RESOLVING RELATIONSHIPS AT THE ANIMAL-FUNGAL DIVERGENCE: A MOLECULAR PHYLOGENETIC APPROACH FOR UNDERSTANDING THE ECOLOGY AND EVOLUTION OF THE PROTIST TRICHOMYCETES

by

Nicole K. Reynolds

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in Biology

Boise State University

August 2015

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the thesis submitted by

Nicole K. Reynolds

Thesis Title: Resolving Relationships at the Animal-Fungal Divergence: A Molecular Phylogenetic Approach for Understanding the Ecology and Evolution of the Protist Trichomycetes

Date of Final Oral Examination: 04 June 2015

The following individuals read and discussed the thesis submitted by student Nicole K. Reynolds, and they evaluated her presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Merlin M. White, Ph.D.	Chair, Supervisory Committee
James F. Smith, Ph.D.	Member, Supervisory Committee
Stephen Novak, Ph.D.	Member, Supervisory Committee
William Bourland, M.D.	Member, Supervisory Committee

The final reading approval of the thesis was granted by Merlin M. White, Ph.D., Chair of the Supervisory Committee. The thesis was approved for the Graduate College by John R. Pelton, Ph.D., Dean of the Graduate College.

DEDICATION

I would like to dedicate this thesis to my parents who told me I could, my advisor Dr. Merlin M. White who believed that I could, and my friend Emma R. Wilson who showed me I could.

ACKNOWLEDGEMENTS

There are many people who provided me valuable assistance over the course of this thesis, but I would like to start by thanking former lab member Eric D. Tretter for training me in all the basic lab procedures, chromatogram editing, software operation, and giving me a great example to follow. Another former lab member, Emma R. Wilson, was responsible for familiarizing me with the trichomycetes and teaching me all the collecting and dissecting skills necessary for the study of such, not to mention providing much inspiration, organization, and fun. Next, I would like to acknowledge the contributions of undergraduate members Dustin Heeney, Justin Gause, and Prasanna Kandel, each of whom was awarded support through academic programs (SRI and INBRE, respectively) for their work in assisting with molecular data collection. Additionally, undergraduates Mason Hinchcliff and Tyler Pickell aided with sifting through slides, photographs, and dissection log books.

My data set was enriched by genomic sample contributions from Dr. Matías J. Cafaro (MJC) (University of Puerto Rico), Dr. Laia Guárdia-Valle (LGV) (Universitat Autònoma de Barcelona), Dr. Doug B. Strongman (DBS) (St. Mary's University, Nova Scotia), Dr. Robert W. Lichtwardt (RWL) (University of Kansas, retired), and Dr. Scott LaPatra (Clear Springs Foods, Idaho), all of whom I would also like to thank for their advice and prompt correspondence. Furthermore, Drs. MJC, LGV, RWL, and DBS provided slide and/or photographic materials and collection information for samples, for

which I am grateful. Details of the nested PCR protocol for the heat shock protein 70 and 90 genes were kindly provided by Dr. Jordi Paps (University of Oxford).

I am very appreciative of the willingness and prompt replies of Dr. Boris C. Kondratieff (Colorado State University), Dr. Peter H. Adler (Clemson University), Dr. Peter Zwick (Max Planck Institute, retired), Dr. Ralph W. Holzenthal (University of Minnesota), and Derek Hennen (PhD candidate, University of Arkansas) for host identifications.

Awards from Sigma Xi and the Society of Systematic Biologists supplemented funding for my project; their aid was an important component of my successful data collection. In addition, I would like to acknowledge the following travel grants: Holz-Conner, awarded to me by the International Society of Protistologists, Student Research Program funds provided through Boise State University, and an interdepartmental award from the Department of Biological Sciences. These travel awards were critical for enabling me to gain invaluable experience in scientific communication and public speaking. As such, I wish to thank the International Society of Protistologists, the Mycological Society of America, and the Botanical Society of America for allowing me to present my work at their annual meetings.

I would like to express my gratitude to my committee members (Drs. James F. Smith, Stephen Novak, and William Bourland, M.D.) for their thoughtful reviews and commentary. I am especially appreciative of all the support and encouragement from my advisor, Dr. Merlin M. White, whose positive attitude helped to buoy me through times of frustration and self-doubt.

ABSTRACT

The Ichthyosporea (= Mesomycetozoea) is a relatively understudied class of unicellular symbionts that molecular phylogenies have placed at the divergence of animals and fungi. Subsumed in this class are the cosmopolitan families Eccrinidae and Amoebidiidae (referred to as “protist trichos” or “trichos” herein), which are considered obligate commensal endobionts of various arthropods, including marine, freshwater and terrestrial hosts. Once thought to be members of the fungal class Trichomycetes due to their hyphal-like growth form and ecological similarity, molecular evidence has necessitated reclassification. However, evolutionary relationships within and between them are still unclear as the number of taxa sampled and/or the amount of gene data gathered have been factors limiting resolution. These organisms are also taxonomically challenging since informative, homologous morphological characters are difficult to discern using only a light microscope (the method by which members of Amoebidiidae and Eccrinidae have traditionally been described), and only a few have been obtained in axenic culture. Most protist trichos reported thus far lack sufficiently detailed morphological parameters to permit ease and confidence in species identification. As such, relatively little is known about the ecology and biology of most members, some of which were originally classified as fungi or algae. As new members were discovered or reclassified, two orders were established: Dermocystida and Eccrinida. Whereas members of the Dermocystida are almost entirely parasites of various metazoan hosts, only three clades within the Eccrinida contain known parasites, with the remaining

members regarded as commensalistic. Interestingly, the putative closest extant relative to both groups is *Ichthyophonus*, an economically relevant fish parasite, which can invade vital host tissues (e.g. heart and liver) via circulating amoeba-like cells, causing disease and potentially death. The most recent molecular systematic study of the protist trichos was published about a decade ago, and there is as yet but one *Paramoebidium* (Amoebidiidae) sequence deposited in GenBank. Currently, based on molecular data, the Amoebidiidae are supported as monophyletic (based on one sample from each of its two genera) while the monophyly of the Eccrinidae is indicated, but not supported. Likewise, the relationship of the protist trichos to *Ichthyophonus* remains unresolved. As such, the first chapter of this thesis addressed the molecular phylogeny of order Eccrinida, with particular emphasis on the protist trichos by first amplifying and sequencing rDNA genes (18S and 28S) for over 100 new samples. Amplification tests were also attempted for several protein-coding genes, including heat shock protein 70. The resulting tree inferences were used in subsequent analyses of ecological and life history traits via ancestral state reconstructions and Bayesian tip-association significance testing (BaTS).

In the second chapter, samples of *Paramoebidium* spp. were morphologically and molecularly assessed as a case study into the utility of traditionally described morphological characters for taxonomic delimitation among protist trichos. Morphological differentiation of *Paramoebidium* spp. has been notoriously problematic due to inter- and intraspecific variability. Host specificity within the genus was early suggested, but later questioned, and has not been subjected to thorough evaluation. Therefore, host and hyphal characters were analyzed via three different methods of

ancestral state reconstruction, as well as with BaTS on a molecular phylogeny of over 70 Amoebidiidae samples.

Results of these studies indicate: 1) contrary to previous hypotheses, the Amoebidiidae may be paraphyletic, 2) relationships among Eccrinidae and between the protist trichos and *Ichthyophonus* remain unresolved, 3) several life history and host characters are significantly associated with both the Eccrinida and Amoebidiidae phylogenies, providing platforms for future hypothesis formulation, 4) the protist trichos and the Eccrinida as a whole are likely much more species rich and widespread than what is currently known, 5) species delimitation within *Paramoebidium* is complicated by cryptic speciation, but there is evidence for possible host specificity, and 6) future studies of the protist trichos will benefit from an integrated approach that shifts away from an emphasis on the morphological species concept but includes both genetic sequence data and traditional morphological approaches.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF ABBREVIATIONS

AI	Association Index statistic
AIC	Akaike Information Criterion
ASR	Ancestral State Reconstruction
ATCC	American Type Culture Collection
AU	Approximately Unbiased test
BaTS	Bayesian Tip-association Significance testing
BEAST	Bayesian Evolutionary Analysis Sampling Trees
BEAUti	Bayesian Evolutionary Analysis Utility
BI	Bayesian Inference
BLAST	Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
Cat.	Catalog
CTAB	Cetyltrimethyl Ammonium Bromide
DAMBE	Data Analysis for Molecular Biology and Evolution
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid
Γ	Gamma distribution
GenBank	National Institutes of Health genetic sequence database
GSI	Genealogical Sorting Index
GTG	Genetic Technology Grade

GTR	Generalized Time Reversible
HSP70	Heat Shock Protein 70 gene
HSP90	Heat Shock Protein 90 gene
I	Invariable sites
ITS	Internal Transcribed Spacer
K80	Kimura 2-parameter model
LG	Le-Gascuel amino acid substitution model
MC	Maximum exclusive single-state Clade size statistic
MCM7	Mini-Chromosome Maintenance 7 gene
MCMC	Markov Chain Monte Carlo
ML	Maximum Likelihood
MP	Maximum Parsimony
MUSCLE	Multiple Sequence Alignment by Log-Expectation
PAA	*Palavascia patagonica, P. sphaeromae, Alacrinella,* and *Astreptonema* sp. clade
PCR	Polymerase Chain Reaction
PS	Parsimony Score statistic
rDNA	Ribosomal DNA
RNA	Ribonucleic Acid
RPB	RNA Polymerase II gene
SH	Shimodaira-Hasegawa test
SIMMAP	Stochastic Mutational Mapping on Phylogenies
SOC	Super Optimal Broth media

TAE	Tris base, Acetic acid, and EDTA buffer
TBR	Tree Bisection and Reconnection
TIM	Transition Model
TMS	*Tenebrio molitor* Symbiont
TNT	Tree analysis using New Technology

PHYLOGENETIC AND ANCESTRAL STATE RECONSTRUCTION ANALYSES OF ORDER ECCRINIDA, WITH EMPHASIS ON THE PROTIST TRICHOMYCETES

Abstract

Trichomycetes traditionally was a class of gut fungi comprising four orders of microorganisms obligately associated with arthropods. Since molecular phylogenies revealed two of those orders ("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. Class Ichthyosporea comprises unicellular animal symbionts associated with hosts ranging from arthropods to bivalves and echinoderms to amphibians and mammals. The most recent classification divides the protist trichos between two families, Amoebidiidae and Eccrinidae, within order Eccrinida. However, no new sequence data for the protist trichos has been published in about a decade, and an evaluation of their characters in the context of this reclassification is lacking. Therefore, 18S and 28S rDNA sequences were generated for 106 protist tricho samples and combined with publicly available sequences of remaining Eccrinida taxa to generate a taxon-rich dataset for phylogenetic analyses. Additionally, PCR amplification tests of protein-coding genes were performed on the protist tricho samples. The trees generated were subsequently used as input for ancestral state reconstruction and Bayesian tip-association significance test (BaTS) analyses of six characters relating to the life history and morphology of taxa. The phylogeny provides evidence that both traditional and current taxonomy of the protist trichos may need

revision. Ancestral state reconstructions and BaTS results indicate several character states of host, habitat, and spore production traits are significantly correlated with the phylogeny. From these results, it is clear that 1) the species diversity of protist trichos and other taxa in the order is undersampled and underestimated, 2) the degree of host specificity among protist trichos may be higher than previously thought, and 3) the morphological species concept as it has been applied to Eccrinida taxa is insufficient for species delimitation in many cases, and a shift to an integrated approach that includes genomic sampling and/or culturing efforts should be considered in future studies.

Introduction

Placed at the animal-fungal divergence by molecular phylogenies (Steenkamp et al. 2006; Ruiz-Trillo et al. 2008; Paps et al. 2013; Cavalier-Smith et al. 2014), the class Ichthyosporea (= Mesomycetozoea) is a recently recognized group of unicellular, symbiotic protists that was only elevated from a clade in 1998 (Cavalier-Smith 1998; Herr et al. 1999; Mendoza et al. 2001). At that time, it consisted of only five taxa, but has grown to include over 40 (Glockling et al. 2013). Membership comprises newly discovered organisms (e.g. Marshall et al. 2008; Lohr et al. 2010; Marshall & Berbee 2011; Lord et al. 2012), phylotypes from environmental surveys (e.g. Takishita et al. 2005; Takishita et al. 2007; Lara et al. 2011; Evans et al. 2012; Heidelberg et al. 2013), and reclassified organisms formerly considered to be fungi or other types of protists (e.g. Baker et al. 1999; Mendoza et al. 2001). These taxa are divided into two orders, Dermocystida and Eccrinida (Cavalier-Smith 2013), based largely on phylogenetic analyses (Mendoza et al. 2002). Eccrinida is the larger order with two suborders and six families (Cavalier-Smith 2013). Morphologically, members have fungal-like characters (e.g. hyphal-like growth, sporangia) and produce endospores. All named taxa are found in symbiotic association with metazoan hosts (Glockling et al. 2013), including such disparate animals as arthropods, bivalves, echinoderms, fish, and frogs. Only three clades within the Eccrinida contain known parasites (just one of which is associated with a vertebrate), with the remaining members regarded as commensalistic or even facultative symbionts (Glockling et al 2013).

Among the latter group are the arthropod-associated endobiont families Amoebidiidae and Eccrinidae (referred to collectively here as “protist trichos” or

"trichos"), included within order Eccrinida, suborder Trichomycina (Cavalier-Smith 2013). These two taxa were previously regarded as orders within the fungal class Trichomycetes: the Eccrinales (with three families: Eccrinaceae, Palavasciaceae, and Parataeniellaceae) and Amoebidiales (with one family, Amoebidiaceae) (Lichtwardt et al. 2001) (Table 1.1). Trichomycetes is now realized as a paraphyletic ecological group, occupying the microhabitat of the arthropod digestive tract (Lichtwardt et al. 2001). Both the fungal and protist taxa form a holdfast structure that anchors individual hyphae the chitinous lining (except *Amoebidium* spp. which attach to the host exoskeleton) and produce walled spores as putative long-distance dispersal units, indicating the possibility of proximate convergence between groups (Leander 2008). Whereas the Eccrinidae produce spores in a basipetal manner, as seen within fungi, the Amoebidiidae produce motile amoebae in a holocarpic manner, which encyst and subsequently release cystospores. In addition to this amoeboid stage, *Amoebidium* may produce spores holocarpically directly, and without an intermediate phase (termed "sporangiospores" in older publications, but considered endospores here). These features, in combination with others such as lack of chitin in the cell wall (Whisler 1963; Trotter & Whisler 1965), and appendage ontogeny (Moss 1999), pointed to an independent evolutionary origin of these taxa. This was later confirmed by molecular phylogenetic work (Benny & O'Donnell 2000; Ustinova et al. 2000; Cafaro 2005) that placed them as the putative closest extant relative to *Ichthyophonus* [family Ichthyophonidae (Cavalier-Smith 2013)], an economically relevant fish parasite, which can invade vital host tissues (i.e. heart, liver, kidney) possibly via circulating amoeboid cells (Kocan et al. 2013), causing disease and potentially death (Rowley et al. 2013).

Table 1.1 Comparison of traditional fungal classification (Lichtwardt et al. 2001) within Class Trichomycetes, and the most recent classification (Cavalier-Smith 2013) within Class Ichthyosporea of the protist trichomycetes. Note that Cavalier-Smith did not include all genera of the Eccrinales in his classification.

	Orders	Suborders	Families	# of Genera
Lichtwardt et al. (2001) Traditional fungal classification	Amoebidiales		Amoebidiaceae	2
	Eccrinales		Eccrinaceae	14
			Palavasciaceae	1
			Parataeniellaceae	2
Cavalier-Smith (2013) classification	Eccrinida	Trichomycina	Amoebidiidae	2
			Eccrinidae	6
			Ichthyophonidae	1

Between these taxa, the Eccrinidae is more species rich, with 65 species, than the Amoebidiidae, which has 21. In part, both the disparity and relatively low numbers of descriptions represents the dearth of taxonomically informative characters. The latter has been problematic within both groups, but has especially precluded new species descriptions for the Amoebidiidae. Although, compared to their simply spherical or ovoid closest relatives (i.e. *Creolimax*, *Pirum*, *Psorospermium*, *Sphaeroforma*), the highly variable hyphal and spore morphotypes of the eccrinids verge on the flamboyant. On the other hand, the difference is probably also a reflection of the greater diversity of hosts eccrinids are associated with compared to the Amoebidiidae. According to current knowledge, the Amoebidiidae are restricted to freshwater habitats, with *Paramoebidium* found only in the guts of immature aquatic insects. In contrast, eccrinids have been found from almost each major group of mandibulate arthropods (crustaceans, millipedes, insects) in habitats ranging from around deep-sea hydrothermal vents (Van Dover & Lichtwardt 1986) to terrestrial caves (Manier 1964; Reeves et al. 2000) and from tropical

forests (White et al. 2000) to freshwater streams or lakes (e.g. *Arundinula orconectis*, *Astreptonema gammari*, *Enterobryus hydrophilorum*) (Lichtwardt et al. 2001). Similar diversity has been observed among other members of the Eccrinida, both in terms of host associations and habitats. For example, *Creolimax fragrantissima* was isolated from the digestive tracts of peanut worms, tunicates, sea cucumbers, and chitons (Marshall et al. 2008). Environmental phylotypes include those sampled from anoxic sediments around a submarine caldera near Japan (Takishita et al. 2005), sludge from a domestic waste water treatment plant in Australia (Evans & Seviour 2011), salt crusts from a hypersaline lake in Australia (Heidelberg et al. 2013), a carbon-rich, low nitrogen peat bog in Switzerland (Lara et al. 2011), and a shallow lake in Greece (Nikouli et al. 2013). Such surveys highlight the biodiversity of ichthyophonids, but unfortunately leave questions unanswered regarding their ecological function, impact on resident flora and fauna, and life histories.

Major challenges face those attempting to study these organisms. Firstly, most are unculturable, and despite numerous attempts by different researchers (Lichtwardt et al. 2001), *Amoebidium parasiticum* and *A. appalachense* are the only protist trichos to be cultured axenically (Whisler 1960; White et al. 2006). Therefore, our understanding of the taxa's life history, morphology, and host-symbiont interactions are restricted to those observations made during or following dissection of the host. When hosts are abundant and infection is prevalent, many data points (e.g. morphometric, physiochemical, etc.) may be relatively easily gathered. However, such ideal circumstances are not reliably encountered, especially for surveys of hard to reach habitats such as deep-sea hydrothermal vents, and prohibit detailed studies of the symbionts over the course of their

life cycle. Secondly, genomic samples are frequently mixed with host and/or other symbiont DNA, complicating molecular data collection. As such, relatively little is known about the ecology, biology, and evolutionary history of most protist tricho members, despite their initial discovery over 150 years ago (Leidy 1849). Indeed, some species have not been recorded since they were diagnosed in the early to mid 20th century, and a review of their biology in the context of their newly recognized phylogenetic position has not yet been undertaken.

Nevertheless, these organisms are sister taxa to the clade comprising choanoflagellates and animals. Understanding their diversity and evolutionary trajectory can help elucidate the path from unicellular protists to multicellular metazoans (Ruiz-Trillo et al. 2008). In particular, *Capsaspora owczarzaki* (suggested as either sister to the Ichthyosporea or to choanoflagellates + animals) and *Creolimax fragrantissima* have been the subject of genome and transcriptome sequencing projects due to their aggregative behavior (Ruiz-Trillo et al. 2006; Sebé-Pedrós et al. 2011; Suga et al. 2013; Sebé-Pedrós et al. 2013; Suga & Ruiz-Trillo 2013; Carr & Suga 2014). Additionally, as the Ichthyosporea is an early-diverging lineage, its diversity is potentially vast (Pawlowski et al. 2012; Glockling et al. 2013). Those that are obligately associated with their host and/or are endemic to a restricted habitat may be susceptible to extinction via abiotic factors such as climate change (Corliss 2004; Cotterill et al. 2008; Vicente 2010). Pathogenic species have the potential to impact populations of fish and amphibians due to spread via anthropogenic routes such as the wildlife trade (Rowley et al. 2013; Gozlan et al. 2014). All of these elements could be considered at ecosystem and global scales in

terms of biodiversity, ecology, and processes driving evolutionary patterns (Hoberg et al. 2015).

Progress in our understanding can only be aided by and well rooted in the firmly formed framework of a comprehensive phylogeny that engages a process of organismal and systematic reconsideration. The last published study to include new sequence data for the protist trichos was by Cafaro (2005). His single gene datasets included two representatives from the Amoebidiidae and 14 from the Eccrinidae. The hypotheses presented by his 18S and 28S rDNA molecular phylogenies suggested 1) a monophyletic Amoebidiidae, 2) a monophyletic but unsupported Eccrinidae, and 3) an unresolved sister relationship of the protist trichos with *Ichthyophonus*. Consequently, the goals of this study were to 1) evaluate the previous molecular phylogenetic hypotheses of the protist trichos (Lichtwardt et al. 2001; Cafaro 2005) using broader taxon sampling and a multigene dataset, 2) assess the traditional and current taxonomy of the protist trichos in light of this phylogeny, and 3) investigate ecological and life history patterns across the order Eccrinida via ancestral state reconstructions.

Materials and Methods

Taxon Sampling

Combined two gene (18S and 28S) and three gene (HSP70 and rDNA) datasets containing 106 unique samples (24 Eccrinidae, 81 Amoebidiidae and one *Ichthyophonus* sp.) were supplemented with data downloaded from GenBank to obtain representatives of every known eccrinid and dermocystid taxon, as well as choanoflagellate, filasterid, and animal outgroup sequences (Table 1.2). Additional environmental clone sequences were included, as they have been tentatively placed within the Eccrinida and reveal additional

diversity. Two versions of the rDNA tree were inferred: one with all successfully sequenced samples and large outgroup sampling ("large rDNA alignment/tree") and another focusing just on Eccrinida taxa with reduced *Paramoebidium* representatives and *Capsaspora owczarzaki* as the outgroup ("small rDNA alignment/tree"). The small rDNA taxon set and resulting tree were formed for use in ancestral state reconstructions to reduce computation time and tree drawing size. Character state coding for all *Paramoebidium* samples was the same, so elimination of some samples did not preclude the representation of character states for the genus. In the small rDNA tree, the protist tricho samples include two of five *Amoebidium* species, 12 of 17 eccrinid genera (includes 14 identified specimens out of 65 total eccrinid species), and six named (out of 17 total described species) and 15 putatively new species of *Paramoebidium*. For these unidentified *Paramoebidium* samples enough slide voucher material was available to distinguish them from published descriptions, but not enough to confidently delineate new species at this time. For the unidentified eccrinid samples, slide voucher material for specimens collected outside our lab was not readily available for review. *Enterobryus* sp. specimens collected locally (Boise, Idaho, USA) putatively represent a new species, but slide materials were not reevaluated for this study. Efforts were made to obtain data for as many named species as possible, but amplifications were attempted for all potentially unique samples, whether named or unnamed.

Cultures

Cultures of *Creolimax fragrantissima*, *Sphaeroforma arctica*, and *Sphaeroforma* sp. were obtained from the American Type Culture Collection (ATCC) (ATCC PRA-284, ATCC PRA-297, and ATCC PRA-283, respectively) and grown on ATCC medium 2673

(www.atcc.org). Several *Ichthyophonus* sp. cultures growing on rainbow trout heart tissue explants (Kocan et al. 2010) were obtained from Dr. Scott LaPatra of Clear Springs Foods, Inc., Idaho. A few of these tissue explants were placed directly in 2X CTAB buffer. The remaining were microdissected in an attempt to separate the *Ichthyophonus* cells from those of the host tissue. These cells were rinsed in successive drops of NanoPure water before CTAB preservation and were subsequently used for DNA extraction.

Sample Collection, DNA Extraction, and PCR Amplification

Unculturable specimens were microdissected from arthropod hosts collected (methods as in White et al. 2001) from various locations and timeframes (Table 1.2). Briefly, hosts are dissected using fine-tipped forceps and jeweler's needles with the aid of a stereomicroscope. Hyphae of presumed morphospecies are physically separated as much as possible from host tissues before being placed in 2X CTAB buffer. Genomic samples may contain an individual cluster of hyphae from a single host dissection, or have multiple hyphae of a putative morphospecies pooled from several dissections of the same host type. At the same time, hyphae of a morphospecies are preserved as slide vouchers for future morphological evaluation. In other words, each genomic sample should have a corresponding slide voucher, given enough material is present for both preparations. Therefore, in the context of this study, collections refer to host sampling and dissecting events from a specific location and date; sample refers to a single genomic and slide-preserved morphospecies from a collection. The term slide voucher refers to a slide or series of slides that correspond to a given morphospecies. Genomic samples preserved in CTAB buffer were kept frozen at -20°C or refrigerated at 4°C until the time

of extraction. Extractions followed standard CTAB freezing/thawing and phenol-chloroform methods from White (2006). Some of the same genomic dilutions of eccrinid specimens studied by Cafaro (2005) were received and incorporated for amplification attempts and testing herein as well.

rDNA Genes

The small subunit (18S) and large subunit (28S) ribosomal rDNA genes were targeted for amplification and sequencing. Different primer combinations were used during amplification attempts (Tables 1.3 and 1.4). Primers NS1AA and NS8AA (18S) and NL1AA and LR7AA (28S) were developed by Wang et al. (2014) so as not to amplify animal DNA, a common contaminant of preparations from microdissections. Amplification of the ITS region was also tested, and limited data was obtained (Table 1.4), but not enough to include in the dataset. Primers NS1PT and NS8PT (18S) are novel (designed by Eric Tretter, BSU) and are meant to be specific to the protist trichos. For 18S reactions, the PCR recipe generally consisted of the following reagents: 11.0 µL Promega GoTaq Green Master Mix (Cat.# M7122), 0.66 µL of each primer at 10 µM concentration, 0.88 µL of 25 mM $MgCl_2$ (to a final concentration of 1.0 mM), 0.35 µL of 50 µg/µL bovine serum albumin (BSA) (to a final concentration of 0.8 µg/µL), 6.45 µL nuclease-free, purified water, and 2 µL of diluted genomic DNA for a total volume of 22 µL. The thermal cycling program included these steps: initial denaturation at 95°C for 2 minutes followed by 45 cycles of denaturation at 95°C for 30 seconds, annealing at 60°C for 45 seconds and elongation at 72°C for 3 minutes, and completed with an elongation hold at 72°C for 10 minutes and a final hold at 4°C. The default 28S PCR recipe was as follows: 11.0 µL Promega GoTaq Hot Start Green Master Mix (Cat.# M5122), 0.66 µL of

each primer at 10 μM concentration, 0.44 μL of 25 mM $MgCl_2$ (to a final concentration of 0.5 mM), 2.20 μL of 5.0 M betaine (to a final concentration of 0.5 M), 5.04 μL nuclease-free, purified water, and 2 μL of diluted genomic DNA for a total volume of 22 μL. The cycling program had an initial denaturation at 95°C for 2 minutes followed by 45 cycles of denaturation at 95°C for 30 seconds, annealing at 56°C for 45 seconds and elongation at 72°C for 3 minutes, a final elongation step at 72°C for 10 minutes and a final hold at 4°C. Samples for which the initial amplification attempt (using the preceding reaction conditions) failed were retried using modifications to the BSA, betaine or $MgCl_2$ concentration in the PCR recipe and/or adjusting the annealing temperature of the thermal cycling program. For a few trials, BSA was added to the 28S PCR cocktail to the same final concentration as for the 18S reactions.

Heat Shock Protein 70 (HSP70) Gene

Ichthyosporea-specific primers designed by Paps et al. (2013) were used with a nested PCR approach to amplify the HSP70 gene. Limited attempts were also made to amplify the HSP90, MCM7 (minichromosome maintenance) and RPB I and II (RNA polymerase II largest and second largest subunits) protein-coding genes (see Table 1.3 for primers used), but were, for the most part, unsuccessful. The HSP70 PCR reaction consisted of two rounds. The reaction cocktail for the first round contained the same quantity and reagents as the 28S reaction, except 0.44 μL of 5.0 M betaine (to a final concentration of 0.1 M) and 6.80 μL nuclease-free, purified water were used (to a total volume of 22 μL). This recipe was also used for the second round, but 5 μL of PCR product from the first round was used as template (rather than diluted genomic DNA), for a total reaction volume of 25 μL. The thermal cycling program for both rounds was: an

initial denaturation step at 94°C for 5 minutes followed by 5 cycles of denaturation at 94°C for 35 seconds, annealing at 50°C for 45 seconds, and elongation at 72°C for 1 minute and 15 seconds, then 35 cycles of denaturation at 94°C for 35 seconds, annealing at 45°C for 45 seconds, and elongation at 72°C for 1 minute and 15 seconds, a final elongation step of 72°C for 8 minutes and a final hold at 4°C.

Electrophoresis and Sequencing

The PCR products were electrophoresed through a 1% Lonza Seaplaque GTG agarose gel (Cat.# 50110) in low EDTA 1X TAE buffer. Lonza GelStar nucleic acid stain (Cat.# 50535) was added to the gels and bands were visualized on a Clare Chemical DR46B transilluminator box. Bands were cored out of the gel using wide-bore pipette tips, and the tips containing the cores were placed in 1.5 mL microcentrifuge tubes and frozen at -20°C. For the majority of samples, DNA was separated from the gel using a "freeze and squeeze" method (Tautz & Renz 1983). The cut tips served as columns to squeeze PCR product and buffer from the gel core. Tubes were frozen and centrifuged twice (15 G for 15 minutes) using a Thermo Scientific Legend Micro21 centrifuge. One to 2 μL of the squeezed product was added to sequencing reactions. Sequencing reactions were performed with the Applied Biosystems BigDye v. 3.1 kit for bidirectional sequencing using 0.55 μL premix, 8.25 μL buffer, 0.66 μL of each primer at 10 μM concentration, and 10.54 μL nuclease-free, purified water to a total volume of 22 μL. The cycling program had an initial denaturation step at 96°C for 1 minute followed by 98 cycles of a 96°C denaturation for 10 seconds, a 50°C annealing step for 10 seconds, and a 60°C elongation for 4 minutes, completed with a 4°C hold. Products of these reactions were shipped to the University of Wisconsin Biotechnology Center for cleanup and

Sanger sequencing on an Applied Biosystems ABI 3730xl capillary DNA analyzer. Chromatograms were visualized using Sequencher 5.0. Ends of sequences were trimmed, aligned into contigs and then checked for obvious sequencing errors. Contig sequences were checked against the National Center for Biotechnology Information database using BLASTn or BLASTx as a preliminary assessment of identity.

Cloning

Samples for which primary sequencing attempts were hampered by secondary structure or mixed or weak signal were cloned. The Promega pGEM T Easy Vector System kit (Cat.# A1360) was used along with the JM109 competent cells (Cat.# L2004) for all cloning trials. The ligation reaction mixture was modified from the manufacturer's manual recipe, using the following regent quantities: 2.5 μL ligation buffer, 0.5 μL pGEM T Easy vector, 0.5 μL T4 ligase, 0.5 μL nuclease-free, purified water and 1 to 2 μL PCR product template. Ligations were stored at 4°C overnight for maximal efficiency. Competent cell transformation procedures were as recommended in the manufacturer's manual except that a reduced volume of competent cells (20-25 μL) and SOC media (200 μL) was used in each reaction. The Promega manual recipes were followed to make the stock solutions and media for culture plates. Plates were incubated at 37°C until colonies were sufficiently large for picking (approximately 16-18 hours). A pipette tip was dipped into opaque, white colonies (although a few blue and blue-white "bull's-eye" colonies were tested as well) and the cells were placed directly into tubes containing this PCR mixture: 10.5 μL Promega GoTaq Green Master Mix, 0.63 μL of each primer (M13 forward and reverse, Cat.# Q5601 and Q5421, respectively) at 10 μM concentration, 0.84 μL of 25 mM $MgCl_2$ (to a final concentration of 1.0 mM), and 7.4 μL

nuclease-free, purified water to a total volume of 20 μL. Small volumes of PCR products were run on a 2% Fisher agarose gel (Cat.# BP160-100) stained with GelStar to check for presence and correct size of bands. Once confirmed, PCR products were either sequenced directly, run and cut from a 1% GTG gel (as described above) or treated with Affymetrix ExoSAP-IT (Cat.# 78200) enzymatic PCR product cleanup reagent prior to sequencing.

Phylogenetic Analyses

Sequences were imported into Mesquite v. 3.01 (Maddison & Maddison 2014) and aligned with MUSCLE (Edgar 2004). Adjustments to the alignment and ambiguously aligned regions for exclusion were determined by eye. Alternate versions of the alignment were made using the program Gblocks (Castresana 2000), but missing data affected which settings could be utilized in the program, and the resulting output was not significantly better for estimating tree topologies. For the HSP70 alignment, the reading frame and translation to the protein sequence was determined using the Swiss Institute of Bioinformatics ExPASy online translation tool (Artimo et al. 2012). Additionally, the translation tool in Mesquite was used to create an amino acid alignment from the nucleotide alignment. Putative HSP70 sequences (after an initial screening using BLASTx) were aligned with sequences from a wide variety of fungal, animal, and protist taxa (downloaded from GenBank) and a tree was estimated from these data to ensure they formed a clade with other Eccrinida taxa. As a result, several dubious sequences were identified as possibly being from either the host or a paralogous gene, and thus were removed from further analyses. Individual gene trees as well as trees based on combined datasets were constructed to compare topologies. To check for

substitution saturation, plots of K80 distance against transitions and transversions (Brown et al. 1982) were drawn in DAMBE5 (Xia 2013). The third codon position was evaluated separately from positions 1 and 2 for HSP70.

Alignments were analyzed with PartitionFinder v. 1.1.0 (Lanfear et al. 2012) and jModelTest 2.0 (Darriba et al. 2012; Guindon & Gascuel 2003) to evaluate appropriate data partitions and for substitution model selection. Models were chosen according to the corrected AIC score. For HSP70, partitions based on codon position were compared and trees were inferred from both the nucleotide and protein alignments to determine which was more informative.

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. small and large rDNA and the three gene). Individual gene trees were also inferred and compared to check congruence. The consensus tree output from Bayesian analyses was used as the base tree for illustration purposes, but the MP and ML support values are shown as well. 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 (Stamatakis 2014). The GTR + Γ + I model was used for both the small and large rDNA alignments and LG + Γ + I for the HSP70 amino acid alignment. The GTR + Γ + I model (without partitioning) had the highest support from PartitionFinder, but was second best in jModelTest, behind TIM2 + Γ + I. 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 were assessed with Tracer v. 1.5.0 (Rambaut & Drummond 2009). Tree files were viewed and drawn for illustration using TreeGraph 2 (Stöver & Müller 2010), FigTree v. 1.4.2 (Rambaut 2014), and Inkscape (https://inkscape.org/en). Support values greater than 70 (MP), 75 (ML), and 0.95 (BI) were considered well supported.

Topology Testing

To test alternative hypotheses of tree topology, the large rDNA tree was used, with constraint trees drawn in and exported from Mesquite. Branch length estimates and bootstrapping followed by estimation of per site log likelihoods for each constraint tree were computed with RAxML. Seven alternative topologies were tested based on preliminary topology results, hypotheses from Cafaro 2005, and the traditional (i.e. fungal) classification of families within the Eccrinidae: 1) monophyletic Amoebidiidae (*Paramoebidium* + *Amoebidium*), 2) protist trichos monophyletic (Amoebidiidae + Eccrinidae), 3) *Parataeniella* with *Paramoebidium* (an arrangement that was noted among alternative topologies), 4) monophyletic Eccrinaceae, 5) monophyletic Palavasciaceae (without samples 1115 and 1121), 6) monophyletic Palavasciaceae with samples 1115 and 1121, and 7) Eccrinidae without *Parataeniella*. To perform SH (Shimodaira & Hasegawa 1999) and AU (Shimodaira 2002) tests, TREE-PUZZLE (Schmidt et al. 2002) and CONSEL (Shimodaira & Hasegawa 2001) were used. The null hypotheses are as follows (from Schmidt 2009):

SH test: All trees $T_x \in T$ (including the ML tree) are equally good explanations of the data.

AU test: The expected value E [log likelihood$_a$] of T_a is larger or equal to the expected values of all trees $T_x \in T$.

Ancestral State Reconstruction (ASR)

Terminology

Whereas it is clear that the true diversity of species within the Eccrinida is as yet unknown, and a full accounting of life history stages, ecological roles, etc. is unclear for many species (and, indeed, entirely unknown for environmental samples) (Glockling et al. 2013; Rowley et al. 2013), a plot of general characteristics on a phylogeny may still be illustrative of trends across clades. It is with this background understanding of the order that ancestral state reconstructions were performed. Characters and character states were chosen based on published descriptions of species (e.g. Vogt & Rug 1999; Lichtwardt et al. 2001; Marshall et al. 2008; Lohr et al. 2010; Marshall & Berbee 2010; with the awareness that several such reports were preliminary. Therefore, interpretation of the results of these analyses should be viewed as an initial attempt and are intended to demonstrate notable evolutionary tendencies and affinities of characters across the group.

In an effort to promote uniformity of terminology within the group, and in recognition of the need of ontological studies to evaluate the homology of spore formation, a standaradized set of terms, as so defined, will be used herein. As such, parent cells for which the entire cytoplasmic content forms walled propagules are termed "endospores" (similar to "holocarpic" spore formation in fungal terminology). The spores formed holocarpically in *Amoebidium* spp. have been referred to as "sporangiospores" in the literature (Lichtwardt et al. 2001), but are termed endospores in this study. For *Ichthyophonus*, Kocan (2013) suggested the term "schizont" to describe

these holocarpic cells, and “merozoites” for the daughter cells produced from them. A distinction between motile (i.e. amoeboid) versus nonmotile daughter cells (i.e. endospores) is made here. Most members of the Eccrinidae, however, appear to have a highly derived ontogeny of propagules in that they are (in some cases) dimorphic and formed in a basipetal manner, similar to some Fungi (Lichtwardt et al. 2001). Likewise, propagules formed from a cyst (i.e. as in the Amoebidiidae) are termed “cystospores”, as their derivation follows encystment of an amoeboid propagule. In this case, the motile amoeboids form the cysts from which the dispersive spores are produced. The term “spore” has various definitions depending on the specific group of organisms in question. For the purposes of this study, “spore” refers broadly to the putative dispersive uninucleate or multinucleate asexually-produced propagules that would either amplify the endogenous infection in a host, be taken up by a new host, or be released to the environment (where a free-living stage has neither been observed nor disproved). Finally, “hypha” as used here also follows Kocan (2013), with the caveat that a holdfast, as a feature of hyphal attachment, was not included in that definition. Hyphae with a holdfast are differentiated for the purposes of these analyses, but still adhere to the chosen definition in that the holdfast appears to be formed by a secretion of extracellular material as opposed to a separate cellular structure (Moss 1979). Undoubtedly, as more species are discovered, ultrastructural, biochemical, and ontological evaluations conducted, and homology reassessed, these terms will need continued refinement and perhaps even parsing into more nuanced vocabulary.

Character State Coding and ASR Analyses

Six characters were reconstructed over the final 5,000 trees (the last 2,500 from each of two runs) of the MrBayes analysis of the small rDNA dataset using SIMMAP v. 1.5.2 (Bollback 2006). All trees were rooted using the outgroup method, with *Capsaspora owczarzaki* as the chosen taxon based on recent multigene phylogenetic work (Cavalier-Smith et al. 2014) and availability of molecular and character state data. Alternative topologies with dermocystid taxa as the outgroup were inferred, but the relationships of the in-group taxa were not affected by this change. *Capsaspora owczarzaki* was included in reconstructions. Characters and their states (Table 1.5) were chosen based on their description in the literature, presumed ecological/life history relevance, and potential homology within clades. Coding for individual samples is indicated on Figs. 1.4 - 1.9. The habitats in which ichthyophonids are found are diverse and broadly categorized into marine (0), freshwater (1), and terrestrial (2). Clone LT37 was coded as marine, but it was collected from a benthic salt crust sample from a hypersaline lake in Australia (Heidelberg et al. 2013). Likewise, *Taeniellopsis* was coded as terrestrial, but its hosts are often located in saline/brackish environments (Lichtwardt et al. 2001). Some species described so far appear to have little host specificity, but similarly can be generally categorized as vertebrate (0), crustacean (1), insect (2), millipede (3) or other invertebrates (4) (e.g. bivalves, tunicates, etc.). The stage of development of the host at the time of infection by an ichthyophonid is not consistently recorded, but may be juvenile (0), adult (1) or both juvenile and adult (2). Multiple growth forms have been observed for culturable species such as *Abeoforma whisleri*, *Creolimax fragrantissima*, and *Sphaeroforma* spp., whereas others have only a single,

dominant form as with the Amoebidiidae, Eccrinidae, and *Pirum gemmata*. States include: spherical/ovoid (0), hyphal with holdfast (1), spherical/ovoid and plasmodial (2), amoeboid (referring to the filose forms for the outgroup taxon *Capsaspora owczarzaki*) (3), and spherical/ovoid, hyphal, and plasmodial (4). As specified above, the process by which spores are formed may be delineated as endospores (0), cystospores (1), basipetal (2), endospores and cystospores (3), endospores and basipetal (4), or amoeboid propagules (5). Finally, the location within the host body where ichthyophonids reside are categorized as foregut (0), hindgut (1) (both applicable to arthropod hosts), organs/tissues (2), attached externally (as with *Amoebidium*) (3), digestive tract (to include other invertebrate hosts whose digestive tracts are structurally unlike to those of arthropods) (4), haemolymph (for the outgroup taxon *Capsaspora owczarzaki*) (5), or organs/tissues and digestive tract (6).

The two-step process for choosing priors for these multistate characters described on the SIMMAP website (Bollback 2009) was followed, using a maximum clade credibility tree. The first step uses an MCMC analysis to sample the gamma and beta priors for multistate characters and the second step samples the posterior distributions of these analyses and plots the best fitting distribution in the R Statistical Package (R Foundation for Statistical Computing 2015). 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 & Drummond 2009) to ensure appropriate sampling was achieved and to compare the log likelihoods. Characters were unordered. Once priors were chosen, analyses were run using k=90, rate=1.00, and 20

samples/tree with 20 priors drawn. Results were plotted as pie charts using the R script "PlotSimMap.R" (available from https://github.com/nylander/PlotSimMap) and further modified using Inkscape (https://inkscape.org/en). Polymorphic states were coded as separate states for these analyses such that they would be considered as true polymorphisms rather than uncertainty, with the understanding that this approach implicitly assumes that the states are correlated and arose simultaneously in those taxa (Millanes et al. 2011).

Table 1.5 Characters and character state coding for SIMMAP and BaTS analyses.

Character	State 0	State 1	State 2	State 3	State 4	State 5	State 6
Habitat	Marine	Freshwater	Terrestrial				
Host type	Vertebrate	Crustacean	Insect	Millipede	Other invertebrate		
Host stage	Juvenile	Adult	Both				
Growth form	Spherical/ ovoid	Hypha with holdfast	Spherical/ plasmodial	Amoeboid	Spherical, plasmodial & hypha		
Spore production	Endospores	Cystospores	Basipetal	Endo- & cystospores	Endo- & basipetal	Amoeboid	
Location in host	Foregut	Hindgut	Organs/ tissues	External	Digestive tract	Haemolymph	Organs & digestive tract

Bayesian Tip-Association Significance Testing (BaTS)

As a further test of correlation between characters and the phylogeny, BaTS beta v. 2 (Parker et al. 2008) was run using the last 1,002 trees of the MrBayes analysis (final 501 trees from each of two runs, using the small rDNA alignment). This method 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 first generating a null distribution by randomizing tip rearrangements and then comparing the observed value to the expected value (generated from the null distribution) to obtain a p-value. The null hypothesis for

evaluating the p-values 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, while the MC will show a positive correlation with the association. Characters and their states were the same as for the SIMMAP analysis and 1,000 null replicates were performed.

Results

Genetic Data

A total of 19 new HSP70 and 195 new rDNA sequences were generated (Table 1.3). Protocols 194 and 183 (Table 1.4) were the most successful and frequently used for 18S and 28S amplifications, respectively. Host sequences were occasionally obtained using 18S and HSP70 amplification procedures (approximately 10% and 20%, respectively). Whereas no host sequences were produced using the 28S protocol 183, nine out of 14 attempts resulted in host sequences for protocol 133 (general primer set). These results highlight the importance of primer specificity when attempting PCR on mixed genomic samples, an unavoidable consequence of collection of these unculturable organisms. In particular, the HSP70 primers for use on protist tricho samples should be modified to avoid animal DNA. Gene fragments ranged in size from 1,613 bp (*Ichthyophonus* sp. 1193) to 2,011 bp (*Palavascia* sp. 402) for 18S, 1,965 bp (*Paramoebidium* sp. 616) to 1,483 bp (*Sphaeroforma arctica* 1242) for 28S, and HSP70 from 750 to 800 bp. 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 3rd HSP70 codon position clearly showed saturation. As the tree topologies and number of supported branches inferred from nucleotide and amino acid alignments were

not substantially different, the latter was chosen for use in the final analyses. The final three gene alignment consisted of 27 taxa and 3,051 characters: 1,546 bp of 18S, 1,251 bp of 28S, and 253 HSP70 amino acids, with a total of 1,544 ambiguously aligned characters excluded. The small rDNA gene alignment consisted of 101 taxa and 2,678 characters: 1,479 bp of 18S characters and 1,199 bp of 28S characters, and 2,442 excluded. Finally, the large rDNA gene alignment had 174 taxa and 2,640 characters; 1,466 bp of 18S and 1,174 bp of 28S, with 2,576 excluded. There was a total of 9% missing data (number of missing characters/total characters) in the three gene final alignment, and 20% missing in the small and 22% in the large rDNA two gene datasets.

Phylogenetic Analyses

The only *Paramoebidium* sp. sequence currently in GenBank (AY336708.1) placed within the *Amoebidium* clade (Fig. 1.1), and this was in contrast to all data subsequently assessed for the placement of *Paramoebidium*, so that sample was not included in the small rDNA tree. Attempts were made to amplify and sequence new data from the genomic stock (collection code KS-61-W6, sample 1175), but each was returned as fungal contaminant (as indicated by sequence length and BLAST searches). Furthermore, the 28S sequence for the TMS sample in GenBank (JN699061.1) did not align with the rest of the taxa, even in highly conserved regions, so it was not included in any dataset. There were no supported topology conflicts between individual genes (18S, 28S, HSP70). The three analysis methods (MP, ML, and BI) likewise recovered no conflicts on the large (Fig. 1.1) and small (Fig. 1.3) rDNA trees. However, MP supported *Ichthyophonus* as sister to the protist trichos in the three gene analysis (Fig. 1.2).

There was an unexpected split of the Amoebidiidae by genus, each forming a well-supported (by at least two methods) clade in all trees, but topology tests (SH and AU, Table 1.6) do not reject the hypothesis of a sister taxa relationship between *Amoebidium* and *Paramoebidium*. BI and MP, but not ML supported *Amoebidium* as the earliest-diverging Trichomycina lineage in the small rDNA tree (Fig 1.3). Similarly, the three gene analyses recovered support from BI and ML for that placement of *Amoebidium* (Fig 1.2). Although *Ichthyophonus* is again (e.g. Ustinova et al. 2000; Benny & O'Donnell 2000; Cafaro 2005) indicated as a sister taxon to the protist trichos, the constraint tree forcing a monophyletic protist tricho clade (and thus placing *Ichthyophonus* as the early diverging lineage) is not rejected by topology tests (Table 1.6).

As found previously with fewer taxa (e.g. Cafaro 2005; Marshall et al. 2008; Marshall & Berbee 2010), the Eccrinidae is monophyletic, but without support from any analysis method. Although herein, taxon sampling within the group is not complete, representative rDNA gene sequences from each family were obtained. Based on inferences from these data, the traditional family structure is not supported, with the Palavasciaceae and the Eccrinaceae being non-monophyletic. The Parataeniellaceae is recovered as monophyletic, but with representatives of just one of its two genera included in the dataset. Topology tests reject the hypothesis of monophyly of the Eccrinaceae, but not that of the Palavasciaceae (which consists of one genus). However, samples of two Eccrinaceae genera (*Alacrinella limnoriae* and *Astreptonema* sp.) are supported by all three methods as a clade with the two *Palavascia* species (identified as the PAA clade) (Figs. 1.1, 1.3). Among alternate topologies, *Parataeniella* placed on an early diverging

branch outside the Eccrinidae (not rejected by topology tests) and sometimes as sister to *Paramoebidium*. This latter topology was found to be significantly worse than the best RAxML tree by SH and AU tests (Table 1.6).

In the large rDNA tree, the Eccrinida and Dermocystida each form well-supported clades and are sister taxa, in keeping with previous findings (Glockling et al. 2013). Environmental clones were associated with the same clades as previous analyses (e.g. Marshall et al. 2008; Marshall & Berbee 2010) and *Psorospermium haeckelii* remains on a separate, unsupported branch (Glockling et al. 2013). *Sphaeroforma* and *Creolimax*, and *Caullerya* and the TMS sample are each recovered as sister taxa pairs (but without support from ML in the latter). However, *Caullerya* and the TMS sample are both on long branches, indicating that false association resulting from long-branch attraction cannot be ruled out.

Ancestral State Reconstruction and BaTS Analyses

The environment where ichthyophonids have been collected is significantly associated with clades, as illustrated by SIMMAP reconstructions (Fig. 1.4) and indicated by BaTS results (Table 1.7). At least one major shift from terrestrial to marine and freshwater hosts occurred among the Eccrinidae. Indeed, the ancestral state of the Eccrinidae is indicated with high probability as terrestrial. On the other hand, all the Amoebidiidae are found in freshwater environments, and the probability of freshwater as the ancestral state for the protist trichos as a whole is greater than 50%. Similarly, there is a clear division between the marine environmental clones, which are included in clades containing *Sphaeroforma* and *Creolimax*, and the freshwater clones that form a clade with *Anurofeca*. However, the ancestral state of the entire Eccrinida is equivocal.

The broad categorizations of hosts also significantly align with the phylogeny, as indicated by both methods (Table 1.7, Fig. 1.5). There have been several host shifts within the Eccrinidae, with at least one from millipedes to insects and at least one from insects to crustaceans. Despite these transitions, the ancestral state for the clade is indicated as a millipede host, whereas insect host is given the majority of support at the node uniting the protist trichos. Outside this clade, reconstructions at deeper nodes become equivocal due, at least in part, to missing data for the environmental clones. The developmental stage of the host at the time of association with ichthyophonids is significant for the states "juvenile" and "adult", but not for "both" in the BaTS analysis (Table 1.7). Only *Ichthyophonus*, *Leidyomyces*, *Psorospermium*, the TMS, and *Caullerya* have been reported to infect both the juvenile and adult stages of their hosts. Although all *Paramoebidium* species described to date are associated with immature aquatic insects, the ancestral state for the protist trichos is supported as "adult" (Fig. 1.6). The final host-related character, location within the host (Fig. 1.9), has four states that are significantly associated with the topology (Table 1.7): foregut, hindgut, external, and organs/tissues. All *Paramoebidium* and eccrinid (excepting *Enteromyces*) samples were collected from the hindgut of their host, and that state is indicated as the most probable ancestral location for the protist tricho group.

The growth form (Fig. 1.7) and type of spore production (Fig. 1.8) recorded so far for ichthyophonids are correlated with clades for most states (Table 1.7). Hyphal growth with a holdfast is the single form observed for the protist trichos, including *Amoebidium*. This state is recovered as the most likely ancestral condition for the protist tricho clade, but spherical/ovoid is indicated for the order (Fig. 1.7). In addition to those two states,

spherical/hyphae/plasmodia has a significant correlation with the phylogeny, logically, as *Ichthyophonus* was the only genus coded with that state. The "amoeboid" state returned a significant result, but it was only represented once, for the outgroup. In contrast, spore production across the protist trichos is markedly different among taxa, and all states are significantly correlated with the phylogeny in the BaTS analyses (Table 1.7). "Basipetal" is indicated as the ancestral state for the Eccrinidae, and even at the split of *Ichthyophonus* from the Eccrinidae and *Paramoebidium* (Fig. 1.8). Beyond that node, "endospores" becomes the heavily favored state, including that for the protist trichos as a whole.

Discussion

Phylogeny and Taxonomy of the Protist Trichos

The phylogenetic analyses (Figs. 1.1, 1.2, 1.3) and topology tests (Table 1.6), though not supported with significant values in all cases, are contradictory to the traditional fungal taxonomy and potentially that of Cavalier-Smith (2013). The monophyly of the Amoebidiidae was not rejected by topology tests (Table 1.6), but the placement of *Amoebidium* as the earliest-diverging lineage in the Trichomycina was supported by at least one analysis method in all three trees. Such a signal suggests that additional taxon and gene sampling might reinforce the split. As such, the family Amoebidiidae is likely paraphyletic, and likewise the traditional order Amoebidiales. The diversity of *Paramoebidium* samples, their division into multiple well-supported clades, and their unification as a well-supported monophyletic clade (Fig. 1.1) demonstrate that the elevation of the group to family level would not be unfounded from a phylogenetic perspective. Indeed, life history characteristics of the genera reinforce the

potential evolutionary distance between them, even if many of the gross morphological characters examined to date do not. *Amoebidium* is the only genus of order Eccrinida known to attach to the exterior of its host and to produce both endo- and cystospores (as they are defined here). Both of these character states are significantly correlated with the phylogeny in the BaTS analyses (Table 1.7) and ASRs (Figs. 1.8, 1.9), indicating that they are likely homologous within clades. Furthermore, *Amoebidium* species have a broad range of host associations (Fig. 1.5) (Lichtwardt et al. 2001) as compared to *Paramoebidium*, for which host specificity may be significant (see Chapter 2). Moreover, ultrastructural differences have been noted. For example, *P. curvum* had cylindrical pits at the apex of the cystospores (from which holdfast material is thought to exude), whereas those of *A. parasiticum* were tapered (Dang & Lichtwardt 1979). However, division of the Amoebidiidae remains, as yet, premature due to the unresolved placement of *Ichthyophonus* and the Eccrinidae.

On the other hand, the monophyly of the Eccrinidae, though not supported, aligns with their unification as a taxonomic unit. Now collapsed to a single family, eccrinid genera were divided among three families in the fungal framework (Table 1.1). These divisions are not borne out by the results presented here (see Appendix A for a detailed discussion of Eccrinidae genera). For example, the Palavasciaceae is the smallest family with one genus and three species, and is represented by samples from two species. A monophyletic Palavasciaceae as traditionally circumscribed is not rejected by topology tests (Table 1.6), but neither is the well-supported clade containing *Palavascia* + *Alacrinella limnoriae* and an unnamed *Astreptonema* (WA-3-C3) sample (PAA clade) (Fig. 1.1). Hibbits Galt (1971) hypothesized that *Alacrinella*, *Astreptonema* sp., and

Palavascia were closely related to one another, as well as to two additional genera, *Paramacrinella* and *Ramacrinella* (traditionally placed within the Eccrinaceae). Although no genomic samples were available for the latter two, the well-supported PAA clade supports part of her hypothesis. However, there are several characters that could indicate the relatedness of all six specimens: distinct "microthalli" that produce uninucleate cells [although the thin filaments described for *Palavascia* spp. have not been termed "microthalli", to date (Cafaro 2000)] and a persistent "spore mother-cell" at the distal or proximal end of the hyphae (Lichtwardt et al. 2001). Additionally, *Paramacrinella* and *Ramacrinella* share host (isopods and amphipods [Crustacea]) and habitat (marine) types with the PAA clade. As these character states are significantly correlated with the phylogeny (Table 1.7, Figs. 1.4, 1.5), the hypothesis is bolstered by these results, but awaits molecular phylogenetic confirmation.

In contrast, the Eccrinaceae is the largest of the traditional families, and its monophyly is rejected by topology tests (Table 1.6). Again, clade formation in relation to host and habitat type is evident across the representative samples (Figs. 1.4, 1.5). As a case-in-point, *Enterobryus* is the largest of the eccrinid genera, but the samples included here are polyphyletic. The crab-associated *E. halophilus* is more closely related to other marine, decapod-associated taxa (e.g. *Enteropogon, Taeniella*) than with other *Enterobryus* species (Figs. 1.1, 1.3). Likewise, an unidentified eccrinid (sample 1067) dissected from a freshwater beetle (noted by MMW as another possible *Enterobryus* sp.) is on a branch nearer to the other Coleoptera clade (*Leidyomyces*) and the freshwater-associated *Arundinula opeongoensis*. The remaining samples were taken from millipede hosts, but even within this single host type, they are divided among several well-

supported clades. The main obstacle to taxonomic assignation within *Enterobryus* mirrors that of the Eccrinidae in general: the paucity of informative morphological characters. The production of different spore types (a total of nine for *Enterobryus*) and cells of unknown function combined with intraspecific variability, even along the length of a single host gut (Lichtwardt 1954, 1958; Hibbits Galt 1978; Lichtwardt et al. 2001) have confounded attempts to evaluate genera and species solely with morphology. Certainly new species descriptions would benefit from such thorough statistical analyses of morphological variability as recently published for *E. luteovirgatus* (Contreras & Cafaro 2013). However, genomic samples of species and their putative conspecific morphotypes should be sought in any future collection effort, to disambiguate the taxonomy and contribute to a robust, integrated morpho-phylogenetic framework.

Evolution of the Protist Trichos

Amoebidium is supported as the earliest-diverging protist tricho by at least one method on all three trees. This is an evolutionary scenario that aligns with that hypothesized by Lichtwardt (1986), who proposed that the ancestral [protist] tricho probably had promiscuous affiliations with the exoskeleton of its hosts, and upon repeated ingestion of spores by the arthropod eventually became adapted to, and took up residence in, the gut leading to a *Paramoebidium*-like ancestor. However, this hypothesis assumes a monophyletic protist tricho clade, and while that relationship is not rejected by topology tests, (Table 1.6) the position of *Ichthyophonus* remains unresolved. Certainly, clarification of the relatedness of *Ichthyophonus* to the protist trichos is essential to understanding the evolution of characters within the Trichomycina. For instance, if *Ichthyophonus* is truly sister to *Paramoebidium* + the Eccrinidae, then an *Amoebidium*-

like ancestor would have had to transition not just from the outside to the inside of the host, but also from invertebrate to vertebrate, commensal to parasitic relationship, as well as have required a loss of the holdfast structure. Ancestral state reconstructions of growth forms (Fig. 1.7) give full probability to "hypha with holdfast" as the state for the Trichomycina ancestor. Furthermore, they support the Trichomycina ancestor as living in freshwater (Fig. 1.4). If so, this would imply that *Ichthyophonus* evolved first in freshwater fishes and then secondarily adapted to marine hosts. *Ichthyophonus* is known to infect anadromous fishes, implicating a possible evolutionary pathway for transmission between habitats and hosts. The *Amoebidium*-like ancestor in this scenario might have gained access to the fish host in a similar manner as outlined above: by repeated ingestion of prey insects and small crustaceans with it attached.

Furthermore, *Ichthyophonus* as sister to *Paramoebidium* + the Eccrinidae would suggest either a reversion to ancestral states (e.g. hypha with holdfast structure, association with arthropod hosts) or independent reacquisition of those states in the *Paramoebidium* + the Eccrinidae ancestors. These explanations are clearly not as parsimonious as a monophyletic protist tricho clade for those traits. Nevertheless, under either topology, *Paramoebidium* and the Eccrinidae remain sister taxa, although this is only supported by BI in the large rDNA tree (Fig. 1.1), and not supported by any method in the three gene or small rDNA trees (Figs. 1.2, 1.3). Interestingly, the life history and phylogenetic position of *Parataeniella* may provide insight to the evolutionary history of the two groups. *Parataeniella* is the earliest-diverging eccrinid of those represented here, and it is the only eccrinid with hyphae that produce spores both basipetally and in a holocarpic manner (i.e. endospores) (Fig. 1.8) that strongly resembles the amoebiids.

Possibly, *Parataeniella* reflects an intermediate form between more derived eccrinids and the hypothetical ancestor.

However, whether or not such characters may be symplesiomorphic is an important consideration as well. For example, the production of motile amoeboid propagules is not unique to the Amoebidiidae, as *Abeoforma* (Marshall & Berbee 2010), *Creolimax* (Marshall et al. 2008) and *Psorospermium* (Vogt & Rug 1999) all produce these forms at certain stages in their life cycle (Fig. 1.8), and the amoebae of the latter two were observed to encyst after a period of active crawling. While the life cycle of *Ichthyophonus* has not been fully resolved, Kocan et al. (2013) indicated the presence of infectious amoeba-like cells in the blood of its hosts. After migrating through the body, these cells settle in the tissue, grow larger and form "schizonts" with thick cell walls. These 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), but a planktonic phase has also been hypothesized, with evidence of infective cells released from epidermal lesions on infected hosts (Kocan et al. 2010). From this point of view, the *Ichthyophonus* life cycle appears to contain the amoeba-cyst stages of the Amoebidiidae, but whether spores (either endo- or cystospores, as defined for the protist trichos) are formed at some point is unclear. Possibly the cells released through the epidermal lesions are spores, but they were not identified as such (see Fig. 3 in Kocan et al. 2010). All character states of spore production were significantly correlated with the phylogeny in both the ASR (Fig. 1.8) and the BaTS analyses (Table 1.7), but the proportion of missing data and incomplete life cycle descriptions of some taxa preclude evolutionary inferences of homology. Despite this caveat, the most

parsimonious explanation is that both cystospores and amoeboid propagules are ancestral characters to at least the prostist tricho clade and order Eccrinida, respectively. Undoubtedly, future addition/emendation of life histories for Eccrinida taxa and discovery and inclusion of new members will clarify the homology versus symplesiomorphy of characters.

Ancestral State Reconstructions

Ancestral state reconstructions suggest that habitat states (Fig. 1.4) have independently arisen within the order more than once. For the marine character there appear three separate origins: at least one among the Eccrinidae, the Piridae (*Abeoforma*, *Pirum*) and the Creolimacidae (*Anurofeca*, *Creolimax*, *Sphaeroforma*) clades. Similarly, there are distinct freshwater and terrestrial clades, (Fig. 1.4), but the unresolved placement of *Psorospermium, Caullerya*, and the TMS, and the potential influence of long-branch attraction complicate interpretation of the reconstruction. Thus far, the Eccrinidae is the only group with multiple terrestrial taxa, and the results implicate that habitat as the ancestral state of the clade. If so, a reversion to freshwater and a transition to marine hosts would have occurred among the remaining eccrinid genera. On the other hand, there are at least two defined freshwater origins: one for the trichos and one for the clade including *Anurofeca* and a few environmental clones. Interestingly, these clones were collected from nutrient-rich freshwater environments such as effluent from a domestic wastewater treatment plant (Evans & Seviour 2011) and a peat bog in the mountains of Switzerland (Lara et al. 2010). Several of the marine clones were sampled from anoxic or low oxygen environments (Takishita et al. 2005; Edgcomb et al. 2011; Takishita et al. 2007), including one from a salt crust sample from a hypersaline lake

(Heidelberg et al. 2013). Clearly, further investigation into the role of ichthyophonids in these extreme habitats is warranted.

There has been little indication of host specificity among non-tricho ichthyophonids to date (Glockling et al. 2013), but the broad categorization of host type used here shows that insect, crustacean, millipede, and vertebrate states are significantly correlated with the phylogeny in the BaTS (Table 1.7) and ASR (Fig. 1.5) analyses. *Ichthyophonus* and *Anurofeca* are the only taxa known from vertebrate hosts within the order, and the topology suggests independent origins of these associations (Fig. 1.5). The arthropod character states are mainly associated with the protist trichos, whereas "other invertebrate" is applied to members of Sphaeroformina, some of which have been isolated from a wide range of hosts. For example, *Creolimax fragrantissima* was collected from peanut worms (*Phascolosoma agassizii*), sea cucumbers (*Leptosynapta clarki*), and chitons (*Corella* sp. and *Katharina tunicata*) (Marshall et al. 2008). This reinforces the suggestion that early-diverging members of Eccrinida (e.g. *Anurofeca*, *Creolimax*, *Sphaeroforma*) likely have generalist host associations, the ancestors of which subsequently diverged into marine, freshwater, and terrestrial specialists (Marshall et al. 2008). Both *Amoebidium* (Lichtwardt et al. 2001) and *Ichthyophonus* (Rowley et al. 2013) have been reported as generalists, therefore only members of the Eccrinidae and *Paramoebidium* are suggested to have host specificity (in terms of the number of hosts a single species associates with). Although, as recently described by Poulin et al. (2011), host specificity may occur at different levels beyond the simple number of species a symbiont associates with. Symbionts may exhibit structural, phylogenetic, or geographic

specificity (Poulin et al. 2011), aspects that can be tested among future collections of protist trichos and other ichthyosporeans.

Stage of the host at the time of infection by these symbionts (Fig. 1.6) is a factor that may influence the evolution of members of the Eccrinida, as well as indicate possible routes of transmission between hosts. The "juvenile" and "adult" states were significantly correlated with the phylogeny (Table 1.7), whereas "both" was not. *Caullerya*, *Ichthyophonus*, *Psorospermium*, and the TMS are the few taxa characterized as affecting both juvenile and adult forms of their hosts. *Paramoebidium* is only recorded from juvenile stages (nymphs and larvae), whereas nearly all eccrinids are found in adult forms. The two exceptions are *Leidyomyces* and *Lajasiella*, which infest both larvae and adults in the former and just the larvae in the latter case (Lichtwardt et al. 2001). Intriguingly, both of these taxa are associated with terrestrial beetles, as is the TMS. During surveys of the TMS, Lord et al. (2012) attempted to sterilize eggs in order to grow an uninfected population of hosts, but were unsuccessful, despite no microscopically visible evidence of infection. As the TMS was found to heavily infest the testes, and was subsequently passed to the female during mating, vertical transmission between hosts cannot be ruled out. If so, these symbionts would presumably have to cope with the metamorphosis of their host, but as the beetles do not transition between environments (i.e. aquatic to terrestrial as with mayflies, stoneflies, black flies etc.) it may be possible that infections of the larvae and adults occur as separate events. Conversely, the Amoebidiidae, having wall-less amoeboid propagules (Whisler 1968; Dang & Lichtwardt 1979), could be restricted to the aquatic habitat, thus only associating with whatever hosts are available in the water column. *Amoebidium*, attaching to the

exterior, has a range of hosts and depending on their life cycle may associate with juvenile (e.g. midges) or both juvenile and adult (e.g. *Daphnia*, which never leaves the water). Whether protist trichos can be transmitted vertically between hosts (e.g. in cysts of the ovaries as with the fungal trichos) is unknown, but seems unlikely as none (except *Enterobryus borariae*, discussed below) have been observed to penetrate into the tissues of the host.

The growth forms of most members of the order are variable (Fig. 1.7) and significantly correlated with the phylogeny (except "spherical/plasmodial/hyphal") (Table 1.7), and range from simple spheres to plasmodia-like to hyphal. The trichos are clearly distinguished from the rest of the taxa by their hyphal growth with a holdfast. This structure has logically been assumed to be an adaptation to allow them to maintain their residence in the vicinity of high nutrient availability (inside the gut for the Eccrinidae and *Paramoebidium*, and on the exoskeleton near the anus or mouth for *Amoebidium*) (Lichtwardt et al. 2001). As *Ichthyophonus* invades the tissues of its host, the holdfast would presumably be unnecessary. *Abeoforma*, *Creolimax*, *Pirum*, and *Sphaeroforma* all were isolated from the digestive tract of their hosts, and their apparent lack of holdfast structures could reflect the physiological differences of their hosts' guts (e.g. no chitinous lining) and/or a facultative association. *Abeoforma* and *Sphaeroforma* cells were able to attach to debris in culture (Marshall & Berbee 2010; Marshall & Berbee 2013), but whether they are capable of adhering to the gut wall or not is unknown. Although most of the non-tricho taxa grow as spherical/ovoid cells with thick walls, *Abeoforma* presented plasmodial and amoeboid forms in culture. Interestingly, transitions between forms for *I. hoferi* have been observed at different pH levels

(Okamoto et al. 1985) and Lichtwardt (1958) suggested that physiological gradations along the digestive tract could contribute to morphological variation and niche partitioning among *Enterobryus* spp. Thus, it is important to consider whether these various forms are homologous within clades, and to what degree the environment affects the growth of different species. Polymorphisms were deliberately coded as separate characters here such that they would be evaluated as true variation, but the underlying assumption is that these polymorphisms arose simultaneously and are correlated with one another. Certainly such an assumption is quite challenging to test, but the presence of these polymorphic forms in several early-diverging clades could suggest such a polymorphic ancestor.

The production of propagules has likewise been linked to the condition of the host. For instance, *Amoebidium* (Whisler 1968), *Paramoebidium* (Dang & Lichtwardt 1979), and *Psorospermium* (Vogt & Rug 1999) all produce motile amoebae upon molting or injury (for the Amoebidiidae) and death (*Psorospermium*) of the host. Furthermore, some eccrinids have been noted to produce thick-walled primary spores only upon molting of their hosts (Lichtwardt et al. 2001). Similarly, passage through the gut of the tadpole host triggered replication of *Anurofeca* cells (Beebee & Wong 1993). These interactions point to varying degrees of host-symbiont interaction, depending on the molecular mechanisms that elicit the responses. For example, Whisler (1968) found that calcium, glucose and some amino acids contributed to amoebagenesis in *Amoebidium parasiticum*, but that whole homogenate of the host provoked the greatest amoebagenic response. Whether such responses occur in the marine non-tricho taxa remains to be investigated. *Abeoforma*, *Creolimax*, *Pirum*, and *Sphaeroforma* were collected from the

stomach content of their host, but if they were simply ingested by the animal or if the gut is the preferred habitat of the protists remains to be established.

Spore production is significantly associated with clades on the phylogeny (Fig. 1.8, Table 1.7). Among those taxa that produce amoebae, morphological variation has been noted that could be further investigated as a source of new taxonomical characters. *Psorospermium* amoebae were shown to have "filose pseudopodia" (Vogt & Rug 1999) that somewhat resemble the "uroidal adhesive filaments" recently reported on amoebae of *Paramoebidium ecdyonuridae* (see Figs. 30-33 in Valle 2014a). No other members of Amoebidiidae are known to have these uroidal filaments, but Valle (2014a) noted that they might be more common than previously thought, as they are difficult to see using light microscopy and the amoeboid stage is not always observed among amoebidiid collections. Possibly, endospore characters could be taxonomically informative at a finer level than simple size dimensions, as traditionally used for members of the Eccrinidae. For example, some members of the Eccrinidae (e.g. *Astreptonema gammari, Palavascia sphaeromae, Taeniella carcini*) have "mucilaginous" or "gelatinous" appendages at the poles of their spores (Moss 1979). Future ultrastructural and ontological studies are warranted to determine the potential taxonomic utility of such features.

Finally, location in the host (Fig. 1.9) where these symbionts occur is a factor that may affect their evolution, as also indicated by the BaTS results (Table 1.7). Residence in the digestive tract is recovered as the ancestral state for the entire order (Fig. 1.9). This result, in combination with the topology, point to independent origins of lineages that invade the host tissues and live externally (i.e. *Amoebidium*). Among the former lineages only *Caullerya* and *Ichthyophonus* have clear pathogenic effects on their hosts (Lohr et

al. 2010; Gozlan et al. 2014). Despite the sometimes abundant growth in the nerve chord and testes, no apparent pathology was observed for beetles infected with the TMS (Lord et al. 2012). In parallel, *Psorospermium* grows in the connective tissues of crayfish and has been associated with mortalities, but whether it contributed to these deaths is still unclear (Bangyeekhun et al. 2001). On the other hand, *Anurofeca* is not known to cause direct pathogenicity, but was demonstrated to contribute to interference competition between tadpoles of different species (Bardsley & Beebee 2001). For the trichos, the default assumption (or null hypothesis) has been commensalism. The only member of the order to solely inhabit the foregut of it host, *Enteromyces callianassae*, is also one of the few whose host interactions have been studied. Kimura et al. (2002) suggested the possibility of a mutualistic interaction with the host via supplementation of digestive enzymes within the host stomach in a comparison of infected and uninfected shrimp populations. In a putative parasitic interaction, Lichtwardt (1958) observed unusual, cyst-like spores of *Enterobryus borariae* located outside of the gut lining of the millipede host, penetrating through the lining as it germinated into the interior of the gut. Together, these previous observations and the results presented here could signify that host-symbiont interactions are decoupled from the location of infection for some Eccrinida taxa. That is, symbionts that have no apparent direct effects on the host, even when invasive to the tissues, could affect indirect consequences, such as that shown for *Anurofeca*. Additional studies comparing infected and uninfected host populations, or employing experimental methods could clarify the nature of these symbiotic associations.

Conclusion

Whereas several studies reporting new collections (Strongman 2007; Valle & Santamaria 2009; Hernández Roa et al. 2009; Hernandez Roa & Cafaro 2012; William & Strongman 2013) or describing new species (White et al. 2006; Strongman & White 2008; Strongman & White 2006; Bench & White 2013; Contreras & Cafaro 2013; Valle 2014a; Valle 2014b) have been published on individual members of the protist trichos, relatively little phylogenetic work or morphological comparisons across the group have been published since they were recognized outside Fungi in 2000 (Benny & O'Donnell 2000; Ustinova et al. 2000) and 2005 (Cafaro 2005). As the taxonomic position of the Amoebidiidae and Eccrinidae has been tenuously linked to Fungi, the terminology used to describe their morphological and life history characters are all rooted in this mycological background. The challenge going forward, therefore, is to reevaluate these features in the context of their relationship to other ichthyosporeans. This will be complicated as relatively little is yet known about the class as a whole, and trichos have unique features even in comparison to their relatives (e.g. basipetal propagule formation, hyphal forms, spore polymorphism). Nevertheless, a paradigm shift has occurred and the results presented here highlight what has been suspected by trichomycetologists for some time: few of the characters used to delineate species thus far are informative of the evolutionary relationships among the trichos as a group. Reevaluation of traits such as spore types and ontogeny, propagule appendages/filaments, host specificity, and hyphal polymorphism is necessary. In other words, a shift from a morphological species concept to an integrative, or even strictly phylogenetic concept if no practically measurable (e.g. nuclear number or functional role of different propagules), homologous traits are found,

is needed. Different concepts may be appropriate for different clades. For example, Marshall and Berbee (2013) found that morphology was not sufficient to distinguish samples of *Sphaeroforma* that were phylogenetically distinct, and therefore used the genealogical concordance phylogenetic species recognition method to delimit species. This method would not work (or at least would require an unrealistic amount of resources) for some *Paramoebidium* spp. and most of the Eccrinidae, however, because it utilizes multiple loci from multiple samples of putative species, an arduous task for these unculturable taxa.

The inability to culture these organisms has made and will continue to make biological evaluation markedly challenging, as observations must be restricted to the moments following host dissection, emphasizing the need and opportunity for renewed cultivation efforts. Not only that, but protist tricho diversity is undersampled and considerably underestimated, potentially fragmenting and distorting the evolutionary signal gleaned from current knowledge. Tracking the development of spores and hyphae, understanding their functional roles, both in regard to the life history of the protist and its relationship to its host, and even determining dispersal mechanisms are all gaps yet to be bridged, despite first being discovered over 150 years ago (Leidy 1849). Nevertheless, the results presented here illustrate that phylogenetic tools can be critical for elucidating evolutionary trends and providing platforms from which to launch explorations of new hypotheses for these taxa. For the trichos in particular: 1) the Amoebidiidae may not by monophyletic and the traditional family structure within the Eccrinidae is not supported; 2) diversity within the Trichomycina is great and sampling has yet to encompass the range of species richness; 3) host type is shown to be a more taxonomically informative

character than previously thought for *Paramoebidium* and the Eccrinidae; and 4) a traditional morphological species concept is insufficient for species delimitation for many Eccrinida taxa. Fortunately, with the current reduced cost and relative ease of DNA sequencing, progress in our understanding of these enigmatic microorganisms may advance rapidly.

Table 1.2 List of samples with their sample code, collection code, host information, collection location, PCR protocol, and GenBank accession numbers. Dashes indicate no sequence data obtained/available for that gene.

Sample	Collection code	Species	Host (Order, Family, Genus)	Collector & Date	Collection Location	PCR protocol/ GenBank Accession# 18S	28S	HSP70
38	ME-2-W3	*Paramoebidium* sp.	Ephemeroptera, Leptophlebiidae, *Leptophlebia*	MMW 4/1999	Salmon Pond, 44°38'N, 68°04'W; Hancock Co., Maine, USA	194[C]	183	-
41	NS-6-W8-10	*Paramoebidium* sp.	Ephemeroptera, Leptophlebiidae, *Paraleptophlebia*	MMW 9/1998	Small, spring-fed brook (Black Brook) at Little Nine Mile River, 45°03.90'N, 63°35.47'W; Halifax, Nova Scotia, Canada	194	183	901[C]
42	NY-3-W7	*Paramoebidium* sp.	Plecoptera	MJC 11/1998	Enfield Creek, Treman Park, temp. 4.5°C, 42°23.83'N, 76°33.07'W; Ithaca, New York, USA	194[C]	-	-
122	RMBL-75-1	*Paramoebidium* sp.	Ephemeroptera, Siphlonuridae	RWL 8/1995	Rocky Mountain Biological Laboratory, East River upstream from bridge near Avery Creek, temp. 17.5°C; Crested Butte, Colorado, USA	177	-	-
197	AFR-9	*Paramoebidium curvum*	Diptera, Simuliidae		South Africa	194	183	-
209	FRA-1-14	*Amoebidium parasiticum*	Cladocera, Daphniidae, *Daphnia*	RWL 6/1968	Small, stagnant pool at S.W. junction of Rt. N. 109 to Courpouiran and D5E1, 43°36.20'N, 3°48.41'E; Montpellier, Herault, France	AF274051.1	DQ273802.1	901
303	NS-34-W17	*Paramoebidium* sp.	Plecoptera, Nemouridae	MMW 10/2000	Small, spring-fed brook (Black Brook) at Little Nine Mile River, 45°03.90'N, 63°35.47'W; Halifax, Nova Scotia, Canada	194[C]	183	-
376	JAP-7-2	*Amoebidium parasiticum*	Diptera, Chironomidae, *Chironomus*	RWL 3/1964	Ditch on the side of the road leading to the Tropical Plant Experiment Station of Tokyo University, algae present in quantity; Shimokamo, Shizuoka Prefecture, Japan	194	183	-
377	A1a	*Amoebidium parasiticum*	Cladocera, Daphniidae, *Daphnia*	HCW 1959	California, USA	194	183	901
400	ARG-D4-C11	*Palavascia patagonica*	Isopoda, Sphaeromatidae	MJC 12/1998	Cabo Blanco, 47°12.18'S, 65°44.38'W; Santa Cruz, Argentina	194	183[C]	-
401 (1154)	ARG-D1-C15	*Palavascia patagonica*	Isopoda, Sphaeromatidae, *Exosphaeroma*	MJC 12/1998	Puerto Deseado, 47°45.39'S, 65°53.42'W; Santa Cruz, Argentina	194	183[C] (AY336695.1 not used)	901[C]

402	SC-1-C26	*Palavascia sphaeromae*	Isopoda, Sphaeromidae, *Sphaeroma*	MJC 7/1998	Folly Beach, rock groin adjacent to Center Street Pier, high tide line, from aggregations found among shell debris beneath rocks, 32°39.272'N, 79°56.396'W; Charleston Co., South Carolina, USA	194[C]	-	-
444	HN-3	*Leidyomyces attenuatus*	Coleoptera, Passalidae, *Mastochilus quaestionis*	HN 8/1998	Springbrook, Lamington National Park, Queensland, Austrailia	203	183	-
446	KS-61-W20	*Paramoebidium* sp.	Ephemeroptera, Siphlonuridae, *Siphlonurus*	MMW 5/1998	Ephemeral stream, University of Kansas Field Station, Nelson Environmental Study Area; Lawrence, Kansas, USA	194	-	-
447	MA-3-W51	*Paramoebidium* sp.	Plecoptera, Taeniopterygidae, *Taeniopteryx*	MMW 12/1997	Winnetuxet River where it drains a small pond near where the river crosses State Highway 58, approximately 0.8 km south of Plympton and 2 km north of the junction with State Highway 44, 41°57.2'N, 70°48.4'W; Plymouth Co., Massachusetts, USA	-	183[*]	-
449	MO-19-W15	*Paramoebidium* sp.	Ephemeroptera, Baetidae, *Fallceon*	MMW 3/1999	East tributary of Turkey Creek, temp. 9°C, 37°24.40'N, 93°32.86'W; Polk Co., Missouri, USA	194[C]	183[C]	-
450	NF-10-W15	*Paramoebidium* sp.	Diptera, Simuliidae, *Simulium venustum/verecundum*	MMW 5/1999	Beachy Cove Brook at Witch Hazel Rd., 47°35.82'N, 52°50.84'W; Newfoundland, Canada	194	183[C]	-
457	RMBL-71-3	*Paramoebidium* sp.	Ephemeroptera, Ameletidae, *Ameletus velox*	RWL 8/1995	East River upstream from bridge near Avery Creek, Rocky Mountain Biological Laboratory; Crested Butte, Colorado, USA	194	183	-
458	RMBL-72-3	*Paramoebidium* sp.	Diptera, Simuliidae	RWL 8/1995	Willow Creek near outlet to Taylor Park Reservoir, temp. 15.5°C, Rocky Mountain Biological Laboratory; Crested Butte, Colorado, USA	194	-	-
459	PA-2-W2	*Paramoebidium* sp.	Plecoptera, Nemouridae, *Shipsa*	MMW 3/2000	Unnamed stream on North side Route 286, 1.3 mi. NE of Hillsdale; Indiana Co., Pennsylvania, USA	194	183[C]	-
466	NS-24-W16	*Paramoebidium* sp.	Ephemeroptera, Leptophlebiidae	MMW 10/2000	Small, pebble-bottom stream near James River (2.1 km off Hwy. 104), temp.10.5°C, 45°36.12'N, 62°11.54'W; Antigonish, Nova Scotia, Canada	-	183	-
467	VT-3-W1	*Paramoebidium* sp.	Ephemeroptera, Ephemerellidae, *Drunella*	MMW 8/2000	Just North of Warren County line along Hwy. 100 S., where Mad River crosses the highway, temp. 16°C, 44°04.25'N, 72°51.65'W; Green Mountain National	194[C]	-	-

					Forest, Vermont, USA			
473	UT-1-W14	*Paramoebidium* sp.	Diptera, Simuliidae, *Simulium arcticum* complex	MMW 8/2001	Provo River, temp. 14.5°C, 40°19.25'N, 111°38.12'W; Utah Co., Utah, USA	194 + 203	183	-
488	UT-2-W5	*Paramoebidium* sp.	Ephemeroptera, Baetidae, *Baetis bicaudatus*	MMW 8/2001	Bridal Veil Falls on Provo River Road, elev. 1562 m., temp. 11.5°C, 40° 20.69'N, 111° 36.38'W; Utah Co., Utah, USA	-	183[C]	-
504	NS-35-W8	*Paramoebidium* sp.	Ephemeroptera, Ephemerellidae, *Ephemerella subvaria*	MMW 12/2001	Cape Breton, Big Intervale Cape North, Cabot Trail, temp. 0.5°C, 46°49.74'N, 60°37.07'W; Nova Scotia, Canada	194[C]	-	-
506	NS-35-W14	*Paramoebidium* sp.	Ephemeroptera, Ephemerellidae, *Eurylophella*	MMW 12/2001	Cape Breton, Big Intervale Cape North, Cabot Trail, temp. 0.5°C, 46°49.74'N, 60°37.07'W; Nova Scotia, Canada	194	183	-
508	NS-35-W18	*Paramoebidium* sp.	Ephemeroptera, Leptophlebiidae, *Paraleptophlebia*	MMW 12/2001	Cape Breton, Big Intervale Cape North, Cabot Trail, temp. 0.5°C, 46°49.74'N, 60°37.07'W; Nova Scotia, Canada	194	183	-
511	NS-35-W22b	*Paramoebidium* sp.	Ephemeroptera, Ephemerellidae, *Eurylophella*	MMW 12/2001	Cape Breton, Big Intervale Cape North, Cabot Trail, temp. 0.5°C, 46°49.74'N, 60°37.07'W; Nova Scotia, Canada	194	183	901
514	LA-9-W2a	*Arundinula* sp.	Decapoda, Astacidae (Freshwater crayfish)	MMW 2/2002	Commercial pond crayfish purchased from Country Corner convenience store; Baton Rouge, Louisiana, USA	-	183	-
525	NOR-3-1	*Paramoebidium* sp.	Plecoptera, Capniidae, *Capnopsis schilleri*	RWL 5/2002	Heggelielva River near Skansebakken parking area, 60°01.19'N, 10°35.20'E; Oslo County, Norway	194	183	-
526	NOR-3-W2	*Paramoebidium* sp.	Plecoptera, Nemouridae, *Protonemura* cf. *meyeri*	MMW 5/2002	Heggelielva River near Skansebakken parking area, 60°01.19'N, 10°35.20'E; Oslo County, Norway	194	183	901
531	NOR-4-W9	*Paramoebidium* sp.	Plecoptera	MMW 5/2002	Small stream next to house at old bridge just N. of newer bridge, temp 6.5°C, 60°01.14'N, 10°33.88'E; Oslo County, Norway	194	183	-
533	NOR-7-W2	*Paramoebidium* sp.	Plecoptera, Nemouridae, *Protonemura* cf. *meyeri*	MMW 5/2002	Stream on curve of Sørkendalsveien Rd., with parking just NW of stream. Heading northward, 0.65 km to church, temp. 7°C, 60°00.72'N, 10°36.84'E; Oslo county, Norway	194	183	901
536	NOR-5-W14	*Paramoebidium* sp.	Plecoptera, Nemouridae, *Nemurella pictetii*	MMW 5/2002	Small, unnamed stream off Kampeveien Rd., reached from Lommedalsveien, at wooden bridge, temp. 6.5°C, 59°59.19'N, 10°28.64'E; Akershus county, Norway	194	183	-

538	NOR-5-2	*Paramoebidium* sp.	Plecoptera, Taeniopterygidae, *Brachyptera risi*	RWL 5/2002	Small, unnamed stream off Kampeveien Rd., reached from Lommedalsveien, at wooden bridge, temp. 6.5°C, 59°59.19'N, 10°28.64'E; Akershus county, Norway	194	183	901
543	NOR-10-W1	*Paramoebidium* sp.	Plecoptera, Capniidae, *Capnopsis schilleri*	MMW 5/2002	Small stream next to house at old bridge just N. of newer bridge, temp 6.5°C, 60°01.14'N, 10°33.88'E; Oslo county, Norway	194	183	-
546	NOR-10-W10a	*Paramoebidium* sp.	Plecoptera, Capniidae, *Capnopsis schilleri*	MMW 5/2002	Small stream next to house at old bridge just N. of newer bridge, temp 6.5°C, 60°01.14'N, 10°33.88'E; Oslo county, Norway	194	183	-
551	NOR-16-W3	*Paramoebidium* sp.	Plecoptera, Taeniopterygidae, *Brachyptera*	MMW 5/2002	Dalbekken stream near bungalow, temp. 6.0°C, 60°29.99'N, 10°08.00'E; Buskerud county, Norway	194	183	901
558	NOR-10-W15	*Paramoebidium* sp.	Plecoptera	MMW 5/2002	Small stream next to house at old bridge just N. of newer bridge, temp 6.5°C, 60°01.14'N, 10°33.88'E; Oslo county, Norway	194	183	-
564	NOR-21-W5	*Paramoebidium* sp.	Plecoptera, Leuctridae, *Leuctra hippopus*	MMW 5/2002	Stream at bottom of Haga foss, a tributary of the Ekso River, about 0.4 km W. of Ekse field station, 60°50.11'N, 06°19.48'E; Hordaland County, Norway	194	183	-
566	NOR-22-W10	*Paramoebidium avitruviense*	Plecoptera, Chloroperlidae, *Siphonoperla burmeisteri*	MMW 5/2002	Small stream under Route E16 W. of Voss, 0.6 km E. of Norske Vandrerlyem parking lot and just W. of sign Vosse Vaugen; going W., 1.3 km from turnoff to Kvåle, temp. 6.5°C, 60°37.52'N, 06°23.65'E; Hordaland County, Norway	203	-	-
591	GUA-X-13	*Leidyomyces* sp.	Coleoptera, Passalidae, *Passalus puntatostriatus*	ACB 10/2001	San Jose Pinula, 1615 m.; Guatemala	194	183	-
593	GUA-X-18	*Leidyomyces* sp.	Coleoptera, Passalidae, *Publius agassizi*	ACB 10/2001	Villa Nueva, 1524 m.; Guatemala	194	183[C]	-
606	NS-35-L2	*Paramoebidium* sp.	Ephemeroptera, Ephemerellidae, *Ephemerella*	L 12/2001	Cape Breton, Big Intervale Cape North, Cabot Trail, temp. 0.5°C, 46°49.74'N, 60°37.07'W; Nova Scotia, Canada	194	183	-
614	CAL-17-L1	*Paramoebidium* sp.	Ephemeroptera	L 6/2002	Hummingbird Creek, on Calistoga Road, temp. 16°C, 38°31.93'N, 122°36.25'W; Sonoma Co., California, USA	194	-	-

616	CAL-17-L6	*Paramoebidium* sp.	Ephemeroptera	L 6/2002	Hummingbird Creek, on Calistoga Road, temp. 16°C, 38°31.93'N, 122°36.25'W; Sonoma Co., California, USA	194	-	-
618	AR-30-C7	*Paramoebidium* sp.	Ephemeroptera, Siphlonuridae	MJC 5/1998	Hock Creek (Hock Creek Rd. at Rte. 74), 36°03.13'N 93°49.22'W, temp. 0°C, Madison Co., Arkansas, USA	194	183	901
619	AR-30-C9	*Paramoebidium* sp.	Ephemeroptera, Siphlonuridae	MJC 1/2002	Hock Creek (Hock Creek Rd. at Rte. 74), 36°03.13'N 93°49.22'W, temp. 0°C, Madison Co., Arkansas, USA	194	183	-
622	AR-31-C31	*Paramoebidium* sp.	Ephemeroptera, Siphlonuridae	MJC 1/2002	Osage Creek (Rd. 927 at Rte. 74, near Dog Branch Cemetery, but from Osage Creek proper rather than the Dog Branch), temp. 5°C, 36°12.18'N, 93°21.56'W; Caroll Co., Arkansas, USA	194	183*	-
664	NOR-35-3	*Paramoebidium* sp.	Plecoptera, Taeniopterygidae, *Taeniopteryx nebulosa*	RWL 8/2002	Trib. of main river. Kanten Rd. running parallel to stream that crosses Kongsvegen Rd. Going N., stream is 0.6 km to E6 highway, temp. 14.5°C, 61°59.61'N, 09°14.23'E; Oppland County, Norway	194	183	-
671	NOR-40-W2	*Paramoebidium* sp.	Plecoptera, Taeniopterygidae, *Taeniopteryx nebulosa*	MMW 8/2002	Greitbekken River, a tributary of Jørstadelva on road S. off Hwy. 763 at dead end road, temp. 17°C, 64°10.86'N, 12°17.07'E; Norge County, Norway	194	183	-
674	NOR-40-W8	*Paramoebidium* sp.	Diptera, Simuliidae	MMW 8/2002	Greitbekken River, a tributary of Jørstadelva on road S. off Hwy. 763 at dead end road, temp. 17°C, 64°10.86'N, 12°17.07'E; Norge County, Norway	194	183C	-
680	NOR-50-W2	*Paramoebidium* sp.	Plecoptera, Taeniopterygidae, *Taeniopteryx nebulosa*	MMW 8/2002	Stream crossing Rte. 74, E. of NOR-49 boggy pond, temp. 18°C, 64°28.81'N, 13°12.69'E; Norge County, Norway	194	183	901
681	NOR-53-W1a	*Paramoebidium curvum*	Diptera, Simuliidae	MMW 8/2002	Aunelva stream draining S., temp. 17.5°C, 64°19.66'N, 13°35.52'E; Norge County, Norway	194	183C	-
690	NOR-54-2	*Paramoebidium* sp.	Ephemeroptera, Baetidae, *Baetis rhodani*	RWL 8/2002	Small stream crossing Rte. 759 SSE of Steinkjer 31.5 km from field station, just before sign to Billakkering when going S., temp. 15.5°C, 63°57.64'N, 11°34.27'E; Norge County, Norway	194	183	-

691	NOR-54-W10	*Paramoebidium* sp.	Ephemeroptera, Baetidae	MMW 8/2002	Small stream crossing Rte. 759 SSE of Steinkjer 31.5 km from field station, just before sign to Billakkering when going S., temp. 15.5°C, 63°57.64'N, 11°34.27'E; Norge County, Norway	194	183	-
703	NOR-54-W17	*Paramoebidium* sp.	Ephemeroptera, Baetidae	MMW 8/2002	Small stream crossing Rte. 759 SSE of Steinkjer 31.5 km from field station, just before sign to Billakkering when going S., temp. 15.5°C, 63°57.64'N, 11°34.27'E; Norge County, Norway	194	183	-
715	NOR-61-W11	*Paramoebidium* sp.	Plecoptera, Taeniopterygidae, *Taeniopteryx nebulosa*	MMW 8/2002	Large waterfall: Kjaekerfossen. Just off Rte. 757, 63°50.25'N, 12°01.68'E; Norge County, Norway	194	183[C]	-
771	CA-22-W2c	*Enteromyces callianassae*	Decapoda, Callianassidae, *Neotrypaea* (=*Callianassa*) (Bay ghost shrimp)	MMW 6/2002	Mud flat in front of housing at Bodega Marine Reserve (along Westshore Rd.), Bodega Harbor, 38°19.05'N, 123°03.38'W; Bodega Bay, California, USA	194	183[C]	-
772	LA-9-W1g	*Arundinula* sp.	Decapoda, Astacidae (Freshwater crayfish)	MMW 2/2002	Commercial pond crayfish purchased from Country Corner convenience store; Baton Rouge, Louisiana, USA	194[C]	183	-
776	NOR-1-W1	*Parataeniella* sp.	Isopoda	MMW 5/2002	Boards harboring isopods on University of Oslo campus; Oslo, Norway	194	183	-
830	NS-X-17	*Paramoebidium* sp.	Plecoptera	DBS 4/2003	Small, spring-fed brook near Black Duck Lake, 44°42.10'N, 64°42.45'W; East Dalhousie, Nova Scotia, Canada	194	183	901
833	ONT-X-10	*Paramoebidium* sp.	Diptera, Simuliidae	DBS 5/2003	Small, spring-fed brook near Black Duck Lake, 44°42.10'N, 64°42.45'W; East Dalhousie, Nova Scotia, Canada	177	183[C]	-
870	CO-16-W12	*Paramoebidium* sp.	Trichoptera, Brachycentridae	MMW 6/2003	Beaver Creek, near Gunnison, temp. 9°C, 38° 29.77N, 107° 01.94W; Colorado, USA	-	183	-
872	FL-2-W7	*Enterobryus* sp.	Polydesmida, Paradoxosomatidae, *Oxidus gracilis*	MMW 2/2003	Archbold Biological Station, within 100 m. of buildings, 27.182843, -81.351786; Venus, Florida, USA	194	183	901
901	TN-27-A3	*Amoebidium appalachense*	Diptera, Chironomidae, *Chironomus*	AS 7/2004	Rock pools above Roaring Fork Creek, just upstream from bridge across river, elev. 560 m., 35°42.55'N' 83°28.65'W; Gatlinburg, Tennessee, USA	194	183	-
903	TN-27-W4	*Amoebidium* sp.	Diptera, Chironomidae, *Chironomus*	MMW 7/2004	Rock pools above Roaring Fork Creek, just upstream from bridge across river, elev. 560 m., 35°42.55'N' 83°28.65'W; Gatlinburg,	194	183	-

					Tennessee, USA			
904	TN-46-A6	*Amoebidium appalachense*	Diptera, Chironomidae, *Chironomus*	AS 8/2004	Rock pools on Little Pigeon River, and River itself, elev. 430 m., 35°44.1'N, 83°24.8'W; Gatlinburg, Tennessee, USA	194	183	-
921	TN-27-W1a	*Amoebidium appalachense*	Diptera, Chironomidae, *Chironomus*	MMW 7/2004	Rock pools above Roaring Fork Creek, just upstream from bridge across river, elev. 560 m., 35°42.55'N' 83°28.65'W; Gatlinburg, Tennessee, USA	194	183	-
935	PEI-X-12	*Paramoebidium* sp.	Plecoptera, Taeniopterygidae, *Taeniopteryx*	DBS 12/2003	Trout River (wide, slow moving stream) at Tyne Valley, 46°34'13.50"N, 063°55'44.18"W; Prince Edward Island, Canada	194	183	-
943	ALG-15-W1a	*Arundinula opeongoensis*	Decapoda, Astacidae (Freshwater crayfish)	MMW 5/2004	Opeongo Lake, Sproule Bay, Algonquin Park, 45°38.05'N, 78°21.84'W; Ontario, Canada	194	183	-
950	ONT-3-W6	*Enterobryus* cf. *euryuri*	Polydesmida, Polydesmidae, *Apheloria virginiensis corrugata*	MMW 5/2004	Webster Falls, Spencer Gorge, 43°16.77'N, 79°58.74'W; Hamilton, Ontario, Canada	194	183	-
1048	TN-38-W15	*Paramoebidium* sp.	Plecoptera	MMW 7/2004	Cades Cove, Tater Branch of Abrams Creek, at Cades Cove Loop Road, elev. 537 m., temp. 20°C, 35°36.49'N, 83°49.78'W; Tallassee, Tennessee, USA	194*	183	-
1049	ALG-9-W10	*Paramoebidium* sp.	Diptera, Simuliidae	MMW 5/2004	Fast flowing stream draining beaver pond beside Opeongo Rd., Algonquin Park, 45°36.19'N, 078°20.19'W; Ontario, Canada	194	183[C]	-
1052	OR-3-W12	*Paramoebidium* sp.	Plecoptera, Peltoperlidae, *Soliperla*	MMW 8/2005	New Belgium Creek, elev. 750 m., 44°13.5'N, 122°10.6'W; Blue River, Oregon, USA	194	183	-
1058	Boi-14-W5	*Parataeniella* sp.	Isopoda	MMW 4/2008	Boise River, Friendship Bridge near BSU campus, 43.605149, -116.203810; Boise, Idaho, USA	194	183	-
1067	CO-16-W6	*Enterobryus* sp.	Coleoptera	MMW 6/2003	Beaver Creek, small drain, near Gunnison, temp. 9°C, 38° 29.77N, 107° 01.94W; Colorado, USA	-	183	-
1115	MA-8-W4	*Alacrinella limnoriae*	Isopoda, Limnoridae, *Limnoria*	MMW 3/1998	Woods Hole, Massachusetts, USA	AY336703.1	-	-

1117	MA-7-W17	*Astreptonema gammari*	Amphipoda, Gammaridae	MMW 4/1998	Winnetuxet River where it drains a small pond near where the river crosses State Highway 58, approximately 0.8 km south of Plympton and 2 km north of the junction with State Highway 44, 41°57.2'N, 70°48.4'W; Plymouth Co., Massachusetts, USA	194	183	-
1118	SET-3-C3	*Astreptonema gammari*	Amphipoda, Gammaridae, *Gammarus*	MJC 8/2002	Source du Lez, temp. 19°C, 43°42.956'N, 3°50.938'E; Saint-Clément-de-Rivière, France	194	183	-
1120	MN-3-W6	*Astreptonema gammari*	Amphipoda, Gammaridae, *Gammarus*	MMW 3/1998	Woods Hole, Massachusetts, USA	AY336709.1	183	-
1121	WA-3-C3	*Astreptonema* sp.	Isopoda, Sphaeromatidae	MJC 3/1999	Eagle Cove, San Juan Island, Washington, USA	AY336706.1	183	-
1126	SPA-10-C2	*Eccrinidus flexilis*	Glomerida, Glomeridae, *Glomeris* (Pill millipede)	MJC 8/2002	St. Llorens del Munt, Barcelona, Spain	AY336698.1	183	-
1128	SPA-11-C45	*Eccrinidus flexilis*	Glomerida, Glomeridae, *Glomeris* (Pill millipede)	MJC 8/2002	Punta de la Mora, 41°7.89'N, 1°20.54'E; Tarragona, Spain.	AY336700.1	183	-
1133	CA-11-C4	*Enterobryus halophilus*	Decapoda, Hippidae, *Emerita* (Pacific sand crab)	MJC 7/2001	Salmon Creek Beach, Bodega Bay, California, USA	-	183 (AY336694.1 not used)	-
1135	KS-79-W2	*Enterobryus oxidi*	Polydesmida, Paradoxosomatidae, *Oxidus gracilis*	MJC 8/1999	Rice Woodland Tract, Kansas University Ecological Reserve; Lawrence, Kansas, USA	AY336710.1	183	-
1137	SPA-10-C6	*Enterobryus* sp.	Julida, Julidae	MJC 8/2002	St. Llorens del Munt, Barcelona, Spain	AY336711.1	183	-
1138	SPA-2-C10	*Enterobryus* sp.	Polydesmida, Polydesmidae, *Brachydesmus*	MJC 8/2002	Near Riera de Santa Fe, 41°46.41'N, 2°27.87'E; Montseny, Spain	AY336712.1	183	-
1139	CR-LS-C1	*Enterobryus* sp.	Diplopoda	MJC 7/1999	La Selva Biological Station, 10°25′19″N, 84°00′54″W; Heredia Province, Costa Rica	203	-	-
1141	MA-11-C1	*Enterobryus* sp.	Diplopoda	MJC 2/2002	Peach's Point up Beacon Rd. in Steep Swamp area, 42°31.096'N, 70°50.495'W; Essex, Massachusetts, USA	AY336701.1	183	-
1145	CA-12-C8	*Enteromyces callianassae*	Decapoda, Callianassidae, *Neotrypaea* (=*Callianassa*) (Bay ghost shrimp)	MJC 7/2001	Walker Creek Marsh, Tomales Bay, Marin, California, USA	AY336702.1*	183	-
1146	WA-1-C5	*Enteropogon sexuale*	Decapoda, Upogebiidae, *Upogebia* (Blue mud shrimp)	MJC 3/1999	False Bay, San Juan Island, Washington, USA	AY336705.1	183	-
1164	SC-4-C6	*Parataeniella dilatata*	Isopoda	MJC 11/1998	South Carolina, USA	194	183	-

1166	KS-48-C18	*Parataeniella armadillidii*	Isopoda, Armadillidae	MJC 1/1998	Rice Woodland Tract, Kansas University Ecological Reserve; Lawrence, Kansas, USA	194	-	-
1170	WA-1-C37	*Taeniella carcini*	Decapoda, Varunidae, *Hemigrapsus* (Purple shore crab)	MJC 3/1999	False Bay, San Juan Island, Washington, USA	AY336707.1	183	-
1172	MA-5-C17	*Taeniellopsis sp.*	Amphipoda, Talitridae, *Orchestia*	MJC 3/1999	Nobska Point, Barnstable Co., Massachusetts, USA	AY336704.1	AY336697.1*	-
1175	KS-61-W6	*Paramoebidium* sp.	Ephemeroptera, Siphlonuridae	MMW 5/1998	Breidenthal Biological Reserve, Douglas Co., Lawrence, Kansas, USA	AY336708.1	-	-
1176	AR-31-C7	*Paramoebidium* sp.	Ephemeroptera, Siphlonuridae	MJC 5/1998	Osage Creek, (Rd. 927 at Rte. 74, near Dog Branch Cemetery, but from Osage Creek proper rather than the Dog Branch), temp. 5°C, 36°12.18'N, 93°21.56'W; Caroll Co., Arkansas, USA	194	183	901
1181	ID-164-G2	*Paramoebidium* cf. *grande*	Diptera, Simuliidae, *Simulium piperi*	JG 6/2013	Cottonwood Creek, Military Reserve Park, 43°38.19'N, 116°14.28'W; Ada Co., Boise, Idaho, USA	194	-	-
1183	ID-157-G11	*Paramoebidium* sp.	Ephemeroptera, Baetidae, *Baetis*	JG 6/2013	Cottonwood Creek, Military Reserve Park, 43°38.19'N, 116°14.28'W; Ada Co., Boise, Idaho, USA	194	-	-
1185	ID-163-G5	*Enterobryus* sp.	Polydesmida, Polydesmidae, *Polydesmus*	JG 6/2013	Garden area of backyard in Northend neighborhood, 43°37.59'N, 116°12.47'W; Ada Co., Boise, ID, USA	194	183*	-
1187	ID-165-G2	*Enterobryus* sp.	Julidae, Parajulidae	JG 6/2013	Garden area of backyard in Northend neighborhood, 43°37.59'N, 116°12.47'W; Ada Co., Boise, ID, USA	194	-	-
1188	ID-165-G4	*Enterobryus* sp.	Polydesmida, Polydesmidae, *Polydesmus*	JG 6/2013	Garden area of backyard in Northend neighborhood, 43°37.59'N, 116°12.47'W; Ada Co., Boise, ID, USA	194	-	-
1193	ID-155-N1-2	*Ichthyophonus* sp.	Salmoniformes, Salmonidae, *Oncorhynchus* (Rainbow trout)	SL 3/2013	Clear Springs Foods, Buhl, Idaho, USA	194	-	-
1194	ID-155-N2-1	*Ichthyophonus* sp.	Salmoniformes, Salmonidae, *Oncorhynchus* (Rainbow trout)	SL 3/2013	Clear Springs Foods, Buhl, Idaho, USA	194	183*	901
1196	Mex-16-W4	*Paramoebidium curvum*	Diptera, Simuliidae, *Simulium*	MMW 11/2005	Xico. Road from Xico to Xico Viejo, Km. 2.5. Puente de la Virgen de Guadalupe, elev. 1667 m., temp. 14°C, 19°26.6779'N, 97°02.7579'W; Veracruz, Mexico	177 + 194	183C	-

1197	Mex-4-C1	*Enterobryus* sp.	Diplopoda	MJC 11/2005	San Andrés Tlalnelhuayocan, Rancho Viejo, Agüita Fría. Spring fed stream, Misty forest, elev. 1440 m., temp. 14.5°C, 19°31.18'N, 96°59.23'W; Xalapa, Veracruz, Mexico	203	-	-
1198	KY-5-P5	*Enterobryus* sp.	Diplopoda	PK 6/2010	Kentucky, USA	194	-	-
1200	DR-16-C8	*Paramoebidium* sp.	Ephemeroptera, Ephemerellidae	MJC 12/2006	La vega, Ébano Verde, Río Arroyazo, elev. 1010 m., temp. 17°C, Dominican Republic	194	183	-
1201	Mex-38-C8	*Leidyomyces* sp.	Coleoptera, Passalidae	MJC 8/2006	Submontane humid evergreen tropical forest of Los Tuxtlas (San Andrés Tuxtlas), Veracruz, Mexico	203 + 194	-	-
1207	ID-166-G6	*Paramoebidium* cf. *hamatum*	Ephemeroptera, Baetidae, *Baetis bicaudatus*	JG 6/2013	Cottonwood Creek, Military Reserve Park, temp. 14°C, 43°38.19'N, 116°14.28'W; Boise, Idaho, USA	194	183[C]	-
1210	NS-X-39	*Paramoebidium stipula*	Plecoptera, Nemouridae	DBS 2/2005	Small, spring-fed brook near Black Duck Lake, 44°42.10'N, 64°42.45'W; East Dalhousie, Nova Scotia, Canada	194	183	-
1214	KS-114-3	*Paramoebidium corpulentum*	Plecoptera, Capniidae, *Allocapnia*	RWL 1/2005	Rock Creek, temp. 1°C, 38°50.89'N, 95°26.71'W; Douglas Co., Kansas, USA	194[C]	183	-
1215	NS-X-29	*Paramoebidium* sp.	Plecoptera, Taeniopterygidae, *Taeniopteryx*	DBS 3/2004	Small river, through urban area on Sackville Dr. adjacent to Downsview Mall, 44°46.15'N, 63°41.22'W; Sackville, Nova Scotia, Canada	194	183	-
1217	OR-13-W1	*Paramoebidium* sp.	Ephemeroptera, Ephemerellidae, *Ephemerella aurivilli*	MMW 9/2005	First Creek on Rd. 12, temp. 12°C, elev. 1079 m., 44°27.57'N, 121°41.35'W; Camp Sherman, Oregon, USA	194	183	-
1218	TN-13-W20	*Paramoebidium* sp.	Ephemeroptera, Baetiscidae, *Baetisea carolina*	MMW 3/2004	Two confluent streams at junction of Little Cove, Mill Creek roads, 35°45.91'N, 83°34.18'W; Pigeon Forge, Tennessee, USA	194	-	-
1219	TN-38-W13	*Paramoebidium* sp.	Plecoptera, Perlidae, *Acromeuria*	MMW 7/2004	Cades Cove, Tater Branch of Abrams Creek, at Cades Cove Loop Road, elev. 537 m., temp. 20°C, 35°36.49'N, 83°49.78'W; Tallassee, Tennessee, USA	203	-	-
1223	Mex-17-W2	*Paramoebidium* sp.	Diptera, Simuliidae, *Simulium parrai*	MMW 11/2005	Xico. Road from Xico to Xico Viejo, Km 2. Cofre de Perote mountain. Little stream, elev. 1650 m., temp. 17.5°C, 19°25.9489N, 97°03.7669W; Veracruz, Mexico	194[C]	-	-
1225	OR-14-W1	*Paramoebidium*	Plecoptera, Peltoperlidae,	MMW	Jack Creek on Rd. 1230, temp. 8°C, elev.	194	183	901[C]

		sp.	*Yoraperla*	9/2005	928 m., 44°29.27'N, 121°41.88'W; Camp Sherman, Oregon, USA			
1227	CO-3-W13	*Paramoebidium* sp.	Ephemeroptera	MMW 9/1997	Onahu Creek, Rocky Mountain National Park, temp. 7.5°C, 40°22.8N, 105°51.0W; Grand Lake, Colordao, USA	194	183	-
1228	BC-5a-W6	*Paramoebidium* sp.	Ephemeroptera, Baetidae, *Callibaetis*	MMW 7/2006	Pond at South Slocan, 5a was still water, elev. 445 m., pH 7, temp. 20°C, 49°27.27'N, 117°28.18'W; Nelson, British Columbia, Canada	194	133	-
1234	NS-16-W2	*Paramoebidium* sp.	Ephemeroptera, Leptophlebiidae, *Leptophlebia* cf. *cupida*	MMW 6/1999	Medium river with boulders (Halfway Brook) in Cape Breton Highlands National Park, 46°48.53'N; 60°20.68'W; Nova Scotia, Canada	203*	-	-
1236	ID-65e-E1	*Paramoebidium* sp.	Diptera, Simuliidae	ERW 2/2010	Cottonwood Creek, Military Reserve Park, temp. 7.2°C; Ada Co., Boise, Idaho, USA	194	-	-
1237	ID-156-T1	*Paramoebidium* sp.	Plecoptera, Nemouridae	TP 5/2013	Cottonwood Creek, Military Reserve Park, temp. 10.5°C; Ada Co., Boise, Idaho, USA	203	-	-
1238	ATCC-PRA-283	*Sphaeroforma* sp. (*nootkatensis*)	Echinoida, Strongylocentrotidae, *Strongylocentrotus franciscanus* (Red sea urchin)			194	183C	-
1239	ATCC-PRA-283	*Sphaeroforma* sp. (*nootkatensis*)	Echinoida, Strongylocentrotidae, *Strongylocentrotus franciscanus* (Red sea urchin)			194	-	901C
1240	ATCC-PRA-284	*Creolimax fragrantissima*	Phleobranchia, Corellidae, *Corella* sp. (Transparent tunicate)			194	183	-
1242	ATCC-PRA-297	*Sphaeroforma arctica*	Amphipoda, Gammaridae, *Gammarus*			194	183	901
1247	SPA-X-74	*Paramoebidium avitruviense*	Plecoptera, Chloroperlidae, *Siphonoperla torrentium*	LGV 4/2013	Montseny Natural Park, Cànoves, Vallforners stream above the homonym marsh, elev. 530 m., 41°43'37.66'N, 2°20'12.31'E; Barcelona, Catalunya, Spain	-	183	-
1249	SPA-X-76	*Paramoebidium ecdyonuridae*	Ephemeroptera, Heptageniidae, *Ecdyonurus forcipula*	LGV 4/2013	Montseny Natural Park, Cànoves, Vallforners stream above the homonym marsh, elev. 530 m. (same site and date as the holotype), 41°43'37.66'N, 2°20'12.31'E; Barcelona, Catalunya, Spain	194	183C	-

1250	SPA-X-77	*Paramoebidium ecdyonuridae*	Ephemeroptera, Heptageniidae, *Ecdyonurus forcipula*	LGV 4/2013	Montseny Natural Park, Cànoves, Vallforners stream above the homonym marsh, elev. 530 m. (same site and date as the holotype), 41°43'37.66'N, 2°20'12.31'E; Barcelona, Catalunya, Spain	177	-	-
	FRA-1-14 (=NRRL20524, ATCC-32708)	*Amoebidium parasiticum*	Cladocera, Daphniidae, *Daphnia*	RWL 6/1968	Small, stagnant pool at S.W. junction of Rt. N. 109 to Courpouiran and D5E1, 43°36.20'N, 3°48.41'E; Montpellier, Herault, France	AF274051.1	DQ273802.1	AY582831.1
	ATCC-PRA-280	*Abeoforma whisleri*	Mytiloida, Mytilidae, *Mytilus* (Mussel)			GU810145.1	-	-
		Acanthoeca spectabilis	-			EU011922.1	EU011933.2	-
	2-04	*Amphibiocystidium ranae*	Anura, Ranidae, *Rana lessonae* (Pool frog)		Switzerland	AY692319.1	-	-
	Cl07	*Amphibiocystidium* sp.	Anura, Ranidae, *Rana italica* (Italian stream frog)		Italy	EU650666.1	-	-
		Amphibiothecum penneri	Anura, Bufonidae, *Bufo americanus* (American toad)		Connecticut, USA	AY772001.1	-	-
		Anurofeca richardsi	Anura, Ranidae, *Rana temporaria* (Common frog)		School of Biological Sciences, University of Sussex, Falmer, Brighton, UK	AF070445.1	-	-
	ATCC-30864	*Capsaspora owczarzaki*	Planorboidea, *Biomphalaria glabrata* (Freshwater snail)		Corvalis, OR	AY363957.1	AY724688.1	XM_004365659.2
	CAUL-BRN01	*Caullerya mesnili*	Cladocera, Daphniidae, *Daphnia longispina* hybrid complex			GU123051.1	-	-
	QM G313693	*Clathrina* sp.	-		Great Barrier Reef, Yonge Reef	AM180960.1	JQ272286.1	-
	IOW94	*Codosiga balthica*	-		Baltic Sea	JQ034424.1	JQ034425.1	-
		Corallochytrium limacisporum	-		Coral reef lagoons of the Lakshadweep Islands of the Arabian Sea	L42528.1	EU011936.1	AY582834.1
	CH2	*Creolimax fragrantissima*	Phleobranchia, Corellidae, *Corella* sp. (Transparent tunicate)		Barkley Sound, Scott's Bay, Bamfield, British Columbia, Canada	EU124915.1	HQ896016.1	HQ896021.1
	Pw1	*Creolimax fragrantissima*	Phascolosomatiformes, Phascolosomatidae, *Phascolosoma* (Peanut worm)		Scott's Bay, Bamfield, British Columbia, Canada	EU124914.1	-	-
	Clone 52	*Dermocystidium percae*	Perciformes, Percidae, *Perca fluviatilis* (European perch)			AF533949.1	-	-

	Dermocystidium salmonis	Salmoniformes, Salmonidae, *Onchorhynchus tschawytscha* (Chinook salmon)		U21337.1	-	-
CM-2002	*Dermocystidium* sp.	Perciformes, Percidae, *Perca fluviatilis* (European perch)		AF533950.1	-	-
LKM51	Ichthyophonida sp.	?	Lake Ketelmeer, Flevoland, The Netherlands	AJ130859.1	-	-
A3	*Ichthyophonus* sp.	Mugiliformes, Mugilidae (Mullet)	Western Mediterranean	FJ887961.1	-	-
D5	*Ichthyophonus* sp.	Mugiliformes, Mugilidae (Mullet)	Western Mediterranean	FJ869836.1	-	-
Clone 1-17	*Ichthyophonus hoferi*	Pleuronectiformes, Pleuronectidae, *Limanda ferruginea* (Yellowtail flounder)		U25637.1	AY026370	-
	Ichthyophonus irregularis	Pleuronectiformes, Pleuronectidae, *Limanda ferruginea* (Yellowtail flounder)	Nova Scotia shelf, Canada	AF232303.1	-	-
Manhattan YLS-1	Ichthyosporea sp. *Tenebrio molitor* symbiont (TMS)	Coleoptera, Tenebrionidae, *Tenebrio molitor* (Mealworm beetle)	Center for Grain and Animal Health Research, Manhattan, KS, USA	JN699060.1*	Not used (JN699061.1)	-
QM G313818	*Levinella prolifera*	-	Great Barrier Reef, Hook Reef	AM180956.1	JQ272292.1	-
ATCC-50519	*Ministeria vibrans*	-		AF271998.1	-	AY582836.1
ATCC-50154	*Monosiga brevicollis*	-	Church Cave, Bermuda	AF100940.1	AY026374.1	-
CCAP 1552/4	*Nuclearia simplex*	-	Freshwater, Heidelberg, West Berlin, Germany	AF484687.1	AY148095.1	AY582835.1
ATCC-PRA-279	*Pirum gemmata*	Phascolosomatiformes, Phascolosomatidae, *Phascolosoma* (Peanut worm)		GU810144.1	-	-
	Psorospermium haeckelii	Decapoda, Astacidae, *Astacus* (Freshwater crayfish)		U33180.1	-	-
	Rhinosporidium cygnus	Anseriformes, Anatidae, *Cygnus* (Swan)	Florida, USA	AF399715.2	-	-
	Rhinosporidium seeberi	Human	Department of Microbiology, University of Peradeniya, Sri Lanka	AF118851.2	-	-

	QM G313668	*Soleneiscus stolonifer*		Great Barrier Reef, Wistari Reef	AM180955.1	JQ272290.1	-
	IE7	*Sphaeroforma (Pseudoperkinsus) tapetis*	Apodida, Synaptidae, *Leptosynapta clarki* (Sea cucumber)	Barkley Sound, Scott's Bay, Bamfield, British Columbia, Canada	GU727527.1	-	-
		Sphaeroforma (Pseudoperkinsus) tapetis	Veneroida, Veneridae, *Ruditapes decussatus* (Carpet shell clam)	Spain	AF192386.1	-	-
	JP610	*Sphaeroforma arctica*	Amphipoda, Gammaridae, *Gammarus setosus*	Littoral zone on the northern coast of Spitsbergen of the high-arctic Svalbard archipelago (79°47'N, 11°53'E)	Y16260.2	-	DQ403166.1
	UK-Cefas1	*Sphaerothecum destruens*	Cypriniformes, Cyprinidae, *Leucaspius delineatus* (Sunbleak)	Freshwater pond in southern England, Park Pond, North Stoneham, Hampshire, UK, (Grid Ref SU43301730)	FN996945.1	-	-
	BML	*Sphaerothecum destruens*	Salmoniformes, Salmonidae, *Onchorhynchus tschawytscha* (Chinook salmon)	Washington, USA	AY267345.1	-	-
	LT37-C21	Uncultured *Amoebidium* isolate	?	Hypersaline Lake Tyrrell, 320 km NW of Melbourne in semi-arid northwestern Victoria, Australia	KC486775.1	-	-
	AI3F14RJ2D12	Uncultured eukaryote clone	?	Micro-oxic water column sample, Cariaco Basin, Venezuela, Caribbean Sea	GU824755.1	-	-
	B47	Uncultured eukaryote clone	?	Domestic waste treatment plant sludge, Bendigo, Victoria, Australia	JN054668.1	-	-
	D3P06H09	Uncultured eukaryote clone	?	Oxygen-depleted intertidal marine sediment, upper 2 cm sediment surface, Arctic, Greenland	EF100301.1*	-	-
	KRL03E12	Uncultured eukaryote clone	?	Lake Karla, central Greece	KC315811.1	-	-
	NAMAKO-33	Uncultured eukaryote clone	?	Anoxic sediments, Namako-ike Lake (31°51'51" N, 129°52'18" E), Kamikoshiki Island, Japan	AB252773.1	-	-
	SGUH1520	Uncultured eukaryote clone	?	Eastern North Pacific, 33.55N, 118.4W, 5m depth	KJ763063.1	-	-
	TAGIRI-25	Uncultured eukaryote clone	?	Anoxic sediment around fumaroles on a submarine caldera floor, off shore from Fukuyama in Kagoshima Bay, Taigiri site (204 m, 31°39.747'N, 130°46.285'E: WGS-84 Datum), Japan	AB191433.1	-	-
	TAGIRI-26	Uncultured	?	Anoxic sediment around fumaroles on a	AB191434.1	-	-

	eukaryote clone		submarine caldera floor, off shore from Fukuyama in Kagoshima Bay, Taigiri site (204 m, 31°39.747'N, 130°46.285'E: WGS-84 Datum), Japan			
PR2-4E-07	Uncultured Ichthyophonida clone	?	Praz-Rodet peat bog, 46°33’N 06°10’E, altitude 1041m., Switzerland	GQ330605.1	-	-

ACB, Anna Cristina Bailey; AS, Augusto Siri; DBS, Doug B. Strongman; ERW, Emma R. Wilson; HCW, Howard C. Whisler; HN, Helen Nahrung; JG, Justin Gause; L, Chris L. Frey; LGV, Laia Guàrdia-Valle; MJC, Matías J. Cafaro; MMW, Merlin M. White; PK, Prasanna Kandel; RWL, Robert W. Lichtwardt; SL, Scott LaPatra; TP, Tyler Pickell

[C] = PCR products from these samples were cloned to obtain sequence data

* = Partial sequence

Table 1.3 Primers used for PCR amplification and sequencing, with citations and primer sequences.

Primer name	Target gene	Source	Direction	Sequence (5'-3')	Length
NS1AA	SSU rDNA	Wang et al. 2014	For	AAGCCATGCATGTCTAAGTATAA	23 bp
NS8AA	SSU rDNA	Wang et al. 2014	Rev	TACTTCCTCTAAATGACCAAGTTTG	25 bp
NL1AA	LSU rDNA	Wang et al. 2014	For	GAGTGAAGCGGGAAIAGCTCAAG	23 bp
LR5	LSU rDNA	Vilgalys & Hester 1990	Rev	TCCTGAGGGAAACTTCG	17 bp
LR7AA	LSU rDNA	Wang et al. 2014	Rev	CCACCAAGATCTGCACTAGA	20 bp
NS3	SSU rDNA	White et al. 1990	For	GCAAGTCTGGTGCCAGCAGCC	21 bp
NS2	SSU rDNA	White et al. 1990	Rev	GGCTGCTGGCACCAGACTTGC	21 bp
NS4	SSU rDNA	White et al. 1990	Rev	CTTCCGTCAATTCCTTTAAG	20 bp
NS8PT	SSU rDNA	New to this study	Rev	TACTTCCTCTAAATGATCAAGTTTG	25 bp
BMB-BR	SSU rDNA	Lane et al. 1986	For	CTTAAAGGAATTGACGGAA	19 bp
NS1PT	SSU rDNA	New to this study	For	AAGCCATGCATGTCCAAGTATAA	23 bp
NL4	LSU rDNA	O'Donnell 1993	Rev	GGTCCGTGTTTCAAGACGG	19 bp
LR6	LSU rDNA	Vilgalys & Hester 1990	Rev	CGCCAGTTCTGCTTACC	17 bp
LR3.1R	LSU rDNA	New to this study	For	GTCTTGAAACACGGACCAAGG	21 bp
LR5.1R	LSU rDNA	New to this study	For	GCCGAAGTTTCCCTCAGGAT	23 bp
LR7	LSU rDNA	Vilgalys & Hester 1990	Rev	TACTACCACCAAGATCT	17 bp
LR0R	LSU rDNA	Vilgalys lab page*	For	ACCCGCTGAACTTAAGC	17 bp
hsp70 Ichthyo F1	HSP70	Paps et al. 2013	For	AAYGAYCARGGHAACCGCACMACYCC	26 bp
hsp70 Ichthyo F2	HSP70	Paps et al. 2013	For	CAGCGYCAGGCYACCAAGGAYGC	23 bp
hsp70 Ichthyo R1	HSP70	Paps et al. 2013	Rev	ATCTGRGGARNTCRAAYTTRCC	22 bp
hsp70 Ichthyo R2	HSP70	Paps et al. 2013	Rev	GTGGGSAYGGTNGTGTTDCGC	21 bp
Primers for protein-coding gene and ITS amplification tests					
hsp90 Ichthyo F1	HSP90	Paps et al. 2013	For	TCYGATGCTYTKGAYAAGATTCG	23 bp
hsp90 Ichthyo F2	HSP90	Paps et al. 2013	For	AACAACCTGGGWACWATTGC	20 bp
hsp90 Ichthyo R1	HSP90	Paps et al. 2013	Rev	GCCTGNGCCTTCATGATRCGCTCC	24 bp

hsp90 Ichthyo R2	HSP90	Paps et al. 2013	Rev	AGSKTCTTGCCCTCGTACTCCTT	23 bp
MCM7-7f	MCM7	New to this study	For	ACIMGIGTITCVGAYGTHAARCC	23 bp
MCM7-7fp	MCM7	New to this study	For	ACIMGIACIACIGAYGTIAARCC	23 bp
MCM7-16r	MCM7	Tretter et al. 2013	Rev	GTYTGYTGYTCCATIACYTCRTG	23 bp
MCM7-16.1r	MCM7	New to this study	Rev	TGYTGYTCCATIACYTCRTGRATIGC	26 bp
MCM7-15r	MCM7	New to this study	Rev	TCCATYTTRTCRAAYTCRTCRATRCA	26 bp
RPB1-AfL	RPB1	Wang et al. 2014	For	GARTGYCCDGGDCAYTTYGGICA	23 bp
RPB1-DrL	RPB1	Wang et al. 2014	Rev	TTCATYTCRTCDCCRTCRAARTCIGC	26 bp
fRPB2-5F	RPB2	Liu et al. 1999	For	GAYGAYMGWGATCAYTTYGG	20 bp
fRPB2-7cR	RPB2	Liu et al. 1999	Rev	CCCATRGCTTGYTTRCCCAT	20 bp
NS7AA	ITS	Tretter et al. 2014	For	GGAAGTTTGAGGCAATAACAGG	22 bp
ITS3	ITS	White et al. 1990	For	GCATCGATGAAGAACGCAGC	20 bp
ITS2	ITS	White et al. 1990	Rev	GCTGCGTTCTTCATCGATGC	20 bp
LR2/LR22 mix	ITS	Vilgalys lab page*	Rev	TTTTCAAAGTTCTTTTC/ CCTCACGGTACTTGTTCGCT	17/20 bp

* Available from http://sites.biology.duke.edu/fungi/mycolab/primers.htm

Table 1.4 PCR protocol details for rDNA and protein-coding gene amplifications. Promega brand mixes were used.

Protocol #	Gene	Forward primer	Reverse primer	Master mix	Program Details: Cycles	Initial denature	Denature	Annealing	Extension	Final extension	Betaine	BSA	Notes
194	18S rDNA	NS1AA	NS8PT	GoTaq Green	45	95°C/2:00	95°C/0:30	60°C/0:45	72°C/3:00	72°C/ 10:00	no	0.8 µg/µL	Worked for Eccrinidae and Amoebidiidae; 35% amp failure
203	18S rDNA	NS1PT	NS8PT	GoTaq Green	45	95°C/2:00	95°C/0:30	58°C/0:45	72°C/3:00	72°C/ 10:00	no	0.8 µg/µL	Worked for Eccrinidae and Amoebidiidae; 64% amp failure
177	18S rDNA	NS1AA	NS8AA	GoTaq Green	45	95°C/2:00	95°C/0:30	58°C/0:45	72°C/3:00	72°C/ 10:00	no	0.8 µg/µL	Only obtained successful sequences for *Paramoebidium*; 57% amp failure
183	28S rDNA	NL1AA	LR7AA	GoTaq Hot Start Green	45	95°C/2:00	95°C/0:30	56°C/0:45	72°C/3:00	72°C/ 10:00	0.5 M	no*	Worked for Eccrinidae and Amoebidiidae; no host sequences, but 39% amp failure
133	28S rDNA	LR0R	LR7	GoTaq Green	45	95°C/2:00	95°C/0:30	51°C/0:45	72°C/3:00	72°C/ 10:00	0.5 M	0.8 µg/µL	Prone to host sequence amplification; obtained successful amplification for *Paramoebidium* and *Enteromyces*; 45 % amp failure
901	HSP70	hsp70 Ichthyo F1/F2	hsp70 Ichthyo R1/R2	GoTaq Hot Start Green	40	94°C/5:00	94°C/0:35	50 & 45°C/ 0:45	72°C/1:15	72°C/ 8:00	0.1 M	no	Nested program; 30% amp failure; prone to host sequence amplification
Test protocols for ITS and protein-coding genes													
131	ITS rDNA	NS7AA	LR2/ LR22 mix	GoTaq Hot Start Green	45	95°C/2:00	95°C/0:30	52-50°C step-down/ 0:45	72°C/3:00	72°C/ 10:00	no	0.8 µg/µL	Obtained partial sequences for samples 377, 533, 872, 1052, 1194, 1240, 1242; secondary structure evident in sequence chromatograms; 13% amp failure out of 3 reaction attempts; prone to host sequence amplification
902	HSP90	hsp90 Ichthyo F1/F2	hsp90 Ichthyo R1/R2	GoTaq Hot Start Green	40	94°C/5:00	94°C/0:35	50 & 45°C/ 0:45	72°C/1:30	72°C/ 10:00	0.1 M	no	Nested program; multiple products observed upon electrophoresis; 54% amp failure out of 3 reaction attempts
303	MCM7	MCM7-7f	MCM7-15r	GoTaq Hot Start Green	45	95°C/2:00	95°C/0:30	50°C/0:45	72°C/1:15	72°C/ 10:00	no	0.8 µg/µL	Multiple products observed upon electrophoresis; 31% amp failure out of 2 reaction attempts
317	MCM7	MCM7-7fp	MCM7-16r	GoTaq Hot Start Green	45	95°C/2:00	95°C/0:30	52°C/1:15	72°C/2:00	72°C/ 10:00	0.1 M	no	Multiple products observed upon electrophoresis; 73% amp failure out of 6 reaction attempts
318	MCM7	MCM7-7fp	MCM7-16.1r	GoTaq Hot Start Green	45	95°C/2:00	95°C/0:30	56°C/0:45	72°C/1:15	72°C/ 10:00	no	no	Multiple products observed upon electrophoresis; 83% amp failure out of 3 reaction attempts

602	RPB1	RPB1-AfL	RPB1-DrL	GoTaq Hot Start Green	49	95°C/2:00	95°C/1:00	57-47°C step-down/ 1:15	72°C/2:45	72°C/ 10:00	no	0.8 µg/µL	81% amp failure out of 3 reaction attempts
82	RPB2	fRPB2-5F	fRPB2-7cR	GoTaq Hot Start Green	49	95°C/2:00	95°C/1:00	55-45°C step-down/ 1:15	72°C/2:45	72°C/ 10:00	no	no	68% amp failure out of 5 reaction attempts

* Added BSA at 0.8 µg/µL for 5 reactions

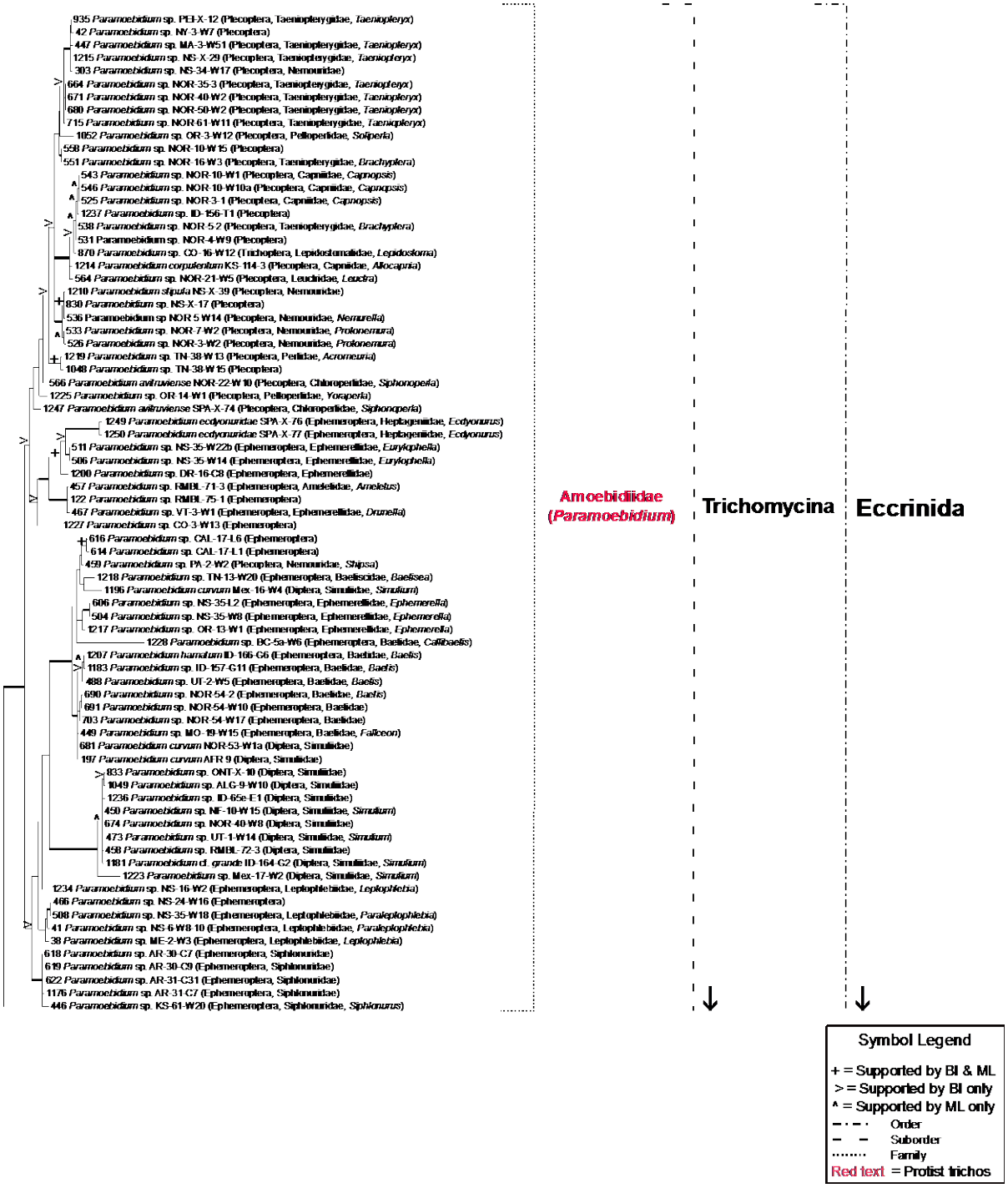
Table 1.6 SH and AU topology test results. Constraint topologies are listed under Hypothesis and were compared to the RAxML best tree. Bold values indicate significant results ($p < 0.05$).

Hypothesis	SH p-value	AU p-value
Amoebidiidae (*Paramoebidum* + *Amoebidium*) monophyletic	0.128	0.118
Protist trichos (Amoebidiidae + Eccrinidae) monophyletic	0.180	0.176
Paramoebidium* + *Parataeniella	**0.032**	**0.024**
Eccrinaceae monophyletic	**0**	**0.0002**
Palavasciaceae monophyletic (no 1121)	0.159	0.139
Palavasciaceae monophyletic (with 1121 & 1115)	0.698	0.75
Eccrinidae without *Parataeniella*	0.587	0.535

Table 1.7 BaTS 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 ASR analyses. Bold values indicate significant results ($p < 0.05$). Results for character states only represented once in the dataset are not shown.

	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**
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
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
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/hypha**	6	6	6	1.119	1	2	**1.00E-03**
	MC **Spherical**	3	3	3	1.4	1	2	**0.01700002**
	MC Spherical/plasmodial/hyphae	1	1	1	1	1	1	1
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**


935 Paramoebidium sp. PEI-X-12 (Plecoptera, Taeniopterygidae, Taeniopteryx)
42 Paramoebidium sp. NY-3-W7 (Plecoptera)
447 Paramoebidium sp. MA-3-W51 (Plecoptera, Taeniopterygidae, Taeniopteryx)
1215 Paramoebidium sp. NS-X-29 (Plecoptera, Taeniopterygidae, Taeniopteryx)
303 Paramoebidium sp. NS-34-W17 (Plecoptera, Nemouridae)
664 Paramoebidium sp. NOR-35-3 (Plecoptera, Taeniopterygidae, Taeniopteryx)
671 Paramoebidium sp. NOR-40-W2 (Plecoptera, Taeniopterygidae, Taeniopteryx)
680 Paramoebidium sp. NOR-50-W2 (Plecoptera, Taeniopterygidae, Taeniopteryx)
715 Paramoebidium sp. NOR-61-W11 (Plecoptera, Taeniopterygidae, Taeniopteryx)
1052 Paramoebidium sp. OR-3-W12 (Plecoptera, Pelloperlidae, Soliperla)
558 Paramoebidium sp. NOR-10-W15 (Plecoptera)
551 Paramoebidium sp. NOR-16-W3 (Plecoptera, Taeniopterygidae, Brachyptera)
543 Paramoebidium sp. NOR-10-W1 (Plecoptera, Capniidae, Capnopsis)
546 Paramoebidium sp. NOR-10-W10a (Plecoptera, Capniidae, Capnopsis)
525 Paramoebidium sp. NOR-3-1 (Plecoptera, Capniidae, Capnopsis)
1237 Paramoebidium sp. ID-156-T1 (Plecoptera)
538 Paramoebidium sp. NOR-5-2 (Plecoptera, Taeniopterygidae, Brachyptera)
531 Paramoebidium sp. NOR-4-W9 (Plecoptera)
870 Paramoebidium sp. CO-16-W12 (Trichoptera, Lepidostomatidae, Lepidostoma)
1214 Paramoebidium corpulentum KS-114-3 (Plecoptera, Capnidae, Allocapnia)
564 Paramoebidium sp. NOR-21-W5 (Plecoptera, Leuctridae, Leuctra)
1210 Paramoebidium stipula NS-X-39 (Plecoptera, Nemouridae)
830 Paramoebidium sp. NS-X-17 (Plecoptera)
536 Paramoebidium sp NOR-5-W14 (Plecoptera, Nemouridae, Nemurella)
533 Paramoebidium sp. NOR-7-W2 (Plecoptera, Nemouridae, Prolonemura)
526 Paramoebidium sp. NOR-3-W2 (Plecoptera, Nemouridae, Prolonemura)
1219 Paramoebidium sp. TN-38-W13 (Plecoptera, Perlidae, Acroneuria)
1048 Paramoebidium sp. TN-38-W15 (Plecoptera)
566 Paramoebidium avitruviense NOR-22-W10 (Plecoptera, Chloroperlidae, Siphonopelia)
1225 Paramoebidium sp. OR-14-W1 (Plecoptera, Pelloperlidae, Yoraperla)
1247 Paramoebidium avitruviense SPA-X-74 (Plecoptera, Chloroperlidae, Siphonopelia)
1249 Paramoebidium ecdyonuridae SPA-X-76 (Ephemeroptera, Heptageniidae, Ecdyonurus)
1250 Paramoebidium ecdyonuridae SPA-X-77 (Ephemeroptera, Heptageniidae, Ecdyonurus)
511 Paramoebidium sp. NS-35-W22b (Ephemeroptera, Ephemerellidae, Eurylophella)
506 Paramoebidium sp. NS-35-W14 (Ephemeroptera, Ephemerellidae, Eurylophella)
1200 Paramoebidium sp. DR-16-C8 (Ephemeroptera, Ephemerellidae)
457 Paramoebidium sp. RMBL-71-3 (Ephemeroptera, Ameletidae, Ameletus)
122 Paramoebidium sp. RMBL-75-1 (Ephemeroptera)
467 Paramoebidium sp. VT-3-W1 (Ephemeroptera, Ephemerellidae, Drunella)
1227 Paramoebidium sp. CO-3-W13 (Ephemeroptera)
616 Paramoebidium sp. CAL-17-L6 (Ephemeroptera)
614 Paramoebidium sp. CAL-17-L1 (Ephemeroptera)
459 Paramoebidium sp. PA-2-W2 (Plecoptera, Nemouridae, Shipsa)
1218 Paramoebidium sp. TN-13-W20 (Ephemeroptera, Baetiscidae, Baetisca)
1196 Paramoebidium curvum Mex-16-W4 (Diptera, Simuliidae, Simulium)
606 Paramoebidium sp. NS-35-L2 (Ephemeroptera, Ephemerellidae, Ephemerella)
504 Paramoebidium sp. NS-35-W8 (Ephemeroptera, Ephemerellidae, Ephemerella)
1217 Paramoebidium sp. OR-13-W1 (Ephemeroptera, Ephemerellidae, Ephemerella)
1228 Paramoebidium sp. BC-5a-W6 (Ephemeroptera, Baetidae, Callibaetis)
1207 Paramoebidium hamatum ID-166-C6 (Ephemeroptera, Baetidae, Baetis)
1183 Paramoebidium sp. ID-157-G11 (Ephemeroptera, Baetidae, Baetis)
488 Paramoebidium sp. UT-2-W5 (Ephemeroptera, Baetidae, Baetis)
690 Paramoebidium sp. NOR-54-2 (Ephemeroptera, Baetidae, Baetis)
691 Paramoebidium sp. NOR-54-W10 (Ephemeroptera, Baetidae)
703 Paramoebidium sp. NOR-54-W17 (Ephemeroptera, Baetidae)
449 Paramoebidium sp. MO-19-W15 (Ephemeroptera, Baetidae, Falceon)
681 Paramoebidium curvum NOR-53-W1a (Diptera, Simuliidae)
197 Paramoebidium curvum AFR-9 (Diptera, Simuliidae)
833 Paramoebidium sp. ONT-X-10 (Diptera, Simuliidae)
1049 Paramoebidium sp. ALG-9-W10 (Diptera, Simuliidae)
1236 Paramoebidium sp. ID-65e-E1 (Diptera, Simuliidae)
450 Paramoebidium sp. NF-10-W15 (Diptera, Simuliidae, Simulium)
674 Paramoebidium sp. NOR-40-W8 (Diptera, Simuliidae)
473 Paramoebidium sp. UT-1-W14 (Diptera, Simuliidae, Simulium)
458 Paramoebidium sp. RMBL-72-3 (Diptera, Simuliidae)
1181 Paramoebidium cf. grande ID-164-G2 (Diptera, Simuliidae, Simulium)
1223 Paramoebidium sp. Mex-17-W2 (Diptera, Simuliidae, Simulium)
1234 Paramoebidium sp. NS-16-W2 (Ephemeroptera, Leptophlebiidae, Leptophlebia)
466 Paramoebidium sp. NS-24-W16 (Ephemeroptera)
508 Paramoebidium sp. NS-35-W18 (Ephemeroptera, Leptophlebiidae, Paraleptophlebia)
41 Paramoebidium sp. NS-6-W8-10 (Ephemeroptera, Leptophlebiidae, Paraleptophlebia)
38 Paramoebidium sp. ME-2-W3 (Ephemeroptera, Leptophlebiidae, Leptophlebia)
618 Paramoebidium sp. AR-30-C7 (Ephemeroptera, Siphlonuridae)
619 Paramoebidium sp. AR-30-C9 (Ephemeroptera, Siphlonuridae)
622 Paramoebidium sp. AR-31-C31 (Ephemeroptera, Siphlonuridae)
1176 Paramoebidium sp. AR-31-C7 (Ephemeroptera, Siphlonuridae)
446 Paramoebidium sp. KS-61-W20 (Ephemeroptera, Siphlonuridae, Siphlonurus)
Amoebidiidae (Paramoebidium)
Trichomycina
Eccrinida
Symbol Legend
+ = Supported by BI & ML
> = Supported by BI only
^ = Supported by ML only
Order
Suborder
Family
Red text = Protist trichos

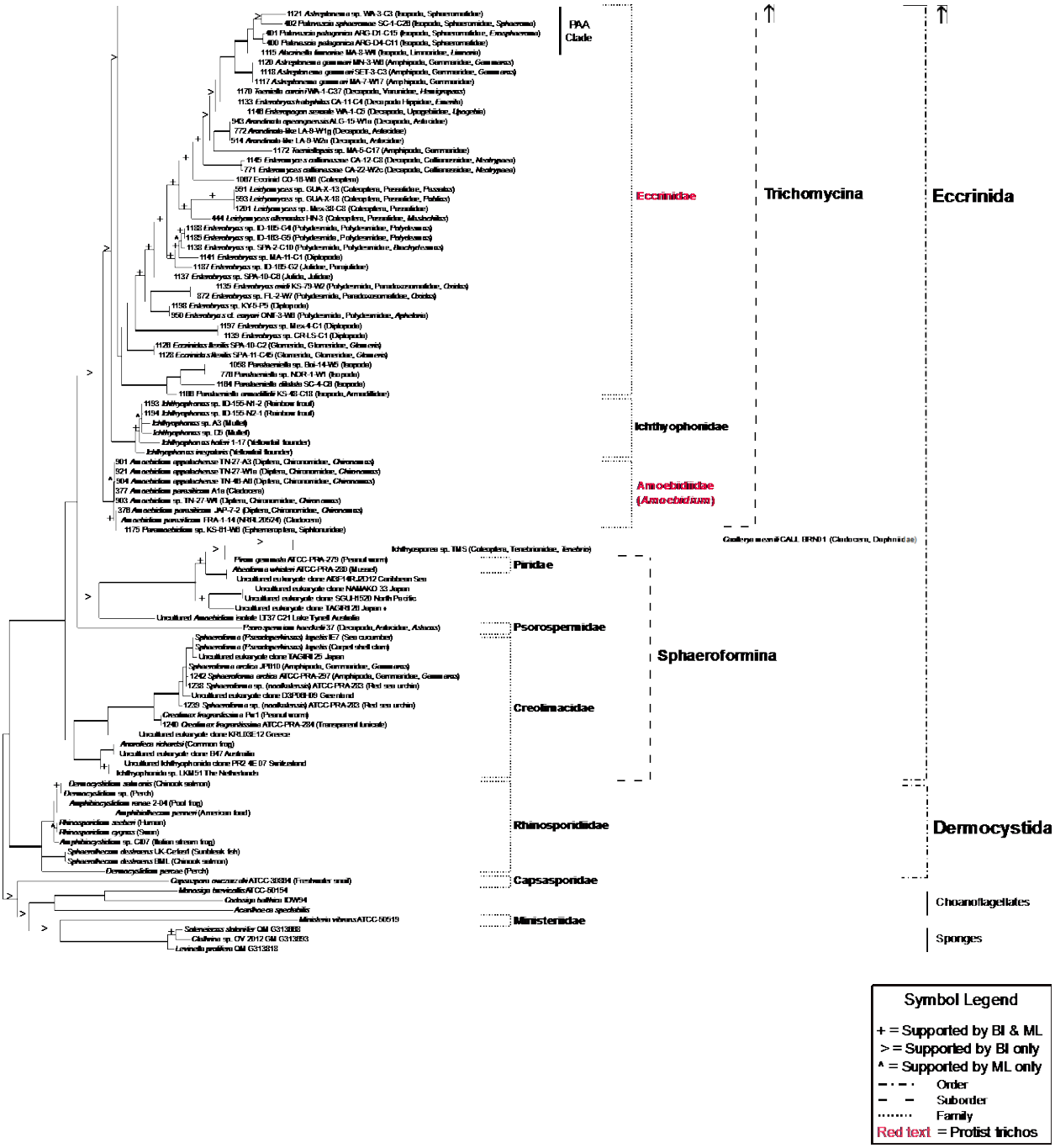


Figure 1.1 "Large rDNA tree", a MrBayes (BI) consensus tree of the 18S and 28S combined dataset, and including all *Paramoebidium* samples and a large outgroup sampling, with maximum likelihood (ML) and maximum parsimony (MP) values indicated symbolically. Red text highlights protist tricho clades. Branches supported by all three methods are bolded. Classification of Cavalier-Smith (2013) is indicated.

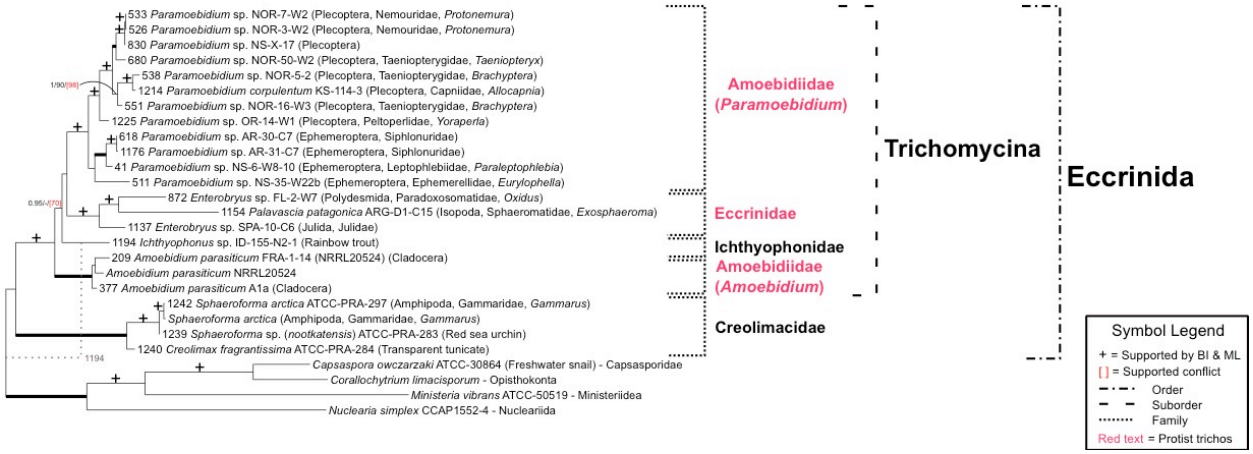


Figure 1.2 MrBayes (BI) consensus tree of the 18S, 28S and HSP70 amino acid combined dataset, with maximum likelihood (ML) and maximum parsimony (MP) values indicated symbolically. Bolded branches indicate support from all three methods. Grey dashed line indicates the alternative placement of *Ichthyophonus* in the MP tree. Classification of Cavalier-Smith (2013) is indicated.

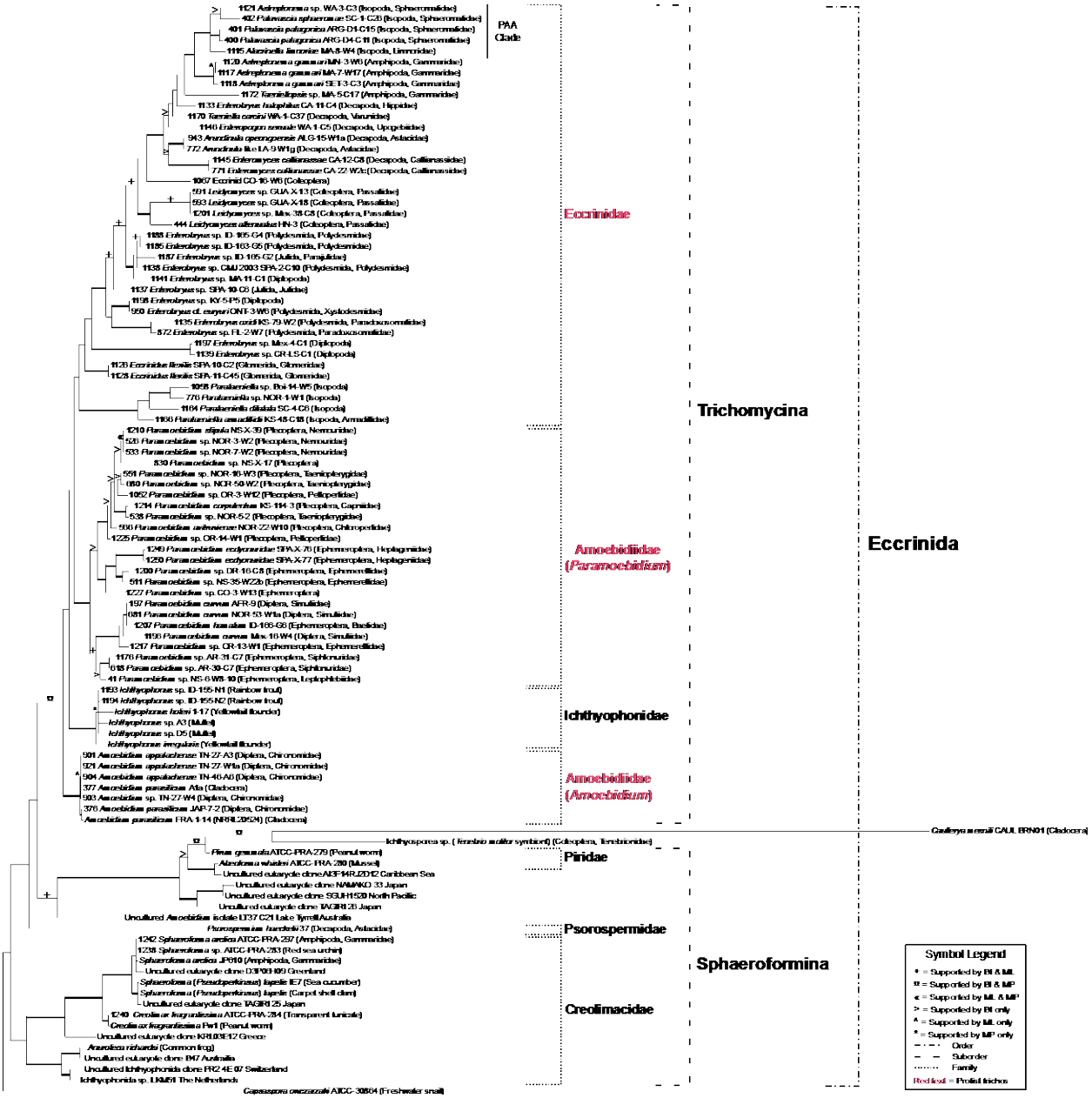


Figure 1.3 "Small rDNA tree", a MrBayes (BI) consensus tree of the 18S and 28S combined dataset, and including fewer *Paramoebidium* samples and *Capsaspora owczarzaki* as the outgroup, with maximum likelihood (ML) and maximum parsimony (MP) values indicated symbolically. Branches supported by all three methods are bolded. Classification of Cavalier-Smith (2013) is indicated.

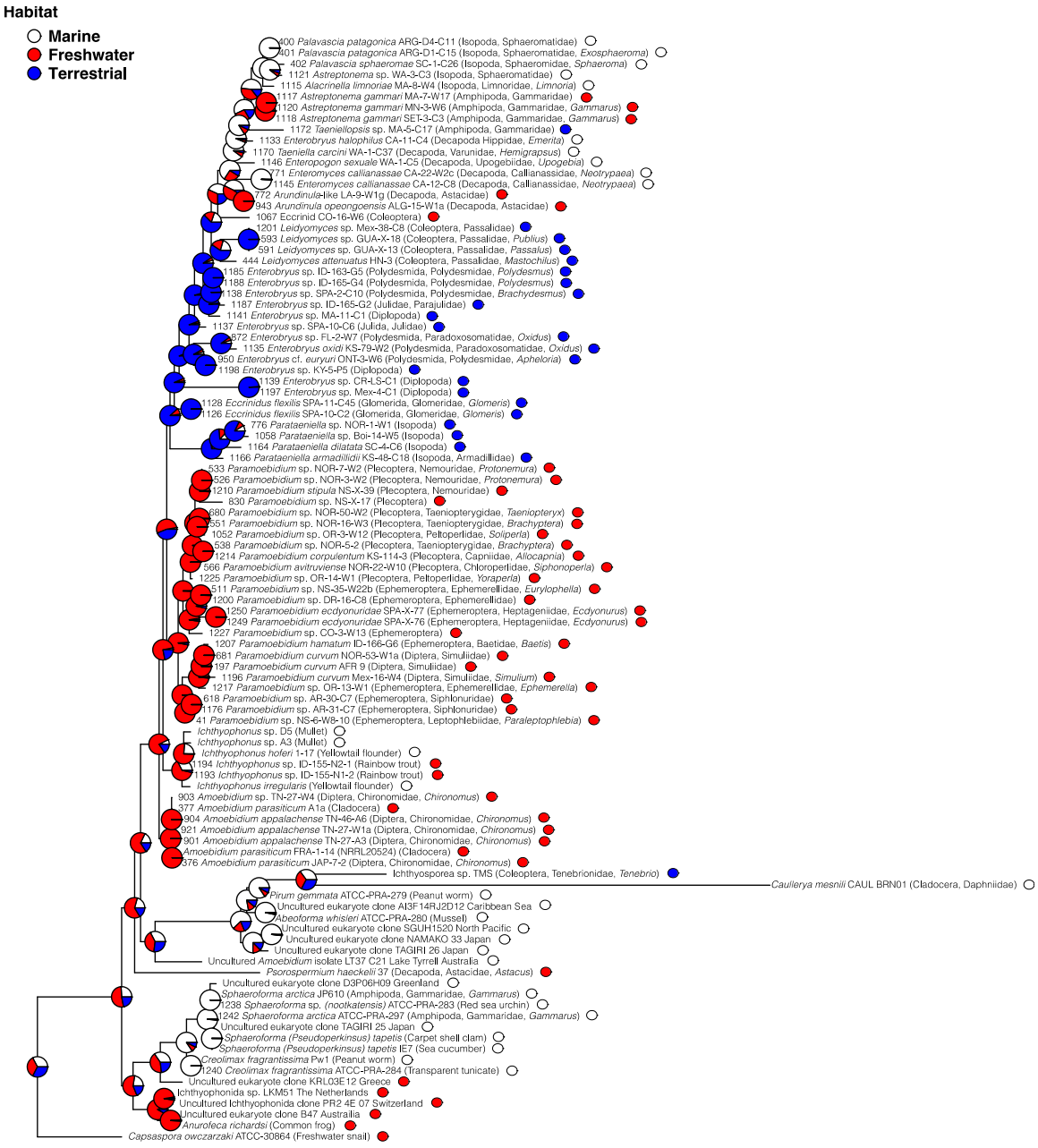


Figure 1.4 SIMMAP output for the Habitat character drawn on the MrBayes small rDNA consensus tree. Pie charts at each node indicate the proportion of probability for each state. Character state coding for each sample is indicated by a colored dot after the name. White represents marine, red represents freshwater, and blue represents terrestrial.

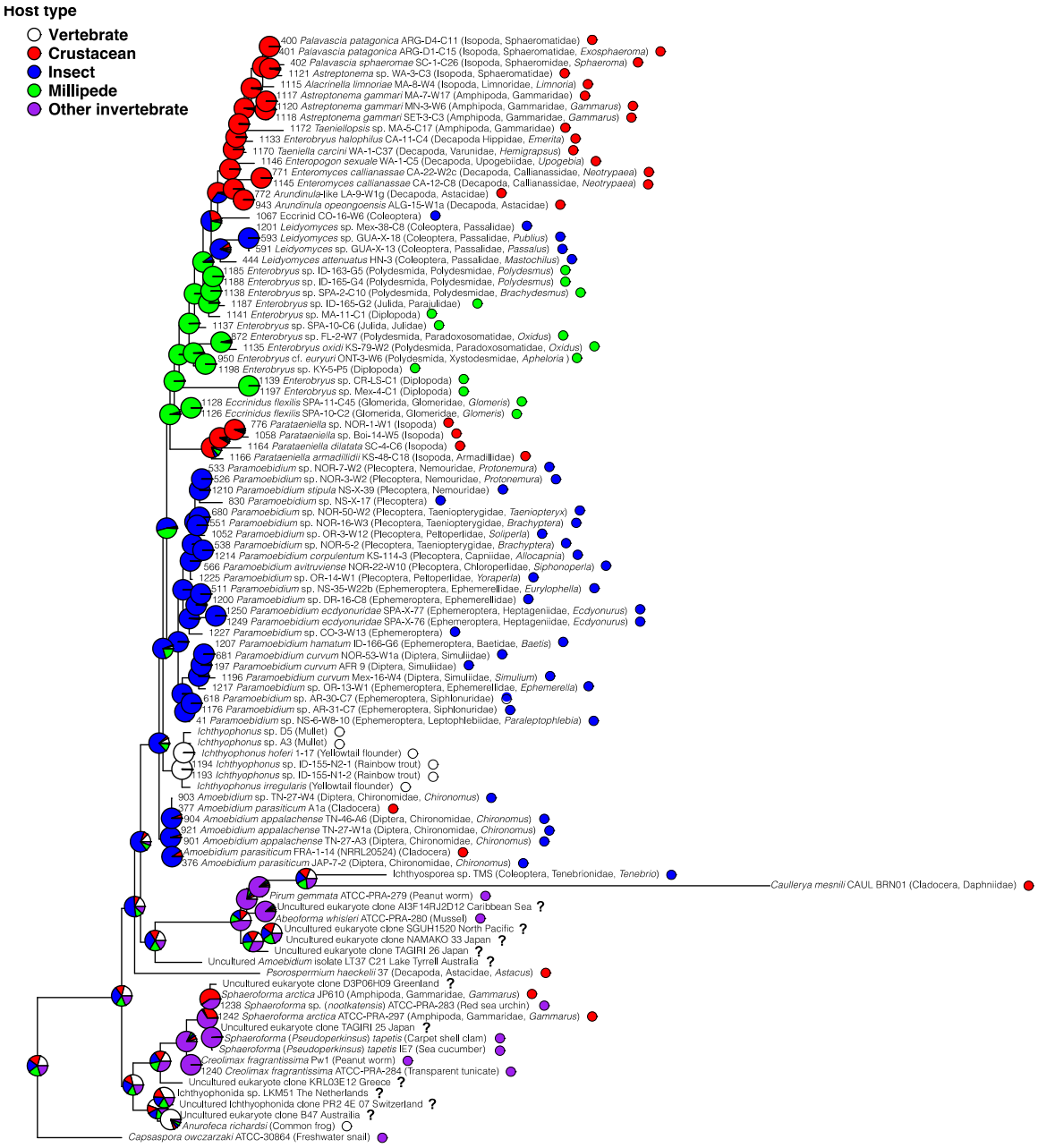


Figure 1.5 SIMMAP output for the Host type character drawn on the MrBayes small rDNA consensus tree. Pie charts at each node indicate the proportion of probability for each state. Character state coding for each sample is indicated by a colored dot after the name. White represents vertebrate, red represents crustacean, blue represents insect, green represents millipede, and purple represents other invertebrates (e.g. bivalves, tunicates, echinoderms). Missing data is indicated by a question mark.

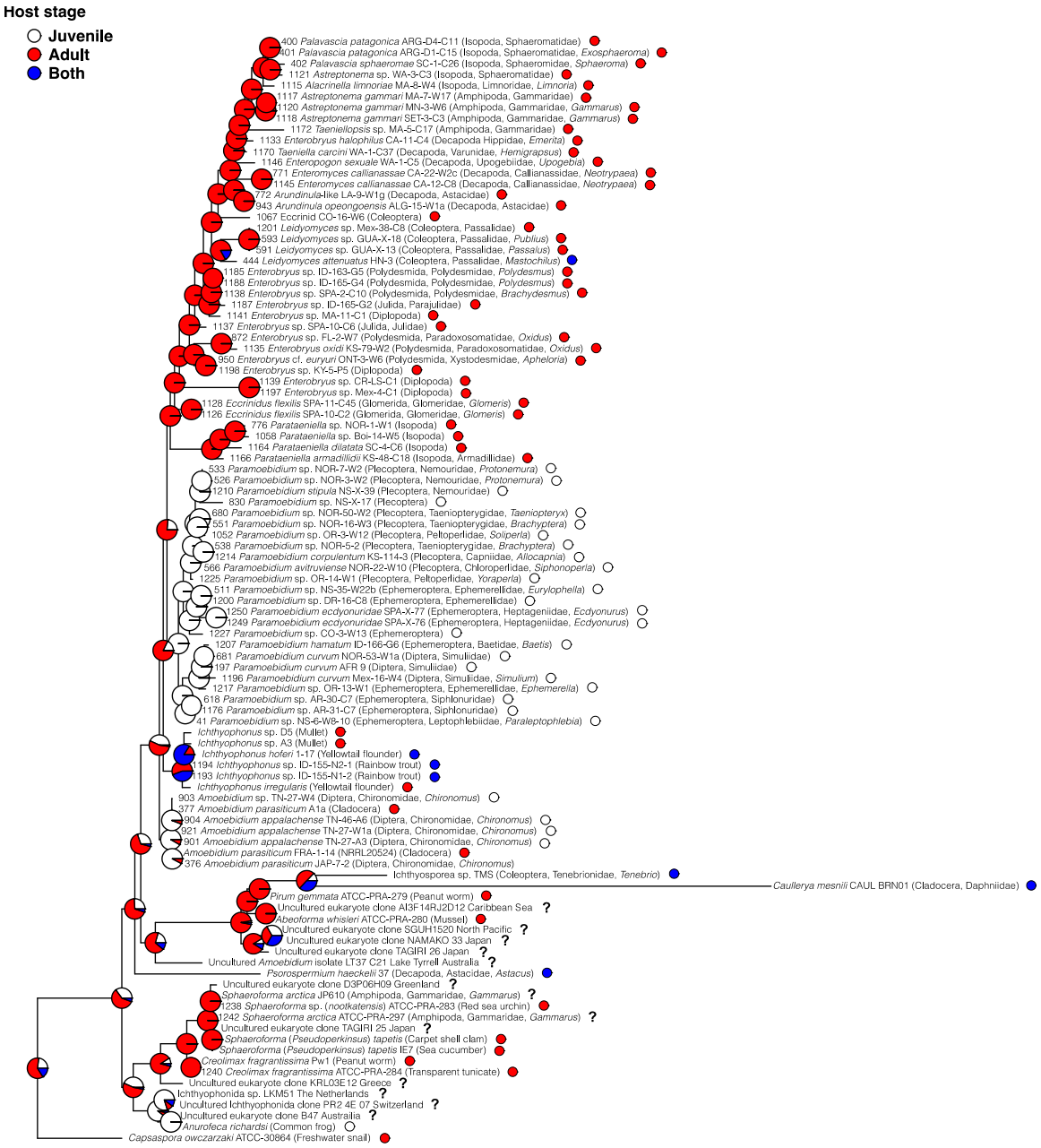


Figure 1.6 SIMMAP output for the Host stage character drawn on the MrBayes small rDNA consensus tree. Pie charts at each node indicate the proportion of probability for each state. Character state coding for each sample is indicated by a colored dot after the name. White represents juvenile, red represents adult, and blue represents both (juvenile & adult). Missing data is indicated by a question mark.

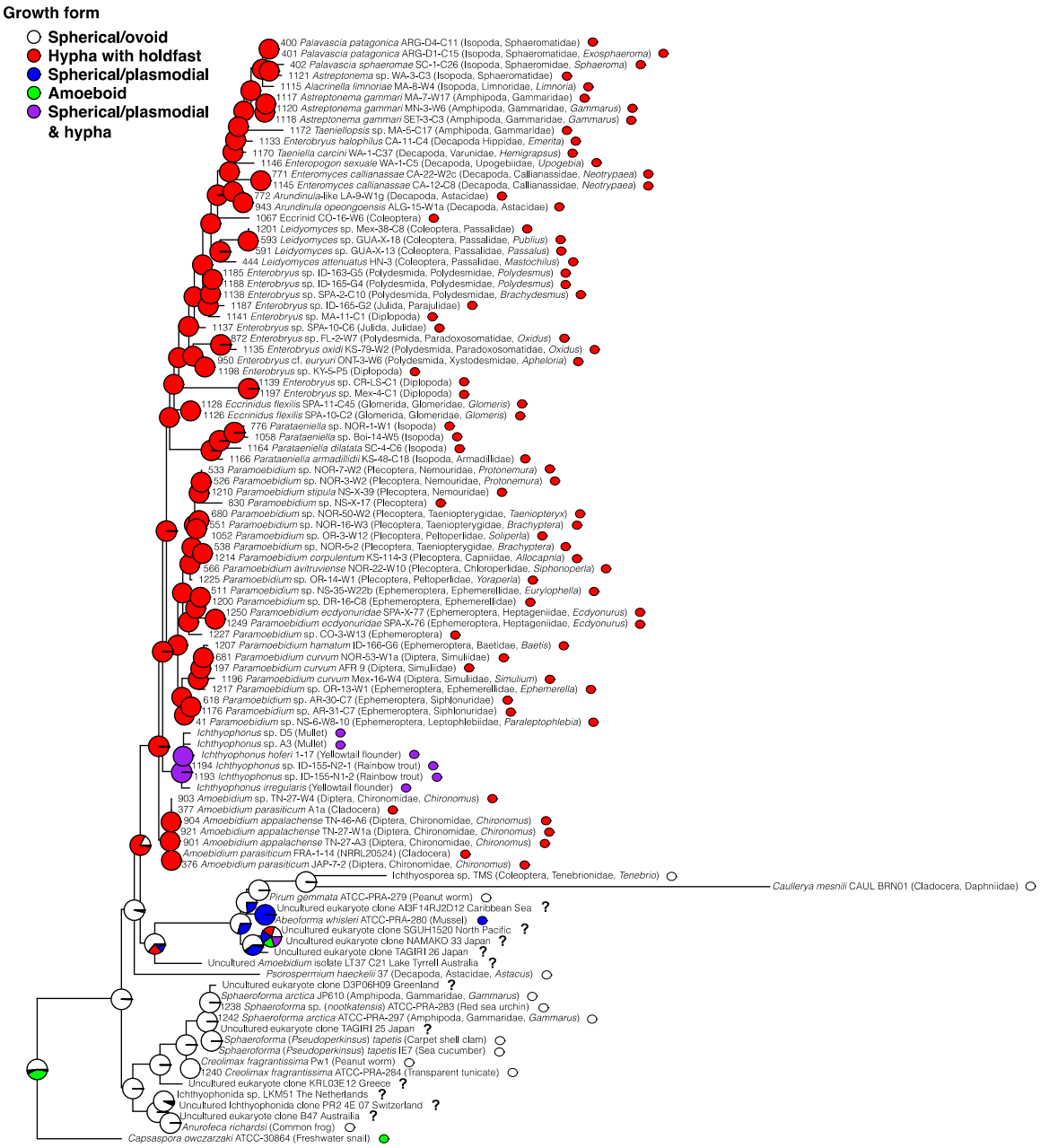


Figure 1.7 SIMMAP output for the Growth form character drawn on the MrBayes small rDNA consensus tree. Pie charts at each node indicate the proportion of probability for each state. Character state coding for each sample is indicated by a colored dot after the name. White represents spherical/ovoid, red represents hypha with holdfast, blue represents spherical/plasmodial, green represents amoeboid, and purple represents spherical/hypha/plasmodial. Missing data is indicated by a question mark.

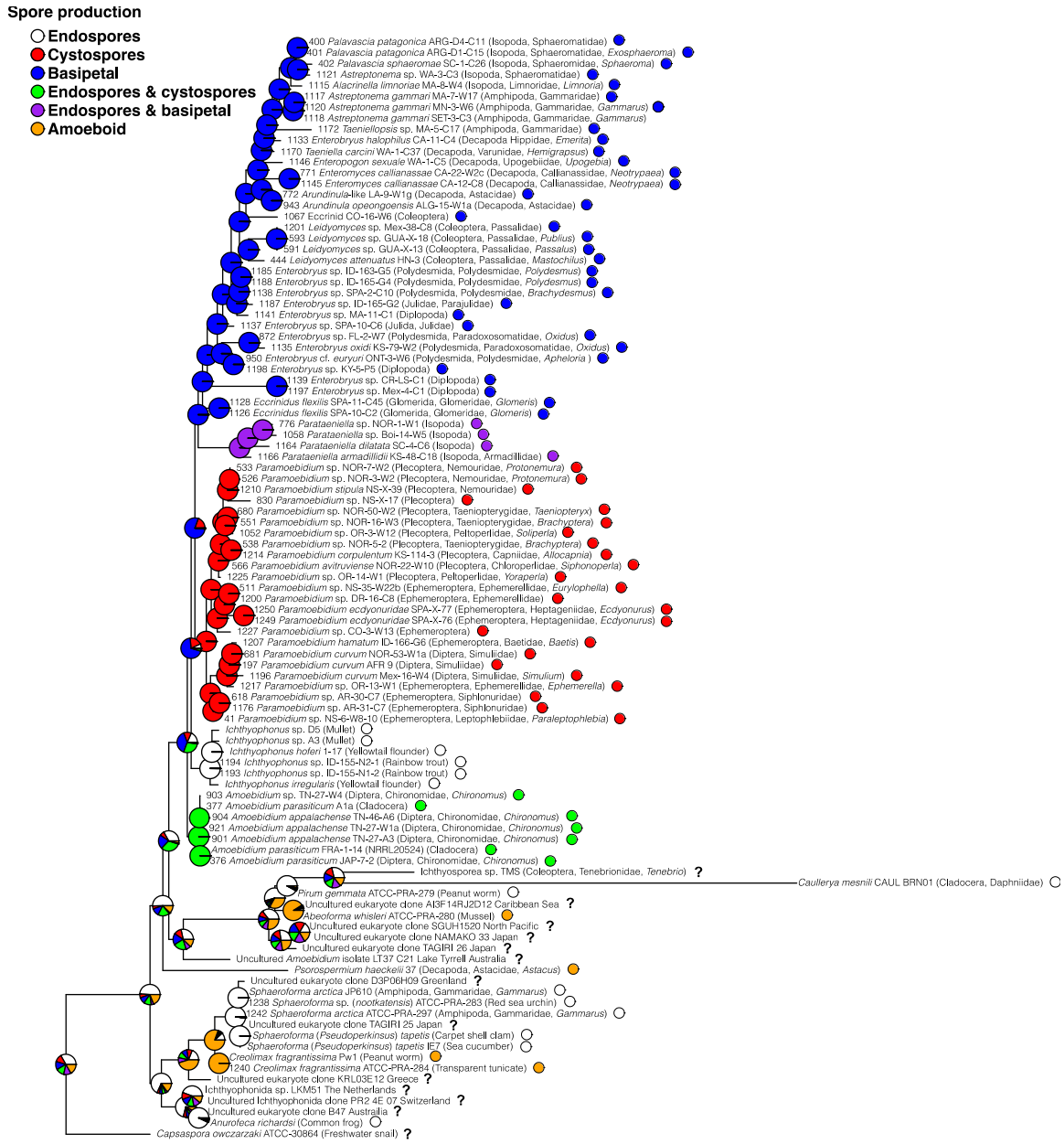


Figure 1.8 SIMMAP output for the Spore production character drawn on the MrBayes small rDNA consensus tree. Pie charts at each node indicate the proportion of probability for each state. Character state coding for each sample is indicated by a colored dot after the name. White represents endospores, red represents cystospores, blue represents basipetal, green represents endospores and cystospores, purple represents endospores and basipetal, and orange represents amoeboid. Missing data is indicated by a question mark.

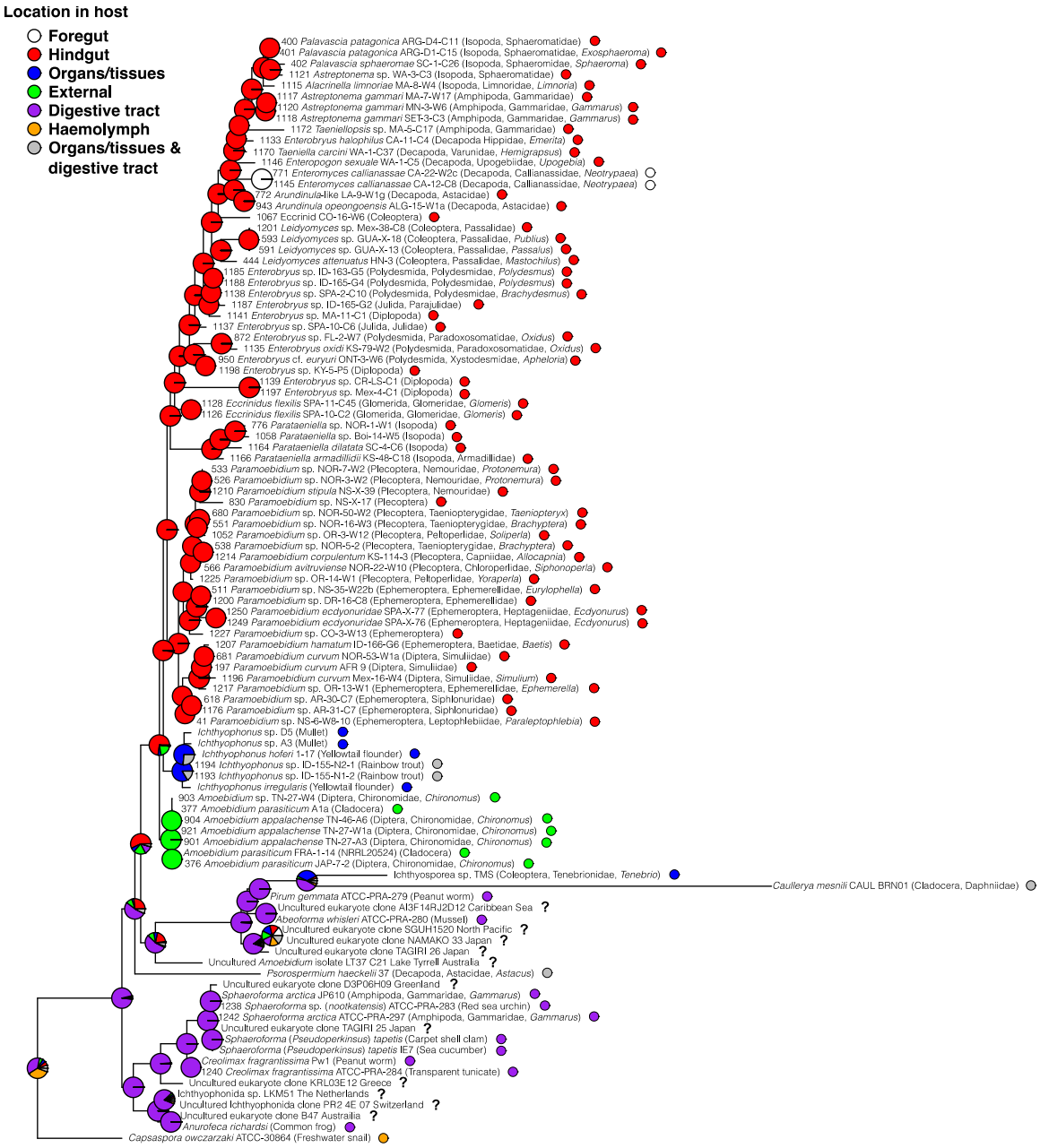


Figure 1.9 SIMMAP output for the Location in host character drawn on the MrBayes small rDNA consensus tree. Pie charts at each node indicate the proportion of probability for each state. Character state coding for each sample is indicated by a colored dot after the name. White represents foregut, red represents hindgut, blue represents organs/tissues, green represents external, purple represents digestive tract, orange represents haemolymph, and grey represents organs/tissues and digestive tract. Missing data is indicated by a question mark.

EVALUATION OF HOST ASSOCIATIONS AND THE UTILITY OF HYPHAL MORPHOLOGY FOR SPECIES DESCRIPTIONS OF *PARAMOEBIDIUM* USING MOLECULAR PHYLOGENETIC ASSESSMENT METHODS

Abstract

Despite first being established as a genus in 1929, *Paramoebidium* circumscribes just 17 species, and its putative sister taxon, *Amoebidium*, includes only five. An oft-cited reason for the lack of species descriptions is the paucity of informative morphological characters in the form of significant inter- and intraspecific variation. As a morphological species concept has been traditionally applied to these taxa, characters to construct an effective taxonomic framework have been elusive. Further compounding these issues has been the uncertain evolutionary relationship of these taxa to Fungi. Originally they were included as order Amoebidiales in the fungal class Trichomycetes, though this was early considered only a tentative placement. Consequently, the group has remained relatively understudied, but molecular systematics has necessitated their reclassification as family Amoebidiidae within the protist class Ichthyosporea. Still, there is only a single 18S rDNA sequence available in GenBank for *Paramoebidium*. The two sister taxa to the Amoebidiidae include another former Trichomycetes group, Eccrinidae, and a parasite of fishes, *Ichthyophonus*. *Amoebidium* and *Ichthyophonus* demonstrate generalist host associations, whereas host specificity among the Eccrinidae remians unclear. On the other hand, *Paramoebidium* spp. were initially thought to be specific to host genus, and perhaps even to species level, although researchers later

questioned this hypothesis. This study generated of a large molecular dataset including 18S and 28S rDNA sequences for 72 *Paramoebidium* samples, as well as morphometric evaluation of their corresponding slide vouchers. From the multiple sequence alignment, phylogenies were inferred and subsequently used as input for ancestral state reconstruction and Bayesian tip-association significance testing analyses of hyphal morphology and host characters. Results of these analyses indicate: 1) some species of *Paramoebidium* may have a high degree of host specificity, even to the genus level, 2) *Paramoebidium* is likely substantially more species rich than previously understood, 3) cryptic speciation is evident, and 4) a morphological species concept is inadequate for species delimitation for many, but not all, members of the genus. Moving forward, an integrated morpho-phylogenetic approach is recommended in order to establish a robust, practically applicable taxonomic framework.

Introduction

What has recently been defined as the protist family Amoebidiidae (comprising *Amoebidium* and *Paramoebidium*) (Cavalier-Smith 2013) was previously classified as order Amoebidiales in the fungal class Trichomycetes, a group of arthropod-associated, obligate endobionts (Lichtwardt et al. 2001). Though tenuous, their historical inclusion in this fungal group was based on residence in the digestive tracts (*Paramoebidium*) or on the exterior (*Amoebidium*) of arthropods, a hyphal growth form with a holdfast, and production of spores (Moss 1979). Whereas the actual relatedness of the Amoebidiidae to Fungi was long suspect due to the formation of amoeboid propagules and lack of chitin in their cell wall (Whisler 1963; Trotter & Whisler 1965), their phylogenetic placement remained uncertain until 2000 (Benny & O'Donnell 2000; Ustinova et al. 2000; Cafaro 2005) when molecular phylogenetic inferences revealed their relationship to members of class Ichthyosporea (= Mesomycetozoea). The family is now included with order Eccrinida, along with their close relatives, the Eccrinidae (also formerly within class Trichomycetes) and *Ichthyophonus* (Cavalier-Smith 2013). Class Ichthyosporea comprises animal-associated, unicellular symbionts and has been placed near the divergence of animals and fungi in multigene phylogenies (Steenkamp et al. 2006; Ruiz-Trillo et al. 2008; Paps et al. 2013; Cavalier-Smith et al. 2014).

Members of the Eccrinida have been found in association with marine invertebrates such as bivalves and tunicates, but the Amoebidiidae are apparently restricted to freshwater hosts such as mayfly (Ephemeroptera) and stonefly (Plecoptera) nymphs, black fly and midge larvae (Diptera), and water fleas (Cladocera) (Moss 1979). These commensal organisms attach to the chitinous digestive tract lining or exoskeleton

via a non-cellular, secreted holdfast, and produce walled spores as putative long-distance dispersal units (Lichtwardt et al. 2001). During their life cycle, the cellular content of the coenocytic Amoebidiidae is divided entirely into motile amoebae that disperse and subsequently encyst. From these cysts are formed "cystospores" that are presumably released to the environment to be ingested by a new host. This amoeba-cyst cycle is the only known method of spore production in *Paramoebidium* spp., but *Amoebidium* may also produce spores directly (termed "sporangiospores") via holocarpic division (Whisler 1968).

Spores and amoebae are not unique characters among the Eccrinida. As sister taxa to the Amoebidiidae, members of the Eccrinidae produce only spores, basipetally, and attach to the digestive tracts of a much broader range of arthropods, including crabs, shrimps, beetles, and millipedes (Lichtwardt et al. 2001). Conversely *Ichthyophonus*, the other sister taxon, is a parasite of fish. Its life cycle is not completely known, but it grows as spherical, multinucleate bodies in the host tissues [termed a "schizont" by Kocan (2013)] from which amoeba-like cells disperse (Spanggaard et al. 1995; Kocan et al. 2013). While the exact relationships between these four taxa are not resolved, morphological delineation of species within groups has also been problematic. Growing as simple spheres in host tissues, but exhibiting a range of forms in culture (e.g. plasmodial, hypha-like, spherical) depending on pH and media recipe (Okamoto et al. 1985), *Ichthyophonus* spp. are challenging to differentiate morphologically. Likewise, members of the Eccrinidae circumscribe a wide spectrum of hyphal morphotypes and spore forms within the host gut, and the interspecific overlap of these features complicates taxonomic efforts to interpret and define them. Further confounding

morphometric assessments of *Paramoebidium* and the Eccrinidae is the intractability of these taxa to grow in culture, requiring descriptions to be based on material dissected from living arthropod hosts.

Collections of the Amoebidiidae are similarly challenging, although *Amoebidium parasiticum* (Whisler 1962) and *A. appalachense* (White et al. 2006) have been isolated in pure culture. Traditionally, voucher slides and photomicrographs of the Amoebidiidae (and Eccrinidae), from both living and fixed specimens, are used to morphometrically define characters and apply a morphological species concept to identify and diagnose species (Lichtwardt et al. 2001). However, it is not uncommon for researchers (e.g. Whisler 1963; White et al. 2000; Lichtwardt et al. 2001b; Strongman & White 2006; White et al. 2006; Hapsari et al. 2009; William & Strongman 2013) to include lists of unnamed *Paramoebidium* spp. morphotypes without further identification, despite their sometimes extensive occurrence within collections.

Since the genus was first established in 1929 (Leger & Duboscq 1929), 17 species have been named and recognized (Lichtwardt et al. 2001). Indeed, most recently described species of *Paramoebidium* have been those that exhibit “unique” or “unusual” extremes of morphologies such as branching, size and/or curvature of the hyphae, presence of a papillum, and/or growing in clusters (e.g. Lichtwardt et al. 1990; Lichtwardt & Williams 1992; Strongman et al. 2010). Interestingly, early French protozoologists (i.e. Duboscq, Leger, Manier, Poisson, and Tuzet) included the host as part of their taxonomic considerations when identifying *Paramoebidium*, with the specific epithet sometimes referring to the host genus (Léger & Duboscq 1929; Duboscq et al. 1948; Manier 1950), a trend that continues today (Valle 2014a). Most of these early

species were later regarded as *nomen nudum*, as their descriptions did not meet the criteria of the International Code of Botanical Nomenclature (Lichtwardt et al. 2001). However, later work, particularly in the United States by both Lichtwardt and Whisler, cast doubt on the degree of host specificity within the genus. In particular, variability of the symbionts' characters, even within a particular host genus or family, has been the main obstacle (Lichtwardt et al. 2001). As Whisler (1963) stated, *Paramoebidium* spp. "present a spectrum of overlapping characters … Separation on the basis of different hosts seems inadequate and it would be meaningless or even misleading to assign a specific epithet … until the taxonomy of the genus has become clarified." This referred to both vegetative characters and reproductive structures (i.e. amoebae, cysts, cystospores). Certainly the possibility of more than one *Paramoebidium* species inhabiting a single host is a consideration, such as *P. chattoni* and *P. curvum* from black flies (Dang & Lichtwardt 1979; Valle 2014b), but no studies have thoroughly analyzed the range of intra- and/or interspecific character variability at different taxonomic host levels.

Similarly, there have been few molecular systematic studies focused on the Amoebidiidae. There is but one *Paramoebidium* sequence in GenBank (Cafaro 2005), although there are several for *Amoebidium parasiticum*. The goals of this study, therefore, were to use sequence and morphometric data for many *Paramoebidium* samples from various geographic locations and hosts to: 1) illustrate the potential for species diversity (or lack thereof), 2) attempt to evaluate any implication of host specificity, 3) gauge evidence of cryptic speciation, and 4) assess the taxonomic utility of

general vegetative characters and, by extension, the practicality of the application of a morphological species concept to the genus.

Materials and Methods

Genetic Data

DNA extraction, PCR amplification, sequencing, and cloning procedures are as outlined in Chapter 1.

Voucher Materials

All collection materials available in our lab were mined for data regarding the *Paramoebidium* samples used here. Dissection log notebooks, photographs, and slides were reviewed to ascertain collection location, host identification, and morphological information. Host vouchers (in ethanol) were sent out for identification. All slide specimens were preserved with lactophenol cotton blue. Slide vouchers corresponding to genomic samples or of the same morphospecies, when present, were examined with a Nikon Eclipse 80i microscope. Images of specimens were taken with a 2 Mp Spot Color Mosaic camera (Diagnostic Instruments, Sterling Heights, Michigan) and measurements taken using the accompanying advanced software (4.6). The Lucid Keys (http://keys.lucidcentral.org), developed through the Trichomycete Monograph (Lichtwardt et al. 2001), were used to compare morphological assessments with described species. The morphological identification of sample 566 as *P. avitruviense* was confirmed with Dr. Laia Guàrdia-Valle (personal communication) based on photographs of fresh specimens and limited slide voucher material. Aspects of the amoebae, cysts, and cystospores are not recorded or are poorly defined for some described species, and

they were not among the slide voucher materials for the samples included here; therefore, those characters were not considered for these analyses.

Phylogenetic Analyses

Multiple sequence alignments and model testing were performed as in Chapter 1, with the exception that Gblocks v. 0.91 (Castresana 2000) output was used in forming the final alignment. The program was run with half gap positions allowed and the output was further modified by eye to exclude any remaining ambiguous sites. Trees were built and analyzed three ways: MP using TNT (Goloboff et al. 2008), ML using RAxML v. 8.0.22 (Stamatakis 2014), and BI using BEAUti and BEAST v. 1.8.1. (to create the xml input file and run the analysis, respectively) (Drummond et al. 2012). For MP, 10,000 random addition sequences using TBR were followed by 5,000 bootstrap replicates to create a 50% majority rule tree (branch lengths of 0 collapsed). The GTR + Γ + I model was used without partitioning the alignment for ML, based on results from PartitionFinder v. 1.1.0 (Lanfear et al. 2012). That model was listed as second best by jModelTest 2.0 (Darriba et al. 2012; Guindon & Gascuel 2003) using AICc, behind TIM1 + Γ + I. Using BEAUti, substitution models were unlinked, *Paramoebidium* and *Amoebidium* were separated into two taxon sets (to create reciprocally monophyletic groups), the GTR + Γ + I model was used along with the lognormal relaxed clock (Drummond et al. 2006) mutation model and Yule process model of speciation (Gernhard 2008; Yule 1925) using a random starting tree. The uncorrelated lognormal relaxed clock mean was set as an exponential distribution with a mean of 10. This analysis was run for 20 million generations in BEAST and the output was checked for effective sampling and convergence in Tracer v.

1.5 (Rambaut & Drummond 2009). Branches with greater than 70 for MP, 75 for ML, and 0.95 for BI are considered well supported.

Ancestral State Reconstruction (ASR)

Character State Coding

Morphological features of the hyphae commonly used in species delineation were evaluated for their utility in that regard via ancestral state reconstructions (Table 2.1). Characters and states were chosen based on published descriptions and features observed among voucher slide materials available for the samples included here. The ancestral host association was also reconstructed at the order level for the deeper, well-supported nodes. Host orders include: Plecoptera (stoneflies) (0), Ephemeroptera (mayflies) (1), Diptera (black flies and midges) (2), Trichoptera (caddis flies) (3), and Cladocera (*Daphnia*) (4). Four morphological characters were chosen, two describing the holdfast and two describing the hyphal shape. Although not always recorded, holdfast types have been loosely categorized as globose (0), cylindrical (1), cylindrical with lateral grasping projections (2), discoid (flat, wide, thin layer of secretion, forming a "suction cup-like" attachment) (3), and a final descriptor defined here: wrapped (4). The "wrapped" holdfast appears as a thin, flat layer of secretion that coats the base of the hypha, without forming a "pedestal" or any visible projections. The second holdfast character is position: basal (0), lateral (here defined as approximately 90° or greater angle away from the proximal end of the hypha) (1), and central (i.e. middle or near middle of the hypha such that the hypha becomes bifurcated or branched) (2). Terms to describe hyphal growth forms so far include "straight and cylindrical", "branched", "hooked", and "curved". These forms were seen among voucher materials, but additional categories

were added to accommodate as-yet undescribed shapes, to include: branched (0), straight to sinuous (1), coiled to looped (coiled = hypha forming a spiral shape, curving up to 360°, but not crossing over itself; whereas looped hyphae curve greater than 360°, to cross back over itself) (2), arched [i.e. hypha with obtuse, gradual curvature, bending to create a half-circle – as illustrated for *P. chattoni* (Valle 2014b)] (3), hooked [hypha forming an approximately 180°, obtuse bend within the lower (=proximal) half of the hypha – as illustrated for *P. hamatum* (Bench & White 2012)] (4), hairpin [hypha with an approximately 180°, acute bend in the middle, such that the length of the hypha is nearly equally divided (i.e. folded in half) – as illustrated for the type species, *P. inflexum* (Léger & Duboscq 1929b)] (5), and bent (differing from "hooked" by an acute bend, and "hairpin" by the folded portion occurring in the lower half of the hypha, rather than in the middle) (6). Lastly, the consistency of the width along the length of the hyphae may be defined as: equal (less than 60% difference from the widest point to the narrowest point along the length) (0), tapering distally (= widest portion of hypha is proximal, defined as greater than 60% difference in width from widest point to narrowest point, with a gradual reduction in width along the length) (1), tapering basally (= widest portion is at the distal end) (2), middle (= widest portion is in the middle) (3), base (widest point is proximal, but there is an abrupt narrowing of the hypha, after which point the remaining length of the hypha has an equal width) (4). Finally, *Paramoebidium* species are recognized to have considerable morphological variation, so samples were coded based on the dominant form seen among the voucher material, but there were some instances of variation that will be discussed. Only "mature" hyphae were evaluated and measured, but much of the noted hyphal variation may be seen among "immature" forms. For

example, many coiled, looped or hairpin hyphal shapes appear hooked when immature. Similarly, when the cellular content is dividing into amoebae prior to dispersal, the width along the length of the hyphae may become distorted, with certain portions (e.g. proximal) becoming somewhat inflated compared to the vegetative state.

Table 2.1 *Paramoebidium* characters and character state coding used for SIMMAP and BaTS analyses.

Character	State 0	State 1	State 2	State 3	State 4	State 5	State 6
Host Order	Plecoptera	Ephemeroptera	Diptera	Trichoptera	Cladocera		
Holdfast type	Globose	Cylindrical	Grasping	Discoid	Wrapped		
Holdfast position	Basal	Lateral	Central				
Hypha curvature	Branched	Straight/ sinuous	Coiled/ looped	Arched	Hooked	Hairpin	Bent
Hypha width	Equal	Tapering distally	Tapering basally	Middle	Base		

Ancestral State Reconstruction Analyses

Reconstructions were recorded for 24 well-supported (by ML and BI) nodes at varying depth in order to compare the strength of the characters at species level and above. Analyses were run three ways: using parsimony and likelihood methods in Mesquite v. 3.01 (Maddison & Maddison 2014) and Bayesian methods in SIMMAP v. 1.5 (Bollback 2009; Bollback 2006). Just as phylogenetic trees are typically inferred and evaluated by more than one method due to the different assumptions and model complexities inherent in each, Ekman et al. (2008) recommended ASR analyses be conducted several ways as well. Concordance among methods could suggest a degree of robustness of signal in the data. For parsimony analyses, states were reconstructed across the set of the last 5,005 trees from the BEAST analysis (Trace Character Over Trees command in Mesquite) using the unordered model. Thus, the resulting pie charts reflect

the proportion of trees (that contain the node) that return a particular state as the best state at each node. Currently, Mesquite cannot perform the Trace Character Over Trees function for a maximum likelihood analysis, so the maximum clade credibility tree generated by Tree Annotator v. 1.8.1 (Drummond et al. 2012) was used. The analysis as implemented in Mesquite also cannot handle missing data, so results were not obtained for some nodes. Reconstructions were performed with a one-parameter Markov k-state model (Lewis 2001). In SIMMAP, priors were chosen using the two-step approach as outlined on the website (Bollback 2009), which consists of first using an MCMC analysis to sample overall rate and bias values followed by sampling the posterior distributions of these parameters to find the best fitting distribution in the R Statistical Package (http://www.r-project.org/ 2015). Both the equal (1/k) and empirical bias priors were tested over 1 million generations, sampling every 200 generations, with a burn-in of 10,000 and the upper bound set at 1,000. The MCMC output was imported into Tracer to compare the log likelihood values and check for effective sampling. All characters were unordered. Using the priors indicated from the R plots, analyses were conducted with k=60, rate=1.00, 20 samples/tree and 20 prior draws. Results were visualized as pie charts with the R script "PlotSimMap.R" (https://github.com/nylander/PlotSimMap) and cut and pasted into table format using Inkscape (https://inkscape.org/en).

Bayesian Tip-association Significance Testing (BaTS) and Genealogical Sorting Index (GSI) Analyses

The BaTS beta v. 2 program (Parker et al. 2008) was used to further test for correlation between morphological characters and the phylogeny. Characters and their states were coded the same way as for SIMMAP and the last 1,002 trees (final 501 trees

from each of two runs) from the MrBayes analysis were used to run 1,000 replicates. However, host family and genus were added into this analysis as these characters include too many states to be analyzed by ASRs, but their possible connection to the phylogeny was of interest. Support for delineation of well-supported (by ML and BI) clades as potentially new species was evaluated with the GSI (Cummings et al. 2008). This analysis holds the topology constant, but randomly shuffles taxa at the tips many times (thus randomizing the ancestry of clades) to generate a null distribution that observed values are compared against. The null hypothesis is: the amount of exclusive ancestry observed is that which might be observed at random (Cummings et al. 2008). The program accepts 100 trees as input, so the last 50 trees from each of the two MrBayes runs were used to create the tree file and 10,000 permutations were conducted. Samples were sorted into putative taxa groups based on their inclusion in strongly supported, monophyletic clades, here referred to as clade 1-6, sp.1-3, and cf. *chattoni*/*grande* complex. Sample 488 was included with the "*hamatum*" group.

Results

Genetic Data and Phylogenetic Analyses

One hundred thirty (as shown in Chapter 1, Table 1.1) new *Paramoebidium* rDNA sequences were generated: 72 18S and 58 for 28S. The final alignment consisted of 1,521 18S bp characters (566 excluded), 1,261 28S bp characters (665 excluded), and approximately 15% missing data. Dr. Laia Guàrdia-Valle (Universitat Autònoma de Barcelona, Spain) graciously provided samples of *P. chattoni*, but unfortunately, amplification, sequencing, and cloning attempts only recovered host data, and further processing of the samples was not possible during the course of this study. Sample 870

was collected from a trichcopteran, providing the first record of a *Paramoebidium* from that host order, family, and genus (Trichoptera, Lepidostomatidae, *Lepidostoma*). Additionally, vouchers represent new records from the following families and genera: mayflies: Baetiscidae (*Baetisea*), Ephemerellidae (*Drunella*, *Ephemerella*, *Eurylophella*); from previously recorded families: *Ameletus* (Ameletidae), *Callibaetis*, *Fallceon* (Baetidae), *Leptophlebia*, *Paraleptophlebia* (Leptophlebiidae), *Siphlonurus* (Siphlonuridae); stoneflies: Peltoperlidae (*Soliperla*, *Yoraperla*), Perlidae (*Acromeuria*); from previously recorded families: *Capnopsis* (Capniidae), *Leuctra* (Leuctridae), *Nemurella*, *Protonemura*, *Shipsa* (Nemouridae), *Taeniopteryx* (Taeniopterygidae). New geographic records for *Paramoebidium* spp. collection presented here include: British Columbia (Canada), Mexico, and Oregon (USA). Furthermore, sample 566 represents a new collection record for *P. avitruviense* from Norway.

There were no topology conflicts supported by more than one analysis method, however several branches supported by ML and BI did not have support above 70 from MP (as indicated by a star in Fig. 2.1). The 24 branches supported by ML and BI that were used for subsequent ASR analyses are labeled in Fig 2.1. Branch 7 supports the clade of all samples from stonefly hosts, except for 459, which forms an unsupported clade with a *P. curvum* voucher (1196) from Mexico. This *P. curvum* voucher is on a longer branch, as is sample 1228, potentially obscuring resolution of their placement via long branch attraction. Other host-associated branches include 13, 15, 18, 23 (mayflies), and 21 (black flies). Multiple vouchers of identified specimens (*P. avitruviense*, remaining *P. curvum*, *P. ecdyonuridae*, *P. hamatum*) formed clades supported by at least two methods. Branch 24 (including all *Paramoebidium* vouchers) was forced to be

monophyletic in the BEAST analyses, but this clade is supported by all three methods (MP, ML, BI) without topology constraints.

Morphology

In many cases voucher material was limited or not available, but morphometric analyses were sufficient to separate samples from known species (except for 566, as mentioned above). Table 2.1 lists the morphological findings for individual samples. Only a few samples had potentially enough slide voucher material to name a new species, but that is beyond the scope of this thesis. Variability within character states, as outlined here, was noted, such as the degree to which the lateral, grasping projections on holdfasts incurved or the acuteness of curvature among individual hyphae. Several morphotypes were observed that are not included with accepted descriptions, but do resemble those in the *nomen nudum* list (Lichtwardt et al. 2001). These unvalidated taxa are still relevant as they afford researchers starting points from which to initiate collection efforts, and possibly lead to validation, as was recently demonstrated for *P. chattoni* (Valle 2014b).

As hypha curvature is the character with the most significant states in the BaTS analyses, this discussion will focus on those results. Within that character, “hairpin”, “sinuous”, “looped”, and “bent” are all unrepresented states in the literature, although *P. inflexum* (the type for the genus) is illustrated such that it evokes the “hairpin” and “hooked” shapes (Léger & Duboscq 1929b). The hairpin type has been illustrated for *P. arcuatum* (Léger & Duboscq 1929b), *P. pavillardi* (Manier 1950), and *P. giganteum* (Duboscq et al. 1948) *nomen nudum* taxa. The first two were collected from *Baetis* mayflies, the former from *B. gemellus* and *B. rhodani*, and the latter from *B. atrebatinus*, whereas *P. giganteum* was collected from *Chloroperla* stoneflies. The hairpin specimens

here were collected from baetid mayflies (449, 691, 703, and 690, the host of which was identified as *B. rhodani*) as well as perlid and peltoperlid stoneflies. Unfortunately, very limited material was available for these samples (state determination for most was based on comments in the dissection logs of the respective researchers), so comparison beyond host types and general shape is not feasible. However, *P. giganteum* hyphae were recorded up to 2,700 µm long and the single hypha measured for 1219 was over 3,500 µm.

Looped hyphae are found in the drawings of *P. arcuatum* and *P. dispersum* (Duboscq et al. 1948), but the latter was taken from a leptophlebiid mayfly (*Habrophlebia*). There was one very distinct example seen here for sample 1200. A couple of hyphae for 1227 superficially had this shape, but it was attributed to artifacts of slide fixation because the voucher slide contained many hyphae, the remainder of which strictly adhered to a coiled curvature. However, both slides had "hooked" shaped hyphae as well. Additionally, both were collected from mayflies, but the host of sample 1227 was only identified to order, and 1200 to family (Ephemerellidae).

Finally, sinuous and bent, as they are defined here, do not appear even among *nomen nudum* taxa, and were not entirely common among examined specimens, as opposed to "hooked". Indeed, the examples of sinuous hyphae seen here (680, 664, 715) are quite distinctive, as they also have a knob-like base, papillae, and a distally positioned, lateral, thumb-like projection. The bent shape was likewise observed in combination with other distinguishing features. For example, sample 551 hyphae had a bulbous base followed by a constriction in the width at the bent portion, and the

photographs of 538 showed extremely thick, consistently wide hyphae (apparently on the verge of rupturing to release amoebae).

Ancestral State Reconstruction

There was general concordance among the three methods for most characters at most nodes (Fig. 2.2). Likelihood could not provide reconstructions for holdfast type at nodes 6, 12, 15, 19, and 22, for holdfast position at nodes 6, 15, 19, and 22, for curvature at nodes 12, 15, and 22, or for width at nodes 12, 15, 19, and 22. Similarly, parsimony found no best states on any tree for holdfast type and position at node 22, curvature at node 19, and width at node 12. These nodes (6, 12, 15, 19, 22) were problematic due to missing character state data. Figure 2.1 shows symbolic representations of the most probable states reconstructed for labeled clades and the deeper nodes for which host orders were analyzed. Character coding is indicated in Table 2.1, with representative drawings of character states in Fig. 2.3.

Plecoptera is given a probability near 1 by all three methods for nodes 4 and 7, whereas Ephemeroptera has the highest probability at the remaining nodes, with the exception of 21, which subtends the cf. *chattoni*/*grande* group. The most commonly reconstructed state for holdfast type is “grasping projections”, but this character also had the highest number of reconstruction disagreements between methods (six total). For example, at node 3 parsimony returns “discoid” as the best state, but likelihood and SIMMAP give “cylindrical” the highest probability. Furthermore, all three methods disagree at node 4, with parsimony favoring “discoid”, likelihood “globose”, and SIMMAP “grasping projections”. Interestingly, there was agreement between methods

for this character at deeper nodes (i.e. 7, 13, 20, 24). A comparable pattern may be observed for hypha curvature, with conflicts arising at "midlevel" nodes 4, 19, and 21.

On the other hand, there was only one conflict for hypha width and holdfast position for reconstructions. The former occurring at node 3, with parsimony suggesting "basal" rather than "central" of the other two. The latter is observed at node 1, with SIMMAP giving less than half probability to "base", the most likely state returned by parsimony and likelihood. "Equal" was the most common state observed overall for hypha width, and this is reflected among the reconstructions; only nodes 1 and 3 give an alternate state.

GSI and BaTS Analyses

The GSI analysis found significant results for clades 1-6, sp. 1-3, and the "cf *chattoni*/*grande*" group. In addition, vouchers identified as *P. avitruviense*, *P. ecdyonuridae*, and *P. hamatum* were supported. BaTS indicated significant correlation of several character states with the phylogeny (Table 2.2). Specifically, for host states, three orders, none families, and five genera returned significant values. Morphologically, fewer states appear informative, with three states regarding holdfast and six states regarding hyphal characters being significant. Of these, hyphal curvature had the most states correlated with the phylogeny.

Discussion

Utility of Morphological Characters

The overall agreement among ancestral state reconstruction methods and the significant result for some states (found with the BaTS program) suggest that there is a degree of phylogenetic signal for the morphological characters analyzed. The broad,

relatively general characters and states used here align with morphotaxonomic platforms for *Paramoebidium* species that have been used to date. *Paramoebidium avitruviense* and *P. curvum* both have been described using the morphological species concept and are sufficiently represented here to provide initial commentary. Firstly, the *P. avitruviense* from Norway (collection code NOR-22-W10, voucher 566), identified morphologically, gives some insight into the possible range of intraspecific morphological variability when compared to the voucher from that species' type locality and host (collection code SPA-X-74, voucher 1247). Upon examination, the Norway specimen matched the type (Valle 2014a) in having branched hyphae, a central, globose holdfast, and a small "pedestal-like" protrusion above the holdfast. However, it differed from the type in its dimensions, being longer and wider. Unfortunately, the only mature material for the Norway specimen was a film photograph taken through the objective of a stereomicroscope, so measurements are somewhat approximate, but could be as much as 200 μm longer and 50 μm wider. Furthermore, both samples were taken in the spring, and the Norway sample is from a different species of the same host genus. Therefore, the postulation that specific morphometric dimensions have a wider range of variation than general shape, overall aspect, and even possibly host organism, does not seem unqualified.

In contrast, the *P. curvum* vouchers from Norway (681) and Mexico (1196) matched the dimensions and hyphal features of the type (Dang & Lichtwardt 1979). No material was available for the Africa sample (197), however. Whereas 197 and 681 form a clade with support from BI and ML, 1196 is on a long branch, likely contributing to the obfuscation of its placement. *Paramoebidium curvum* is easily recognized by its consistent position in the posterior hindgut, even located as far outside as the anal

papillae, and short, coiled hyphae (Dang & Lichtwardt 1979). A review of slide vouchers did not reveal any morphological features to suggest that 1196 is not *P. curvum*, so a number of possible explanations for its position in the phylogeny exist, including cryptic speciation. The other two species vouchers, though collected from geographically distant locations, are recovered as sister taxa. Therefore, factors such as environment and host may be important to consider. The collection location of sample 1196 in Mexico was a high elevation (1,667 m) tropical uplift area, and the genetic divergence indicated by the long branch could be a reflection of the effects of this environment on the symbiont and its host (e.g limited dispersal, endemism, etc.). Indeed, voucher 1223 (collected from a nearby location in Mexico) is similarly diverged from its relatives, but its morphology was not, at least at the broad level examined here. Future phylogeographic studies will help elucidate whether, and to what degree, allopatric speciation drives such trends among *Paramoebidium* populations.

Although *P. ecdyonuridae* and *P. hamatum* have two vouchers each, they were collected by the same individual, from the same locality, at a single time point. Therefore, the conclusions that may be drawn from their placement in the phylogeny are somewhat more limited. In particular, *P. hamatum* was described from not only two different host families (Ameletidae and Baetidae), but also two different host orders (Ephemeroptera and Chironomidae) (Bench & White 2012), an unusual extension beyond observed host boundaries in traditional trichomycete taxonomy. Based on the prevalence of the "hooked" morphotype found in this study and correlation of host type with the phylogeny, it seems doubtful that the *P. hamatum* Bench & White (2012) observed in chironomid larvae is closely related to the *P. hamatum* from mayfly nymphs. Especially

considering that *P. ecdyonuridae* also has this hooked curvature, a size range that overlaps that of *P. hamatum*, and a mayfly host, but they are separated by considerable genetic distance in the phylogeny. Of course, a definitive resolution to this problem awaits phylogenetic analysis with vouchers from each *P. hamatum* host type to identify whether their observed specimens are all "true" *P. hamatum* representatives, in a genealogical sense. Indeed, Bench & White (2012) suggested the possibility of cryptic speciation and the utility of genetic sequencing to help unravel host specificity versus promiscuity in the species.

Morphological consistency among the remaining unnamed vouchers would indicate that the analyzed morphological features have some taxonomic signal. However, they do present varying degrees of inferred utility, as only a few states are significantly correlated with the phylogeny. Firstly, curvature of the hypha is principally important as it encompasses the primary features that have been used to delineate species to date. The most problematic of these is the hooked morphotype, which was observed among various mayfly and stonefly samples. Vouchers 506, 511, and 1200 (Plate 2.1) are closely related to *P. ecdyonuridae*, for example, and the first two bear significant resemblance to that species except that they lack a papillum and their maximum length and width measurements are only half of the maximum reported (Valle 2014a). Without a molecular phylogeny, these samples could readily be ascribed to *P. hamatum* in light of that species' host affiliations, described above. Sample 1200 is more easily differentiated, however, as many of the hyphae are looped, but it remains to be determined if such a feature falls within the acceptable range of intraspecific variability. Additionally, 504 is another example of a *P. hamatum* doppelganger, but it is positioned

distantly from that species and *P. ecdyonuridae* in the tree, and itself forms a distinct clade.

On the other hand, sample 488 from Utah was collected from a baetid mayfly and morphologically matched the type of *P. hamatum* from Idaho mayflies (see Fig. 79 in Bench & White 2012). Indeed, 488 forms a well-supported clade with the *P. hamatum* samples, a relationship that is reinforced by a significant GSI value for the *P. hamatum* clade. Interestingly, the hyphae depicted in Bench & White (2012) from chironomid larvae (Figs. 74-78 in that publication) appear to have a more tightly coiled bend at the proximal end than those from mayflies. This slight difference creates a hyphal shape analogous to a "sewing needle" rather than the "candy cane" appearance of those found in mayflies. While this distinction was not noted by the authors, the mixed results of the utility of hyphal morphology found here cannot suggest whether this feature falls within the range of intraspecific variability or represents morphological evidence of a different species. Possibly, the actual degree measurement of hyphal curvature of mature specimens could prove to be a character used in future morphotype investigations, and even species descriptions.

Another illustrative example of morphotaxonomic complexity is the cf. *chattoni*/*grande* clade (Plate 2.2). Not only are both species found in the same hosts (simulids), but also the description of *P. grande* (Lichtwardt & Arenas 1996) is very brief and lacks mention of the holdfast, and the hyphal measurements overlap those of *P. chattoni* (Valle 2014b). In fact, *P. grande* is a rare example of a species for which dimensions of the amoebae were considered relevant, being comparatively quite large. Some of the original slide materials from the *P. grande* type collection from Chile were

received (from R.W. Lichtwardt, now with M.M. White, BSU) and reviewed, but a clear determination of which hyphae belonged to the *P. grande* designation could not be easily made. There were slides that had longer hyphae (i.e. greater than 800 μm), and thus were more *P. chattoni*-like, but there were also other shorter, wider hyphae that would fit either the *P. grande* or *P. chattoni* description. Further confounding these efforts is that only a single hypha was imaged in the *P. grande* publication (Lichtwardt & Arenas 1996). The specimens for the vouchers in the cf. *chattoni*/*grande* clade are consistent in having an arched curvature, a holdfast with grasping projections, and measurements that fall within the range of both species. However, samples 1049 and 833 have distally tapering widths, whereas the remaining samples are either wider in the middle or are more equal in width along the length. Furthermore, these specimens had lateral holdfasts, while the rest were basal. Despite these differences, their sequences are not very diverged from others in the clade in this tree, but in the large rDNA phylogeny (see Chapter 1), BI supports them as a separate group. Overall, members of this clade are more *P. chattoni*-like in their morphology, but their relatedness to or identity with *P. grande* cannot be ruled out. Certainly genetic samples from established and confirmed identifications of both species, especially *P. grande* from Chile, are needed to clarify their positions and relationships.

All three of these examples illustrate cryptic speciation as a real challenge, but also potential boon for *Paramoebidium* taxonomy and scope of biodiversity. It is perhaps not unfair to suggest that there are some inherent taxonomic pitfalls in the way species traditionally have been described. Reproductive characters that may have been earlier discounted (Lichtwardt et al. 2001; Whisler 1963), could be reconsidered for their utility,

or at least be better tested to confirm or refute their effectiveness. This is suggested while noting that, practically speaking, the reproductive structures are often more difficult to obtain in routine collections and likely would demand longer-term efforts and even modified laboratory methods. Nonetheless, descriptions of the amoebae, cysts, and cystospores have been reported for many species, and close inspection of amoebae of *P. ecdyonuridae* recently discovered uroidal filaments (Valle 2014a), indicating the prospect of a new character to investigate. Clearly, a combination of characters is required for any adequate delimitation, but those used to date should be reevaluated and supplemented with methods that include genomic samples for molecular study, to generate a more robust, integrative taxonomic framework.

Paramoebidium Diversity and Host Specificity

As noted, host organism was originally a character used for *Paramoebidium* species considerations and delimitation. Certain host types are significantly correlated with the phylogeny in the BaTS analyses, even to the genus level. *Paramoebidium* samples collected from stonefly hosts are clearly distinct from those collected from mayflies and black flies, forming a well-supported clade. Sample 459 is the only stonefly specimen that falls outside that clade. The host voucher for this sample was not preserved, but it was field identified as *Shipsa*. A misidentification cannot be ruled out, but if correct, its placement could signal a host-switching event or it could represent an example of a generalist *Paramoebidium* species. Similarly, sample 870 was collected from a trichopteran, but it is included in the well-supported stonefly clade. Surely there could be varying levels of host specificity among these symbionts, but such an evaluation requires additional samples. How old is the relationship between *Paramoebidium* spp.

and their hosts? Has coevolution occurred? The Ichthyosporea, located near the divergence point of animals and fungi (Cavalier-Smith et al. 2014), is an ancient group, thus the amoebidiids could have a long, shared evolutionary history with these arthropods. The preliminary analyses presented here cannot address these questions, but they do suggest the possibility of host specificity, which may lead future investigations of coevolutionary hypotheses.

What is clearly demonstrated by this phylogeny is that *Paramoebidium* is much more diverse than previously understood, with at least three putatively new species and six potential species clusters. A long-standing debate among protistologists has been whether protist species are globally distributed (i.e. everything is everywhere) or if there are many endemic species that are masked by cryptic speciation (e.g. Foissner 1999; Slapeta et al. 2005; Foissner 2006). Recent studies have provided evidence for a combination of the two, with some cosmopolitan species, and some that are endemic (Cotterill et al. 2008; Weisse 2008; Caron et al. 2012). Surveying the, albeit limited, topology and taxa presented here, there is evidence to suggest the latter hypothesis. For example, there is little genetic divergence among samples in the cf. *chattoni*/*grande* clade and *P. curvum* from Africa and Norway, but the *Paramoebidium* spp. from Mexico (1196 and 1223) are diverged from their close relatives. However, whether and how much microhabitat versus broad-scale geographic partitioning impact these observed patterns remains to be determined. Whatever the case, the number of distinctive clades and new host associations, and even the level of differentiation of vouchers within clades (e.g. clade 3) point to unrealized species richness. With thousands of potential host species distributed globally, and records of *Paramoebidium* spp. from locations such as

Argentina (Lichtwardt et al. 1990), China (Strongman et al. 2010), and New Zealand (Lichtwardt & Williams 1992), it is clear that much remains to be discovered.

Future Species Descriptions and Species Concept

Going forward, a shift from an emphasis on a strictly morphological species concept to an integrative approach that combines morphology, ultrastructure, and especially, gene sequences, should be implemented, as has been advocated for, and applied to, other taxonomically challenging organisms (e.g. Azevedo et al. 2015; Katz et al. 2015; Li et al. 2015; Lecocq et al. 2015). At the same time, additional informative characters should be sought, such as protein-coding genes, especially considering the difficulty in assessing homology and the subjective nature of morphological evaluations. Future species descriptions would benefit from an approach that included: 1) sequence data or at least a genomic voucher that can be processed later, 2) host organism identification, and if the species is thought to occur in more than one genus of host, then morphological and genomic vouchers of the putative species from each different host type should be obtained, 3) a detailed, statistically analyzed report of intraspecific variation when deemed necessary for differentiation from other species, as recently published for *Enterobryus luteovirgatus* (Contreras & Cafaro 2013), and 4) assessments of amoebae, cysts, and cystospores, whenever possible. Hoberg et al. (2015) highlighted the need for and importance of taxonomic clarification in the field of parasitology, and the implications such would have in terms of broad understanding of biodiversity, evolution, and ecology. An invigorated interest in, and study of, *Paramoebidium* and other members of class Ichthyosporea as whole would contribute to, and advance

considerably, our understanding of these symbiotic, unicellular relatives of multicellular animals and fungi.

Table 2.2 Character state coding for *Paramoebidium* and *Amoebidium* samples. Grey lines and dashes indicate no data was available for that sample/character. Length and width measurements of the hyphae, and presence of a papillum and other projections are given, but these were not analyzed with BaTS or ASR methods. See Figure 2.3 for representative drawings of characters.

Sample and Collection code	Hypha Length range (µm)	Hypha Width range (µm)	Holdfast type	Holdfast position	Papillum	Hypha curvature	Hypha width	Other projections
1048 *Paramoebidium* sp. TN-38-W15	3597	103	-	-	-	hairpin	equal	-
1049 *Paramoebidium* sp. ALG-9-W10	650-1188	29-61	cylindrical	basal	n	arched	tapering distally	n
1052 *Paramoebidium* sp. OR-3-W12	-	-	-	-	-	hairpin	-	-
1176 *Paramoebidium* sp. AR-31-C7	634-672	25-26	-	-	-	hooked	equal	n
1181 *Paramoebidium* cf. *grande* ID-164-G2	-	-	-	basal	n	arched	-	n
1183 *Paramoebidium hamatum* ID-157-G11	204	14	cylindrical	basal	n	hooked	equal	n
1196 *Paramoebidium curvum* Mex-16-W14	141-201	19-34	cylindrical	lateral	n	coiled/ looped	equal	n
1200 *Paramoebidium* sp. DR-16-C8	232-353	13-21	cylindrical	basal	n	hooked	tapering distally	n
1207 *Paramoebidium hamatum* ID-166-G6	260-800	10-60	cylindrical	basal	n	hooked	equal	n
1210 *Paramoebidium stipula* NS-X-39	650-700	45	cylindrical	central	n	branched	tapering basally	y
1214 *Paramoebidium corpulentum* KS-114-3	300	30-90	globose	basal	n	arched	tapering distally	n
1215 *Paramoebidium* sp. NS-X-29	725-965	10-25	discoid	lateral	y	straight	equal	n
1217 *Paramoebidium* sp. OR-13-W1								
1218 *Paramoebidium* sp. TN-13-W20	2070-4006	23-47	cylindrical	basal	n	straight	equal	n
1219 *Paramoebidium* sp. TN-38-W13	3597	103	-	-	-	hairpin	equal	n
122 *Paramoebidium* sp. RMBL-75-1								
1223 *Paramoebidium* sp. Mex-17-W2	850-1443	21-59	cylindrical	basal	n	arched	middle	n
1225 *Paramoebidium* sp. OR-14-W1	225-316	8-11	discoid	basal	n	hooked	tapering distally	n
1227 *Paramoebidium* sp. CO-3-W13	255-541	13-31	cylindrical	basal	n	coiled/ looped	equal	n
1228 *Paramoebidium* sp. BC-5a-W6	870-1321	10-27	cylindrical	basal	n	straight	tapering basally	n
1234 *Paramoebidium* sp. NS-16-W2	290-462	15-19	cylindrical	basal	n	hooked	equal	n
1236 *Paramoebidium* sp. ID-65e-E1	1396	33	cylindrical	basal	n	arched	equal	n

1237 *Paramoebidium* sp. ID-156-T1	98-202	45-78	cylindrical	basal	n	straight	tapering distally	n
1247 *Paramoebidium avitruviense* SPA-X-74	670-770	20-30	globose	central	n	branched	tapering distally	y
1249 *Paramoebidium ecdyonuridae* SPA-X-76	1000	80	cylindrical	basal	y	hooked	equal	n
1250 *Paramoebidium ecdyonuridae* SPA-X-77	1000	80	cylindrical	basal	y	hooked	equal	n
197 *Paramoebidium curvum* AFR-9	140-280	20-60	cylindrical	lateral	n	coiled/ looped	equal	n
303 *Paramoebidium* sp. NS-34-W17	-	-	discoid	basal	n	hooked	equal	n
376 *Amoebidium parasiticum* JAP-7-2	32-500	6-11	discoid	-	n	straight	equal	n
377 *Amoebidium parasiticum* A1a	32-500	6-11	discoid	-	n	straight	equal	n
38 *Paramoebidium* sp. ME-2-W3	204-448	7-14	cylindrical	lateral	n	hooked	equal	n
41 *Paramoebidium* sp. NS-6-W8-10	289-299	10-11	discoid	basal	n	hooked	equal	n
42 *Paramoebidium* sp. NY-3-W7								
446 *Paramoebidium* sp. KS-61-W20	345-600	12-22	cylindrical	basal	y	arched	equal	n
447 *Paramoebidium* sp. MA-3-W51	833-1106	19-37	cylindrical	lateral	y	straight	tapering basally	n
449 *Paramoebidium* sp. MO-19-W15	-	-	-	-	-	hairpin	-	-
450 *Paramoebidium* sp. NF-10-W15								
457 *Paramoebidium* sp. RMBL-71-3								
458 *Paramoebidium* sp. RMBL-72-3								
459 *Paramoebidium* sp. PA-2-W2	138	10	cylindrical	basal	n	hooked	equal	n
466 *Paramoebidium* sp. NS-24-W16								
467 *Paramoebidium* sp. VT-3-W1	750	37	-	basal	n	hooked	middle	y
473 *Paramoebidium* sp. UT-1-W14	968-1265	39-53	cylindrical	basal	n	arched	middle	n
488 *Paramoebidium* sp. UT-2-W5	379-434	31-39	cylindrical	basal	n	hooked	middle	n
504 *Paramoebidium* sp. NS-35-W8	996	54	cylindrical	basal	n	hooked	equal	n
506 *Paramoebidium* sp. NS-35-W14	308-549	18-41	cylindrical	basal	n	hooked	equal	n
508 *Paramoebidium* sp. NS-35-W18	154-495	6-12	discoid	lateral	n	hooked	equal	n
511 *Paramoebidium* sp. NS-35-W22b	279-504	14-23	cylindrical	basal	n	hooked	equal	n
525 *Paramoebidium* sp. NOR-3-1	180-210	60-70	discoid	basal	n	arched	equal	n
526 *Paramoebidium* sp. NOR-3-W2	609-1026	58-64	discoid	basal	n	hooked	tapering basally	n
531 *Paramoebidium* sp. NOR-4-W9								
533 *Paramoebidium* sp. NOR-7-W2								
536 *Paramoebidium* sp. NOR-5-W14								
538 *Paramoebidium* sp. NOR-5-2	220-380	55-90	globose	basal	n	bent	equal	n
543 *Paramoebidium* sp. NOR-10-W1	300	75	discoid	basal	n	straight	equal	n
546 *Paramoebidium* sp. NOR-10-W10a								

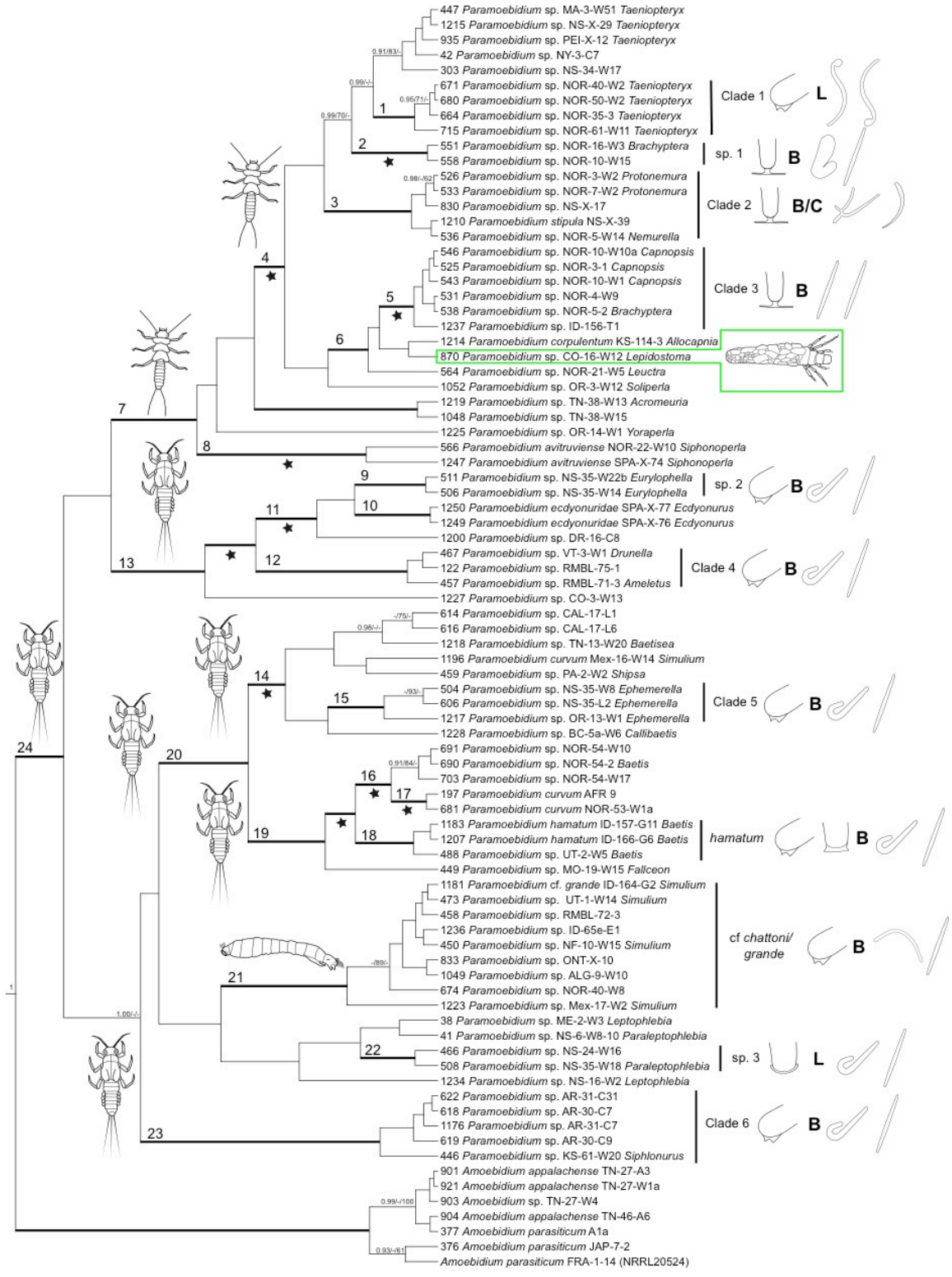
551 *Paramoebidium* sp. NOR-16-W3	449-521	40-46	discoid	basal	n	bent	equal	n
558 *Paramoebidium* sp. NOR-10-W15	-	-	discoid	basal	n	bent	equal	n
564 *Paramoebidium* sp. NOR-21-W5	3000	50	-	-	n	straight	equal	n
566 *Paramoebidium avitruviense* NOR-22-W10	800-1000	80-90	globose	central	n	branched	equal	y
606 *Paramoebidium* sp. NS-35-L2	-	-	-	-	-	hooked	equal	
614 *Paramoebidium* sp. CAL-17-L1								
616 *Paramoebidium* sp. CAL-17-L6								
618 *Paramoebidium* sp. AR-30-C7	242-460	10-13	discoid	basal	y	hooked	equal	n
619 *Paramoebidium* sp. AR-30-C9								
622 *Paramoebidium* sp. AR-31-C31								
664 *Paramoebidium* sp. NOR-35-3								
671 *Paramoebidium* sp. NOR-40-W2								
674 *Paramoebidium* sp. NOR-40-W8	1080	31	discoid	basal	n	arched	equal	n
680 *Paramoebidium* sp. NOR-50-W2	485-791	17-40	cylindrical	lateral	y	straight	base	y
681 *Paramoebidium curvum* NOR-53-W1a	140-280	20-60	cylindrical	lateral	n	coiled/ looped	equal	n
690 *Paramoebidium* sp. NOR-54-2	-	-	-	-	-	hairpin	-	-
691 *Paramoebidium* sp. NOR-54-W10	-	-	-	-	-	hairpin	-	-
703 *Paramoebidium* sp. NOR-54-W17	375-750	20-30	cylindrical	basal	n	hairpin	equal	n
715 *Paramoebidium* sp. NOR-61-W11	417-552	30-49	cylindrical	lateral	-	straight	base	-
830 *Paramoebidium stipula* NS-X-17	650-700	45	cylindrical	central	n	branched	tapering basally	y
833 *Paramoebidium* sp. ONT-X-10	685-883	29-51	cylindrical	lateral	n	arched	tapering distally	n
870 *Paramoebidium* sp. CO-2-W12	190-494	42-143	globose	basal	n	straight	equal	n
901 *Amoebidium appalachense* TN-27-A3	55-80	4-8	discoid	basal	n	straight	equal	n
903 *Amoebidium* sp. TN-27-W4	-	-	-	-	n	straight	equal	n
904 *Amoebidium appalachense* TN-46-A6	55-80	4-8	discoid	basal	n	straight	equal	n
921 *Amoebidium appalachense* TN-27-W1a	55-80	4-8	discoid	basal	n	straight	equal	n
935 *Paramoebidium* sp. PEI-X-12	907	21	discoid	basal	n	arched	equal	n
Amoebidium parasiticum FRA-1-14 (NRRL20524)	32-500	6-11	discoid	-	n	straight	equal	n

Table 2.3 Bayesian Tip-association Significance testing (BaTS) results. The statistics (Association Index, Parsimony Score and maximum exclusive single-state clade size), observed means, null means, confidence intervals and p-values are given for all states and characters. Character state coding was the same as for ASR analyses. Bold numbers indicate significant results ($p < 0.05$). Results for character states only represented once in the dataset are not shown.

Host Order	Statistic	Observed mean	Lower 95% CI	Upper 95% CI	Null mean	Lower 95% CI	Upper 95% CI	Significance
	AI	0.84932822	0.42211232	1.26890361	5.83471251	4.91656923	6.72862387	0
	PS	9	9	9	39.1255264	35.3882217	42.5988007	0
	MC **Diptera**	9	9	9	1.74796236	1.15668666	2.4381237	**1.00E-03**
	MC Cladocera	1	1	1	1.00659776	1	1	1
	MC **Ephemeroptera**	9	9	9	2.77016902	2.00399208	4.43712568	**1.00E-03**
	MC **Plecoptera**	15.6257486	11	16	2.6916225	1.98103797	4.83732557	**1.00E-03**
Host Family	AI	2.81444359	2.24810338	3.41229606	8.33708382	7.79773092	8.80851555	0
	PS	31.6846313	31	32	63.6851463	60.5898209	66.6786423	0
	MC Chironomidae	1.5	1	3	1.06354761	1	1.32934129	1
	MC Daphniidae	1	1	1	1.00589132	1	1	1
	MC **Leptophlebiidae**	1.91616762	1	2	1.14188659	1	1.9371258	**0.028**
	MC **Siphlonuridae**	2.273453	1	4	1.07536173	1	1.54291415	**0.01099998**
	MC **Baetidae**	5	5	5	1.10325861	1	1.62175643	**0.00099999**
	MC **Capniidae**	3.00399208	3	3	1.20860493	1	1.96806383	**0.00199997**
	MC Chloroperlidae	1.6357286	1	3	1.04566026	1	1.27644706	1
	MC **Simuliidae**	1.99001992	2	2	1.00649202	1	1	**0.00099999**
	MC **Heptageniidae**	9	9	9	1.44116902	1	2.05489016	**0.00099999**
	MC **Peltoperlidae**	2	2	2	1.0029732	1	1	**0.00099999**
	MC Nemouridae	1	1	1	1.00736809	1	1	1
	MC **Ephemerellidae**	5	5	5	1.2019484	1	1.96407187	**0.00099999**

	MC **Taeniopterygidae**	2.72954082	2	3	1.14984417	1	1.91916168	**0.00099999**
Host Genus	AI	3.88455892	3.12670112	4.68431854	7.89598894	7.33976746	8.39205933	0
	PS	39.0309372	37	41	55.1049957	53.2774467	56.7235527	0
	MC *Chironomus*	1.5	1	3	1.06979287	1	1.37624753	1
	MC *Daphnia*	1	1	1	1.00731432	1	1	1
	MC ***Baetis***	3.96606779	4	4	2.37427163	1.74251497	3.39520955	**0.02200001**
	MC *Siphonoperla*	1	1	1	1	1	1	1
	MC ***Simulium***	1.99001992	2	2	1.00363779	1	1	**1.00E-03**
	MC *Ecdyonurus*	1.37225544	1	3	1.06130719	1	1.34331334	1
	MC *Ephemerella*	1	1	1	1	1	1	1
	MC ***Taeniopteryx***	2.58283424	1	3	1.02422631	1	1.15169656	**1.00E-03**
	MC *Capnopsis*	1	1	1	1	1	1	1
	MC *Brachyptera*	1.6357286	1	3	1.01953495	1	1.08982038	1
	MC *Protonemura*	1	1	1	1	1	1	1
	MC ***Eurylophella***	1.96307385	2	2	1.00694513	1	1	**0.00199997**
	MC ***Paraleptophlebia***	2	2	2	1.010445	1	1	**1.00E-03**
	MC *Leptophlebia*	1	1	1	1.00673544	1	1	1
Holdfast type	AI	4.31549406	3.49616003	5.12116909	6.69734478	5.89303732	7.43649197	0
	PS	38.1916161	36	40	46.1521873	42.6477051	49.4451103	0
	MC Cylindrical	1.73053896	1	3	1.27935398	1	2	0.05500001
	MC Discoid	2.01397204	2	2	1.65500224	1.02894211	2.46007991	0.21399999
	MC Globose	3	3	3	2.42031217	1.79940116	3.69061875	0.153
	MC **Grasping**	1.99900198	2	2	1.06341076	1	1.34331334	**0.00700003**
	MC **Wrapped**	4.97305393	5	5	2.19617701	1.52395213	3.02894211	**0.005**

Holdfast position	AI	3.99607873	3.20100331	4.73894978	5.36370373	4.49562311	6.17108488	0.00599998
	PS	31.047905	29	33	34.3312225	31.285429	36.9191628	0.04000002
	MC Basal	5.00199604	5	6	3.67770863	2.4261477	5.21656704	0.12099999
	MC Central	2.26746511	2	3	2.32362962	1.66367269	3.4311378	0.74199998
	MC **Lateral**	1.99101794	2	2	1.03565979	1	1.2275449	**0.005**
Hypha curvature	AI	3.93399239	3.23814893	4.63009644	7.33377838	6.61065817	7.99542904	0
	PS	36.8083839	35	38	52.1415367	48.6497002	55.235527	0
	MC Straight/sinuous	2.79341316	2	4	1.90325093	1.22854292	2.9431138	0.046
	MC Hooked	1.97604787	2	2	1.76432991	1.13173652	2.58283424	0.27399999
	MC **Branched**	4.88423157	3	5	1.99117339	1.34131742	2.9311378	**0.00199997**
	MC **Coiled/looped**	1.99101794	2	2	1.04131126	1	1.27644706	**0.00700003**
	MC **Arched**	2	2	2	1.06929862	1	1.4141717	**0.014**
	MC **Hairpin**	1.68662679	1	3	1.18974769	1	1.99001992	**0.04400003**
	MC **Bent**	2.38922167	2	3	1.13208866	1	1.81636727	**0.01700002**
Hypha width	AI	4.4713273	3.72924018	5.26312208	5.57150984	4.77452421	6.36616707	0.01300001
	PS	31.4910183	30	33	35.5085335	32.9221573	37.9680634	0.02200001
	MC Equal	3.99201608	3	4	3.69091487	2.48403192	5.21656704	0.28399998
	MC Distally tapering	2.26447105	2	3	2.14254498	1.45708585	3.00199604	0.59799999
	MC **Middle**	1.54291415	1	2	1.10293818	1	1.68762469	**0.01599997**
	MC Basally tapering	1.00399196	1	1	1.04601145	1	1.27644706	1
	MC Base	1.21457088	1	2	1.06769252	1	1.37624753	1


447 Paramoebidium sp. MA-3-W51 Taeniopteryx
1215 Paramoebidium sp. NS-X-29 Taeniopteryx
935 Paramoebidium sp. PEI-X-12 Taeniopteryx
42 Paramoebidium sp. NY-3-C7
303 Paramoebidium sp. NS-34-W17
671 Paramoebidium sp. NOR-40-W2 Taeniopteryx
680 Paramoebidium sp. NOR-50-W2 Taeniopteryx
664 Paramoebidium sp. NOR-35-3 Taeniopteryx
715 Paramoebidium sp. NOR-61-W11 Taeniopteryx
Clade 1
L
551 Paramoebidium sp. NOR-16-W3 Brachyptera
558 Paramoebidium sp. NOR-10-W15
sp. 1
B
526 Paramoebidium sp. NOR-3-W2 Protonemura
533 Paramoebidium sp. NOR-7-W2 Protonemura
830 Paramoebidium sp. NS-X-17
1210 Paramoebidium stipula NS-X-39
536 Paramoebidium sp. NOR-5-W14 Nemurella
Clade 2
B/C
546 Paramoebidium sp. NOR-10-W10a Capnopsis
525 Paramoebidium sp. NOR-3-1 Capnopsis
543 Paramoebidium sp. NOR-10-W1 Capnopsis
531 Paramoebidium sp. NOR-4-W9
538 Paramoebidium sp. NOR-5-2 Brachyptera
Clade 3
B
1237 Paramoebidium sp. ID-156-T1
1214 Paramoebidium corpulentum KS-114-3 Allocapnia
870 Paramoebidium sp. CO-16-W12 Lepidostoma
564 Paramoebidium sp. NOR-21-W5 Leuctra
1052 Paramoebidium sp. OR-3-W12 Soliperla
1219 Paramoebidium sp. TN-38-W13 Acromeuria
1048 Paramoebidium sp. TN-38-W15
1225 Paramoebidium sp. OR-14-W1 Yoraperla
566 Paramoebidium avitruviense NOR-22-W10 Siphonoperla
1247 Paramoebidium avitruviense SPA-X-74 Siphonoperla
511 Paramoebidium sp. NS-35-W22b Eurylophella
506 Paramoebidium sp. NS-35-W14 Eurylophella
sp. 2
B
1250 Paramoebidium ecdyonuridae SPA-X-77 Ecdyonurus
1249 Paramoebidium ecdyonuridae SPA-X-76 Ecdyonurus
1200 Paramoebidium sp. DR-16-C8
467 Paramoebidium sp. VT-3-W1 Drunella
122 Paramoebidium sp. RMBL-75-1
457 Paramoebidium sp. RMBL-71-3 Ameletus
Clade 4
B
1227 Paramoebidium sp. CO-3-W13
614 Paramoebidium sp. CAL-17-L1
616 Paramoebidium sp. CAL-17-L6
1218 Paramoebidium sp. TN-13-W20 Baetisea
1196 Paramoebidium curvum Mex-16-W14 Simulium
459 Paramoebidium sp. PA-2-W2 Shipsa
504 Paramoebidium sp. NS-35-W8 Ephemerella
606 Paramoebidium sp. NS-35-L2 Ephemerella
1217 Paramoebidium sp. OR-13-W1 Ephemerella
Clade 5
B
1228 Paramoebidium sp. BC-5a-W6 Callibaetis
691 Paramoebidium sp. NOR-54-W10
690 Paramoebidium sp. NOR-54-2 Baetis
703 Paramoebidium sp. NOR-54-W17
197 Paramoebidium curvum AFR 9
681 Paramoebidium curvum NOR-53-W1a
1183 Paramoebidium hamatum ID-157-G11 Baetis
1207 Paramoebidium hamatum ID-166-G6 Baetis
488 Paramoebidium sp. UT-2-W5 Baetis
hamatum
B
449 Paramoebidium sp. MO-19-W15 Fallceon
1181 Paramoebidium cf. grande ID-164-G2 Simulium
473 Paramoebidium sp. UT-1-W14 Simulium
458 Paramoebidium sp. RMBL-72-3
1236 Paramoebidium sp. ID-65e-E1
450 Paramoebidium sp. NF-10-W15 Simulium
833 Paramoebidium sp. ONT-X-10
1049 Paramoebidium sp. ALG-9-W10
674 Paramoebidium sp. NOR-40-W8
1223 Paramoebidium sp. Mex-17-W2 Simulium
cf chattoni/ grande
B
38 Paramoebidium sp. ME-2-W3 Leptophlebia
41 Paramoebidium sp. NS-6-W8-10 Paraleptophlebia
466 Paramoebidium sp. NS-24-W16
508 Paramoebidium sp. NS-35-W18 Paraleptophlebia
sp. 3
L
1234 Paramoebidium sp. NS-16-W2 Leptophlebia
622 Paramoebidium sp. AR-31-C31
618 Paramoebidium sp. AR-30-C7
1176 Paramoebidium sp. AR-31-C7
619 Paramoebidium sp. AR-30-C9
446 Paramoebidium sp. KS-61-W20 Siphlonurus
Clade 6
B
901 Amoebidium appalachense TN-27-A3
921 Amoebidium appalachense TN-27-W1a
903 Amoebidium sp. TN-27-W4
904 Amoebidium appalachense TN-46-A6
377 Amoebidium parasiticum A1a
376 Amoebidium parasiticum JAP-7-2
Amoebidium parasiticum FRA-1-14 (NRRL20524)

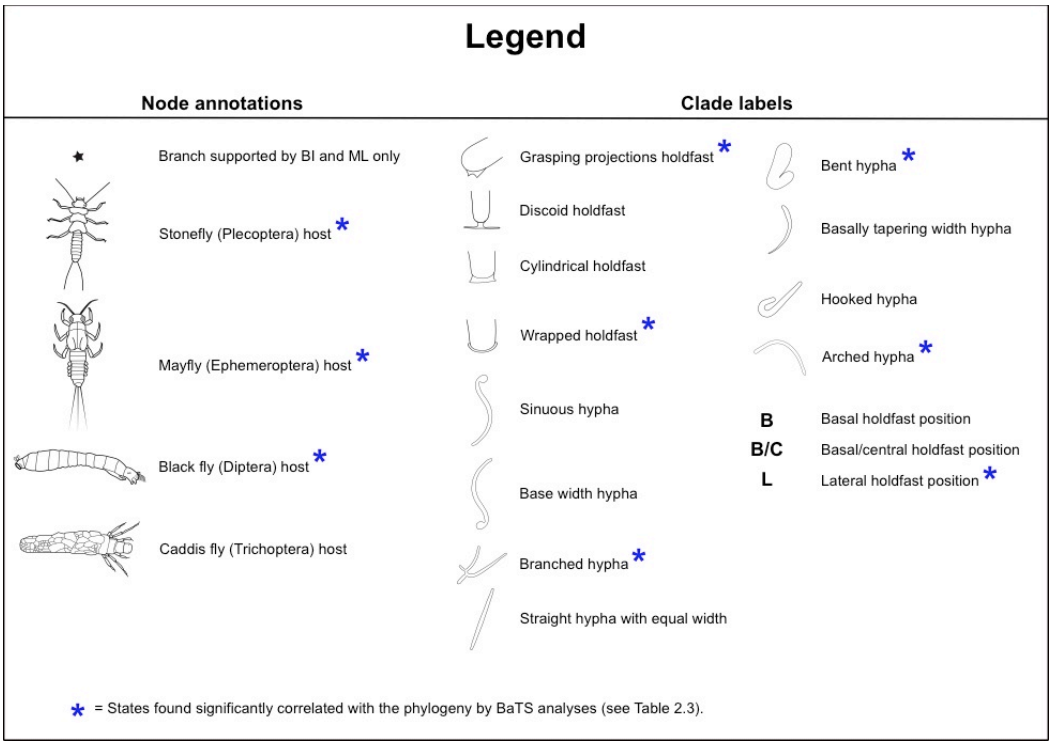


Figure 2.1 BEAST (BI) maximum clade credibility tree of the 18S and 28S combined dataset, including *Amoebidium* as the outgroup, with maximum likelihood (ML) and maximum parsimony (MP) values given (BI/ML/MP). Values below 0.90 for BI, 70 for ML and 60 for MP are indicated by a dash (-). Nodes with BI and ML support that were used in ASR analyses are numbered 1-24. Well-supported clades are labeled: Clade 1-6, sp. 1-3, *hamatum*, and cf. *chattoni/grande*. Character states given the highest values by ASR analyses are represented symbolically next to clades or nodes (for hosts), with the legend shown above. A drawing of a trichopteran host is located next to sample 870 and highlighted with a green box to mark that unique sample, but that state (Trichoptera) was not significant among ASR or BaTS analyses.

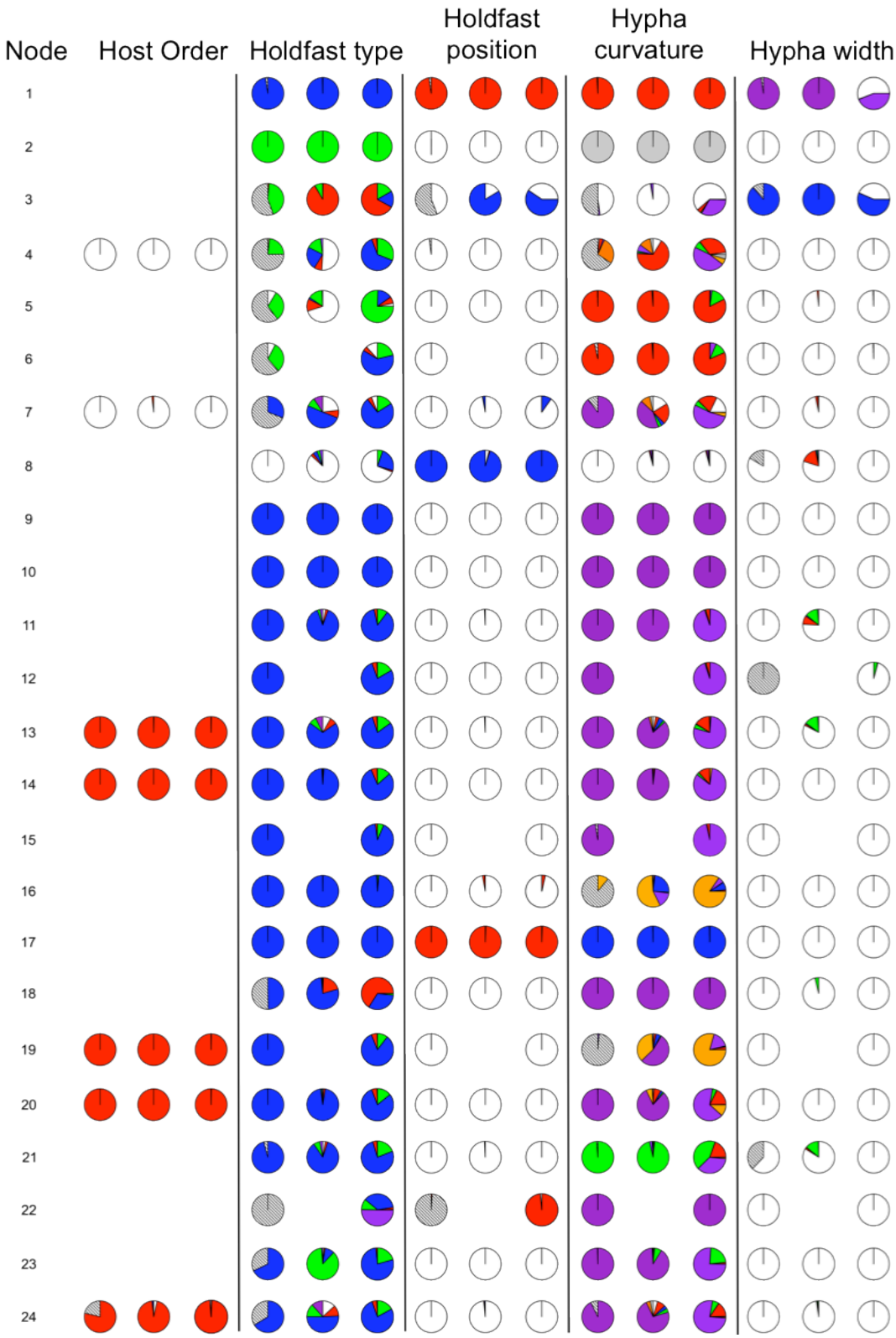


Figure 2.2 Anscestral state reconstruction pie chart output from parsimony, maximum likelihood, and SIMMAP, respectively, in each column for nodes 1-24 (Fig. 2.1). Parsimony and ML were performed in Mesquite. The last 5,005 trees from the BEAST analysis were used for

parsimony, so pie charts represent the proportion of trees returning a particular state as the best state at each node.

The characters and character states are as follows:

Host order: white = Plecoptera, red = Ephemeroptera, blue = Diptera, green = Trichoptera, purple = Cladocera
Holdfast type: white = globose, red = cylindrical, blue = grasping projections, green = discoid, purple = wrapped
Holdfast position: white = basal, red = lateral, blue = central
Hypha curvature: white = branched, red = straight/sinuous, blue = coiled/looped, green = arched, purple = hooked, orange = hairpin, grey = bent
Hypha width: white = equal, red = tapering distally, blue = tapering basally, green = middle, purple = base
Hatched indicates equivocal results in parsimony.

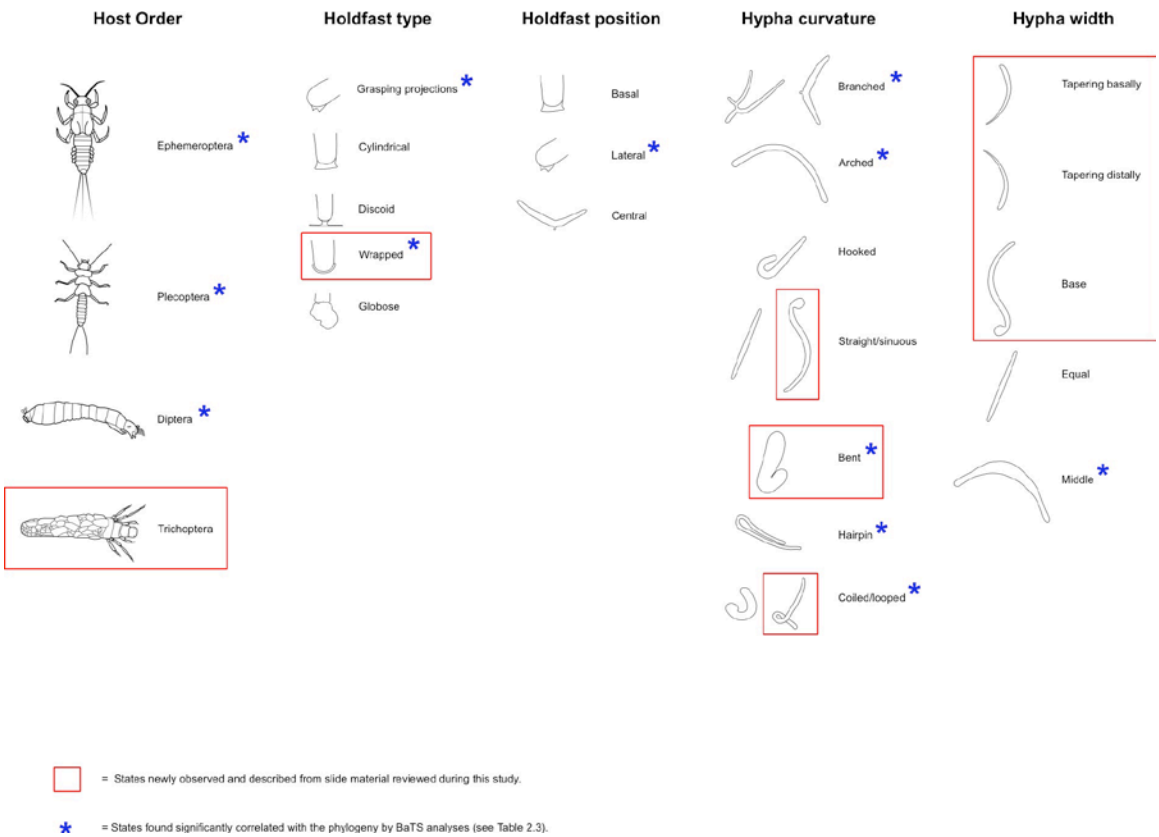


Figure 2.3 Representative drawings of character states, as observed among voucher slides and from published species descriptions. Red boxes indicate new character states observed during this study; asterisks indicate states that returned significant values in the BaTS analyses (Table 2.3).

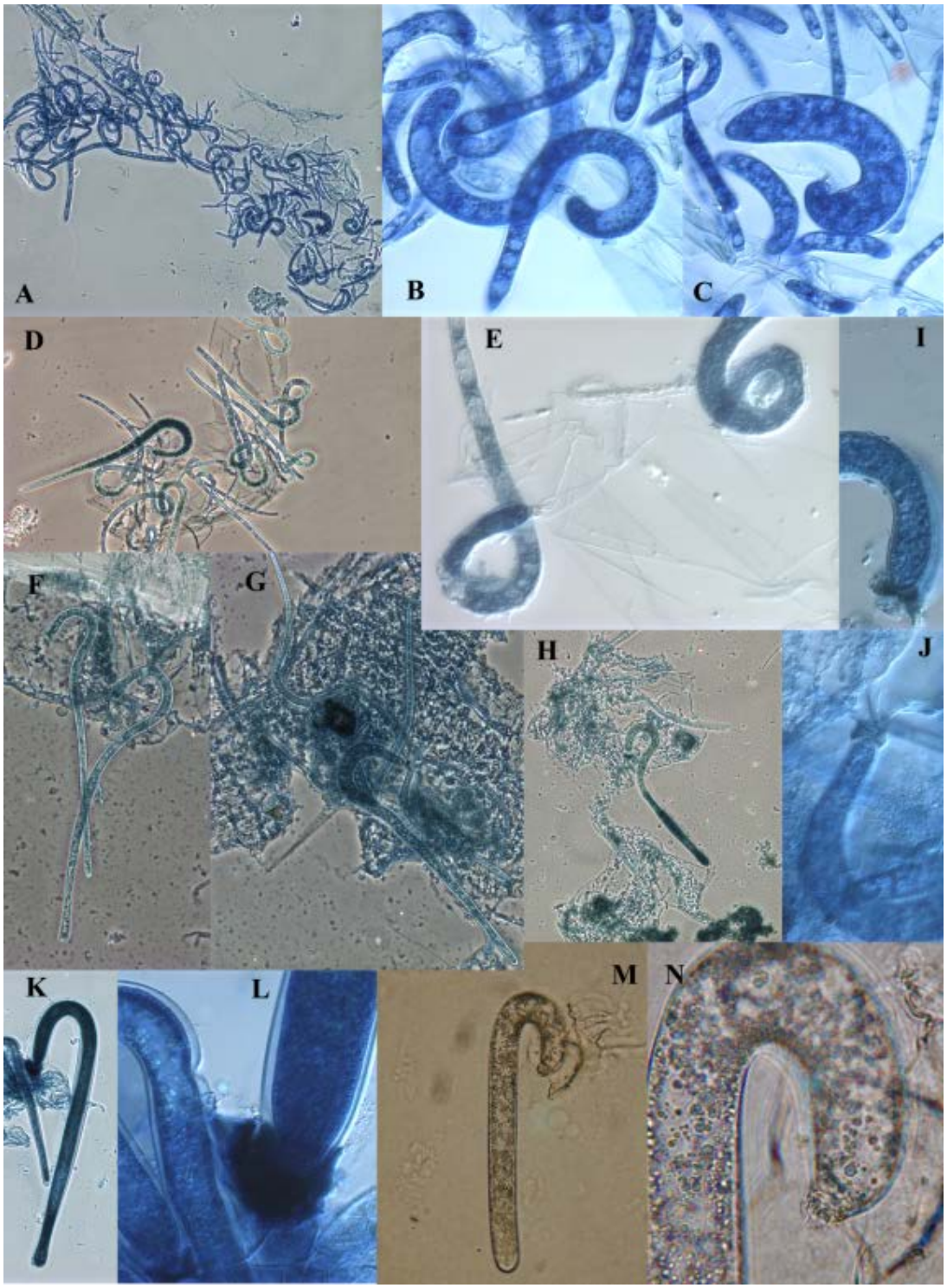


Plate 2.1 Representative examples of *Paramoebidium* spp. highlighting variation in hyphae and holdfasts for representative vouchers of the "coiled" (A-C), "hooked" (D, F-N), and "looped" (D-E) shapes. These were dissected from mayflies collected from various geographic regions including (A-C) Colorado, USA; (D,E) Dominica; (F-K) Nova Scotia, Canada and (M, N) Pennsylvania, USA. Images are from microscope slide voucher codes: CO-3-W11 (A-C); DR-16-C8 (D-E); NS-35-W22 (F-J); NS-35-W9 (K-L); PA-2-W1 (M-N).

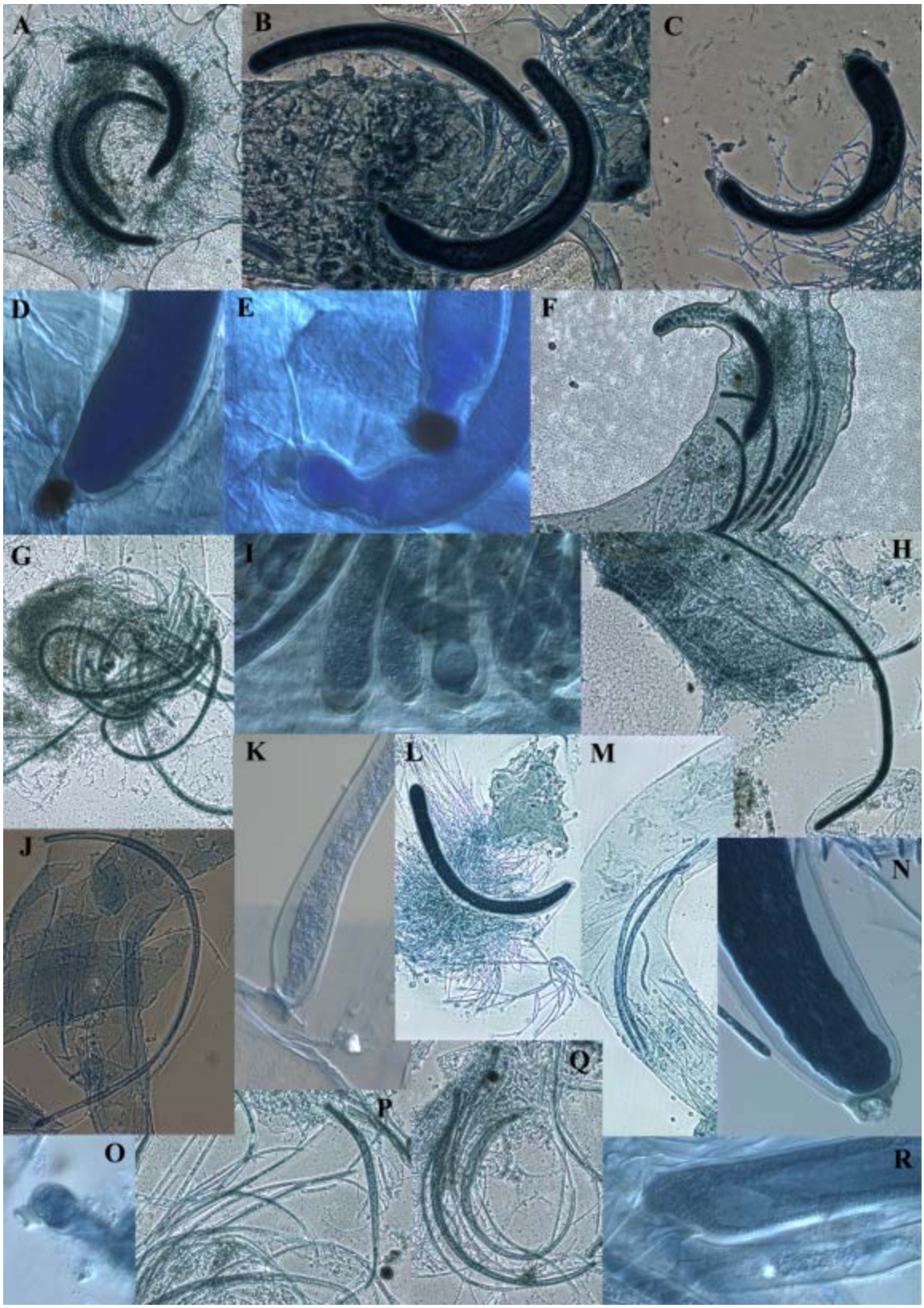


Plate 2.2 Images of *Paramoebidium grande* (C) and representatives of the morphologically overlapping *Paramoebidium* cf. *chattoni/grande* complex. Large, curved

thalli of putative *Paramoebidium grande* from Chile (A-E) mixed with *P. chattoni*-like (in F) and in comparison with *P. chattoni*-like specimens (G-I) imaged from slide voucher collection of R.W. Lichtwardt. Other images are from specimens considered to be *Paramoebidium* cf. *chattoni/grande* from three regions in North America: Idaho (J-K) and Colorado (L-M), USA and Ontario (O-R), Canada. See text for discussion of species overlap. These were dissected from black flies and images are from microscope slide voucher codes: CHI-1-1 (A, F); CHI-1-2(B-E); CHI-5-12 (G); CHI-20-8 (H); CHI-5-16 (I); ID-65-E1 (J-K); RMBL-78-6 (L-N) and ALG-9-W6a (O-R).

CONCLUSION

On a broader scale, this project represents one of the first steps and studies toward a higher-resolution picture of the evolutionary history of species within the Eccrinida. Especially useful is the attempt to shift the evaluation of the protist trichos from a mycological standpoint to a protistological one, which may reveal a whole different suite of taxonomically informative characters. Whereas previous studies have focused on discovery and placement of species, this project combined expanded taxon sampling with a multigene approach to probe larger-scale relationships. From this strengthened phylogenetic backbone, further hypotheses can be proposed toward ongoing investigations into the evolution of symbiosis within the group (e.g. coevolution between host and symbionts, shifts between commensalism and parasitism, etc.). Finally, as these organisms are unicellular relatives of animals and fungi, these data will also augment other research initiatives into the origins of multicellularity such as the recently published works involving transcriptome analyses (Sebé-Pedrós et al. 2011) and utilizing gene silencing and genetic transformation of ichthyosporeans (Suga & Ruiz-Trillo 2013).

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., Xenarios, I., Stockinger, H., 2012. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Research*, 40(W1), pp.597–603.

Azevedo, F., Cóndor-Luján, B., Willenz, P., Hajdu, E., Hooker, Y., Klautau, M., 2015. Integrative taxonomy of calcareous sponges (subclass Calcinea) from the Peruvian coast: morphology, molecules, and biogeography. *Zoological Journal of the Linnean Society*, 173(4), pp.787–817.

Baker, G.C., Beebee, T.J.C., Ragan, M. A., 1999. *Prototheca richardsi*, a pathogen of anuran larvae, is related to a clade of protistan parasites near the animal-fungal divergence. *Microbiology*, 145(7), pp.1777–1784.

Bangyeekhun, E., Ryynanen, H.J., Henttonen, P., Huner, J.V., Cerenius, L., Söderhäll, K., 2001. Sequence analysis of the ribosomal internal transcribed spacer DNA of the crayfish parasite *Psorospermium haeckeli*. *Diseases of aquatic organisms*, 46(3), pp.217–222. Available at: http://www.int-res.com/articles/dao/46/d046p217.pdf.

Bardsley, L. & Beebee, T.J., 2001. Non-behavioural interference competition between anuran larvae under semi-natural conditions. *Oecologia*, 128(3), pp.360–367.

Beebee, T.J.C. & Wong, A.L.C., 1993. Stimulation of cell-division and DNA-replication in *Prototheca richardsi* by passage through larval amphibian guts. *Parasitology*, 2(107), p.119.

Bench, M.E. & White, M.M., 2012. New species and first records of trichomycetes from immature aquatic insects in Idaho. *Mycologia*, 104(1), pp.295–312.

Benny, G.L. & O'Donnell, K., 2000. *Amoebidium parasiticum* is a protozoan, not a Trichomycete. *Mycologia*, 92(6), pp.1133–1137.

Bollback, J.P., 2009. SIMMAP Software. Available at: http://www.simmap.com/index.html [Accessed January 1, 2010].

Bollback, J.P., 2006. SIMMAP: stochastic character mapping of discrete traits on phylogenies. *BMC bioinformatics*, 7, p.88.

Brown, W.M., Prager, E.M., Wang, A., Wilson, A.C., 1982. Mitochondrial DNA sequences of Primates: Tempo and mode of evolution. *Journal of Molecular Evolution*, pp.225–239.

Byzov, B.A., 2006. Intestinal microbiota of millipedes. In H. König & A. Varma, eds. *Intestinal microorganisms of termites and other invertebrates*. Berlin, Heidelberg: Springer Verlag, pp. 89–114.

Cafaro, M.J., 2005. Eccrinales (Trichomycetes) are not fungi, but a clade of protists at the early divergence of animals and fungi. *Molecular Phylogenetics and Evolution*, 35(1), pp.21–34.

Caron, D.A., Countway, P.D., Jones, A.C., Kim, D.Y., Schnetzer, A., 2012. Marine Protistan Diversity. *Annual Review of Marine Science*, 4(1), pp.467–493.

Carr, M. & Suga, H., 2014. The Holozoan *Capsaspora owczarzaki* possesses a diverse complement of active transposable element families. *Genome Biology and Evolution*, 6(4), pp.949–963.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular biology and evolution*, 17(4), pp.540–552.

Cavalier-Smith, T., 2013. Early evolution of eukaryote feeding modes, cell structural diversity, and classification of the protozoan phyla Loukozoa, Sulcozoa, and Choanozoa. *European Journal of Protistology*, 49(2), pp.115–178.

Cavalier-Smith, T., Chao, E.E., Snell, E.A., Berney, C., Fiore-Donno, A.M., Lewis, R., 2014. Multigene eukaryote phylogeny reveals the likely protozoan ancestors of opisthokonts (animals, fungi, choanozoans) and Amoebozoa. *Molecular phylogenetics and evolution*, 81, pp.71–85. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25152275.

Chien, C.Y. & Hsieh, L.H., 2001. Occurrence of Trichomycetes in Taiwan. In *Trichomycetes and other fungal groups*. Enfield, New Hampshire: Science Publishers, Inc., p. 55.

Contreras, K. & Cafaro, M.J., 2013. Morphometric studies in *Enterobryus luteovirgatus* sp. nov. (Ichthyosporea: Eccrinales) associated with yellow-banded millipedes in Puerto Rico. *Acta Protozoologica*, 52(4), pp.291–297.

Corliss, J.O., 2004. Why the World Needs Protists! *Journal of Eukaryotic Microbiology*, 51(1), pp.8–22.

Cotterill, F.P.D., Al-Rasheid, K. & Foissner, W., 2008. Conservation of protists: Is it needed at all? *Biodiversity and Conservation*, 17(2), pp.427–443.

Cummings, M.P., Neel, M.C. & Shaw, K.L., 2008. A genealogical approach to quantifying lineage divergence. *Evolution*, 62(9), pp.2411–2422.

Dang, S.-N. & Lichtwardt, R.W., 1979. Fine Structure of *Paramoebidium* (Trichomycetes) and a new species with virus-like particles. *American Journal of Botany*, 66(9), p.1093.

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(8), p.772. Available at: http://dx.doi.org/10.1038/nmeth.2109.

David, J.F. & Handa, I.T., 2010. The ecology of saprophagous macroarthropods (millipedes, woodlice) in the context of global change. *Biological Reviews*, 85(4), pp.881–895.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, 29(8), pp.1969–1973.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology*, 4(5), pp.699–710.

Duboscq, O., Léger, L. & Tuzet, O., 1948. Contribution à la connaissance des Eccrinides: les Trichomycètes. *Archives de Zoologie Expérimentale et Générale*, 86, pp.29–144.

Edgar, R.C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), pp.1792–1797.

Evans, T.N. & Seviour, R.J., 2012. Estimating biodiversity of Fungi in activated sludge communities using culture-independent methods. *Microbial Ecology*, 63(4), pp.773–786.

Fitch, W.M., 1971. Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Zoology*, 20(4), pp.406–416.

Foissner, W., 2006. Biogeography and dispersal of micro-organisms: A review emphasizing protists. *Acta Protozoologica*, 45(2), pp.111–136.

Foissner, W., 1999. Protist diversity: estimates of the near-imponderable. *Protist*, 150(4), pp.363–368.

Gernhard, T., 2008. The conditioned reconstructed process. *Journal of Theoretical Biology*, 253(4), pp.769–778.

Glockling, S.L., Marshall, W.L. & Gleason, F.H., 2013. Phylogenetic interpretations and ecological potentials of the Mesomycetozoea (Ichthyosporea). *Fungal Ecology*, 6(4), pp.237–247. Available at: http://dx.doi.org/10.1016/j.funeco.2013.03.005.

Goloboff, P.A., Farris, S. & Nixon, K., 2008. TNT, a free program for phylogenetic analysis. *Cladistics*, 24, pp.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? *Frontiers in Microbiology*, 5(Feb), pp.1–16.

Guindon, S. & Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic biology*, 52(5), pp.696–704.

Heidelberg, K.B., Nelson, W.C., Holm, J.B., Eisenkolb, N., Andrade, K., Emerson, J.B., 2013. Characterization of eukaryotic microbial diversity in hypersaline Lake Tyrrell, Australia. *Frontiers in Microbiology*, 4(May), pp.1–14.

Hernandez 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(2), pp.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(6), pp.896–903.

Herr, R.A., Ajello, L., Taylor, J.W., Arseculeratne, S.N., Mendoza, L., 1999. Phylogenetic analysis of *Rhinosporidium seeberi*'s 18S small-subunit ribosomal DNA groups this pathogen among members of the protoctistan Mesomycetozoa clade. *Journal of Clinical Microbiology*, 37(9), pp.2750–2754.

Hibbits Galt, J., 1978. Marine Eccrinales (Trichomycetes) found in crustaceans of the San Juan Archipelago, Washington. *Syesis*, 11(11), pp.213–261.

Hibbits Galt, J., 1971. *Studies on some protists associated with Crustacea: the Ellobiopsidae and the Trichomycetes*. University of Washington.

Katz, A.D., Giordano, R. & Soto-Adames, F.N., 2015. Operational criteria for cryptic species delimitation when evidence is limited, as exemplified by North American *Entomobrya* (Collembola: Entomobryidae). *Zoological Journal of the Linnean Society*, 173(4), pp.818–840.

Kocan, R., LaPatra, S. & Hershberger, P., 2013. Evidence for an amoeba-like infectious stage of *Ichthyophonus* sp. and description of a circulating blood stage: a probable mechanism for dispersal within the fish host. *The Journal of parasitology*, 99(2), pp.235–40. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22924915.

Kocan, R.M., Gregg, J.L. & Hershberger, P.K., 2010. Release of infectious cells from epidermal ulcers in *Ichthyophonus* sp.-infected Pacific herring (*Clupea pallasii*): evidence for multiple mechanisms of transmission. *The Journal of parasitology*, 96(2), pp.348–352.

Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29(6), pp.1695–1701.

Lecocq, T., Dellicour, S., Michez, D., Dehon, M., Dewulf, A., De Meulemeester, T., Brasero, N., Valterová, I., Rasplus, J.-Y., Rasmont, P., 2015. Methods for species delimitation in bumblebees (Hymenoptera, Apidae, *Bombus*): towards an integrative approach. *Zoologica Scripta*, p.n/a–n/a. Available at: http://doi.wiley.com/10.1111/zsc.12107.

Léger, L. & Duboscq, O., 1929a. *Eccrinoïdes henneguyi* n. g. n. sp. et la systématique des Eccrinides. *Archives d' Anatomie Microscopique*, (25), pp.309–324.

Léger, L. & Duboscq, O., 1929b. L'évolution des *Paramoebidium*, nouveau genre d'Eccrinides, parasite des larves aquatiques d'Insectes. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences Paris*, (189), pp.75–77.

Leidy, J., 1849. *Enterobrus*, a new genus of Confervaceae. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 4, pp.225–233.

Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic biology*, 50(6), pp.913–925.

Li, X.Q., Zuo, Q., Li, M., He, S., Jia, Y., Wang, Y.F., 2015. Phylogenetic relationships and species delimitation among the taxa with propagula in *Pseudotaxiphyllum* (Plagiotheciaceae, Bryophyta). *Journal of Systematics and Evolution*, 53(2), pp.179–190.

Lichtwardt, R.W., White, M.M., Cafaro, M.J., Misra, J.K., 1999. Fungi associated with Passalid beetles and their mites. *Mycologia*, 91(4), pp.694–702.

Lichtwardt, R.W., 1954. Three species of Eccrinales inhabiting the hindguts of millipeds, with comments on the Eccrinids as a group. *Mycologia*, 46(5), pp.564–585.

Lichtwardt, R.W. & Arenas, J.M., 1996. Trichomycetes in aquatic insects from Southern Chile. *Mycologia*, 88(5), pp.844–857.

Lichtwardt, R.W., Cafaro, M.J. & White, M.M., 2001. The Trichomycetes: Fungal associates of arthropods. Available at: www.nhm.ku.edu/~fungi [Accessed January 1, 2015].

Lohr, J.N., Laforsch, C., Koerner, H., Wolinska, J., 2010. A daphnia parasite (*Caullerya mesnili*) constitutes a new member of the Ichthyosporea, a group of protists near the animal-fungi divergence. *Journal of Eukaryotic Microbiology*, 57(4), pp.328–336.

Lord, J.C., Hartzer, K.L. & Kambhampati, S., 2012. A nuptially transmitted ichthyosporean symbiont of *Tenebrio molitor* (Coleoptera: Tenebrionidae). *Journal of Eukaryotic Microbiology*, 59(3), pp.246–250.

Maddison, W.P. & Maddison, D.R., 2014. Mesquite: a modular system for evolutionary analysis. Available at: http://mesquiteproject.org.

Manier, J.F., 1964. Endophytes parasites d'Arthropodes cavernicoles récoltés dans des grottes de l'Ariège et de la Haute-Garonne. *Annales de Spéléologie*, 19, pp.803–812.

Manier, J.F., 1950. Recherches sur les Trichomycètes. *Annales des Sciences Naturelles Botanique et Biologie Végétale Série*, 11(11), pp.53–162.

Manier, J.F. & Grizel, H., 1971. *Paramacrinella microdeutopi* n. g., n. sp., Trichomycète parasite de *Microdeutopus anomalus* H. Rathke (Amphipode). *Annales des Sciences Naturelles Botanique et Biologie Végétale, Paris Série*, 12(12), pp.1–7.

Manier, J.F. & Lichtwardt, R.W., 1968. Révision de la systématique des Trichomycètes. *Annales des Sciences Naturelles Botanique et Biologie Végétale, Paris Série*, 12(9), pp.519–532.

Marshall, W.L., Celio, G., McLaughlin, D.J., Berbee, M.L., 2008. Multiple isolations of a culturable, motile ichthyosporean (Mesomycetozoa, Opisthokonta), *Creolimax fragrantissima* n. gen., n. sp., from marine invertebrate digestive tracts. *Protist*, 159(3), pp.415–433.

Marshall, W.L. & Berbee, M.L., 2013. Comparative morphology and genealogical delimitation of cryptic species of sympatric isolates of *Sphaeroforma* (Ichthyosporea, Opisthokonta). *Protist*, 164(2), pp.287–311. Available at: http://dx.doi.org/10.1016/j.protis.2012.12.002.

Marshall, W.L. & Berbee, M.L., 2011. Facing Unknowns: Living cultures (*Pirum gemmata* gen. nov., sp. nov., and *Abeoforma whisleri*, gen. nov., sp. nov.) from invertebrate digestive tracts represent an undescribed clade within the unicellular opisthokont lineage Ichthyosporea (Mesomycetozoea). *Protist*, 162(1), pp.33–57. Available at: http://dx.doi.org/10.1016/j.protis.2010.06.002.

Marshall, W.L. & Berbee, M.L., 2010. Population-level analyses indirectly reveal cryptic sex and life history traits of *Pseudoperkinsus tapetis* (Ichthyosporea, Opisthokonta): A unicellular relative of the animals. *Molecular Biology and Evolution*, 27(9), pp.2014–2026.

Mendoza, L., Ajello, L. & Taylor, J.W., 2001. The taxonomic status of *Lacazia loboi* and *Rhinosporidium seeberi* has been finally resolved with the use of molecular tools. *Revista iberoamericana de micologia : organo de la Asociacion Espanola de Especialistas en Micologia*, 18(3), pp.95–98.

Mendoza, L., Taylor, J.W. & Ajello, L., 2002. The class Mesomycetozoea: a heterogeneous group of microorganisms at the animal-fungal boundary. *Annual review of microbiology*, 56, pp.315–344.

Millanes, A.M., Diederich, P., Ekman, S., Wedin, M., 2011. Phylogeny and character evolution in the jelly fungi (Tremellomycetes, Basidiomycota, Fungi). *Molecular Phylogenetics and Evolution*, 61(1), pp.12–28. Available at: http://dx.doi.org/10.1016/j.ympev.2011.05.014.

Moss, S.T., 1979. Commensalism of Trichomycetes. In L. R. Batra, ed. *Insect-Fungus Symbiosis Nutrition, Mutualism, and Commensalism*. Montclair: Allanheld, Osmun & Co. Publishers, Inc., pp. 175–227.

Okamoto, N., Nakase, K., Suzuki, H., Nakai, Y., Fujii, K., Sano, T., 1985. Life History and Morphology of *Ichthyophonus hoferi* in vitro. *Fish Pathology*, 20(2), pp.273–285.

Paps, J., Medina-Chacón, L.A., Marshall, W., Suga, H., Ruiz-Trillo, I., 2013. Molecular phylogeny of Unikonts: New insights into the position of Apusomonads and Ancyromonads and the internal relationships of Opisthokonts. *Protist*, 164(1), pp.2–12.

Parker, J., Rambaut, A. & Pybus, O.G., 2008. Correlating viral phenotypes with phylogeny: Accounting for phylogenetic uncertainty. *Infection, Genetics and Evolution*, 8(3), pp.239–246.

Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., Bowser, S.S., Cepicka, I., Decelle, J., Dunthorn, M., Fiore-Donno, A.M., Gile, G.H., Holzmann, M., Jahn, R., Jirků, M., Keeling, P.J., Kostka, M., Kudryavtsev, A., Lara, E., Lukeš, J., Mann, D.G., Mitchell, E.A.D., Nitsche, F., Romeralo, M., Saunders, G.W., Simpson, A.G.B., Smirnov, A.V., Spouge, J.L., Stern, R.F., Stoeck, T., Zimmermann, J., Schindel, D., de Vargas, C., 2012. CBOL Protist Working Group: Barcoding Eukaryotic Richness beyond the Animal, Plant, and Fungal Kingdoms. *PLoS Biology*, 10(11).

Poisson, R., 1931. Recherches sur les Eccrinides, deuxième contribution. *Archives de Zoologie Expérimentale et Générale*, 74, pp.53–68.

Poisson, R., 1929. Recherches sur quelques Eccrinides parasites de Crustacés Amphipodes et Isopodes. *Archives de Zoologie Expérimentale et Générale*, (69), pp.179–216.

Poulin, R., Krasnov, B.R. & Mouillot, D., 2011. Host specificity in phylogenetic and geographic space. *Trends in Parasitology*, 27(8), pp.355–361.

R Foundation for Statistical Computing, 2015. Available at: www.r-project.org [Accessed March 2, 2015].

Rambaut, A., 2014. FigTree. Available at: http://tree.bio.ed.ac.uk/software/figtree/.

Rambaut, A. & Drummond, A.J., 2009. Tracer 1.5.0. Available at: http://beast.bio.ed.ac.uk/.

Reeves, W.K., Jensen, J.B. & Ozier, J.C., 2000. New faunal and fungal records from caves in Georgia, USA. *Journal of Cave and Karst Studies*, 62(3), pp.169–179.

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61(3), pp.539–542.

Ronquist, F., Huelsenbeck, J. & Teslenko, M., 2011. *Draft MrBayes Version 3.2 Manual: Tutorials and Model Summaries*, pp.1-103.

Rowley, J.J.L., Gleason, F.H., Andreou, D., Marshall, W.L., Lilje, O., Gozlan, R., 2013. Impacts of mesomycetozoean parasites on amphibian and freshwater fish populations. *Fungal Biology Reviews*, 27(3-4), pp.100–111. Available at: http://dx.doi.org/10.1016/j.fbr.2013.09.002.

Ruiz-Trillo, I., Roger, A.J., Burger, G., Gray, M.W., Lang, B.F., 2008. A phylogenomic investigation into the origin of Metazoa. *Molecular Biology and Evolution*, 25(4), pp.664–672.

Ruiz-Trillo, I., Lane, C.E., Archibald, J.M., Roger, A.J., 2006. Insights into the evolutionary origin and genome architecture of the unicellular opisthokonts *Capsaspora owczarzaki* and *Sphaeroforma arctica*. *Journal of Eukaryotic Microbiology*, 53(5), pp.379–384.

Schmidt, H.A., Strimmer, K., Vingron, M., von Haeseler, A., 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics (Oxford, England)*, 18(3), pp.502–504.

Schmidt, H.A., 2009. Testing tree topologies. In P. Lemey, M. Salemi, & A. M. Vandamme, eds. *The Phylogenetic Handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge: Cambridge University Press, pp. 381–404.

Schuster, J.C. & Schuster, L.B., 1985. Social behavior in passalid beetles (Coleoptera: Passalidae): cooperative. *The Florida Entomologist*, 68(2), pp.266–272.

Sebé-Pedrós, A., Irimia, M., del Campo, J., Parra-Acero, H., Russ, C., Nusbaum, C., Blencowe, B.J., Ruiz-Trillo, I., 2013. Regulated aggregative multicellularity in a close unicellular relative of metazoa. *eLife*, 2013(2).

Sebé-Pedrós, A., De Mendoza, A., Lang, B.F., Degnan, B.M., Ruiz-Trillo, I., 2011. Unexpected repertoire of metazoan transcription factors in the unicellular holozoan *Capsaspora owczarzaki*. *Molecular Biology and Evolution*, 28(3), pp.1241–1254.

Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. *Systematic biology*, 51(3), pp.492–508.

Shimodaira, H. & Hasegawa, M., 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics (Oxford, England)*, 17(12), pp.1246–1247.

Shimodaira, H. & Hasegawa, M., 1999. Multiple Comparisons of Log-Likelihoods with Applications to Phylogenetic Inference. *Molecular Biology and Evolution*, pp.1114–1116.

Slapeta, J., Moreira, D. & López-García, P., 2005. The extent of protist diversity: insights from molecular ecology of freshwater eukaryotes. *Proceedings. Biological sciences / The Royal Society*, 272(1576), pp.2073–2081.

Stamatakis, A., 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), pp.1312–1313.

Steenkamp, E.T., Wright, J. & Baldauf, S.L., 2006. The protistan origins of animals and fungi. *Molecular Biology and Evolution*, 23(1), pp.93–106.

Stöver, B.C. & Müller, K.F., 2010. TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics*, 11, p.7.

Strongman, D.B. & White, M.M., 2006. New species of *Lancisporomyces*, *Orphella*, and *Paramoebidium*, endosymbionts of stonefly nymphs from streams in Nova Scotia, Canada. *Canadian Journal of Botany*, 84(9), pp.1478–1495.

Strongman, D.B. & White, M.M., 2008. Trichomycetes from lentic and lotic aquatic habitats in Ontario, Canada. *Botany*, 86(12), pp.1449–1466.

Suga, H., Chen, Z., de Mendoza, A., Sebé-Pedrós, A., Brown, M.W., Kramer, E., Carr, M., Kerner, P., Vervoort, M., Sánchez-Pons, N., Torruella, G., Derelle, R., Manning, G., Lang, B.F., Russ, C., Haas, B.J., Roger, A.J., Nusbaum, C., Ruiz-Trillo, I., 2013. The *Capsaspora* genome reveals a complex unicellular prehistory of animals.

Nature Communications, 4, p.2325. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3753549&tool=pmcentrez&rendertype=abstract.

Suga, H. & Ruiz-Trillo, I., 2013. Development of ichthyosporeans sheds light on the origin of metazoan multicellularity. *Developmental Biology*, 377(1), pp.284–292.

Takishita, K., Miyake, H., Kawato, M., Maruyama, T., 2005. Genetic diversity of microbial eukaryotes in anoxic sediment around fumaroles on a submarine caldera floor based on the small-subunit rDNA phylogeny. *Extremophiles*, 9(3), pp.185–196.

Takishita, K., Tsuchiya, M., Kawato, M., Oguri, K., Kitazato, H., Maruyama, T., 2007. Genetic diversity of microbial eukaryotes in anoxic sediment of the saline meromictic Lake Namako-ike (Japan): On the detection of anaerobic or anoxic-tolerant lineages of eukaryotes. *Protist*, 158(1), pp.51–64.

Tautz, D. & Renz, M., 1983. An optimized Freeze-Squeeze method for the recovery of DNA fragments from agarose gels. *Analytical Biochemistry*, 132(1), pp.14–19.

Trotter, M.J. & Whisler, H.C., 1965. Chemical composition of the cell wall of *Amoebidium parasiticum*. *Canadian Journal of Botany*, (43), pp.869–876.

Tuzet, O. & Manier, J.F., 1950. *Lajassiella aphodii* n.g., n.sp. Palavascide parasite d'une larve d'Aphodius (Coléoptère, Scarabaeidae). *Annales des Sciences Naturelles Zoologie et Biologie Animale*, 11(12), pp.465–470.

Ustinova, I., Krienitz, L. & Huss, V.A., 2000. *Hyaloraphidium curvatum* is not a green alga, but a lower fungus; *Amoebidium parasiticum* is not a fungus, but a member of the DRIPs. *Protist*, 151(3), pp.253–262.

Valle, L.G., 2014a. New species of *Paramoebidium* (trichomycetes, Mesomycetozoea) from the Mediterranean, with comments about the amoeboid cells in Amoebidiales. *Mycologia*, 106(3), pp.481–490. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24895422.

Valle, L.G., 2014b. Validation of the trichomycete *Paramoebidium chattoni* (Amoebidiales, Mesomycetozoea), a common and cosmopolitan black fly endosymbiont. *Mycologia*, 106(3), pp.573–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24871601.

Valle, L.G. & Santamaria, S., 2009. Arthropod gut symbionts from the Balearic Islands: Majorca and Cabrera. Diversity and biogeography. *Anales del Jardín Botánico de Madrid*, 66(1), pp.109–120.

Van Dover, C.L. & Lichtwardt, R.W., 1986. A new Trichomycete commensal with a Galatheid squat lobster from deep-sea hydrothermal vents. *Biological Bulletin*, 171(2), pp.461–468. Available at: http://www.jstor.org/stable/1541687.

Vicente, F., 2010. Micro-invertebrates conservation: Forgotten biodiversity. *Biodiversity and Conservation*, 19(13), pp.3629–3634.

Vogt, G. & Rug, M., 1999. Life stages and tentative life cycle of *Psorospermium haeckeli*, a species of the novel DRIPs clade from the animal-fungal dichotomy. *Journal of Experimental Zoology*, 283(1), pp.31–42.

Wang, T.H., Donaldson, Y.K., Brettle, R.P., Bell, J.E., 2001. Identification of shared populations of Human Immunodeficiency Virus Type 1 infecting microglia and tissue macrophages outside the central nervous system. *Journal of Virology*, 75(23), pp.11686–11699.

Wang, Y. Tretter, E.D., Johnson, E.M., Kandel, P., Lichtwardt, R.W., Novak, S.J., Smith, J.F., White, M.M., 2014. Using a five-gene phylogeny to test morphology-based hypotheses of *Smittium* and allies, endosymbiotic gut fungi (Harpellales) associated with arthropods. *Molecular Phylogenetics and Evolution*, 79(1), pp.23–41.

Weisse, T., 2008. Distribution and diversity of aquatic protists: An evolutionary and ecological perspective. *Biodiversity and Conservation*, 17(2), pp.243–259.

Whisler, H.C., 1962. Culture and nutrition of *Amoebidium parasiticum*. *American Journal of Botany*, 49(3), p.193.

Whisler, H.C., 1968. Developmental control of *Amoebidium parasiticum*. *Developmental Biology*, 17(5), pp.562–570.

Whisler, H.C., 1966. Host-integrated development in the Amoebidiales. *The Journal of Protozoology*, 13(1), pp.183–188.

Whisler, H.C., 1963. Observations on some new and unusual enterophilous phycomycetes. *Canadian Journal of Botany*, 41(6), pp.887–900.

White, M.M., 2006. Evolutionary implications of a rRNA-based phylogeny of Harpellales. *Mycological Research*, 110(9), pp.1011–1024.

White, M.M., Cafaro, M.J. & Gottlieb, A.M., 2001. Taxonomy and Systematics of Trichomycetes - Past, Present, and Future. In J. K. Misra & B. W. Horn, eds. *Trichomycetes and other fungal groups*. Enfield, NH: Science, pp. 27–37.

White, M.M., Cafaro, M.J. & Lichtwardt, R.W., 2000. Arthropod gut fungi from Puerto Rico and summary of tropical Trichomycetes worldwide. *Carribean Journal of Science*, 36(3-4), pp.210–220.

White, M.M., Siri, A. & Lichtwardt, R.W., 2006. Trichomycete insect symbionts in Great Smoky Mountains National Park and vicinity. *Mycologia*, 98(2), pp.333–352.

William, R. & Strongman, D., 2013. Trichomycetes from Governor's Lake and Lake Micmac within the Halifax Regional Municipality, Nova Scotia, Canada. *Botany*, 91(February), pp.382–402. Available at: http://www.nrcresearchpress.com/doi/abs/10.1139/cjb-2012-0236.

Xia, X., 2013. DAMBE5: A comprehensive software package for data analysis in molecular biology and evolution. *Molecular Biology and Evolution*, 30(7), pp.1720–1728.

Yule, G.U., 1925. A Mathematical Theory of Evolution based on the Conclusions of Dr. J.C. Willis, F.R.S. *Journal of the Royal Statistical Society*, 88(3), pp.433–436.

APPENDIX

Review of the Morphology and Classification of Eccrinidae genera According to the Traditional Family Structure

This review is provided as a more in-depth discussion of the morphological features, host associations, and traditional classification of Eccrinidae genera. A few points from this review are mentioned in the discussion section of Chapter 1, but the additional information included here gives an historical perspective, as well as offers hypotheses regarding the placement of some as-yet unsequenced taxa.

Palavasciaceae

Representatives from each of the three former Trichomycete families (currently collapsed into family Eccrinidae) were obtained, and though taxon sampling is not complete, some of the traditional taxonomic framework within the group is not supported. Palavasciaceae, the smallest family with one genus and three species, is represented by vouchers from two species. These form a well-supported clade with *Alacrinella limnoriae* and an unnamed *Astreptonema* voucher (termed PAA clade here). Both are currently classified within the Eccrinaceae, so their inclusion with Palavasciaceae is contradictory, although not unexpected (Lichtwardt et al. 2001). The putative *Astreptonema* voucher (WA-3-C3) was thought to be the same organism as that described by Hibbits Galt (1971) (*Astreptonema* sp. 2), whose observed characters were intermediate between that genus and *Alacrinella* (i.e. a hooked base vs. lobed and one spore appendage rather than two). The specimen collected for the sample used in this study (and Cafaro 2005) was immature, but did have the wide, hooked base and an abrupt constriction of the hypha, just proximal to the sporulating region, as previously illustrated (Hibbits Galt 1971). In fact, Hibbits Galt (1971) suggested that *Alacrinella*, *P. sphaeromae* and *Astreptonema* could be closely related. Indeed, there are several morphological features that could unite them: distinct “microthalli” that produce

uninucleate cells (although the thin filaments described for *Palavascia* spp. have not been termed “microthalli”, to date), a persistent “spore mother-cell” at the distal end of the hyphae, and similarly shaped and sized spores. Two other genera, *Ramacrinella* and *Paramacrinella*, have been reported to produce microthalli and share persistent “spore mother-cells” with the PAA clade. These genera were morphologically distinguished based on the proximal position of the mother-cells, and the branched hyphae of *Ramacrinella* (Manier & Grizel 1971). Furthermore, Palavasciaceae was formally established as a separate family due, at least in part, to the absence of secondary spore formation; however, *Lajasiella* (Parataeniellaceae) (Manier & Lichtwardt 1968) is not known to produce these spore types either. Such morphological variation within a clade is not uncommon among the protist trichos (e.g. *Arundinula, Enterobryus,* and *Enteropogon,* discussed below) and should not, alone, rule out the possible relatedness of *Ramacrinella* and *Paramacrinella* to the PAA group or preclude dissolution/emendation of the Palavasciaceae.

Besides morphology, *Astreptonema* sp. 2 and *Ramacrinella* and *Paramacrinella* share the same host and habitat affiliations as *Palavascia* and *Astreptonema*, respectively. The significant correlation of habitat and host type found in this study provide additional support for Hibbits Galt’s (1971) hypothesis that *Ramacrinella* and *Paramacrinella* belong with the PAA clade, but tend to refute the placement of her *Astreptonema* sp. 2 in that genus. Similarly, predictions regarding the unrepresented *Astreptonema* species would place *A. longispora* (Hauptfleisch 1895) and *A. typica* (Manier 1968) with *A. gammari* (Manier 1964) due to their association with gammarid hosts in freshwater habitats. On the other hand, the affiliation of *A. corophii* (Manier 1968) and *A. pacificum*

(Hibbits Galt 1978) are viewed with less certainty due to their association with different host families from marine habitats, making them desirable targets for future collecting/vouchering efforts.

Beyond the apparent host-related splits, geographic partitioning does not appear to have an obvious influence within the clade. Only two vouchers were collected from the Southern Hemisphere (Argentina) (ARG-D4-C11 and ARG-D1-C15), giving this dataset an obvious bias toward Northern Hemisphere samples. However, *Palavascia patagonica* is not significantly diverged from *P. sphaeromae* from South Carolina or from the other North American samples it forms a clade with (at least in the rDNA gene trees, Figs. 1.1, 1.3). Comparing the Northern Hemisphere examples, Canadian, European, and USA samples form clades without indication of geographic bias. *Astreptonema gammari* from Minnesota and Massachusetts, for example, have little divergence from the sample from France.

Parataeniellaceae

Although *Parataeniella* species are found in isopods (Poisson 1929), they are terrestrial hosts as opposed to the marine/halophilic species associated with *Palavascia* and *Alacrinella*. Out of the two genera (*Parataeniella* and *Lajasiella*) and seven species described in the Parataeniellaceae, one genus and two species are represented here. However, these samples form a distinct clade from the rest of the eccrinids, falling out separately in alternate topologies. Indeed, SH and AU tests do not reject this hypothesis (Table 1.6). Evaluation of their described life history seems to align with the transitional nature of the molecular signal. *Parataeniella* hyphae are reported to holocarpically form spores or to release propagules basipetally. Interestingly, some *Lajasiella* hyphae also

divide along most of their length, except for the proximal portion (Tuzet & Manier 1950); other hyphae produce propagules in a basipetal manner. Homology of this holocarp-like trait between the two genera remains to be determined, but *Lajasiella* is especially of interest as it is the only protist tricho found (so far) in scarab beetles (Coleoptera, Scarabaeidae), the only eccrinid genus reported from just the larval host form and not the adult, and it is not known to produce "secondary" spores (Lichtwardt et al. 2001). *Lajasiella* also is described as having a persistent spore mother-cell, but this has not been observed in *Parataeniella*. While the confirmed schizont-like formation of propagules for *Parataeniella* is unique among the eccrinids (but see Eccrinaceae discussion below), it is, at least superficially, similar to the way [sporangio]spores (in *Amoebidium*) and amoebae are formed in the Amoebidiidae. No basipetal forms have been reported for accepted Amoebidiidae species, but a *nomen nudum* (i.e. not validated according to the International Code of Botanical Nomenclature) (Lichtwardt et al. 2001) *Paramoebidium* was given the epithet *eccriniformis* based on the linear arrangement of amoebae within the distal portion of the hypha (Duboscq et al. 1948), pointing to the possibility of other species with intermediate forms.

Geographically, most *Parataeniella* records are from North America and Europe, except for *P. flavospora* (Taiwan) (Chien & Hsieh 2001) and *P. latrobi* (Australia) (Lichtwardt & Williams 1990). *Parataeniella scotonisci* was collected from cavernicolous hosts in France (Manier 1964). Collections of species have been recorded from a range of isopod families, some of which have broad distributions (David & Handa 2010), implying wider surveys for *Parataeniella* are warranted. The consistency of morphological characters recorded for these species (i.e. production of uni-/binucleate

spores and both holocarpic and basipetal hyphal forms) across geographic boundaries and the strongly supported clade seen here suggests that *Parataeniella* will hold as a genus in future analyses.

Eccrinaceae

The third family, and the largest with 14 genera, is the Eccrinaceae. Topology tests reject its monophyly (Table 1.6) and again clade formation appears linked to host organism. For example, *Enterobryus halophilus* places closer to other marine, decapod-associated taxa than with other *Enterobryus* species. Likewise, an unidentified eccrinid (sample 1067) (noted by MMW as another possible *Enterobryus* sp.) is on a branch positioned nearer to the other Coleoptera clade (*Leidyomyces*) and freshwater-associated *Arundinula opeongoensis*. The only described *Enterobryus* from a beetle host is *E. hydrophilorum*. Cursory review of slide material from the unnamed sample superficially resembled that *E. hydrophilorum*, but the measurements were again outside the prescribed range. Additionally, the host was not identified, but an 18S sequencing attempt for the sample returned BLAST results as a *Helophorus* beetle (with 99% coverage and identity), of the same host family (Hydrophilidae) as that for *E. hydrophilorum*. Two illegitimate specimens were recorded from hosts of that family by Poisson (1931), but given the genus *Trichella*, which has been rejected (Lichtwardt et al. 2001). However, due to the morphological complexities outlined below, the specimen will remain unnamed here.

Enterobryus is the largest genus of the Eccrinidae, and aside from those mentioned above, all other species have been found in association with millipede hosts. Twenty-six species descriptions have been accepted, but more than 20 are considered

illegitimate or incomplete (Lichtwardt et al. 2001; Contreras & Cafaro 2013). Host specificity has not been determined to date, and, as species have been described from crabs and beetles, was not thought to be an influential factor based on morphological character similarities. The main obstacle to *Enterobryus* (and the Eccrinidae in general) taxonomy has been the paucity of informative morphological characters. Their production of different spore types (a total of nine) and cells of unknown function combined with intraspecific variability, even along the length of a single host gut (Lichtwardt 1954, 1958; Lichtwardt et al. 2001) have confounded attempts to evaluate the genus solely with morphology. Therefore, the genus has been functioning as a "catch-all" pending better methods (e.g. culturing, genetic barcoding) and characteristics with which to differentiate them. For example, future species descriptions should include such thorough analyses of morphological variability as recently published for *E. luteovirgatus* (Contreras & Cafaro 2013). However, even with the somewhat limited taxon sampling presented here, *Enterobryus* samples separate into defined, well-supported clades, indicating the non-homology of characters and the requirement for genetic evaluation of species.

While the samples used here were coded as "hindgut" dwellers for the purpose of analysis, the (in many cases) very long millipede gut (Byzov 2006) could be subpartitioned according to physiochemical gradations, not only in terms of microhabitats that the gut symbionts are exposed to, but also in terms of states for future character mapping analyses. The dividing of *Enterobryus* samples into different well-supported clades suggests that molecular data would clarify species boundaries and likely lead to a splitting up of the genus. The hint of association of these clades with host family or even

genus is tantalizing, but certainly there is not enough data, as yet, to support any delineation on that basis alone.

On the other hand, the morphological distinctiveness (i.e. bilocular spores) and host association described for *Eccrinidus* matches its placement in the tree. This genus is monotypic and the representative samples split off on a branch between *Parataeniella* and the rest of the Eccrinidae, thus being more closely related to other terrestrial, millipede-associated taxa. Only two other species are recorded from glomerid hosts: *Eccrinoides henneguyi* and *E. broelemanni*. Hibbits Galt (1971) hypothesized that *Eccrinoides* is closely related to the PAA clade. However, the present analyses would suggest that the genus could be split along host lines, with the two glomerid-associated taxa being closely related to *Eccrinidus* and the two terrestrial isopod-associated taxa presumably falling out separately or aligning more closely with *Parataeniella*. In fact, *E. henneguyi* is noted (Lichtwardt et al. 2001) to be very similar morphologically to *Eccrinidus flexilis*, but the spores of the former are not bilocular and have channels at each pole (Léger & Duboscq 1929a). The *Eccrinoides* species found in terrestrial isopod hosts (*E. helleriae* and *E. monticolae*) both have the polar channels. It seems questionable whether this single morphological character truly represents a homologous trait given the phenotypic plasticity among eccrinids. Would the host-based/habitat clade-sorting patterns hold up and divide these species in a future analysis? Could the polar channels in the spores of *Eccrinoides* be a morphological intermediary between undivided and bilocular spores? An in-depth, combined molecular and morphological analysis of these genera is warranted to elucidate these lingering questions.

The remaining genera (*Arundinula*, *Enteromyces*, *Enteropogon*, *Leidyomyces*, *Passalomyces*, *Taeniella* and *Taeniellopsis*) branch separately from one another, except for *Passalomyces* (for which no genomic samples were obtained) and those discussed above. All but *Arundinula*, *Enteropogon*, and *Taeniellopsis* are monotypic. In the case of *Taeniella*, three species were described at one point, but Lichtwardt et al. (2001) felt that character overlap prohibited proper delineation, and so they were all collapsed into *T. carcini*. On the other hand, Hibbits Galt (1971) acknowledged the similarity and intraspecies variability of characters, but felt each species was distinguishable. However, she sometimes observed *T. carcini* in the hindgut of the same hosts harboring *Arundinula washingtoniensis*. *Taeniella* and *Arundinula* have morphological similarity to each other and produce spores with appendages (four on the former and two on the latter), but also exhibit variability of their hyphal morphotypes. Combined with the overlap in hosts, morphological differentiation of these species could be problematic. Certainly, as with *Enterobryus*, future combined morphometric/molecular studies covering the range of hosts of these genera are needed to disentangle possible phenotypic plasticity and niche partitioning of the host gut.

Taeniella and the other crustacean-associated taxa (minus the terrestrially associated *Parataeniella*) form unsupported clades together while *Eccrinidus* is on a branch positioned near the millipede-associated *Enterobryus* spp. Within the crustacean clade, subclade formation may be at least partially attributable to freshwater vs. marine habitat. For example, *Arundinula opeongoensis* (White & Strongman 2008) and *Astreptonema gammari* are found in freshwater hosts and each form well-supported clades apart from marine taxa. However, both of these genera contain marine taxa that

are not represented here. *Arundinula orconectis* (Lichtwardt 1962) shares the same host as *A. opeongoensis*, and together they are the only two species in the genus from freshwater crayfish, rather than marine anomurids. Based on the patterns observed here, the marine *Arundinula* could be expected to form a separate clade from these freshwater specimens.

In contrast, *Taeniellopsis* (Poisson 1927) might be expected to align more closely with the marine *Astreptonema pacificum*. Not only have described *Taeniellopsis* species been collected from the same amphipod host genus as *A. pacificum* (*Orchestia* spp.) (Lichtwardt et al. 2001), but also the representative voucher here is an early diverging lineage to the PAA clade (Fig. 1.1). The sequence data for that sample is incomplete and its placement is unsupported, but the intermediate position could reflect its simultaneous affinity for the amphipod and marine clades. Morphological dimensions of the three *Taeniellopsis* species overlap (Poisson 1927; Poisson 1929; Manier 1970), and the main distinguishing features to separate them from *Astreptonema* are lack of microthalli and spore mother-cells. While there was no suggestion of a microthallus-like structure for any species, Poisson (1929) did mention "enigmatic protuberances" on some *T. flexilis* hyphae and the mother-cells of some *Astreptonema* spp. are not as conspicuous or well defined as those of *Alacrinella* or *Arundinula* (Lichtwardt et al. 2001).

Another correlated character among eccrinids is location within the host. *Arundinula* (Leger & Duboscq 1906) and *Enteromyces* (Lichtwardt 1961) are the only two genera found to inhabit the foregut (Lichtwardt et al. 2001), *Enteromyces* exclusively so and *Arundinula* in both fore- and hindgut (but *A. opeongoensis* has only been recorded from the hindgut so far). *Enteromyces* hyphae are distinct in that they have a multiple

holdfast system giving them a “tuft”-like appearance (*Leidyomyces* is the only other eccrinid with this generalized growth form), but they have a spore mother-cell similar to *Arundinula*. The foregut hyphae of *Arundinula* are morphologically distinguishable from those in the hindgut, having thicker cell walls as well as variability in hyphal length and curvature, and spores produced. As the most complex, Hibbits Galt (1978) described *A. hapalogaster* with six different hyphal forms. She also reported rare holocarpic cleavage for hyphal morphotypes of that species and *A. washingtoniensis* (1971), some cells of which rounded up into a spherical form, but their function was not clear. Duboscq et al. (1948) hypothesized that these rounded cells were gametes, but no evidence has been acquired for or against this idea. Similarly, *Enteropogon* spp. have four hyphal morphotypes with spore mother-cells, some of which produce “rounded up” cells suggested to play a role in a (as yet unconfirmed) sexual process (Hibbits Galt 1978; Chien & Hsieh 2001). *Enteropogon* hyphae are different from *Arundinula* in that they form scalariform fusions with one another and are only found in the hindgut (Hibbits Galt 1978). The function of these cells and confirmation of these hyphal morphotypes as true conspecifics remains to be determined. It is interesting to note that Lichtwardt et al. (2001) questioned the legitimacy of *Enteropogon* (but not the distinctiveness of *E. sexuale*), as its morphology strongly resembles that of *Enterobryus*. The type species (*Enterobryus elegans*) was taken from a millipede, and considering the apparent polyphyly of the genus and the significant correlation of host type with the phylogeny, *Enteropogon* should remain a separate genus at least until more specimens of anomurid-associated taxa are obtained.

Finally, both *Leidyomyces* and *Passalomyces* are found in the guts of passalid beetles and, indeed, may cohabit a single individual host. Beetle guts are convoluted as the length of the gut is more than that of the insect itself, and these genera are physically separated along it, with *Leidyomyces* found in the anterior hindgut and *Passalomyces* in the posterior hindgut (Lichtwardt et al. 1999). In addition to this partitioning within the host, *Passalomyces* has only been recorded from tropical areas, whereas *Leidyomyces* has been found in those same areas plus at sites in North America (Lichtwardt et al. 1999; Lichtwardt et al. 2001). Morphologically, *Leidyomyces* grows in "tufts" of hyphae attached to a multiple holdfast structure or individually, and *Passalomyces* produces only flattened, disk-like, thick-walled spores. The three *Leidyomyces* sp. samples from Guatemala and Mexico form a well-supported clade, but the *L. attenuatus* sample from Australia is sister to that clade with support only from BI and MP. Whether the divergence between these specimens is more attributable to geographic isolation, a difference in host genus, or is indicative of *Leidyomyces* sp. as a new species (or a combination of these factors), is unclear. Future phylogeographic evaluations of these genera will help explicate these issues, but are worthy of consideration in these social beetles (Schuster & Schuster 1985).