Resolving relationships at the animal-fungal divergence: A molecular phylogenetic study of the protist trichomycetes (Ichthyosporea, Eccrinida)

Nicole K. Reynolds^a, Matthew E. Smith^b, Eric D. Tretter^a, Justin Gause^a, Dustin Heeney^a, Matías J. Cafaro^c, James F. Smith^a, Stephen J. Novak^a, William A. Bourland^a, Merlin M. White^{a*}

^a *Boise State University, Department of Biological Sciences, Boise, Idaho 83725-1515*

^b *University of Florida, Department of Plant Pathology, Gainesville, Florida 32611*

^c *University of Puerto Rico, Department of Biology, Mayaguez, Puerto Rico 00681*

*Corresponding author email: merlinwhite@boisestate.edu

Abstract

Trichomycetes is a group of microorganisms that was considered a class of fungi comprising four orders of commensal, gut-dwelling endosymbionts obligately associated with arthropods. Since molecular phylogenies revealed two of those orders (Amoebidiales and Eccrinales = “protist trichos”) to be closely related to members of the protist class Ichthyosporea (= Mesomycetozoea), trichomycetes have been considered an ecological association of both early-diverging fungi and protists. Understanding of the taxonomy, evolution, and diversity of the protist trichos is lacking largely due to the difficulties inherent in species collection that have contributed to undersampling and understudy. The most recent classification divides the protist trichos between two families, Amoebidiidae and Eccrinidae (suborder Trichomycina, order Eccrinida). However, there is no comprehensive molecular phylogeny available for this group and major questions about the systematics of protist trichos remain unanswered. Therefore, we generated 18S and 28S rDNA sequences for 106 protist tricho samples and combined them with publicly available Eccrinida sequences for phylogenetic analyses. We also sequenced a conserved protein-coding gene (heat-shock 70 protein) to obtain a

multigene data set. We conducted ancestral state reconstruction (ASR) and Bayesian tip-association significance test (BaTS) analyses by mapping six morphological and ecological characters onto the resulting phylogenetic trees. Our results demonstrate: 1) several ecological and morphological character states (habitat, host type, host stage at time of infestation, location within host, spore production, and growth form) are significantly correlated with the phylogeny, and 2) two additional protist tricho families should be incorporated into the taxonomy to reflect phylogenetic relationships. Our data suggest that an integrated strategy that combines morphological, ecological, and molecular characters is needed to further resolve and clarify the systematics of the Eccrinida.

Keywords: Eccrinida, evolution, Ichthyosporea, protists, symbiosis, trichomycetes

1. Introduction

Trichomycetes, previously considered a class of arthropod gut-dwelling fungi, has undergone radical taxonomic rearrangements with the advent of molecular systematics. Characters such as habitat, hyphal morphology, and the production of spores originally placed the microscopic trichomycetes into a single class. Variations in spore development, host association, and location of attachment in the digestive tract divided trichomycetes into four different fungal orders (Amoebidiales, Asellariales, Eccrinales, and Harpellales) (Léger and Duboscq, 1929a; Lichtwardt and Manier, 1978). However, appendage ontogeny (Moss, 1999) and the lack of chitin in the cell wall (Trotter and Whisler, 1965; Whisler, 1963) pointed to an independent evolutionary origin of two orders (Amoebidiidae and Eccrinidae). Later molecular studies supported this division between Amoebidiidae and Eccrinidae and also confirmed that these taxa are divergent from the fungal trichomycetes (Benny and O'Donnell, 2000; Cafaro, 2005; Ustinova et al., 2000). Trichomycetes is now considered a polyphyletic ecological group comprised of fungi and protists, with the protist trichomycetes ("protist trichos") belonging

to the suborder Trichomycina, order Eccrinida, class Ichthyosporea (= Mesomycetozoea) (Cavalier-Smith, 1998; Cavalier-Smith, 2013; Mendoza et al., 2001). The protist tricho group includes Amoebidiidae and Eccrinidae but is not equivalent to the Trichomycina. The Trichomycina also includes a third family (Ichthyophonidae) that does not share the same morphological or ecological traits as the protist trichos, and was not previously thought to be evolutionarily related to the protist trichos.

Protist trichos are found in association with a diverse range of hosts and habitats. Members of the family Eccrinidae have been found from the foreguts or hindguts of nearly every major group of mandibulate arthropods (crustaceans, millipedes, insects) in habitats ranging from deep-sea hydrothermal vents (VanDover and Lichtwardt, 1986) to terrestrial caves (Manier, 1964; Reeves et al., 2000) and from tropical forests (White et al., 2000) to freshwater streams or lakes (Lichtwardt et al., 2001). In contrast, species of the family Amoebidiidae are apparently restricted to freshwater habitats and are mainly found in the hindguts of immature aquatic insects such as stonefly and mayfly nymphs and blackfly larvae (Lichtwardt et al., 2001). Major differences in life cycles also divide the two orders. Amoebidiidae have an amoeboid phase and produce endospores (i.e. the entire “hyphal cell” divides to become spores) whereas Eccrinidae do not have an amoeboid phase and they produce spores directly from the tips of the hyphae (i.e. basipetally) (Fig. 1) (Moss, 1979).

Another character that separates the two families of protist trichos is maturity of the host at the time of infestation. The majority of the Eccrinidae have been collected from the guts of the adult form of their host. On the other hand, species of *Paramoebidium* (Amoebidiidae) have only been found in the nymph or larval forms of their hosts. *Amoebidium* species (Amoebidiidae) may be associated with both immature and mature hosts and they are unique among the protist trichos in that they attach externally onto the exoskeleton.

While the protist trichos are considered commensalistic (Moss, 1979), some of their putatively close relatives are parasites of various animals (Mendoza et al., 2002; Glockling et al., 2013). The monogeneric family Ichthyophonidae (*Ichthyophonus*), inferred sister to the protist trichos (Benny and O'Donnell, 2000; Ustinova et al., 2000), are pathogens of fishes and amphibians, and outbreaks within an aquaculture setting can have economic impacts (Glockling et al., 2013; Rowley et al., 2013). Pathogenesis occurs when the parasite forms cyst-like structures within the tissues of the host's vital organs such as the kidney or heart. Species of *Ichthyophonus* were also previously considered fungi and may have hyphal-like growth in culture (Okamoto et al., 1985). However, one species of *Ichthyophonus* has recently been suggested to have a cryptic amoeboid phase in its life cycle (Kocan et al., 2013). *Ichthyophonus* species or strains may infect either freshwater or marine fishes, but dispersal and transmission of the parasites remains poorly understood (Gozlan et al., 2014; Kocan et al., 2013; Rowley et al., 2013).

The study of protist trichos is difficult for several reasons. Firstly, most are unculturable. Despite attempts by different researchers (Dang and Lichtwardt, 1979; Lichtwardt et al., 2001; Whisler, 1965), *Amoebidium parasiticum* and *A. appalachense* are the only protist trichos to be cultured axenically (Whisler, 1960; White et al., 2006). Consequently, studies are dependent on obtaining material directly from inside the bodies of host organisms. This means that collection is dependent on abundant hosts with high infection rates. Observations on life history, morphology, and host-symbiont interactions are restricted to those moments during dissection of the host. Secondly, molecular studies are often hampered by contamination with host and/or other symbiont DNA. Finally, identification of samples is a challenge due to the inter- and intraspecific variability of morphological characters and the fact that samples may be collected at different life stages when the most important diagnostic characters may not always be

present. Coupled with unknown host specificity and the possible cohabitation of multiple protist tricho species in a single gut, the result is that morphological species identification is of limited utility (e.g. Lichtwardt, 1986; Valle, 2014a).

The most recent molecular phylogenetic treatment of the protist trichos was published over a decade ago and included only 16 species and single gene datasets (Cafaro, 2005). As a result, there are many open questions regarding the basic relationships and evolutionary patterns among the protist trichos. The most recent taxonomy of the protist trichos by Cavalier-Smith (2013) divides this group among two families (Amoebidiidae, and Eccrinidae, with the Ichthyophonidae erected for *Ichthyophonus*). However, this taxonomic scheme has not been confirmed by a molecular phylogenetic analysis.

This study aimed to reconstruct a robust multigene phylogeny of the protist trichos in order to: 1) examine evolutionary patterns in ecology and life-history across the group, including habitat, host type, host maturity at the time of infection, and location of symbiont in the host, 2) investigate morphological variation across the group, including patterns in the growth form and spore production, and 3) evaluate the congruence of Cavalier-Smith's (2013) classification with a molecular phylogeny. We also used two different types of ancestral state reconstruction analyses to gain further insight into the evolution of these traits across taxa of protist trichos.

2. Materials and Methods

2.1 Taxon Sampling

We included sequences from all known Eccrinida taxa, with particular emphasis on genera in the suborder Trichomycina. Of five named *Amoebidium* species, two are represented here, as well as six of 17 *Paramoebidium* species and representatives from 12 of 17 Eccrinidae genera. Our dataset is the first to include representative rDNA gene sequences from the genera *Arundinula*, *Leidyomyces*, and *Parataeniella* within

Eccrinidae. Representatives of named species were included when possible but some specimens not identified to species were included to maximize phylogenetic diversity. Publicly available sequences for *Ichthyophonus*, environmental sequences, taxa in the order Dermocystida, and other outgroup representatives were downloaded from GenBank (Table S1).

We obtained axenic cultures of several species in order to amplify protein-coding genes. *Creolimax fragrantissima* (ATCC PRA-284), *Sphaeroforma arctica* (ATCC PRA-297) and *Sphaeroforma* sp. (ATCC PRA-283) were obtained from the American Type Culture Collection (ATCC) and grown on ATCC medium 2673 (www.atcc.org). Several isolates of *Ichthyophonus* sp. growing on rainbow trout heart tissue explants (Kocan et al., 2010) were obtained from Dr. Scott LaPatra of Clear Springs Foods, Inc., Idaho. Several of these tissue explants were placed directly in 2X CTAB buffer whereas the remaining tissues were microdissected in order to separate the *Ichthyophonus* cells from the host tissue. These *Ichthyophonus* cells were rinsed in successive drops of NanoPure water before CTAB preservation and were subsequently used for DNA extraction.

2.2 Sample collection, DNA Extraction, and PCR Amplification

Unculturable specimens were microdissected directly from arthropods using the methods of White et al. (2001) (Table S1) and placed in 2X CTAB buffer. Remaining tissues were preserved on microscope slides using the methods of Lichtwardt (1986). Samples preserved in CTAB buffer were stored at -20°C or 4°C until extraction. DNA extractions followed the methods of White (2006). Some DNA extractions of Cafaro (2005) were also included in this study.

The small subunit (18S) and large subunit (28S) ribosomal rDNA genes were PCR-amplified and sequenced. Several primer combinations were used (Tables S2 and S3). Primers NS1AA and NS8AA (18S) and NL1AA and LR7AA (28S) were developed

by Wang et al. (2014) to exclude animal DNA, a common contaminant. Primers NS1PT and NS8PT (18S) are taxon-specific primers designed for this study that only amplify protist trichos. The PCR recipes utilized Promega GoTaq Green Master Mix and GoTaq Hot Start Green Master Mix, with reagent concentrations following Wang et al. (2014) and thermal cycling programs listed in Table S3. *Ichthyosporea*-specific primers designed by Paps et al. (2013) were used with a nested PCR approach to amplify the HSP70 gene.

2.3 Electrophoresis and Sequencing

For most samples, amplicons were isolated with the “freeze and squeeze” method of Tautz and Renz (1983). In the “freeze and squeeze” method, bands are excised from agarose gels, frozen, and centrifuged. The resultant liquid is then used for direct Sanger sequencing. Products were cleaned and sequenced at the University of Wisconsin Biotechnology Center on an Applied Biosystems ABI 3730xl capillary DNA analyzer. Chromatograms were visualized, trimmed, aligned into contigs and edited for errors using Sequencher 5.0. Edited sequences were checked against the NCBI database using BLASTn or BLASTx as a preliminary assessment of identity.

2.4 Cloning

Samples for which direct sequencing was unsuccessful were cloned using the Promega pGEM T Easy Vector System kit with JM109 cells. The ligation reaction mixture was modified using the following reagent quantities: 2.5 µl ligation buffer, 0.5 µl pGEM T Easy vector, 0.5 µl T4 ligase, 0.5 µl water and 1 to 2 µl PCR product. Ligations were stored at 4°C overnight. Transformation procedures followed manufacturer protocols except that we used a reduced volume of cells (20-25 µl) and SOC media (200 µl). Plates were incubated at 37°C for 16-18 hours and then amplified via PCR. PCR products were electrophoresed to confirm fragment sizes and either sequenced directly or cut from gels as described above.

2.5 Phylogenetic Analyses

Sequences were imported into Mesquite v. 3.01 (Maddison and Maddison, 2014) and aligned with MUSCLE (Edgar, 2004). Alignments were adjusted and ambiguously aligned regions were excluded manually. Alternate versions of the alignment were created using Gblocks (Castresana, 2000) but missing data affected which settings could be utilized. Since the resulting output was not significantly better for estimating tree topologies, we relied on the manually edited data matrix. For the HSP70 alignment, the reading frame and translation to the protein sequence was determined using the ExPASy translation tool (Artimo et al., 2012). The translation tool in Mesquite was also used to create an amino acid alignment from the nucleotide alignment. After an initial screening using BLASTx, putative HSP70 sequences were aligned with sequences from a wide variety of fungal, animal, and protist taxa and a tree was estimated from these data to ensure they formed a clade with other Eccrinida. As a result, several sequences were identified as either host DNA or a paralogous gene and were thus removed from further analyses. Individual gene trees as well as trees based on combined datasets were constructed to check for congruence. To check for substitution saturation, plots of K80 distance against transitions and transversions (Brown et al., 1982) were drawn in DAMBE5 (Xia, 2013). For this saturation analysis, the third codon position was evaluated separately from the first and second codon positions for HSP70. Distance versus transitions and transversions plots indicated no significant saturation for rDNA or the first two codon positions of HSP70. On the other hand, the third HSP70 codon position was saturated. We inferred similar tree topologies and number of supported branches from the nucleotide and amino acid datasets. Accordingly, we used only the HSP70 nucleotide alignment for the final analyses.

Alignments were analyzed with PartitionFinder v. 1.1.0 (Lanfear et al., 2012) and jModelTest 2.0 (Darriba et al., 2012; Guindon and Gascuel, 2003) to evaluate

appropriate data partitions and for substitution model selection. Models were chosen based on corrected AIC scores. Two versions of the rDNA tree were inferred: one with all successfully sequenced samples and large outgroup sampling ("large rDNA dataset") and another focusing only on Eccrinida taxa with *Capsaspora owczarzaki* as the outgroup ("small rDNA dataset"). Only the small rDNA dataset was used for ancestral state reconstructions to reduce computation time and enhance tree visualization. Combined two gene (18S, 28S) and three-gene (18S, 28S rDNA, HSP70) datasets containing 53 unique Trichomycina samples (24 Eccrinidae, 28 Amoebidiidae and one *Ichthyophonus* sp.) were analyzed.

Tree inferences and evaluations of support were conducted using maximum parsimony (MP), maximum likelihood (ML) and Bayesian inference (BI) methods for all datasets (i.e., the large rDNA dataset, the small rDNA dataset, and the three-gene dataset). Support values from ML and MP were added to the BI consensus tree for illustration. The software package TNT (Goloboff et al., 2008) was used for parsimony analyses. Tree searches were conducted using 10,000 random addition sequences and TBR swapping, followed by 5,000 bootstrap replicates. Maximum likelihood calculations were performed with RAxML v. 8.0.22 using the GTR + Γ + I model (Stamatakis, 2014). Finally, BI was conducted with MrBayes v. 3.2.2 and v. 3.2.5 (Ronquist et al., 2011; Ronquist et al., 2012) using the partitioned dataset, four chains, 20 million generations for the two gene alignments and 10 million for the three-gene, with half of each discarded as burn-in. Convergence and effective sampling of parameters were assessed with Tracer v. 1.5.0 (Rambaut and Drummond, 2009). Tree files were viewed and illustrated using TreeGraph 2 (Stöver and Müller, 2010), FigTree v. 1.4.2 (Rambaut, 2014), and Inkscape (<https://inkscape.org/en>). Nodes with bootstrap values >70% (MP) or >75% (ML) and posterior probabilities >0.95 (BI) were considered well supported.

Individual gene trees were assessed for topological congruence. Any conflicts that were not well-supported (e.g. < 70% MP, 75% ML, and 0.95 PP) were considered as congruent.

2.6 Topology testing

Alternative hypotheses of tree topology for the large rDNA dataset were tested using unresolved polytomy constraint trees drawn in Mesquite. These trees were then imported into RAxML to generate ML constraint trees, followed by estimation of per site log likelihoods for each tree (-f g command). Four alternative topologies were tested based on the hypotheses of Cafaro (2005) and Cavalier-Smith (2013). The alternative topologies are: 1) monophyly of Amoebidiidae (*Paramoebidium* + *Amoebidium*), 2) monophyly of all protist trichos (Amoebidiidae, including *Paramoebidium* + Eccrinidae, including *Parataeniella*), 3) sister taxa relationship of the genera *Parataeniella* and *Paramoebidium* (an arrangement found in the second best RAxML tree), and 4) monophyly of Eccrinidae excluding *Parataeniella*. The software package TREE-PUZZLE (Schmidt et al., 2002) was used to perform Shimodaira-Hasegawa (SH) tests (Shimodaira and Hasegawa, 1999) and the software package CONSEL (Shimodaira and Hasegawa, 2001) was used to conduct Approximately Unbiased (AU) tests (Shimodaira, 2002). The null hypothesis for the SH test is that all trees $T_x \in T$ (including the ML tree) are equally good explanations of the data (Schmidt, 2009). The null hypothesis for the AU test is that the negative log likelihoods of the constrained trees and best unconstrained maximum likelihood tree do not differ significantly from each other.

2.7 Ancestral State Reconstruction (ASR)

Ancestral state reconstructions were performed based on morphological, ecological, and life-history characters described in published papers and species descriptions (e.g. Hibbits Galt, 1978; Lichtwardt et al., 2001; Lohr et al., 2010; Lord et al., 2012; Marshall et al., 2008; Marshall and Berbee, 2010; Marshall and Berbee, 2011;

Vogt and Rug, 1999). A full list of characters and character states is in Table 1. Six characters were reconstructed using the last 5,000 trees of the BI analysis of the small rDNA dataset using SIMMAP v. 1.5.2 (Bollback, 2006). A set of 5,000 trees was chosen in order to incorporate phylogenetic uncertainty while also minimizing computation time. All trees were rooted using *Capsaspora owczarzaki* as the outgroup based on Cavalier-Smith (2013) and availability of character state data. We also performed alternative topologies with species of order Dermocystida as the outgroup. However, since relationships within the protist trichos were unaffected by this change we only present analyses with *C. owczarzaki* as the outgroup.

We used the two-step process described on the SIMMAP website (Bollback, 2009) to choose priors for multistate characters using a maximum clade credibility tree. The first step uses an MCMC analysis to sample the gamma and beta rate values for multistate characters and the second step uses the samples from the posterior distributions of these analyses to plot the best fitting distribution in the R Statistical Package (R Foundation for Statistical Computing 2015). These plots were drawn with the R script included with the SIMMAP software download. Both an equal (1/k) and empirical bias prior were tested with this two-step approach and output of the MCMC runs (one million generations each, sampling every 200 generations with a burn-in of 10,000 and upper bound of 1,000) were visualized with Tracer (Rambaut and Drummond, 2009) to ensure appropriate sampling was achieved and to compare the log likelihoods. Character states were unordered. Once priors were chosen, analyses were run using k=90, rate=1.00, and 20 samples/tree with 20 priors drawn from the prior distributions. Results were plotted as pie charts using the R script "PlotSimMap.R" (<https://github.com/nylander/PlotSimMap>) and Inkscape (<https://inkscape.org/en>). Polymorphic states were coded as separate states for these analyses so they are considered as polymorphisms rather than uncertainty. This approach implicitly assumes

that the states are correlated and arose simultaneously in those taxa with polymorphisms (Millanes et al., 2011). For example, *Ichthyophonus* species exhibit three different growth forms (hyphal, plasmodial, spherical) so taxa with this set of features are coded with a single state (Table 1).

2.8 Bayesian Tip-association Significance Testing (BaTS)

The software package BaTS beta v. 2 (Parker et al., 2008) was used to test correlations between characters and the phylogeny. The last 1,002 trees of the BI analysis were used to incorporate phylogenetic uncertainty. We used 1,002 trees based on BaTS beta v. 2 software recommendations that datasets significantly >1,000 trees are computationally unwieldy. The BaTS analysis generates three statistics: the association index (AI) (Wang et al., 2001), parsimony score (PS) (Fitch, 1971), and maximum exclusive single-state clade size (MC) (Parker et al., 2008). These statistics are computed by generating a null distribution of phylogeny-trait associations that is then used to compare with observed values and calculate significance. The null hypothesis is that characters at the tips are randomly distributed across the phylogeny. For both the AI and PS, lower values indicate stronger phylogeny-trait associations whereas higher values of MC indicate stronger phylogeny-trait associations. Characters and character states were the same as for the ASR analysis and 1,000 null replicates were performed with p -value of < 0.05 considered significant.

3. Results

3.1 Molecular Data

We generated a total of 19 new HSP70 and 121 new rDNA sequences (Table S1). PCR protocol 194 was the best for amplifying 18S whereas protocol 183 was the best for amplifying 28S (Table S3). Host sequences were occasionally sequenced (10–20% of the time) with the 18S and HSP70 protocols and with 28S protocol 133. The 28S protocol 183 never amplified host sequences. These results highlight the importance of

primer specificity when attempting PCR on metagenomic samples recovered from host dissections.

The 18S rDNA fragment ranged from 1,613 – 2,011 bp, the 28S rDNA fragment ranged from 1,483 – 1,965 bp, and HSP70 from 750 – 800 bp. The final three-gene alignment consisted of 27 taxa and 3,400 characters: 1,546 bp of 18S, 1,251 bp of 28S, and 759 bp of HSP70, with a total of 1,544 unalignable characters that were excluded. The “small rDNA dataset” alignment consisted of 101 taxa and 2,678 characters: 1,479 bp of 18S, and 1,199 bp of 28S characters, and 2,442 unalignable characters that were excluded. Finally, the “large rDNA dataset” had 121 taxa and 2,640 characters; 1,466 bp of 18S and 1,174 bp of 28S, and 2,576 unalignable characters that were excluded. There was a total of 9% missing data (number of missing characters/total characters) in the three-gene final alignment, 20% missing data in the “small rDNA dataset”, and 22% missing data in the “large rDNA dataset.”

3.2 Phylogenetic Analyses

The only *Paramoebidium* sp. rDNA sequence in GenBank (AY336708) was excluded from ASR, the small rDNA, and three-gene datasets because BI analyses placed it within the *Amoebidium* clade (Fig. S1), suggesting that this isolate was misidentified. Attempts to amplify and sequence from the stored genomic DNA (KS-61-W6, sample 1175) were unsuccessful. The 28S sequence of the *Tenebrio molitor* symbiont (TMS) in GenBank (JN699061) was excluded because it could not be aligned with the rest of the dataset. However, the TMS 18S sequence was retained because it was alignable and apparently is a valid sequence.

There were no supported topology conflicts between phylogenies of the three individual genes (18S, 28S, HSP70). Trees based on the three analysis methods (MP, ML, and BI) were congruent for the small rDNA dataset (Fig. 2). However, MP analysis produced trees that were different from ML and BI trees in the other two datasets. In the

three-gene MP analysis (Fig. S2), *Ichthyophonus* was placed outside the Trichomycina and sample 551 was sister to sample 538. In the large rDNA MP dataset (Fig. S1), *Pirum gemmata* was sister to the four environmental sequences as opposed to being sister to *Caullerya* and the TMS.

There was an unexpected split of the Amoebidiidae by genus, with *Amoebidium* and *Paramoebidium* each forming well-supported clades that were phylogenetically separated from one another in all the best ML and BI trees (Figs. 2, S1, and S2). However, topology tests do not reject the hypothesis of a sister relationship between these two clades (Table 2). The BI and MP (but not ML) analyses supported *Amoebidium* as the earliest-diverging Trichomycina lineage in the small rDNA dataset (Fig. 2). Similarly, the three-gene analyses recovered support from BI and ML for the same placement of *Amoebidium* (Fig. S2). Our best ML and BI trees resolved *Ichthyophonus* inside the Trichomycina, i.e., among the protist trichos (Figs. 2, S1, S2), as found in previous studies (Benny and O'Donnell, 2000; Cafaro, 2005; Ustinova et al., 2000). Nonetheless, the constraint tree forcing a monophyletic protist tricho clade (and thus placing *Ichthyophonus* outside of the protist trichos) is not rejected by topology tests (Table 2).

All of our analyses recovered the family Eccrinidae as a monophyletic clade but without statistical support (Figs. 2, S1, and S2). This result is consistent with previous studies that included fewer taxa and genes (Cafaro, 2005; Marshall et al., 2008; Marshall and Berbee, 2011). Among alternate topologies, *Parataeniella* was resolved on an early diverging branch outside the Eccrinidae (not rejected by topology tests) and sometimes as sister to *Paramoebidium*. The topology of *Parataeniella* and *Paramoebidium* as sister taxa was a significantly worse explanation of the data than the best RAxML tree by SH and AU tests (Table 2).

In the large rDNA dataset (Fig. S1), orders Eccrinida and Dermocystida each form well-supported clades but are not sister taxa. Thus, the relationship of these two orders remains unresolved (Cafaro, 2005; Marshall et al., 2008). Environmental sequences were placed with the same clades as in previous analyses and *Psorospermium haeckelii* remained on a long, unsupported branch in both the small and large rDNA datasets (Marshall et al., 2008; Marshall and Berbee, 2011).

3.3 Ancestral State Reconstruction and BaTS Analyses

The environment where protist trichos (and close relatives) have been collected is significantly associated with phylogenetic clades, as illustrated by SIMMAP reconstructions (Fig. 3) and BaTS results (Table 3). The ancestral habitat of the family Eccrinidae is recovered as terrestrial but there has been at least one major shift from terrestrial to aquatic habitats (marine and freshwater). On the other hand, all Amoebidiidae are found in freshwater, so this is inferred as the ancestral habitat for this clade. The most likely ancestral habitat of the suborder Trichomycina is recovered as freshwater (ca. 75% probability). Similarly, there is a clear division between the marine environmental sequences that are closely related to *Creolimax* and *Sphaeroforma* and the freshwater environmental sequences that are closely related to *Anurofeca*. However, the ancestral state of the entire order Eccrinida was not resolved in our analyses.

Both types of ancestral analyses indicate that host is an important evolutionary factor that is significantly correlated with the phylogeny (Fig. 4, Table 3). There have been several host shifts within the family Eccrinidae, with at least one potential host shift from millipedes to insects and at least one host shift from insects to crustaceans. Despite these transitions, our analyses suggest that the ancestral host for the family Eccrinidae was a millipede. In contrast, our analyses suggest that the ancestral host for the suborder Trichomycina was likely an insect (Fig. 4). Ancestral state reconstructions

at deeper nodes become equivocal due to uncertainty and to missing data for environmental sequences.

The developmental stage of the host at the time of symbiotic association is significant for the states “juvenile” and “adult”, but not for “both” in the BaTS analysis (Fig. 5, Table 3). Only *Caullerya*, *Ichthyophonus*, *Leidyomyces*, *Psorospermium*, and the TMS have been reported to infect both the juvenile and adult stages of their hosts. All *Paramoebidium* species described to date are associated with immature aquatic insects, but *Amoebidium* species may associate with either juvenile or adult stages. Therefore, the ancestral state for the Trichomycina is equivocal (Fig. 5). The final host-related character, location within the host (Fig. 6), has four states that are significantly associated with the topology (Table 3): foregut, hindgut, external, and digestive tract. All *Paramoebidium* and Eccrinidae were collected from the hindgut of their hosts except for *Enteromyces callianassae* (which inhabits the foregut) and species of *Arundinula* (which may attach to either the foregut or hindgut). Not surprisingly, our ancestral state reconstruction suggests that the most recent common ancestor of the Trichomycina likely colonized the hindgut of its hosts.

The growth form (Fig. 7) and type of spore production (Fig. 8) are phylogenetically correlated (Table 3). Hyphal growth with a holdfast is the growth form that is found across all protist trichos, including *Amoebidium*. This state is recovered as the most likely ancestral condition for the Trichomycina, but spherical/ovoid is likely the ancestral condition for the entire order Eccrinida (Fig. 7). In addition to those two states, “spherical/plasmodial and hypha” is significantly correlated with the phylogeny because the monophyletic *Ichthyophonus* is the only genus with this growth form.

In contrast to many other traits, spore production across the protist trichos differs markedly among taxa, and all states are significantly correlated with the phylogeny (Fig. 8, Table 3). Basipetal spore production is reconstructed as the ancestral state for the

family Eccrinidae (ca. 99% probability), the ancestor of *Paramoebidium* and the Eccrinidae (ca. 80% probability), and the ancestor of *Ichthyophonus*, *Paramoebidium*, and the Eccrinidae (ca. 75% probability) (Fig. 8). However, the ancestral spore production for deeper nodes is equivocal.

3.4 Taxonomy of the protist trichos

Based on our phylogenetic results we propose a taxonomic reorganization of the protist trichos. We describe a new family, Paramoebidiidae fam. nov., to accommodate species of *Paramoebidium*, which have a unique combination of morphological and ecological features not found in any other lineage. We also rename the (fungal) family Parataeniellaceae to Parataeniellidae to reflect the inclusion of this group in the protists rather than in the fungi. Descriptions of the remaining families are in the following references: *Amoebidium* (Amoebidiidae) and Eccrinidae (Kirk et al., 2001; Lichtwardt et al., 2001), Ichthyophonidae (Cavalier-Smith, 2013).

Family: Paramoebidiidae fam. nov. N. K. Reynolds and M. M. White

Type genus: *Paramoebidium* Léger and Duboscq, 1929b

Diagnosis: Family Paramoebidiidae is monogeneric. Species of *Paramoebidium* produce hyphae attached to the hindgut cuticle (or anal gills) of freshwater larvae or nymphs of Insecta and form uninucleate, teardrop-shaped amoeboid cells that encyst and form elongate cystospores. Hyphae of Paramoebidiidae may be straight, curved, coiled, branched, or sigmoidal and are anchored to the host by secreted holdfast structures. This family is distinguished from Amoebidiidae by the absence of endospores and the predominant location within the digestive tract of the host rather than attachment to the exoskeleton. The Paramoebidiidae is differentiated from the Parataeniellidae because species of Parataeniellidae are only associated with terrestrial isopod hosts and lack the motile amoeboid phase of the life cycle.

Commentary: *Amoebidium* and *Paramoebidium* have traditionally been considered as sister taxa, but the genera are reciprocally monophyletic and their different life history characteristics support their division into two separate families.

Family: Parataeniellidae Manier and Lichtwardt, 1968

Type genus: *Parataeniella* Poisson, 1929

Description: Family Parataeniellidae includes species of *Parataeniella*. Members of Parataeniellidae are morphologically distinguished from Eccrinidae by hyphae that basipetally produce binucleate (occasionally uninucleate) endospores or where the entire thallus becomes one sporangium containing many uni- or binucleate endospores (Tuzet and Manier, 1950). Species in this group are found in the hindguts of terrestrial Isopoda (Crustacea).

Commentary: Sequences of *Parataeniella armadillidiia* and *P. dilatata* form a well-supported, monophyletic clade distinct from Eccrinidae. Four additional species in the genus *Parataeniella* are provisionally included in the family but require future molecular analyses to confirm their phylogenetic position. This phylogenetic placement of *P. armadillidiia* and *P. dilatata* in combination with the method of spore formation (basipetal and endospores) and host associations (terrestrial Isopoda) are sufficient to recognize this monophyletic clade as a separate family within the Trichomycina.

4. Discussion

4.1 The phylogeny of Trichomycina reveals strong host, habitat, and morphology trends

This research constitutes the first molecular phylogenetic analysis of the protist trichos within suborder Trichomycina as described by Cavalier-Smith (2013). In addition to >100 new rDNA sequences, we also obtained the first sequences of an independent nuclear protein-coding gene (HSP70) for 16 protist trichos. This work significantly increased the molecular sampling for this group and included a wider array of taxa from more diverse hosts and a broader geographic range. We also included several

previously unsampled genera (i.e. species of *Arundinula*, *Leidyomyces*, and *Parataeniella*). Previous analyses included only two representative sequences of Amoebidiidae and our expanded taxon sampling in this group determined that the family is polyphyletic.

Habitat and host type are tightly correlated in our phylogenetic reconstructions and we infer that these two factors have probably constrained the evolutionary opportunities of individual protist tricho lineages (Figs. 3 and 4). In other words, the combination of host and habitat has reduced opportunities to disperse and associate with new host organisms. This tight coupling of hosts and habitats suggests that the critical evolutionary events in the past have been colonization of new hosts or emergence into new habitats, followed by rapid evolutionary radiations to related host species in similar habitats. Based on our results we have reorganized the taxonomy of the Trichomycina to recognize five monophyletic families (Figs. 1, S1, and S2 and see above). In addition to our molecular phylogenetic data, each of these five families has clear morphological, habitat, and host associations that can be used to consistently differentiate them from one another.

4.2 Ichthyophonidae is nested within the protist trichos

The most surprising result was that the monogeneric family Ichthyophonidae (*Ichthyophonus*), an aquatic family of fish and amphibian parasites, is phylogenetically nested within the protist trichos. Although phylogenetic support is low and varies between methods and datasets, all of the best ML phylogenetic trees from the two- and three-gene datasets resolve the same topology (Figs. 1, S1, and S2). The MP phylogeny of the three-gene dataset was the only analysis to place Ichthyophonidae outside of the Trichomycina (Fig. S2). Although topology tests (Table 2) do not reject a sister relationship between Amoebidiidae and the clade that includes Paramoebidiidae +

Parataeniellidae + Ecrinidae, the best available evidence nonetheless suggests that Ichthyophonidae evolved from within the larger protist tricho lineage.

Regardless of whether Ichthyophonidae or Amoebidiidae was the earliest diverging lineage, we infer from the strong correlations between the character states and the phylogeny that the common ancestor of the Trichomycina underwent major evolutionary transitions along several phylogenetic branches. If our phylogenetic hypothesis is correct and Ichthyophonidae evolved from within the protist trichos, this would require an *Amoebidium*-like ancestor to transition from the outside to the inside of the host, from an invertebrate to a vertebrate host, and from a commensalistic to a parasitic relationship (Fig. 4).

Ancestral state reconstructions of habitat (Fig. 3) and growth form (Fig. 7) suggest that the ancestor of the Trichomycina likely lived in freshwater and had a hyphal growth form with a holdfast. Our best phylogenetic hypothesis would also require a secondary loss of the holdfast structure in the genus *Ichthyophonus* and would imply that this group evolved first in freshwater fishes and migrated secondarily to marine hosts. Some species of *Ichthyophonus* are known to infect anadromous fishes that move back and forth between freshwater and marine habitats (e.g., rainbow trout, *Oncorhynchus mykiss*), indicating a possible evolutionary pathway for transmission between habitats and hosts. Similarly, the host analysis (Fig. 4) suggests that an *Amoebidium*-like ancestor (i.e., most recent common ancestor of Ecrinidae + Parataeniellidae + Paramoebidiidae + Ichthyophonidae) may have had promiscuous, facultative arthropod host affiliations (also hypothesized by Lichtwardt, 1986). Upon repeated ingestion of spores and infested insect prey by fishes, the ancestral *Amoebidium*-like organism may have evolved the ability to invade deeper into animal tissues.

Although some details of the *Ichthyophonus* life cycle are uncertain, recent research (Kocan et al., 2013) suggests similarities in spore production to Amoebidiidae and Paramoebidiidae. Spore production characteristics were significantly correlated with the phylogeny in both our ASR (Fig. 8) and BaTS analyses (Table 3). Kocan et al. (2013) found evidence that an *Ichthyophonus* species that attacks sculpin and trout has infectious amoeba-like cells that circulate in the blood of its hosts. After migrating through the body, these cells settle in the tissue, enlarge, and form thick-walled resting structures called “schizonts”. The schizonts subsequently release amoeba-like cells into the stomach upon ingestion by a new host (ingestion of infected tissue is thought to be the main route of transmission of the parasite). However, a planktonic phase has also been hypothesized and there is some evidence that infective cells can be released from epidermal lesions on infected hosts (Kocan et al., 2010). Taken together this evidence suggests that the *Ichthyophonus* life cycle includes amoeba and cyst stages that are similar to those in the arthropod commensal members of Amoebidiidae and Paramoebidiidae. However, it is unclear whether spores (e.g. endospores or cystospores, as defined for the protist trichos) are formed at some point in the life cycle or not. It is possible that the cells released through host epidermal lesions are actually spores, but more research is needed to confirm this hypothesis (Kocan et al., 2010).

Resolving the relationship of Ichthyophonidae to the protist trichos will be critical in the future to understand the evolution of characters in Trichomycina. The analyses and datasets presented here are more robust than in any previous studies but we acknowledge that the strong signal in rDNA may be misleading. Previous studies within the Kingdom Fungi provide an illuminating example. The phylum Glomeromycota (arbuscular mycorrhizal fungi) was inferred as sister to the Dikarya (Ascomycota + Basidiomycota) based on rDNA analyses (Schüßler et al., 2001), but recent genome-scale analyses instead place Glomeromycota in a monophyletic lineage with other

“zygomycete” fungi (Mortierellomycotina and Mucoromycotina – Tisserant et al., 2013, Spatafora et al., 2016). This revised topology has led to the hypothesis that fungal-land plant associations are more ancient than previously thought (Spatafora et al., 2016). The case of the Glomeromycota provides a cautionary example of how more data are needed to confirm or reject our current phylogenetic hypotheses for the placement of Ichthyophonidae.

4.3 The protist tricho families in Trichomycina

4.3.1 Eccrinidae

Our ancestral state reconstructions demonstrate that transitions between habitats and host types are rare among protist tricho clades (Figs. 3 and 4). These characters (habitats and hosts) were previously considered of questionable importance in protist tricho systematics, particularly within the Eccrinidae (Glockling et al., 2013). For example, species of *Enterobryus* have been described from terrestrial millipedes, freshwater beetles, and marine crabs (i.e. *E. halophilus*) (Lichtwardt et al., 2001). However, *E. halophilus* from crabs is more closely related to other Eccrinidae from marine crustaceans and distantly related to morphologically similar *Enterobryus* species from millipedes. Likewise, an unidentified, *Enterobryus*-like sample from a freshwater beetle (1067) was phylogenetically nested among a beetle-associated clade (*Leidyomyces*) and a freshwater clade (*Arundinula*). This suggests that there may be a unique but previously unrecognized clade of freshwater beetle-associated protist trichos that have eluded detection. The fact that host and habitat are apparently critical evolutionary forces also suggests that morphology-based generic and species concepts will need to be critically reevaluated for all protist trichos.

Another character that was likely important for the diversification of the family Eccrinidae is the type of spore production (Fig. 8). Members of the Eccrinidae lack an amoeboid phase and release only walled spores. The loss of an amoeboid stage may

have reduced their dependence on water for dispersal and reproduction, thereby facilitating the colonization of terrestrial environments. Furthermore, the basipetal production of spores allows for continuous release of propagules over time. This is in contrast to Amoebidiidae and Paramoebidiidae, which rupture their hyphae and simultaneously release all of their propagules at once. Several genera of Eccrinidae (e.g., *Arundinula*, *Enterobryus*, *Palavascia*) are known to produce different spore forms that can vary in size, shape, and number of nuclei, which has confounded morphological species concepts (Cafaro, 2000; Hibbits Galt, 1978; Lichtwardt, 1954; Lichtwardt et al., 2001). Further detailed analyses and evaluation of these features in a phylogenetic context could provide useful morphotaxonomic characters for future species identifications and descriptions.

Once established in the gut environment, protist trichos have been considered commensal with their hosts, yet few experimental or observational studies have tested this hypothesis. One member of the family Eccrinidae, *Enteromyces callianassae*, exclusively inhabits the foregut of its host and this is also the only taxon whose host interactions have been studied in detail. After comparing infected and uninfected shrimp populations, Kimura et al. (2002) suggested that *E. callianassae* may benefit its host via supplementation of digestive enzymes. However, Lichtwardt (1958) observed unusual, cyst-like spores of *Enterobryus borariae* outside of the gut lining of the millipede host. He considered this a parasitic interaction because *E. borariae* penetrated through the lining of the gut as it germinated into the gut interior.

Our analyses determined that the location in the host (foregut vs. hindgut) was significantly associated with the phylogeny (Fig. 6, Table 3). This suggests that functional and physiological constraints of the digestive tract impact the type of interaction these symbionts have with their hosts. Additional studies that compare

infected and uninfected host populations or employ experimental methods in the lab could clarify the nature of these associations.

4.3.2 Paramoebidiidae

Amoebidium and *Paramoebidium* are reciprocally monophyletic, and their different life history characteristics support their division into two separate families. *Amoebidium* species are the only taxa within the suborder Trichomycina known to attach to the exterior of their hosts (Fig. 6) and all of the described species have well defined cystospore and endospore phases in their life cycles (Figs. 1 and 8). Furthermore, *Paramoebidium* species have only been recovered from immature forms of their hosts, whereas *Amoebidium* species may attach to either juvenile or adult forms (Fig. 5). Limited ultrastructural analyses have been conducted on these genera, but Dang and Lichtwardt (1979) suggest that the cystospores of the two families are probably structurally different at a microscopic level. However, a sister relationship between *Amoebidium* and *Paramoebidium* could not be rejected by topology tests (Table 2) and the two genera do share some features in common. First, both are restricted to freshwater habitats (Fig. 3) and have some overlap in host associations (*Amoebidium* is commonly associated with midge larvae (Chironomidae) whereas *Paramoebidium* are only rarely found with these hosts (Lichtwardt et al., 2001)). Second, both genera form similarly sized and teardrop-shaped dispersal amoebae and commonly encyst on the shed exoskeletons of their hosts (Lichtwardt et al., 2001).

4.3.3 Parataeniellidae

The original description of the family Parataeniellaceae within Kingdom Fungi (Manier and Lichtwardt, 1968) included a second genus, *Lajasiella*. The genus *Lajasiella* was circumscribed based on uninucleate spores and hyphae of a similar size and shape to *Parataeniella* (Manier and Lichtwardt, 1968). No DNA samples have been obtained for *Lajasiella*, so the phylogenetic placement of the genus remains unknown.

Based on our ancestral state reconstruction analyses, however, we believe that *Lajasiella* is probably not closely related to other members of Parataeniellidae. First, the strong correlation of host associations in the phylogeny (Fig. 4, Table 3) implies that *Lajasiella* is probably more closely related to other terrestrial, beetle-associated taxa (e.g. *Leidyomyces* in the family Eccrinidae) than to the isopod-associated *Parataeniella*. Second, *Lajasiella* apparently only produces endospores, rather than both basipetal and endospore production as is found in other members of Parataeniellidae (Tuzet and Manier, 1950) (Fig. 1). Spore production types are significantly associated with the phylogeny so these differences are likely important in the phylogenetic placement (Fig. 8, Table 3).

Topology tests (Table 2) reject Parataeniellidae as sister to Paramoebidiidae, but do not reject it as sister to the Eccrinidae, thereby reinforcing the topology recovered in the best ML trees (Fig. 2). The method of spore production in species of Parataeniellidae could represent a transitional form between Paramoebidiidae and the Eccrinidae. Its placement as the sister clade to Eccrinidae suggests that the ancestral *Parataeniella* may have been among the first protist trichos to have lost the amoeboid phase of the life cycle.

5. Conclusion

Recent studies on the protist trichos have reported a number of morphologically or geographically unique collections (Hernández Roa et al., 2009; Hernández Roa and Cafaro, 2012; Valle and Santamaria, 2009; William and Strongman, 2013) and have described interesting new species (Bench and White, 2013; Contreras and Cafaro, 2013; Strongman and White, 2006; Strongman and White, 2008; Valle, 2014a; Valle, 2014b; White et al., 2006). However, there have been only limited phylogenetic analyses and morphological comparisons of this group since they were excluded from the Kingdom Fungi (Benny and O'Donnell, 2000; Cafaro, 2005; Ustinova et al., 2000).

Significant interspecific and intraspecific variability of the morphological characters traditionally used to describe species warrants reexamination of other traits such as spore types and ontogeny, propagule appendages, host specificity, and hyphal polymorphism. In other words, a shift is needed from a morphological species concept to an integrative or even strictly phylogenetic concept unless clear homologous and easily measurable traits can be found to accurately diagnose species. It is possible that different species concepts may be appropriate for different clades. For example, Marshall and Berbee (2013) found that morphology was not sufficient to distinguish phylogenetically distinct isolates of *Sphaeroforma*. They suggested that the genealogical concordance phylogenetic species recognition method (Baum and Shaw, 1995) was most appropriate to delimit species. The inability to culture most species in *Trichomycina* makes them challenging to study and observations are necessarily restricted to the moments following host dissection. This emphasizes the need for renewed cultivation efforts in order to study ultrastructure, sequence genomes, and characterize these organisms more completely.

Our results suggest that undersampling leads to significant underestimation of protist tricho diversity, potentially fragmenting and distorting the evolutionary signal and obscuring evolutionary patterns. Despite being discovered >150 years ago (Leidy, 1849), basic life history features such as dispersal mechanisms, development of hyphae, and spore formation are still only well understood and documented for a handful of protist tricho species. Our success using protist tricho-specific primers to exclude host animal DNA illustrates that molecular phylogenetic tools are critical for building taxonomic infrastructure and elucidating evolutionary patterns. Fortunately, with the decreasing cost and relative ease of DNA sequencing and the expanding suite of analytical tools available, progress in our understanding of these enigmatic microorganisms may advance rapidly. Future efforts to sequence many independent

loci from across the genomes of diverse species of Trichomycina will be critical for testing our phylogenetic hypotheses and resolving deeper evolutionary questions. In addition, we expect that future sampling of more protist tricho taxa from diverse hosts and geographical locations will likely to lead to the discovery of significant hidden diversity and undocumented evolutionary lineages in this understudied group of organisms.

Acknowledgements

The authors acknowledge financial support from NSF REVSYS Awards DEB-9521811, -0108110, -0344722, and DEB-0918182 (to MMW, either as a trainee, participant or PI), graduate research awards to NR from Sigma Xi and the Society of Systematic Biologists provided financial support, a Boise State Student Research Program (SRP) Award (to DH), and an INBRE Summer Research Fellowship (to JG). Additional funding was provided by NSF grant DEB-1441677 (to MES) and the Institute of Food and Agricultural Sciences (IFAS) at the University of Florida. Host identifications were provided by P. H. Adler, R. W. Holzenthal, B. C. Kondratieff, P. Zwick and D. Hennen. We thank Jordi Paps for advice with PCR primers and protocols. S. LaPatra graciously provided heart tissues explants of *Ichthyophonus*. The support of J. Niece and staff at the University of Wisconsin Biotechnology Center is gratefully appreciated. Many collaborators provided samples: A. C. Bailey, C. L. Frey, P. Kandel, H. Nahrung, T. Pickell, A. Siri, D. Strongman, L. G. Valle, and E. R. Wilson. Dr. Robert W. Lichtwardt provided samples but also many years of guidance, interest, and hard work.

References

Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., De Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., Grosdidier, A., Hernandez, C., Ioannidis, V., Kuznetsov, D., Liechti, R., Moretti, S., Mostaguir, K., Redaschi, N., Rossier, G.,

- 696 Xenarios, I., Stockinger, H., 2012. ExPASy: SIB bioinformatics resource portal. *Nucleic*
697 *Acids Res.* 40, 597–603.
- 698
- 699 Baum, D.A., Shaw, K.L., 1995. Genealogical Perspectives on the Species Problem. In P.
700 C. Hoch , A. G. Stephensen, eds. *Experimental and Molecular Approaches to Plant*
701 *Biosystematics*. St. Louis: Missouri Botanical Garden, pp. 289–303.
- 702 Bench, M.E., White, M.M., 2012. New species and first records of trichomycetes from
703 immature aquatic insects in Idaho. *Mycologia* 104, 295–312.
- 704
- 705 Benny, G.L., O'Donnell, K., 2000. *Amoebidium parasiticum* is a protozoan, not a
706 Trichomycete. *Mycologia* 92, 1133–1137.
- 707
- 708 Bollback, J.P., 2009. SIMMAP Software. Available at:
709 <http://www.simmap.com/index.html> [Accessed January 1, 2010].
- 710
- 711 Bollback, J.P., 2006. SIMMAP: stochastic character mapping of discrete traits on
712 phylogenies. *BMC Bioinformatics* 7, 88.
- 713
- 714 Brown, W.M. Prager, E.M., Wang, A., Wilson, A.C., 1982. Mitochondrial DNA Sequences
715 of Primates: Tempo and Mode of Evolution. *J. Mol. Evol.* 18, 225–239.
- 716
- 717 Cafaro, M.J., 2000. Gut fungi of isopods: the genus *Palavascia*. *Mycologia*. 92 (2), 361-
718 369.
- 719

Cafaro, M.J., 2005. Eccrinales (Trichomycetes) are not fungi, but a clade of protists at the early divergence of animals and fungi. *Mol. Phylogenet. Evol.* 35, 21–34.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552.

Cavalier-Smith, T. Neomonada and the origin of animal and fungi. 1998. In Coombs GH, Vickerman K, Sleigh MA, Warren A, eds. *Evolutionary Relationships Among Protozoa*, Vol. 56. Dordrecht, The Netherlands: Kluwer, pp. 375–407.

Cavalier-Smith, T., Chao, E.E., Snell, E.A., Berney, C., Fiore-Donno, A.M., 2013. Early evolution of eukaryote feeding modes, cell structural diversity, and classification of the protozoan phyla Loukozoa, Sulcozoa, and Choanozoa. *Eur. J. Protistol.* 49, 115–178.

Contreras, K., Cafaro, M.J., 2013. Morphometric studies in *Enterobryus luteovirgatus* sp. nov. (Ichthyosporea: Eccrinales) associated with yellow-banded millipedes in Puerto Rico. *Acta Protozool.* 52, 291–297.

Dang, S.-N., Lichtwardt, R.W., 1979. Fine Structure of *Paramoebidium* (Trichomycetes) and a new species with virus-like particles. *Am. J. Bot.* 66, 1093–1104.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2 : more models, new heuristics and parallel computing CircadiOmics: integrating circadian genomics, transcriptomics, proteomics. *Nature Methods* 9, 772.

Edgar, R.C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.

Fitch, W.M., 1971. Toward defining the course of evolution: minimum change for a specific tree topology. *Syst. Zool.* 20, 406–416.

Glockling, S.L., Marshall, W.L., Gleason, F.H., 2013. Phylogenetic interpretations and ecological potentials of the Mesomycetozoea (Ichthyosporea). *Fungal Ecol.* 6, 237–247.

Goloboff, P.A., Farris, S., Nixon, K., 2008. TNT, a free program for phylogenetic analysis. *Cladistics* 24, 774–786.

Gozlan, R.E., Marshall, W.L., Lilje, O., Jessop, C.N., Gleason, F.H., Andreou, D., 2014. Current ecological understanding of fungal-like pathogens of fish: What lies beneath? *Front. Microbiol.* 5, 1–16.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704.

Hernández Roa, J.J., Cafaro, M.J., 2012. Seasonality and prevalence of the protistan trichomycete *Enterobryus halophilus* (Ichthyosporea: Eccrinales) in the mole crab *Emerita portoricensis*. *Mycologia* 104, 337–344.

Hernández Roa, J.J., Virella, C.R., Cafaro, M.J., 2009. First survey of arthropod gut fungi and associates from Vieques, Puerto Rico. *Mycologia* 101, 896–903.

- 771 Hibbits Galt, J., 1978. Marine Eccrinales (Trichomycetes) found in crustaceans of the
 772 San Juan Archipelago, Washington. Syesis 11, 213–261.
 773
- 774 Kimura, H., Harada, K., Hara, K., Tamaki, A., 2002. Enzymatic approach to fungal
 775 association with arthropod guts: A case study for the crustacean host, *Nihonotrypaea*
 776 *harmandi*, and its foregut fungus, *Enteromyces callianassae*. Marine Ecol. 23, 157-183.
 777
- 778 Kirk, P.M., Cannon, P.F., David, J. C., Stalpers, J.A. 2001. Dictionary of the Fungi 9th
 779 Edition. CAB International, Wallingford, 655 pp.
 780
- 781 Kocan, R., LaPatra, S., Hershberger, P., 2013. Evidence for an amoeba-like infectious
 782 stage of *Ichthyophonus* sp. and description of a circulating blood stage: a probable
 783 mechanism for dispersal within the fish host. J. Parasitol. 99, 235–240.
 784
- 785 Kocan, R.M., Gregg, J.L., Hershberger, P.K., 2010. Release of infectious cells from
 786 epidermal ulcers in *Ichthyophonus* sp.-infected Pacific herring (*Clupea pallasii*): evidence
 787 for multiple mechanisms of transmission. J. Parasitol. 96, 348–352.
 788
- 789 Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: Combined
 790 selection of partitioning schemes and substitution models for phylogenetic analyses.
 791 Mol. Biol. Evol. 29, 1695–1701.
 792
- 793 Léger, L., Duboscq, O. 1929a. *Eccrinoïdes henneguyi* n. g. n. sp. et la systématique des
 794 Eccrinides. Archives d' Anatomie Microscopique 25, 309-324.
 795

- 796 Léger, L., Duboscq, O. 1929b. L'évolution des *Paramoebidium*, nouveau genre
 797 d'Ecclinides, parasite des larves aquatiques d'Insectes. Comptes Rendus
 798 Hebdomadaires des Séances de l'Académie des Sciences Paris 189, 75-77.
 799
- 800 Leidy, J., 1849. *Enterobryus*, a new genus of Confervaceae. P. Acad. Nat. Sci. Phila. 4,
 801 225–233.
 802
- 803 Lichtwardt, R.W., 1954. Three species of Eccliniales inhabiting the hindguts of millipeds,
 804 with comments on the Ecclinids as a group. Mycologia 46, 564–585.
 805
- 806 Lichtwardt, R.W., 1958. An *Enterobryus* from the milliped *Boraria carolina* (Chamberlin).
 807 Mycologia 50, 550-561.
 808
- 809 Lichtwardt, R. W. 1986. The Trichomycetes: Fungal Associates of Arthropods. Springer-
 810 Verlag, New York. 343 pp.
 811
- 812 Lichtwardt, R.W., Cafaro, M.J., White, M.M., 2001. The Trichomycetes: Fungal
 813 associates of arthropods. Available at: www.nhm.ku.edu/~fungi [Accessed January 1,
 814 2015].
 815
- 816 Lichtwardt, R. W., Manier, J.-F. 1978. Validation of the Harpellales and
 817 Asellariales. Mycotaxon 7, 441-442.
 818
- 819 Lohr, J.N., Laforsch, C., Koerner, H., Wolinska, J., 2010. A daphnia parasite (*Caullerya*
 820 *mesnili*) constitutes a new member of the Ichthyosporea, a group of protists near the
 821 animal-fungi divergence. J. Eukaryot. Microbiol. 57, 328–336.

- 822 Lord, J.C., Hartzer, K.L., Kambhampati, S., 2012. A nuptially transmitted ichthyosporean
823 symbiont of *Tenebrio molitor* (Coleoptera: Tenebrionidae). J. Eukaryot. Microbiol. 59,
824 246–250.
- 825
- 826 Maddison, W.P., Maddison, D.R., 2014. Mesquite: a modular system for evolutionary
827 analysis. Available at: <http://mesquiteproject.org>.
- 828
- 829 Manier, J.F., 1964. Endophytes parasites d'Arthropodes cavernicoles récoltés dans des
830 grottes de l'Ariège et de la Haute-Garonne. Annales de Spéléologie 19, 803–812.
- 831
- 832 Manier, J.-F., Lichtwardt, R. W. 1968. Révision de la systématique des
833 Trichomycètes. Annales des Sciences Naturelles Botanique, Paris, Série 12, 9: 519-532.
- 834
- 835 Marshall, W.L., Celio, G., McLaughlin, D.J., Berbee, M.L., 2008. Multiple isolations of a
836 culturable, motile ichthyosporean (Mesomycetozoa, Opisthokonta), *Creolimax*
837 *fragrantissima* n. gen., n. sp., from marine invertebrate digestive tracts. Protist 159, 415–
838 433.
- 839
- 840 Marshall, W.L., Berbee, M.L., 2010. Population-level analyses indirectly reveal cryptic
841 sex and life history traits of *Pseudoperkinsus tapetis* (Ichthyosporea, Opisthokonta): A
842 unicellular relative of the animals. Mol. Biol. Evol. 27, 2014–2026.
- 843
- 844 Marshall, W.L., Berbee, M.L., 2011. Facing Unknowns: Living cultures (*Pirum gemmata*
845 gen. nov., sp. nov., and *Abeoforma whisleri*, gen. nov., sp. nov.) from invertebrate
846 digestive tracts represent an undescribed clade within the unicellular opisthokont lineage
847 Ichthyosporea (Mesomycetozoea). Protist 162, 33–57.

- 848 Marshall, W.L., Berbee, M.L., 2013. Comparative morphology and genealogical
849 delimitation of cryptic species of sympatric isolates of *Sphaeroforma* (Ichthyosporea,
850 Opisthokonta). *Protist* 164, 287–311.
851
- 852 Mendoza, L., Ajello, L., Taylor, J.W., 2001. The taxonomic status of *Lacazia loboi* and
853 *Rhinosporidium seeberi* has been finally resolved with the use of molecular tools.
854 *Revista iberoamericana de micologia : organo de la Asociacion Espanola de*
855 *Especialistas en Micologia* 18, 95–98.
856
- 857 Mendoza, L., Taylor, J.W., Ajello, L., 2002. The class Mesomycetozoea: a
858 heterogeneous group of microorganisms at the animal-fungal boundary. *Annu. Rev.*
859 *Microbiol.* 56, 315–344.
860
- 861 Millanes, A.M., Diederich, P., Ekman, S., Wedin, M., 2011. Phylogeny and character
862 evolution in the jelly fungi (Tremellomycetes, Basidiomycota, Fungi). *Mol. Phylogenet.*
863 *Evol.* 61, 12–28.
864
- 865 Moss, S.T., 1979. Commensalism of the Trichomycetes. *Insect-fungus symbiosis:*
866 *nutrition, mutualism, and commensalism.* Montclair: Allanheld, Osmun & Co. pp.175-
867 227.
868
- 869 Moss, S.T., 1999. *Astreptonema gammari*: An eccrinid with appendaged spores. *Kew*
870 *Bulletin.* 54, 637-650.
871
- 872 Okamoto, N., Nakase, K., Suzuki, H., Nakai, Y., Fujii, K., Sano, T., 1985. Life History
873 and Morphology of *Ichthyophonus hoferi* in vitro. *Fish Pathol.* 20, 273–285.

- 874 Paps, J., Medina-Chacón, L.A., Marshall, W., Suga, H., Ruiz-Trillo, I., 2013. Molecular
875 phylogeny of Unikonts: New insights into the position of Apusomonads and
876 Ancyromonads and the internal relationships of Opisthokonts. *Protist* 164, 2–12.
877
- 878 Parker, J., Rambaut, A., Pybus, O.G., 2008. Correlating viral phenotypes with
879 phylogeny: Accounting for phylogenetic uncertainty. *Infection, Genetics Evol.* 8, 239–
880 246.
881
- 882 Poisson, R. 1929. Recherches sur quelques Eccrinides parasites de Crustacés
883 Amphipodes et Isopodes. *Archives de Zoologie Expérimentale et Générale* 69, 179-216.
884
- 885 R Foundation for Statistical Computing, 2015. Available at: www.r-project.org [Accessed
886 March 2, 2015].
887
- 888 Rambaut, A., 2014. FigTree. Available at: <http://tree.bio.ed.ac.uk/software/figtree/>.
889
- 890 Rambaut, A. , Drummond, A.J., 2009. Tracer 1.5.0. Available at:
891 <http://beast.bio.ed.ac.uk/>.
892
- 893 Reeves, W.K., Jensen, J.B., Ozier, J.C., 2000. New faunal and fungal records from
894 caves in Georgia, USA. *Journal Cave Karst Stud.* 62, 169–179.
895
- 896 Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S.,
897 Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient
898 bayesian phylogenetic inference and model choice across a large model space. *Syst.*
899 *Biol.* 61, 539–542.

- 900 Ronquist, F., Huelsenbeck, J., Teslenko, M., 2011. Draft MrBayes Version 3.2 Manual:
901 Tutorials and Model Summaries, pp.1-103.
902
- 903 Rowley, J.J.L., Gleason, F.H., Andreou, D., Marshall, W.L., Lilje, O., Gozlan, R., 2013.
904 Impacts of mesomycetozoean parasites on amphibian and freshwater fish populations.
905 Fungal Biol. Rev. 27, 100–111.
906
- 907 Schmidt, H.A., Strimmer, K., Vingron, M., von Haeseler, A., 2002. TREE-PUZZLE:
908 maximum likelihood phylogenetic analysis using quartets and parallel computing.
909 Bioinformatics 18, 502–504.
910
- 911 Schmidt, H.A., 2009. Testing tree topologies. In P. Lemey, M. Salemi, A. M. Vandamme,
912 eds. The Phylogenetic Handbook: a practical approach to phylogenetic analysis and
913 hypothesis testing. Cambridge: Cambridge University Press, pp. 381–404.
914
- 915 Schüßler, A., Schwarzott, D., Walker, C. 2001. A new fungal phylum, the
916 Glomeromycota: phylogeny and evolution. Mycol Res 105:1413–1421.
917
- 918 Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection.
919 Syst. Biol. 51, 492–508.
920
- 921 Shimodaira, H., Hasegawa, M., 1999. Multiple Comparisons of Log-Likelihoods with
922 Applications to Phylogenetic Inference. Mol. Biol. Evol. 16, 1114–1116.
923
- 924 Shimodaira, H., Hasegawa, M., 2001. CONSEL: for assessing the confidence of
925 phylogenetic tree selection. Bioinformatics 17, 1246–1247.

- 926 Spatafora, J.W., Chang, Y., Benny, G.L., Lazarus, K., Smith, M.E., Berbee, M.L., Bonito,
927 G., Corradi, N., Grigoriev, I.V., Gryganskyi, A., James, T.Y., O'Donnell, K., Roberson,
928 R.W., Taylor, T.N., Uehling, J., Vilgalys, R., White, M.M., Stajich, J.E. 2016. A phylum-
929 level phylogenetic classification of zygomycete fungi based on genome-scale data.
930 *Mycologia* 108, 1028-1046.
- 931
- 932 Stamatakis, A., 2014. RAxML version 8: A tool for phylogenetic analysis and post-
933 analysis of large phylogenies. *Bioinformatics* 30, 1312–1313.
- 934
- 935 Stöver, B.C., Müller, K.F., 2010. TreeGraph 2: combining and visualizing evidence from
936 different phylogenetic analyses. *BMC Bioinformatics* 11, 7.
- 937
- 938 Strongman, D.B., White, M.M., 2006. New species of *Lancisporomyces*, *Orphella*, and
939 *Paramoebidium*, endosymbionts of stonefly nymphs from streams in Nova Scotia,
940 Canada. *Can. J. Bot.* 84, 1478–1495.
- 941
- 942 Strongman, D.B., White, M.M., 2008. Trichomycetes from lentic and lotic aquatic
943 habitats in Ontario, Canada. *Botany* 86, 1449–1466.
- 944
- 945 Tautz, D., Renz, M., 1983. An optimized Freeze-Squeeze method for the recovery of
946 DNA fragments from agarose gels. *Anal. Biochem.* 132, 14–19.
- 947
- 948 Tisserant, E., Malbreil, M., Kuo, A., Kohler, A., Symeonidi, A., Balestrini, R., Charron, P.,
949 Duensing, N., Frey, N.F.D., Gianinazzi-Pearson, V., Gilbert, L.B., Handa, Y., Herr, J.R.,
950 Hijri, M., Koul, R., Kawaguchi, M., Krajinski, F., Lammers, P.J., Masclaux, F.G., Murat,
951 C., Morin, E., Ndikumana, S., Pagni, M., Petitpierre, D., Requena, N., Rosikiewicz, P.,

952 Riley, R., Saito, K., Clemente, H.S., Shapiro, H., Van Tuinen, D., Bécard, G., Bonfante,
 953 P., Paszkowski, U., Shachar-Hill, Y.Y., Tuskan, G.A., Young, P.W., Sanders, I.R.,
 954 Henrissat, B., Rensing, S.A., Grigoriev, I.V., Corradi, N., Roux, C., and Martin, F., 2013.
 955 Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant
 956 symbiosis. *P Natl Acad Sci USA* 110, 20117–20122.
 957
 958 Trotter, M.J., Whisler, H.C., 1965. Chemical composition of the cell wall of *Amoebidium*
 959 *parasiticum*. *Can. J. Bot.* 43, 869–876.
 960
 961 Tuzet, O., Manier, J.-F. 1950. *Lajassiella aphodii* n.g., n.sp. Palavascide parasite d'une
 962 larve d'*Aphodius* (Coléoptère Scarabaeidae). *Annales des Sciences Naturelles*
 963 *Zoologie, Série* 11, 12: 465-470.
 964
 965 Ustinova, I., Krienitz, L., Huss, V.A., 2000. *Hyaloraphidium curvatum* is not a green alga,
 966 but a lower fungus; *Amoebidium parasiticum* is not a fungus, but a member of the
 967 DRIPs. *Protist* 151, 253–262.
 968
 969 Valle, L.G., 2014a. New species of *Paramoebidium* (trichomycetes, Mesomycetozoea)
 970 from the Mediterranean, with comments about the amoeboid cells in Amoebidiales.
 971 *Mycologia* 106, 481–490.
 972
 973 Valle, L.G., 2014b. Validation of the trichomycete *Paramoebidium chattoni*
 974 (Amoebidiales, Mesomycetozoea), a common and cosmopolitan black fly endosymbiont.
 975 *Mycologia* 106, 573–579.
 976

- 977 Valle, L.G., Santamaria, S., 2009. Arthropod gut symbionts from the Balearic Islands:
978 Majorca and Cabrera. Diversity and biogeography. *Anales del Jardín Botánico de Madrid*
979 66, 109–120.
980
- 981 VanDover, C.L., Lichtwardt, R.W., 1986. A new Trichomycete commensal with a
982 Galatheid squat lobster from deep-sea hydrothermal vents. *Biological Bulletin* 171, 461–
983 468.
984
- 985 Vogt, G., Rug, M., 1999. Life stages and tentative life cycle of *Psorospermium haeckeli*,
986 a species of the novel DRIPs clade from the animal-fungal dichotomy. *J. Exp. Zool.* 283,
987 31–42.
988
- 989 Wang, T.H., Donaldson, Y.K., Brettell, R.P., Bell, J.E., 2001. Identification of shared
990 populations of Human Immunodeficiency Virus Type 1 infecting microglia and tissue
991 macrophages outside the central nervous system. *J. Virol.* 75, 11686–11699.
992
- 993 Wang, Y. Tretter, E.D., Johnson, E.M., Kandel, P., Lichtwardt, R.W., Novak, S.J., Smith,
994 J.F., White, M.M., 2014. Using a five-gene phylogeny to test morphology-based
995 hypotheses of *Smittium* and allies, endosymbiotic gut fungi (Harpellales) associated with
996 arthropods. *Mol. Phylogenet. Evol.* 79, 23–41.
997
- 998 Whisler, H.C., 1960. Pure culture of the Trichomycete, *Amoebidium parasiticum*. *Nature*
999 186, 732-733.
1000
- 1001 Whisler, H.C., 1965. Host-integrated development in the Amoebidiales. *J. Protozool.* 13,
1002 183–188.

- 1003 Whisler, H.C., 1963. Observations on some new and unusual enterophilous
1004 phycomycetes. *Can. J. Bot.* 41, 887–900.
1005
- 1006 White, M.M., 2006. Evolutionary implications of a rRNA-based phylogeny of Harpellales.
1007 *Mycol. Res.* 110, 1011–1024.
1008
- 1009 White, M.M., Cafaro, M.J., Gottlieb, A.M., 2001. Taxonomy and Systematics of
1010 Trichomycetes - Past, Present, and Future. In Misra, J. K., Horn, B. W., eds.
1011 Trichomycetes and other fungal groups. Enfield, NH: Science, pp. 27–37.
1012
- 1013 White, M.M., Cafaro, M.J., Lichtwardt, R.W., 2000. Arthropod gut fungi from Puerto Rico
1014 and summary of tropical Trichomycetes worldwide. *Caribbean J. Science* 36, 210–220.
1015
- 1016 White, M.M., Siri, A., Lichtwardt, R.W., 2006. Trichomycete insect symbionts in Great
1017 Smoky Mountains National Park and vicinity. *Mycologia* 98, 333–352.
1018
- 1019 William, R., Strongman, D., 2013. Trichomycetes from Governor's Lake and Lake
1020 Micmac within the Halifax Regional Municipality, Nova Scotia, Canada. *Botany* 91, 382–
1021 402.
- 1022 Xia, X., 2013. DAMBE5: A comprehensive software package for data analysis. *Mol. Biol.*
1023 *Evol.* 30, 1720–1728.

1024 **Figures**

1025

FIG. 1. Generalized, illustrative drawings of typical growth form, life cycle, and spore production types of the Trichomycina families. Undifferentiated cytoplasm is indicated by grey shading. The holdfast structure and different spore types are labeled.

FIG. 2. Bayesian (BI) consensus tree based on the small combined 18S and 28S rDNA dataset ("small rDNA dataset") of the order Eccrinida. Newly recognized protist tricho families are shown in grey boxes. *Capsaspora owczarzaki* was used as the outgroup. In contrast to the large rDNA dataset depicted in Supplemental Figure 2, this analysis excluded members of order Dermocystida, *Ministeria vibrans*, choanoflagellate, and sponge taxa. Statistical support values are indicated with symbols. Branches supported by all three methods are shown in bold and host organisms are shown in parentheses.

FIG. 3. Ancestral state reconstruction of habitat for the protist trichos and related taxa overlaid on the Bayesian consensus tree of the small rDNA dataset. Pie charts at each node indicate the probability for each state. Character state coding for each sample is indicated by a colored dot after the sample name. Asterisks indicate character states that were significantly associated with the phylogeny in the BaTS analysis (Table 3).

FIG. 4. Ancestral state reconstruction of host type for the order Eccrinida overlaid on the Bayesian consensus tree of the small rDNA dataset. Pie charts at each node indicate the probability for each state. Character state coding for each sample is indicated by a colored dot after the sample name. Question marks indicate missing data. Asterisks indicate character states that were significantly associated with the phylogeny in the BaTS analysis (Table 3).

FIG. 5. Ancestral state reconstruction of host stage for the order Eccrinida overlaid on the Bayesian consensus tree of the small rDNA dataset. Pie charts at each node indicate the probability for each state. Character state coding for each sample is indicated by a colored dot after the sample name. Question marks indicate missing data. Asterisks indicate character states that were significantly associated with the phylogeny in the BaTS analysis (Table 3).

FIG. 6. Ancestral state reconstruction of location within the host for the order Eccrinida overlaid on the Bayesian consensus tree of the small rDNA dataset. Pie charts at each node indicate the probability for each state. Character state coding for each sample is indicated by a colored dot after the sample name. Question marks indicate missing data. Asterisks indicate character states that were significantly associated with the phylogeny in the BaTS analysis (Table 3).

FIG. 7. Ancestral state reconstruction of growth form for the order Eccrinida overlaid on the Bayesian consensus tree of the small rDNA dataset. Pie charts at each node indicate the probability for each state. Character state coding for each sample is indicated by a colored dot after the sample name. Question marks indicate missing data. Asterisks indicate character states that were significantly associated with the phylogeny in the BaTS analysis (Table 3).

FIG. 8. Ancestral state reconstruction of spore production for the order Eccrinida overlaid on the Bayesian consensus tree of the small rDNA dataset. Pie charts at each node indicate the probability for each state. Character state coding for each sample is indicated by a colored dot after the sample name. Question marks indicate missing data.

Asterisks indicate character states that were significantly associated with the phylogeny in the BaTS analysis (Table 3).

Tables

TABLE 1. Characters and character states used for ancestral state reconstruction and correlation analyses using the SIMMAP and BaTS software packages. Explanations of the biological and ecological importance of the characters and character states are provided in the far right column.

TABLE 2. Shimodaira-Hasegawa (SH) and Approximately Unbiased (AU) topology test results. Constraint topologies are listed for each hypothesis and were compared to the best Maximum Likelihood tree. Bold values indicate cases where the recovered topology of the Maximum Likelihood tree was inferred as significantly better than the constraint tree ($p < 0.05$).

TABLE 3. Bayesian Tip-association Significance (BaTS) phylogeny-trait correlation testing results. Association Index (AI), Parsimony Score (PS), and maximum exclusive single-state clade size (MC) statistics are shown. Character state coding was the same as for the ancestral state reconstruction analyses. Bold MC values indicate character states that are significantly correlated with the phylogeny ($p < 0.05$). For both the AI and PS, lower values indicate stronger phylogeny-trait associations. Character states that are only represented once in the dataset are not shown.

Figure1

Eccrinidae

Parataeniellidae

Paramoebidiidae

Ichthyophonidae

Amoebidiidae

Figure2

- Marine *
- Freshwater *
- Terrestrial *

Eccrinidae

Parataeniellidae

Paramoebidiidae

Ichthyophonidae

Amoebidiidae

- Vertebrate *
- Crustacean *
- Insect *
- Millipede *
- Other invertebrate

Eccrinidae

Parataeniellidae

Paramoebidiidae

Ichthyophonidae

Amoebidiidae

Caulerya mesnii CAUL BRN01 (Cladocera, Daphniidae)

○ Juvenile *
● Adult *
● Both

Eccrinidae

Parataeniellidae

Paramoebidiidae

Ichthyophonidae

Amoebidiidae

Caullerya mesnii CAUL BRN01 (Cladocera, Daphniidae)

- Foregut *
- Hindgut *
- Organs/tissues
- External *
- Digestive tract *
- Haemolymph
- Organs/tissues & digestive tract

Eccrinidae

Parataeniellidae

Paramoebidiidae

Ichthyophonidae

Amoebidiidae

○ Spherical/ovoid *

● Hypha with holdfast *

● Spherical/plasmodial

● Amoeboid

● Spherical/plasmodial

& hypha *

- Endospores *
- Cystospores *
- Basipetal *
- Endospores & cystospores *
- Endospores & basipetal *
- Amoeboid *

Eccrinidae

Parataeniellidae

Paramoebidiidae

Ichthyophonidae

Amoebidiidae

Caullerya mesnili CAUL BRN01 (Cladocera, Daphniidae) ○

Tables

Table1

Character	States	Explanation
Habitat	Marine Freshwater Terrestrial	Protist trichos and their relatives have been collected from a wide variety of habitats, including open ocean, intertidal, lentic and lotic freshwater habitats, and a range of terrestrial ecosystems.
Host type	Vertebrate Crustacean Insect Millipede Other invertebrate	The “Other invertebrate” state refers to organisms such as bivalves, tunicates, echinoderms, and peanut worms. The protist trichos have only been found in association with arthropods hosts, so the “Other invertebrate” state only applies to outgroup taxa. Vertebrate hosts include fish and amphibians. Crustacean includes shrimps, crabs, water fleas, crayfish, amphipods, and isopods.
Host stage	Juvenile Adult Both	Most members of the Eccrinida have only been recorded from either the juvenile or adult forms of their host, but a few taxa have been noted to infest both immature and mature forms. For example, <i>Leidyomyces attenuatus</i> has been collected from the hindgut of both larval and adult passalid beetles.
Growth form	Spherical/ovoid Hypha with holdfast Spherical/plasmodial Amoeboid Spherical/plasmodial & hypha	Descriptors refer to the form of the cells as observed in axenic culture or in freshly collected specimens. The term “hypha” describes a filamentous growth form similar to that found in the true fungi. “Holdfast” refers to an anchoring structure that adheres the cells to a host surface. For the protist trichos, the holdfast appears to be formed from an extracellular adhesive material extruded through pits at the base of the hypha.
Spore production	Endospores Cystospores Basipetal Endospores & cystospores Endospores & basipetal Amoeboid	Endospores are defined here as spores formed from parent cells for which the entire cytoplasmic content divides to form these propagules (= holocarpic). Cystospores are spores formed from a cyst (e.g. subsequent to encystment of an amoeboid stage). Basipetal refers to spores that are only released from the distal tip of the hyphae in succession towards the base. The basipetal spore formation type only applies to species of Eccrinidae and does not include a motile amoeboid phase in the life cycle. Amoeboid indicates species for which production of motile amoebae (with subsequent encystment) have been observed but production of spores from the cysts has not been confirmed.
Location in host	Foregut Hindgut Organs/tissues External Digestive tract Haemolymph Organs/tissues & digestive tract	The protist tricho taxa <i>Paramoebidium</i> and most of the Eccrinidae attach to the lining of the hindgut of their hosts. A few species of Eccrinidae are found only in the foregut (e.g. <i>Enteromyces</i>) or both the foregut and hindgut (e.g. <i>Arundinula</i>). Amoebidium is the only member of the suborder Trichomycina known to grow on the exterior of the host. Parasitic taxa (e.g. <i>Caullerya</i> , <i>Ichthyophonus</i> , <i>Psorospermium</i> , the <i>Tenebrio molitor</i> symbiont) have been observed to invade the connective and/or organ tissues (= “Other tissues”) of their hosts whereas putatively commensal species (e.g. <i>Abeforma</i> , <i>Creolimax</i> , <i>Sphaeroforma</i>) have been collected from the content of the host digestive tracts. Only the outgroup species <i>Capsaspora owczarzaki</i> is known to live in the haemolymph of its gastropod host.

Table2

Hypothesis	SH p-value	AU p-value
<i>Amoebidium</i> + <i>Paramoebidum</i> monophyletic	0.128	0.118
Protist trichos (<i>Amoebidiidae</i> + <i>Eccrinidae</i> + <i>Paramoebidiidae</i>) monophyletic	0.180	0.176
<i>Paramoebidium</i> + <i>Parataeniella</i>	0.032	0.024
<i>Eccrinidae</i> without <i>Parataeniella</i>	0.587	0.535

Table3

	Statistic	Observed mean	Lower 95% CI	Upper 95% CI	Null mean	Lower 95% CI	Upper 95% CI	Significance
Habitat	AI	0.913208	0.91320461	0.91320461	7.061775	5.825057983	8.244171143	0
	PS	13	13	13	44.621	40	49	0
	MC Freshwater	24	24	24	3.407	2	5	1.00E-03
	MC Marine	9	9	9	2.377	2	4	1.00E-03
	MC Terrestrial	5	5	5	2.074	1	3	0.005
Host type	AI	2.082251	2.082234859	2.082234859	8.689934	7.476430893	9.769218445	0
	PS	19	19	19	57.097	53	61	0
	MC Insect	24	24	24	2.665	2	4	1.00E-03
	MC Crustacean	16	16	16	2.183	1	3	1.00E-03
	MC Millipede	5	5	5	1.492	1	2	1.00E-03
	MC Vertebrate	6	6	6	1.153	1	2	1.00E-03
	MC Other invertebrate	2	2	2	1.16	1	2	0.15799999
Host stage	AI	2.409571	2.409572601	2.409572601	6.748516	5.552778721	7.929811954	0
	PS	17	17	17	41.068	37	44	0
	MC Juvenile	24	24	24	2.353	2	4	1.00E-03
	MC Adult	16	16	16	4.055	2	6	1.00E-03
	MC Both	2	2	2	1.071	1	2	0.06999999

Location in host	AI	1.824103	1.824117064	1.824117064	6.727517	5.708010197	7.720700741	0
	PS	16	16	16	38.64	37	40	0
	MC External	7	7	7	1.127	1	2	1.00E-03
	MC Hindgut	24	24	24	4.808	3	8	1.00E-03
	MC Foregut	2	2	2	1.002	1	1	0.00300002
	MC Organs/digestive tract	1	1	1	1.042	1	1	1
	MC Digestive tract	3	3	3	1.262	1	2	0.01200002
	MC Organs/tissues	1	1	1	1.057	1	2	1
Growth form	AI	0.932438	0.932446301	0.932446301	5.329536	4.292477608	6.333039284	0
	PS	10	10	10	29.752	28	31	0
	MC Hypha with holdfast	70	70	70	5.862	4	10	1.00E-03
	MC Spherical/plasmodial/hyphae	6	6	6	1.119	1	2	1.00E-03
	MC Spherical/ovoid	3	3	3	1.4	1	2	0.01700002
Spore production	AI	1.278151751	1.278145194	1.278145194	8.800624847	7.691408634	9.868835449	0
	PS	14	14	14	57.8660011	54	62	0
	MC Endo- & cystospores	7	7	7	1.136000037	1	2	1.00E-03
	MC Cystospores	24	24	24	2.066999912	1	3	1.00E-03
	MC Basipetal	35	35	35	2.614000082	2	4	1.00E-03
	MC Basipetal & endospores	4	4	4	1.034000039	1	1	1.00E-03
	MC Endospores (non-motile)	6	6	6	1.480000019	1	2	1.00E-03
	MC Amoeboid (motile)	2	2	2	1.037999988	1	1	0.037

Order Eccrinida