The Environmental Clade LKM11 and *Rozella* Form the Deepest Branching Clade of Fungi

Enrique Lara, David Moreira, and Purificación López-García

**Introduction**

During the last twenty years, environmental DNA surveys based on the analysis of small subunit ribosomal RNA (SSU rRNA) gene have revealed the existence of a wide microbial diversity, including many deep-branching prokaryotic clades lacking cultured members (Barns et al. 1996; Hugenholtz et al. 1998; López-García and Moreira 2008). More recently, similar studies on eukaryotic diversity have also shown that, although eukaryotes are far more diverse than previously thought, most novel sequences determined from environmental surveys actually cluster into major acknowledged eukaryotic “kingdoms” (Berney et al. 2004). To date, in contrast to the situation in prokaryotes, only very few environmental sequences could not be associated with any of the major eukaryotic kingdoms, for example the clade DH148-5-EKD18 (López-García et al. 2001; López-García and Moreira 2008). At the next lower taxonomic level, however, the situation is far from being resolved. Several groups of environmental sequences within alveolates, heterokonts, and a variety of other eukaryotic phyla cannot be linked to any described organism. For instance, within the heterokonts or stramenopiles, thirteen undescribed clades initially detected in marine surface waters have been reported, i.e. the so-called MAST groups (Massana et al. 2004; Massana and Pedrós-Alió 2008).

Previous environmental surveys of eukaryotic diversity using the small subunit ribosomal RNA (SSU rRNA) gene have revealed many clone sequences that branch near the base of Fungi. In this work, we demonstrate that many of these sequences, including those of the environmental clade LKM11, form a monophyletic and strongly supported group that also includes two sequences derived from the parasitic genus *Rozella*. This novel clade, called here “Rozellida”, is the deepest branch of true fungi so far identified, and appears to be extremely diverse in the environment.

© 2009 Elsevier GmbH. All rights reserved.

**Key words:** environmental clades; fungi; LKM11 group; phylogeny; *Rozella*; SSU rRNA gene.
sequences from environmental Groups I and II, which are highly abundant in marine clone libraries (Epstein and López-García 2008; Guillou et al. 2008), have only recently had their taxonomic identification recognised as Syndiniales. Group I includes the genera Duboscquella (Harada et al. 2007), Ichthyodinium (Mori et al. 2007) and a Red Plasmodial parasite (Skovgaard and Daugbjerg 2008), whereas Group II comprises sequences of the genera Syndinium, Amoebophrya and Hemitadinum (Skovgaard et al. 2005). All of the known species of Group I and II alveolates are parasites/parasitoids of marine organisms (radiolaria, dinoflagellates but also copepods and teleost fishes) and this lifestyle is likely extensive to the two groups.

Opisthokonts also include a group of environmental sequences to which no described organism has yet been assigned, the LKM11 group. LKM11 was the first representative sequence, retrieved from a freshwater engineered system ten years ago (Van Hannen et al. 1999). This sequence has since been shown to represent a group of unidentified eukaryotes proposed on the basis of SSU rRNA environmental surveys. At the time when this pioneering study was first published, the small number of eukaryotic sequences present in the database did not allow the authors to place these sequences within the Opisthokonts or the Amoebozoa with significant phylogenetic support. Later, Berney et al. (2004) demonstrated that these sequences were probably related to Fungi. However, the exact phylogenetic position of the LKM11 group with respect to the Fungi remained unclear.

In this work, we analysed the phylogenetic position of the group LKM11 and relatives from the public database together with clones affiliated with this group retrieved from an environmental survey of eukaryotic SSU rRNA sequences from a peat bog. We show the existence, after the inclusion of representative sequences of major deep-branching fungal groups, of a large and diverse monophyletic clade of SSU rRNA sequences including LKM11 and Rozella as either the most basal branch of Fungi or their closest known sister group.

Results and Discussion

In the course of a general study of peat bog protist diversity, we identified a few sequences related to the environmental group LKM11; two of these were retrieved from peat samples (PRS2_4E_31 and PRS2_4E_06) and one from water samples (PR5_4E_71). We incorporated these sequences in an alignment of SSU rRNA gene sequences comprising both environmental and culture-derived sequences from several fungal clades, especially from the most basal-branching groups. Only nearly full-length SSU rRNA gene sequences were included in the alignment and long branching sequences were excluded based on initial phylogenetic analyses. We then constructed phylogenetic trees that were rooted using other opisthokont groups as outgroup sequences, including Nucleariida, Metazoa, Mesomycetozoa and Choanoflagellata. Bayesian and maximum likelihood phylogenetic analyses of these data revealed the existence of a robust clade composed mostly of environmental sequences and supported with posterior probability, PP, of 0.96 and bootstrap value, BV, of 100%. The only exceptions to this are sequences from two isolates belonging to the genus Rozella: R. allomycis and Rozella sp. ex. Rhizoclosmatium (see Fig. 1). The LKM11 (AJ130849) sequence branched within this group, as well as other environmental sequences classified in other studies as related to it (Lefranc et al. 2005; Luo et al. 2005; Šlapeta et al. 2005; Van Hannen et al. 1999). Given the high statistical support in our phylogenetic analyses for the Rozella/LKM11 clade, we propose here to name this group “Rozellida”, in reference to the genus Rozella. We suggest keeping this name between quotation marks until morphological and/or ultrastructural synapomorphies are defined to diagnose and validate this entire group. “Rozellida” appears to be the sister clade of all other true fungi (PP of 1, BV of 97%). The Nucleariida, the group of phagotrophic eukaryotes that has been found to be the closest relatives of Fungi (Steenkamp et al. 2005), appears as the only deeper branch in holofungi than “Rozellida”.

All 26 described species in the genus Rozella are parasites. The host range of Rozella spp. comprises primarily Chytridiomycetes, Blastocladiomycetes and Oomycetes (“zoosporic fungi”). However, the fact that Rozella coleochaetes has been found in the filamentous green alga Coleochaete (Sparrow et al. 1965) suggests additional potential hosts. Each Rozella morphospecies seems to be specific to a restricted number of hosts (Held 1981). Their life cycle comprises a uniflagellate motile stage that allows them to disperse in search of a new host, and a trophic wall-less intracellular stage, which develops inside a host cell (Held 1981). At this point, the parasite is amoeboid and phagocytoses the organelles of the
Figure 1. Phylogenetic tree illustrating the position of LKM11 within fungi, as obtained with a Bayesian approach, upon an alignment of 109 sequences and 1176 characters. Numbers at nodes represent, respectively, Bayesian posterior probabilities (PP) and maximum likelihood bootstrap values (BV). Black circles represent full support for the node with the two approaches (PP=1.00; BV=100%). Only support values above 0.70 PP and 50% BV are indicated. A – sign indicates that the node was not recovered with maximum likelihood. Clones obtained in this study are shown in bold.
host (Powell 1984); no filamentous growth (formation of hyphae/rhizoid) has been observed in this genus. Subsequently, the parasite eventually induces the host to create a cell wall that will surround the parasite’s future sporangium; the parasite never builds its own cell wall. Flagellated cells are produced within these walled sporangia (Held 1980). Alternatively, Rozella spp. can produce resting forms, which are surrounded by a thick wall. Although Rozella species have been reported to be parasites of soil and freshwater hosts mainly, a marine species has been also described, *Rozella marina*, which is associated with the chytrid *Chytridium polysiphonae* (Sparrow 1936).

Originally, the genus *Rozella* was placed into the Chytridiomycota, and specifically within the Spizellomycota, where it was classified inside the family Olpidiaceae (Adl et al. 2005). Its placement inside chytrids was partly justified on the basis of plesiomorphic characters, such as the existence of a flagellated stage (James et al. 2006a). The family Olpidiaceae originally comprised the genera *Olpidium, Entophlyctis, Rhizophlyctis* and *Rozella*. These genera were grouped together on the basis of the behaviour of the nucleus during development (Barr 1990). However, molecular phylogeny based on SSU rRNA and several other markers invalidated this taxon, and placed *Rozella* spp. at the base of the fungal tree with strong support (in agreement with our analysis), whereas *Olpidium* branched with zygomycetes and *Rhizophlyctis* occupied an unstable position within chytrids (Hibbett et al. 2007; James et al. 2006a). The phylogenetic position of “Rozellida” found in our analyses raises the question whether these organisms should be considered true Fungi or not. In our tree, the support for having “Rozellida” outside the remaining fungi is strong (PP=0.97%, BV=92%, Fig. 1). Following Adl et al.’s list of synapomorphies that define Fungi (Adl et al. 2005), *Rozella* spp. should logically not be included within this taxon since true Fungi are defined as organisms that have lost phagocytosis, which is retained in the trophic stage of *Rozella* spp. (Powell 1984). However, whether other uncultured members of “Rozellida” also have a phagotrophic stage in their life cycle remains to be determined. As the Nuclearida, the closest known sister group of Fungi are phagotrophic, it is probable that this feeding strategy was the ancestral state for the Fungi. Consequently, the secondary loss of phagocytosis might not be a good diagnostic character. Other features of *Rozella* spp. advocate for their placement into Fungi: flagellated zoospores with fungal-like ultrastructure of the flagellar apparatus (Held 1975), thick-walled resting spores and penetration of the host via a germ tube (Held 1973). Further prospecting on more representatives of the “Rozellida” will be crucial for circumscribing the true fungi.

The original environments from which the clones considered in this study derived were mainly freshwater (like LKM11; Van Hannen et al. 1999) but also soil (such as Elev_18S_791; Lesaulnier et al. 2008) and marine sediments (for instance TAGIRI 23; Takishita et al. 2005). Many were retrieved from anoxic or potentially anoxic environments (like, for instance, CCW48 and BAQA254; Dawson and Pace 2002; CV1_B2_34; Slapeta et al. 2005). This corresponds to the environments where *Rozella* spp. have been reported, but also to those inhabited by some of their potential hosts. Chytrids, for instance, are typical inhabitants of freshwater and soil systems, although some species can be found in marine environments as well (Küpper and Müller 1999). Additionally, many chytrid species live in anoxic conditions, as some are facultative anaerobes with obligate fermentation (Emerson and Natvig 1981). Conversely, although the most extensively sampled environment in DNA surveys is clearly the oceanic euphotic zone (Epstein and López-García 2008), we did not find any sequence related to LKM11 in published molecular surveys from this environment, suggesting that their potential chytrid and oomycete hosts may be rare or even absent in surface marine waters. Furthermore, the frequent presence of long branches inside the “Rozellida” agrees with the hypothesis that this group might be composed to a large extent (if not entirely) of parasites, as parasites, in particular intracellular ones, very often have accelerated evolutionary rates (Itoh et al. 2002).

However, it would be premature to draw any conclusion on the lifestyle and ecology of these organisms based only on environmental sequences. Further work is required to identify and characterise additional members of this group. Their phylogenetic position as the most basal branch of fungi, or as their sister group, depending on how fungi are to be defined, gives them special importance. This is further reinforced by the suggestion, on the basis of protein-coding genes, that Microsporidia could be related to *Rozella allomyces* (James et al. 2006b). This hypothesis is particularly attractive given the likely parasitic nature of the two groups. Both groups would also share the absence of a cell wall when
invading host cells, so that the plasma membrane of the parasite makes direct contact with the cytoplasm of the host cell. However, no sequences from this group were included in our analyses, because microsporidian SSU rRNA sequences have extremely long branches being one of the most striking examples of a long branch attraction artefact in eukaryotic trees (Philippe and Adoutte 1998). An increase in the sampling of “Rozellida” taxa sampled in a pool where vegetation was dominated by half submerged Sphagnum cuspidatum, with presence of Drosera rotundifolia. The water was prefiltered at 50 μm, and the biomass was gathered on a 0.2 μm nitrocellulose GTP Millipore filter. DNA was extracted from the filters as described elsewhere (Lara et al. 2009). In addition, peat was also collected for DNA extraction, which was performed using a MoBio Power SoilTM DNA extraction kit (Carlsbad, CA USA).

Methods

Sample collection, DNA extraction, PCR and cloning: Samples were taken at the Praz-Rodet peat bog, located in Switzerland (46°33′ N; 06° 10′ E; altitude 1041 m). Water was sampled in a pool where vegetation was dominated by half submerged Sphagnum cuspidatum, with presence of Drosera rotundifolia. The water was prefiltered at 50 μm, and the biomass was gathered on a 0.2 μm nitrocellulose GTP Millipore filter. DNA was extracted from the filters as described elsewhere (Lara et al. 2009). In addition, peat was also collected for DNA extraction, which was performed using a MoBio Power SoilTM DNA extraction kit (Carlsbad, CA USA) following the manufacturer’s instructions. DNA was amplified using the eukaryotic-specific primers EK-82F (GAAACTGC- GAATGGCTC) and EK-1492R (CACCTACGGAAACCTTGTTA). Subsequently, two clone libraries per sample were built. PCR reactions were carried out in 25 μl of reaction buffer containing 1 μl DNA template (~1-5 ng), 1.5 mM MgCl2, dNTPs (10 nmol each), 20 pmol of each primer, and 1 U Taq DNA polymerase (Promega). PCR reactions were performed under the following conditions: 35 cycles (denaturation at 94 °C for 15 s, annealing at 50 °C for 30 s, extension at 72 °C for 2 min) preceded by 2 min denaturation at 94 °C, and followed by 8 min extension at 72 °C. Amplicons were cloned into pCR2.1 Topo TA cloning vector (Invitrogen) and transformed into Escherichia coli TOP10’ One Shot Cells (Invitrogen) according to the manufacturer’s instructions. Clone inserts were amplified with vector T7 and M13R primers, and inserts of the expected size were sequenced directly using vector primers by Cogenics (Meylan, France). The obtained sequences can be retrieved from GenBank under accession codes FJ976648–FJ976650.

Sequence and phylogenetic analyses: LKM11 clone sequences were aligned with a representation for fungal SSU rRNA sequences obtained from the GenBank data base using MUSCLE v. 3.6 (Edgar 2004). After excluding ambiguously aligned positions and gaps, the resulting alignment was subjected to Bayesian phylogenetic analysis with the software MrBayes v. 3.1.2 (Huelsenbeck and Ronquist 2001). The model chosen was GTR (Lanave et al. 1984; Rodríguez et al. 1990) with the number of invariable sites being estimated, and a gamma-shaped distribution of variable sites with four rate categories (GTR+I+Γ), four chains were run up to 1,000,000 generations from a random starting tree well beyond convergence (the value of Ln L stabilised around -15000; PSRF parameter=1.006). Trees were sampled every 1000 generations. The first 5000 trees were discarded as the burn in. In addition, a maximum likelihood analysis was carried out using the program TREEFINDER (Jobb et al. 2004), applying a GTR+I+Γ model of nucleotide substitution (Rodríguez et al. 1990). All necessary parameters were estimated from the data sets. Bootstrap values were calculated from 1000 replicates.

Acknowledgements

This work was supported by an ATIP Plus grant of the French Centre National de la Recherche Scientifique (CNRS), section “Dynamique de la biodiversité” to P.L.G. E.L benefited from an ATIP-associated CNRS research assistant contract.

References


