that inhibition of TSA-FISH binding by a chitin cell wall was not a factor, and again this protocol step was dropped from further experiments.

Negative control tests were devised to check the specificity of the probes and eliminate the possibility of autofluorescence in the sample. Filter pieces from each sample were hybridized in identical conditions to those described above, but without a TSA-FISH probe, which was replaced with distilled water. Furthermore, hybridizations conducted using a TSA-FISH probe specific to a 16S plastid SSU rRNA gene (see Supplementary Table 2) for a novel algal group[14] and probes for further cryptomycota groups were used as additional negative controls. Positive cells were never observed.

Co-staining for flagella and cell wall apparatuses. To investigate whether a flagellum was present, filter pieces subjected to TSA-FISH hybridization were re-permeabilized with 0.1% v/v nonidet P-40 in PBS (10 mM Na_2HPO_4, 2 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.2), blocked with 1% w/v bovine serum albumin in PBS and incubated for 1 h with the monoclonal antibody TAT1[16] against α-tubulin, followed by fluorescein isothiocyanate (FITC)-conjugated goat anti-mouse immunoglobulins (Jackson ImmunoResearch/Stratech).

To determine the presence and composition of a cell wall, filter pieces pre-hybridized with TSA-FISH probes were counter-stained with cell wall markers: calcofluor white (Sigma, 1 mg ml^{-1}, fluorochrome for chitin and/or cellulose) and wheat germ agglutinin (Invitrogen, Alexa 555–580 tetramethylrhodamine conjugate, 5 μg ml^{-1}, for chitin). These were then viewed under epifluorescent ultraviolet radiation and red light respectively. These markers readily detected cell wall structures in a *Blastocladiella emersonii* culture (Supplementary Fig. 5) and among the total microbial population sampled on our environmental filters (Fig. 2b, c and Supplementary Figs 6 and 7).

26. Bass, D. *et al.* Yeast forms dominate fungal diversity in the deep oceans. *Proc. R. Soc. Lond. B* **274,** 3069–3077 (2007).
27. Vandenkoornhuyse, P., Baldauf, S. L., Leyval, C., Straczek, J. & Young, J. P. W. Extensive fungal diversity in plant roots. *Science* **295,** 2051 (2002).
28. Ludwig, W. *et al.* ARB: a software environment for sequence data. *Nucleic Acids Res.* **32,** 1363–1371 (2004).
29. Pruesse, E. *et al.* SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* **35,** 7188–7196 (2007).
30. Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Wheeler, D. L. GenBank. *Nucleic Acids Res.* **34,** D16–20 (2006).
31. James, T. Y. *et al.* A molecular phylogeny of the flagellated fungi (Chytridiomycota) and description of a new phylum (Blastocladiomycota). *Mycologia* **98,** 860–871 (2006).
32. Keane, T. M. *et al.* Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol. Biol.* **6,** 29 (2004).
33. Lanave, C., Preparata, G., Saccone, C. & Serio, G. A new method for calculating evolutionary substitution rates. *J. Mol. Evol.* **20,** 86–93 (1984).
34. Foster, P. G. & Hickey, D. A. Compositional bias may affect both DNA-based and protein-based phylogenetic reconstructions. *J. Mol. Evol.* **48,** 284–290 (1999).
35. Swofford, D. L. *PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods).* Version 4 (Sinauer Associates, 2002).
36. Baschien, C., Manz, W., Neu, T. R. & Szewzyk, U. Fluorescence in situ hybrization of freshwater fungi. *Int. Rev. Hydrobiol.* **86,** 371–381 (2001).