environmental microbiology reports

Environmental Microbiology Reports (2022) **14**(4), 538–548 doi:10.1111/1758-2229.13060

Brief Report

Alveolates (dinoflagellates, ciliates and apicomplexans) and Rhizarians are the most common microbial eukaryotes in temperate Appalachian karst caves

A. Bruce Cahoon * **and Robert D. VanGundy**
Department of Natural Sciences, The University of Virginia's College at Wise, Wise, VA, 24293.

Summary

The purpose of this study was to survey the eukaryotic microbiome of two karst caves in the Valley and Ridge physiographic region of the Appalachian Mountains. Caves are known to harbour eukaryotic microbes but their very low densities and small cell size make them difficult to collect and identify. Microeukaryotes were surveyed using two methodologies, filtering water and submerging glass microscope slides mounted in periphytometers in cave pools. The periphyton sampling yielded 13.5 times more unique amplicon sequence variants (ASVs) than filtered water. The most abundant protist supergroup was Alveolata with large proportions of the ASVs belonging to dinoflagellate, ciliate and apicomplexan clades. The next most abundant were Rhizarians followed by Stramenopiles (diatoms and chrysophytes) and Ameobozoans. Very few of the ASVs, 1.5%, matched curated protist sequences with greater than 99% identity and only 2.5% could be identified from surface plankton samples collected in the same region. The overall composition of the eukaryotic microbiome appears to be a combination of bacterial grazers and parasitic species that could possibly survive underground as well as cells, cysts and spores probably transported from the surface.

Introduction

Caves are extremely oligotrophic environments yet harbour a multitude of cave-adapted organisms, from charismatic macroscopic fauna to extremophilic prokaryotes (Culver and Pipan, 2009; Engel, 2010; Tomczyk-Żak and Zielenkiewicz, 2016). These survive without the benefit of sunlight or abundant primary producers so biological energy is limited to dissolved organic material and microbes transported from the surface by water movement (Laiz *et al*., 1999, Simon *et al*., 2003, Engel and Northup, 2008) and chemolithoautotrophic cave-adapted microbes (Northup and Lavoie, 2001; Chen *et al*., 2009; Ortiz *et al*., 2014).

Cave-dwelling prokaryotic microbes have been the focus of numerous surveys, beginning with culturable species and most recently molecular metagenomic barcoding techniques. These have revealed a wide array of prokaryotes including many potential endemic species, suggesting caves may act as refugia for various prokaryotic microbes, some of which are pathogenic or have unusual physiologies (Engel, 2010; Igreja, 2011). Recent metabarcoding surveys focused on the microbial ecology of caves has revealed that only 11.2%–21.4% of the prokaryotic barcodes found on speleothems (cave formations) and 53.8% of the microbes in cave sediments were found in nearby surface soils (Ortiz *et al*., 2013; Lavoie *et al*., 2017; Thompson *et al*., 2019). The use of barcoding has also revealed the movement of microbes from the surface and epikarst into karst caves (Morse *et al*., 2021). These results suggest that, despite the constant transfer of microbes from the surface, the cave ecosystem, especially speleothems, are inhabited by many cave-endemic microbes.

The eukaryotic microbiome of cave systems has received less attention than macroscopic fauna and prokaryotic microbiomes. Surveys of these organisms are challenging as they are difficult to find due to their low population density and they are often much smaller than their surface-dwelling counterparts, making them difficult to identify. When systematic microscopic screens of protists have been conducted, numerous uncharacterized species were found (Coppellotti and Guidolin, 2003; Bastian *et al*., 2009; Sigala-Regalado *et al*., 2011; Baković *et al*., 2019). Metabarcoding has recently been used in some eukaryotic surveys, including cave walls in a freshwater algae biodiversity survey of the Hawaiian

Received 24 June, 2021; accepted 10 March, 2022. *For correspondence. E-mail: abc6c@uvawise.edu; Tel. (276) 328-0201.

© 2022 The Authors. *Environmental Microbiology Reports* published by Society for Applied Microbiology and John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

archipelago (Sherwood *et al*., 2014), surveys of green algae and diatoms in show cave Lampenflora (Pfendler *et al*., 2018; Burgoyne *et al*., 2021), soils from the Manao-Pee Cave in Thailand (Wisechart *et al*., 2019), and a survey of microbial mats growing in low oxygen airbells in Movile Cave Romania (Reboul *et al*., 2019). Microeukaryotes are essential components of every food web (Corliss, 2002) as they link bacteria and small metazoans so describing them is important for a full understanding of the underground food web. The purpose of this study was to attempt a survey of the eukaryotic microbiome of two relatively shallow temperate caves in the Appalachian mountain chain using environmental DNA metabarcoding. Our goals were to refine sampling techniques to enhance the recovery of these organisms and attempt to identify as many barcodes as possible. We hope this study can serve as a productive step in the exploration of cave eukaryotic microbial ecology.

Results and discussion

Sampling strategy

Our first attempts to detect microeukaryotes using metabarcoding utilized DNA extracted from 49 L of filtered water collected from Panel and Bolling Caves from 2016 to 2019. These were screened for microeukaryotes using the 'universal' 18S rDNA primer set and collectively, 54 unique amplicon sequence variants (ASVs) were detected from all samples from both caves. Hoping to increase yields, we attempted a different year-long assay using periphytometers and DNA extracted from glass slides submerged in cave water, which yielded 729 unique ASVs. For Bolling Cave 261 ASVs were detected using the periphyton sampler and 16 ASVs from scooping water. For Panel Cave 468 ASVs were detected using the periphyton sampler and 38 from scooping water. There was no overlap in ASVs detected from the two sampling methods so the two were combined for a total of 783 that were used for all subsequent analyses.

Our yields using periphytometers were comparable to another 18S rDNA metabarcoding survey from Movile cave, Romania (Reboul *et al*., 2019) which reported microeukaryotes from hypoxic air pockets and noted that they had higher yields from 'fresh' samples, meaning microbial mats whose DNA was extracted very soon after collection as opposed to mats that had been cultured for 2 months. In our study, the fact that the glass slides were immersed in DNA extraction buffer immediately after removal from the periphytometer in the cave may have also contributed to the higher ASV yield than those from filtered water.

ASV identification

GenBank records were used to create a reference tree from a combination of taxa predicted to be the most closely related to the cave ASVs as well as other taxa from microeukaryotic groups (Supplemental File 1). When a phylogenetic tree was produced using the reference sequences and the cave ASVs, 502 of the ASVs were broadly identifiable (summarized in Fig. 1A, all listed in Supplemental File 2). This approach provided taxonomic identities for 394 of the ASVs (Supplemental File 3). Only 1.5% of the ASVs matched a named GenBank record with greater than 99% identity, others matched records named 'environmental sample' or 'uncultured'. We also used the PR2 reference database to make taxonomic assignments and 13.2% could be identified to the genus level with 99% certainty, 17.4% with 95% certainty, or 20.4% at 90% certainty. This is similar to Reboul *et al*. (2019) who found that few of their OTUs matched curated/cultured 18S microeukaryotic rDNA sequences.

This lab has been collecting plankton from surface water samples from various sites in and around Southwestern Virginia (the location of Panel and Bolling Caves) since 2015 and has an archive of freshwater 18S rDNA sequences which, to date, has 3283 unique ASVs (Cahoon *et al*., 2018, Fawley *et al*., 2021, and unpublished data). Only 2.5% of the cave samples matched one of these sequences with greater than 99% identity.

Temporal differences

The microscope slides were installed/removed at three time points over the course of a year so we were able to roughly estimate temporal changes in the most common groups in each cave (Fig. 1B). Notably, Rhizarians were the only group detected in both caves at all three time points. Dinoflagellates were detected in both caves but only from slides left submerged from May to September or September to January. A larger diversity of microeukaryotes for the three time points was found in Panel Cave. The temporal differences in microeukaryote variety and abundance are intriguing but it is not possible to associate them with abiotic factors since none were measured during his project.

Alveolates

The microeukaryotic supergroup with the greatest ASV representation were the alveolates. Dinoflagellates were well represented in the cave samples with 72 ASVs forming a clade with Dinoflagellate reference sequences (Fig. 2 and Supplemental File 4). When compared to the PR2 protist database, the same 72 ASVs were classified as Dinoflagellates with >70% certainty and 51 of the

© 2022 The Authors. *Environmental Microbiology Reports* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

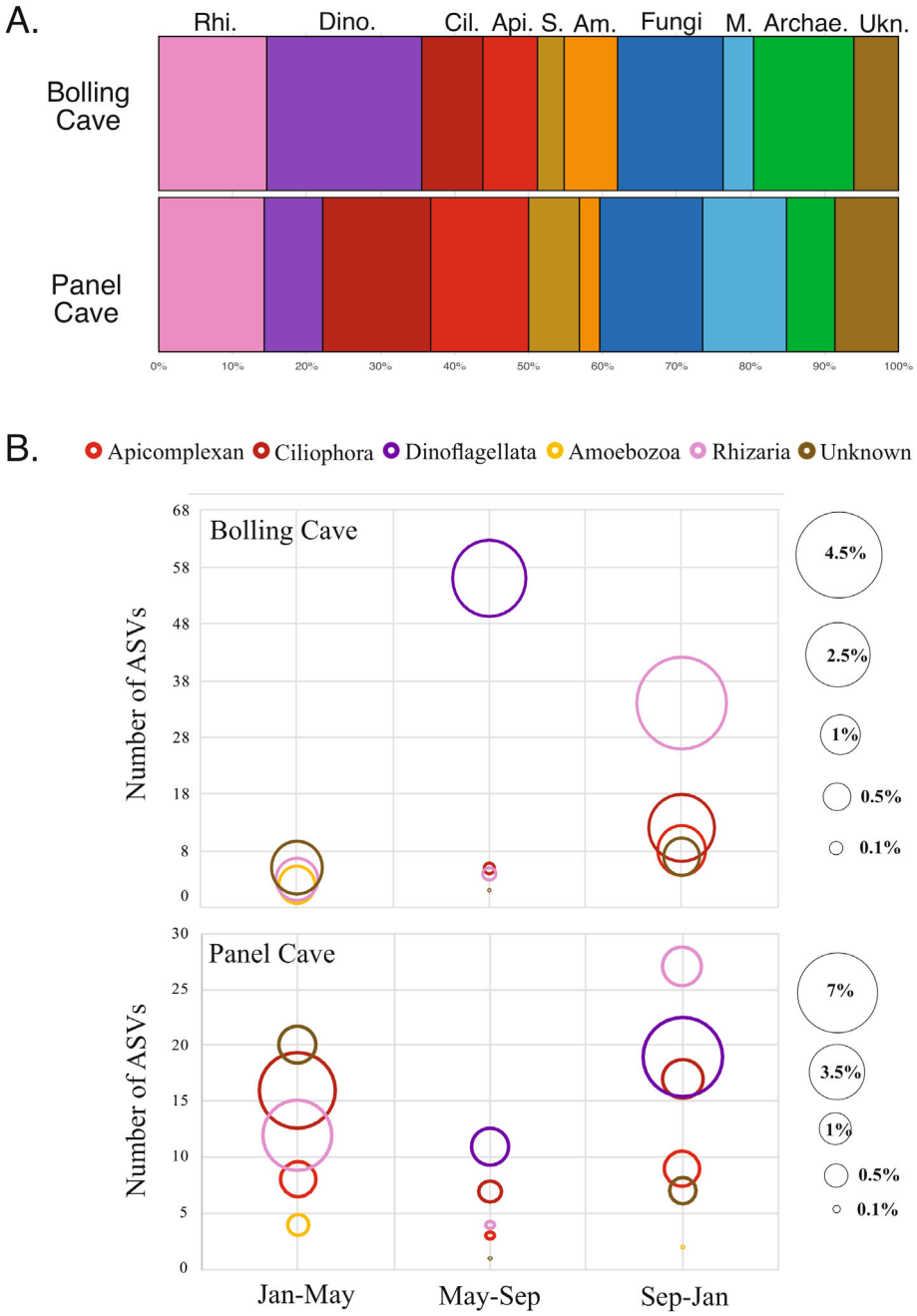

Fig. 1. Distribution of 18S rDNA ASVs found in Panel and Bolling caves.A. Proportion of ASVs from each taxonomic group found in each cave from filtered water and periphytometers combined. Am. = Amoebozoa, Api. = Apicomplexa, Archae. = Archaeplastida, Cil. = Ciliophora, Dino. = Dinoflagellata, M. = Metazoa, Rhi. = Rhizaria, S. = Stramenopiles, Ukn. = Unknowns.B. Distribution of the number of periphytometer collected ASVs and the relative proportion of each group, based on read counts, collected at each time point from each cave.

ASVs with >90% certainty (Supplemental File 3). Dinoflagellates have only been recorded in one other cave survey and it was a minor component (Reboul *et al*., 2019). About 88% of described freshwater dinoflagellates are photosynthetic with very few being benthic or heterotrophic (Tang, 1996; Stoecker, 1999; Gómez, 2012), attributes that would be useful for cave-dwelling species. The majority of the ASVs, 68, formed a clade with the genus *Peridinium*, which was confirmed by the PR2 taxonomic predictions which identified some ASVs as *Peridinium volzii* and *Peridinium willei*. *Peridinium* is a common freshwater genus (Carty and Parrow, 2015) but it is unknown whether they can heterotrophically maintain homeostasis as would be required in a cave. These data do suggest that conditions either in the caves or in the surface hydrological environment have been favourable

© 2022 The Authors. *Environmental Microbiology Reports* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

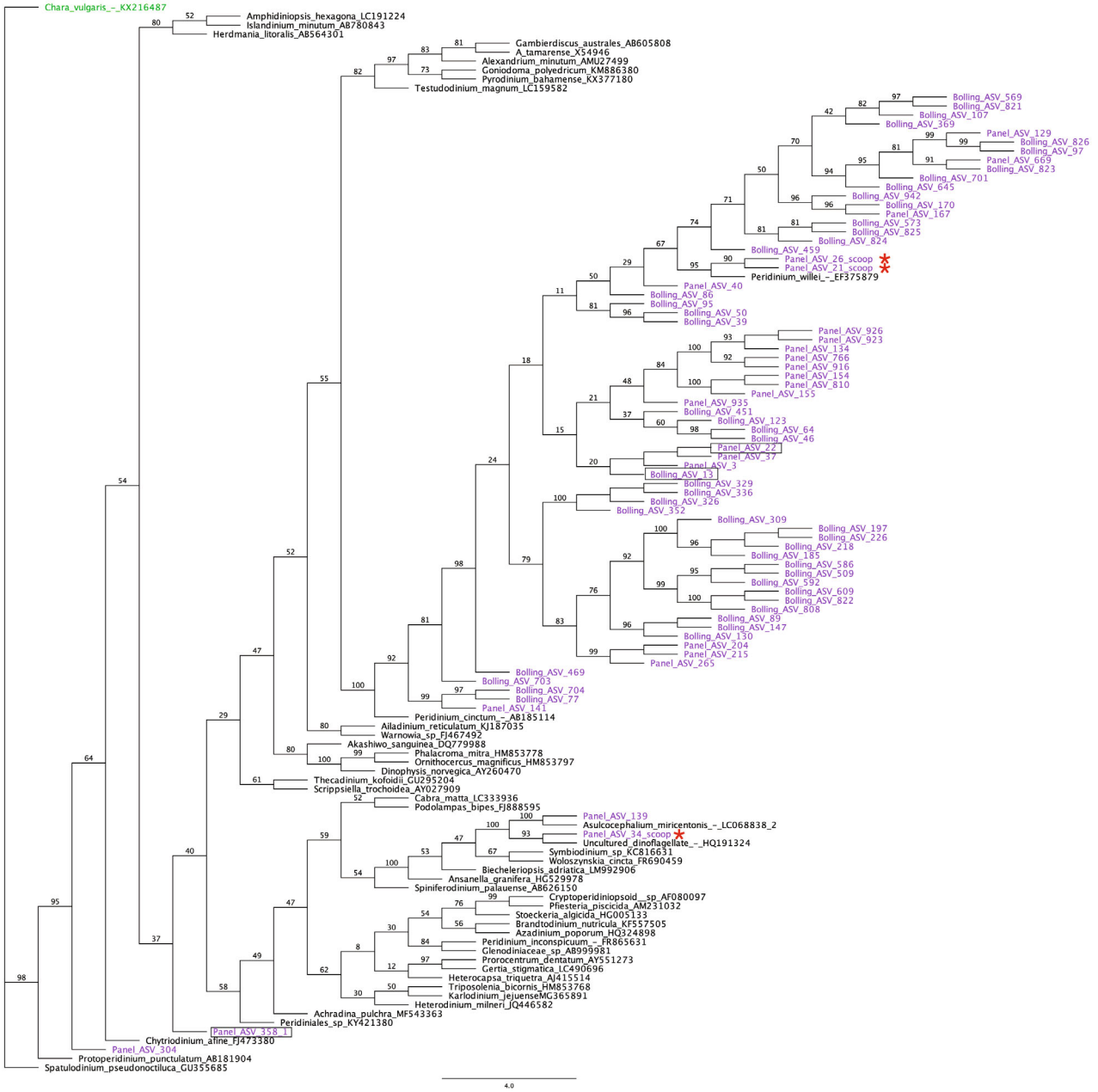

Fig. 2. Cave 18S rDNA ASVs classified as dinoflagellates. A maximum likelihood phylogeny was generated using IQTREE and the ModelFinder option. Reference sequences are black, cave ASVs purple, ASVs collected from scooped water are marked with asterisks and an outgroup sequence is green. ASVs found in surface samples are enclosed by rectangles. Bootstrap values are listed at each node. The scale represents number of substitutions.

for the radiation of this genus. Eleven of the ASVs matched *Asulcocephalium miricentonis* which was first characterized in a temperate freshwater pond in Japan (Takahashi *et al*., 2015).

Reference sequences from Ciliophora formed a clade with 65 cave ASVs (Supplemental Fig. 1 and Supplemental File 5) all of which were classified as Ciliophora with >70% certainty (Supplemental File 3). The majority of those were within the Orders Hymenostomatida and *Peritrichia*. Comparisons with the PR2 database suggested 50 of the ASVs matched a record with greater than 90% certainty. In both microscopic and metabarcoding cave studies, ciliates were the most common microeukaryote (Sigala-Regalado *et al*., 2011; Reboul *et al*., 2019). The most common ciliate in Panel and Bolling caves appear to be within the genus *Tetrahymena*. Three ASVs from Bolling and Panel caves had greater than 99% identity to *Tetrahymena* (*T. eeiotti*, *T.*

© 2022 The Authors. *Environmental Microbiology Reports* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

bergeri and *T. farleyi*) and as many as 15 others formed a clade that may represent unnamed members of this genus. *Tetrahymena* sp. are ubiquitously present in aquatic environments where they graze on bacterial and viral particles (Verni and Gualtieri, 1997; Pinheiro *et al*., 2007). They are also a valuable environmental bio-indicator representative of healthy aquatic environments (Maurya and Pandey, 2020). Their presence in our survey could mean this species may be useful in evaluating the conservation status of a cave ecosystem. Four of these *Tetrahymena* ASVs were also found in surface samples. Two ASVs closely matched the genus *Choreotrichia* and four other ASVs formed a clade consistent with the class Spirotrichea which are found in freshwater environments but are best known in the marine benthosphere where they consume bacteria and microalgae and are prey for small metazoans (Pierce and Turner, 1992; Calbet and Saiz, 2005; Santoferrara *et al*., 2017). One other ASV closely matched a *Peritrichia* sp. record which along with 19 other ASVs grouped within the class Oligohymenophorea, a species-rich but undersequenced group of ciliates (Sun *et al*., 2021). To the best of our knowledge, there are no records of cave endemic *Tetrahymena*, *Choreotrichia*, or *Peritrichia* but their lifestyles and physiology (Lynn, 2017) would be suitable to life in a lightless environment as long as there are sufficient numbers of prokaryotic and microeukaryotic prey available

Fifty nine ASVs formed a clade with Apicomplexan reference sequences (Fig. 3 and Supplemental File 6). Most were in the order Eugregarinorida. Two ASVs, Boll_ASV_155 and Boll_ASV_197 formed a monophyletic branch within the Apicomplexan phylogeny without a closely associated GenBank record but PR2 based taxonomy suggested they were in the Order Colpodellida, a free-living non-parasitic branch of Apicomplexans. No individual ASVs closely matched reference sequences and only 29 were classified within the PR2 database (Supplemental File 3). The taxonomic predictions based on PR2 did suggest that several of the Eugregarinorida ASVs were assigned to the genus level with greater than 90% certainty and included *Monocystis* sp., *Syncystis mirabilis* and *Paraschneideria metamorphosa*. The ASV predicted to be closely related to *Syncystis mirabilis* was also found in surface samples. Apicomplexans are a widespread (Cavalier-Smith, 2014) but under-characterized group of microeukaryotes best-known as parasites (Simdyanov *et al*., 2017; Janouškovec *et al*., 2019). A recent metabarcoding screen of surface-dwelling apicomplexans found very low diversity in freshwater samples (del Campo *et al*., 2019) but they were the most common microeukaryote found in the soils of the Manao-Pee Cave in Thailand (Wisechart *et al*., 2019) and three were found in anoxic microbial mats (Reboul *et al*., 2019). Our data along with other cave studies suggest these environments could be a source of apicomplexan diversity. We attempted to produce apicomplexan specific 18S rDNA barcodes using published primer sequences (Huggins *et al*., 2019) but failed to produce amplicons. We believe this failure was due to these primers having been designed to recognize Apicomplexan parasites using available GenBank records where there is a paucity of sequences similar to those found in these cave environments.

Rhizarians

Within Rhizaria, 78 ASVs were identified using the phylogenetic approach (Supplemental Fig. 2 and Supplemental File 7) and 71 using the PR2-based taxonomy approach (Supplemental File 3). Two of the barcode sequences had greater than 99% similarity to GenBank archived sequences suggesting they may be the species *Orciraptor agilis* – KF207875 and *Neocercomonas* sp. – AY884313, which was confirmed by PR2-based taxonomy (Supplemental File 3). PCR amplifications using primers targeting foraminifera 18S rDNA were attempted but no amplicons were produced. The majority of the Rhizaria ASVs were Cercozoans which are most likely bacterial grazers (Burki and Keeling, 2014) which could survive in an oligotrophic environment. There were also ASVs classified as Vampyrellida, also known as predatory amoebae that parasitize other microeukaryotes (Hess, 2017; More *et al*., 2019) and Plasmodiophorida best known as plant parasites (Hwang *et al*., 2012). Rhizaria were reported as a significant component of microbial mats collected in the Movile cave but they were not the predominant group, except in a cultured sample (Reboul *et al*., 2019). Freshwater foraminifera have been microscopically identified from European karst caves (Mazei *et al*., 2012; Baković *et al*., 2019) but were not detected in our survey using either universal 18S rDNA or foraminifera targeted primers.

Stramenopiles

Fifty two ASVs were predicted to be Stramenopiles according to PR2-based taxonomy (Supplemental File 3) while 30 formed a clade with stramenopile sequences using the phylogenetic approach (Supplemental Fig. 3 and Supplemental File 8) with all grouping in either Bacillariophyta (diatoms) or Chrysophyceae. Six of the 16 Bacillariophyta ASVs were found in surface samples while six of the 14 Chrysophyceae were also found on the surface. Two of the diatom ASVs had greater than 99% identity with *Gomphonema micropus* – JN790282 and *Achnanthidium pyrenaicum* – KY863466, while one of the chrysophytes matched the record for *Uroglenopsis*

© 2022 The Authors. *Environmental Microbiology Reports* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

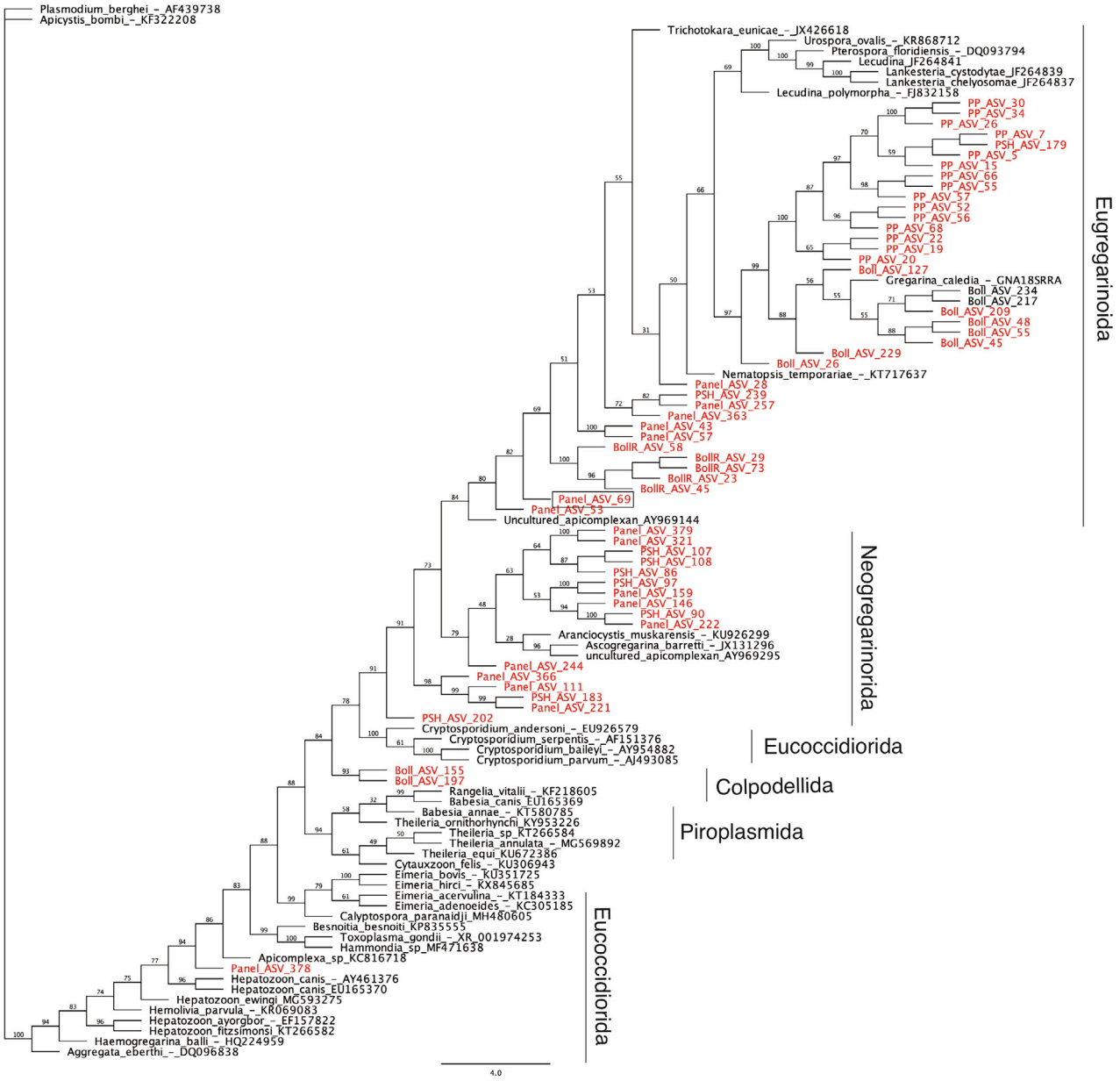

Fig. 3. Cave 18S rDNA ASVs classified as apicomplexans. A maximum likelihood phylogeny was generated using IQTREE and the ModelFinder option. Reference sequences are black and cave ASVs red. ASVs found in surface samples are enclosed by rectangles. Dashed lines represent unknown taxonomy within Apicomplexa. Bootstrap values are listed at each node. The scale represents number of substitutions.

americana – MK153242. PR2-based taxonomic predictions agreed with these identities at the genus level. The additional Stramenopiles identified by PR2-based taxonomy were predicted to be in the Classes MAST-12, Oomycota and Labyrinthulomycetes (Supplemental File 3). Since these species have chloroplasts, a primer pair targeting plastid 23S rRNA was used to confirm their presence in the cave samples. Ninety eukaryotic 23S sequences were identified (Supplemental Fig. 4 and Supplemental File 9). The vast majority of these appeared to be embryophytes and green algae. Nine were grouped within the diatom clade but none matched *Gomphonema* or *Achnanthidium* similar to the 18S rDNA barcodes, but four did group with reference sequences from the genus *Synedra*. The remaining five were not identifiable with the 23S barcode. Fifteen were chrysophytes. Diatoms are well documented in caves as they are a component of show cave lampenflora and a cave conservation concern (reviewed in Falasco *et al*., 2014; Piano *et al*., 2015; Pfendler *et al*., 2018; Burgoyne *et al*., 2021). Chrysophytes are not a common component of lampenflora or of cave surveys but are mixotrophs (Nicholls and Wujek,

© 2022 The Authors. *Environmental Microbiology Reports* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

2015) so it is conceivable they could survive in a lightless environment. The predominating hypothesis regarding diatoms is that they are regularly introduced into caves as living cells or cysts which can opportunistically colonize under artificial light and that there are no endemic cave species. We think it likely that the diatom sequences detected in our screen were from spores or cysts that had been transported from the surface. We believe this since we found no evidence of non-photosynthetic diatoms (Kamikawa *et al*., 2018) which would be the only ones capable of colonizing a cave. Also, a high proportion of them, 40%, were also found in planktonic surface samples.

Amoebozoans

Twenty five ASVs grouped with amoebozoan reference sequences (Supplemental Fig. 5 and Supplemental File 10). All appeared to be within the phylum Discosea but none had greater than 99% identity with a reference sequence and none were found in surface samples. We only found non-testate Discosea amoebae in our screen and none matched a GenBank record with greater than 99% similarity and none were found among surface ASVs. Non-testate amoebae have been reported from Lascaux Cave in France but the largest number of amoebae identified from caves is testate amoebae (Mazei *et al*., 2012; Baković *et al*., 2019).

Are there endemic cave microeukaryotes?

The majority of the ASVs detected in our cave survey did not occur in surface water samples, except for photosynthetic stramenopiles. Based on this alone it is tempting to conclude that most of these ASVs were cave endemic species but there are limitations in our comparisons. For example, the library of local surface planktonic microeukaryotes may not represent a true comparison to the benthic cave samples collected on the glass slides. Therefore we cannot estimate the proportion of the microeukaryotes found in our survey that are truly cave endemics, and it may not be possible until a surface periphyton survey is completed. Another limitation is the lack of archived microeukaryotic freshwater barcode sequences, which reflects the lack of sequence information from many branches of microeukaryotes (del Campo *et al*., 2014) and is even more problematic in undersampled environments. Metabarcoding technology allows us to detect the diversity of microeukaryotes from caves and make relative comparisons but a major challenge will be to isolate and/or culture these organisms to expand our sequence databases.

Baković *et al*. (2019) argue that the hydrological connections between the surface and karst cave systems prevent spatial isolation which would be essential for speciation and endemism of water-dwelling microeukaryotes. The caves we sampled could align with this hypothesis since these are relatively shallow caves and it was previously established that 86.9% of the planktonic prokaryotic microbes overlap with those in surface water, which along with dye trace experiments of Panel Cave shows that surface water is continuously entering and cycling through what is presumed to be an epikarst layer (Morse *et al*., 2021). We also detected photosynthetic eukaryotes such as Stramenopiles, green algae and embryophytes, which would only be present if introduced from the surface by way of percolating water, air movement, or animal vectors. Although both caves are wild and have no permanent artificial lighting, they are in a state park and guided tours of the caves are offered to visitors during summer months who could transport spores, cysts and pollen into these caves.

Conclusions

We used two sampling methods to survey microeukaryotes in two karst caves and found periphytometers with glass slides provided a greater number and diversity of ASVs than scooping and filtering water when sampling this low-density fraction of the microbiome. Using both methods we identified 784 unique microeukaryotic ASVs from two caves representing a wide array of bacterial grazers and parasites that could survive in the cave environment as well as other transient cells and/or spores and cysts transported from the surface. This study provides a baseline survey of microeukaryotes for the relatively common karst caves found in temperate regions.

Experimental procedures

Samples were collected from Panel and Bolling Caves in the Natural Tunnel State Park in Scott County, VA, USA using two different methodologies, scooping water and submerging periphytometers in cave streams and pools. The attributes of each cave, hydrogeology and sample sites were described in Thompson *et al*. (2019) and Morse *et al*. (2021). Briefly, they are gated wild caves located approximately 1.2 km apart on opposite ridges along Stock Creek. To our knowledge, the caves are hydrologically separated and developed in lower to middle Ordovician-aged carbonate rock (Miller and Brosge, 1954; Brent, 1963). Permission to perform this research and collect samples was granted by the Commonwealth of Virginia Department of Conservation and Recreation (Research and Collecting Permit NT-RCP-121819).

The scooping method consisted of collecting three litres of water in three sterile 1-L screw-cap bottles at each collection site at each time. Water was collected

© 2022 The Authors. *Environmental Microbiology Reports* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

from still pools approximately 25 cm deep and running streams 5–10 cm deep. Each litre of water was filtered through a disposable microfunnel with a 0.45 μm mesh filter (Daigger & Co., Vernon Hills, IL, USA) within 24 h of collection in a lab at UVA-Wise. Filters were stored at −20°C until DNA extractions could be completed. Filters were thawed and suspended in the extraction buffer and bead-beater tubes provided with Qiagen's DNeasy PowerWater Kit (Germantown, MD, USA). Using this method, Panel Cave was sampled once in April 2016, once in April 2017, and monthly for 14 months from April 2018 to May 2019 for a total of 16 sampling events. Bolling Cave was sampled once in April 2016 and once in April 2017. The 2017 samples from Panel and Bolling Caves and the monthly 2018–2019 Panel Cave samples were used in previous studies exploring the prokaryotic microbial ecology of these caves (Thompson *et al*., 2019; Morse *et al*., 2021).

The second method utilized periphytometers, devices that provide an artificial substrate for waterborne microbes to adhere to and colonize. These are commonly used to monitor microbial biomass and diversity for environmental bioassessment assays of surface water (Aloi, 1990; Barbour *et al*., 1999). We used plexiglass periphytometers designed to hold standard sterilized glass microscope slides. Beginning in January 2020 the samplers were submerged in water in the same two locations that had been sampled by scooping in Panel Cave and the two locations in Bolling Cave. Water depth ranged from 10 to 50 cm depending upon the collection site and water levels within the cave which varied throughout the year. The substrate supporting the periphytometers was solid rock, pebbles, or sand. The glass slides were collected and replaced approximately every 4 months for one calendar year, three collection events. At each retrieval time, the glass slides were removed from the samplers, placed directly into DNeasy PowerWater Kit bead-beater tubes containing DNA extraction buffer at the sampling site, shaken vigorously by hand, and then transported out of the cave. Samples were stored in light tight bags for transport out of the caves. Tubes were taken to a laboratory at UVA-Wise and DNA extractions were completed the day of each collection.

Samples of DNA from all the replicates collected from each cave at each time (e.g. six per cave for each periphyton collection event) were individually used as templates for three PCR reactions. These replicates were combined to minimize PCR bias. Phusion DNA polymerase (Thermo Fisher, Waltham, MA, USA) was used for all amplifications. The V4 region of the 18S rDNA gene was amplified using 'universal' microeukaryote primers, TAReuk454FWD1 (5′CCAGCASCYGCGGTAATTCC) and TAReukREV3 (5′ACTTTCGTTCTTGATYRA) (Stoeck *et al*., 2010), as well as the 'universal' 23S rDNA primers, p23SrV_f1 (5′ GGACAGAAAGACCCTATGAA) and p23SrV_r1 (5′ TCA GCCTGTTATCCCTAGAG) (Sherwood and Presting, 2007), apicomplexan-targeted 18S rDNA primers, ApicomplexF: (5′-CRAGGAAGTTTRAGGCAATAA CAG) and ApicomplexR: (5′-CTAGGCATTCCTCGTTHAH GATT) (Huggins *et al*., 2019), and foraminifera-targeted 18S rDNA primers, S14F1 (5′-CCATCTCATCCCTGCGTG TCTCCGAC) and S19F (5′-GTACRAGGCATTCCTRGTT) (Morard *et al*., 2018). Primers were synthesized by Integrated DNA Technologies (Coralville, IA, USA). Amplicons were purified using a Select-a-Size DNA Clean and Concentrator kit (Zymo Research, Irvine, CA, USA). Amplicon mixtures were paired-end sequenced using Genewiz's Amplicon EZ Illumina MiSeq service (South Plainfield, NJ, USA). Paired-end reads were processed and ASVs were produced using DADA2 using default parameters except for truncLen = c(250,250) (Callahan *et al*., 2016). The ASV outputs were transferred to Geneious Prime (Biomatters, Auckland, NZ) where primer sequences and duplicate ASVs were identified and removed. Barcode data from this study have been archived in Mendeley (DOI: 10.17632/r5fv5txtfs.1).

For phylogenetic analyses a reference set of 18S rDNA sequences was constructed with GenBank records (https://www.ncbi.nlm.nih.gov/genbank/, listed in Supplemental File 1). First, accessions most similar to cave ASVs were identified using BLAST searches to roughly identify the taxonomic groups present in the samples. Second, once a taxonomic group was identified all GenBank records of 18S rDNA sequences from that taxonomic group (e.g. Rhizaria) with a defined species were downloaded and included in the reference set. The overall reference set and the reference sets used for each taxonomic group's phylogeny were refined over numerous cycles until a suitable reference set containing key named taxa, accessions with the most similarity to the cave ASVs, and only as much taxonomic redundancy as necessary was formed. ASV and the GenBank derived reference sequences were aligned using MUSCLE with default parameters (Edgar, 2004) and sequence ends trimmed within the Geneious environment. Maximum Likelihood analyses were completed using IQTREE 1.6.12 (Trifinopoulos *et al*., 2016) using either ModelFinder (Kalyaanamoorthy *et al*., 2017) or the HKY substitution model (Hasegawa *et al*., 1985). Ultrafast bootstrapping was used to support the ML topology (Hoang *et al*., 2018). These data are represented as cladograms in this report for the simplest expression of the hypothetical classifications of the unknowns. Identity and Similarity scores were calculated using LALIGN (Huang and Miller, 1991) using the web-based version at (https://molbiol-tools.ca/Alignments.htm). ASVs were also classified based on the PR2 protist database, version 4.14.0 SSU (Guillou *et al*., 2013; del Campo *et al*., 2018) using QIIME2 (Bolyen *et al*., 2019) functions to trim the database

© 2022 The Authors. *Environmental Microbiology Reports* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

to the V4 SSU barcode region [feature-classifier extract-reads], train the classifier [feature-classifier fit-classifier-naïve-bayes] and assign taxonomy to the ASVs [feature-classifier classify-sklearn] using default parameters.

Acknowledgements

The authors wish to thank Natural Tunnel State Park in Duffield, VA for permission to sample in the Panel and Bolling Caves, the John C. Buchanan chair of biology endowment for funding, and James A. Nienow of Valdosta State University for supplying periphytometers. We also thank Teresa L. Brown of the University of Virginia's College at Wise for critical reading of the manuscript.

References

Aloi, J. (1990) A critical review of recent freshwater periphyton field methods. *Can J Fish Aquat Sci* **47**: 656–670.

Baković, N., Siemensma, F.J., Baković, R., and Rubinić, J. (2019) Testate amoebae in karst caves of the Dinaric Arc (South-Eastern Europe) with a description of Centropyxis bipilata sp. nov. *Acta Protozool* **58**: 201–215.

Barbour, M.T., Gerritsen, J., Snyder, B.D., and Stribling, J.B. (1999) *Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates, and Fish. EPA 841/B-99/022*. Washington, DC: Office of Water, US Environmental Protection Agency.

Bastian, F., Alabouvette, C., and Saiz-Jimenez, C. (2009) Bacteria and free-living amoeba in the Lascaux Cave. *Res Microbiol* **160**: 38–40.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., *et al*. (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**: 852–857.

Brent, W.B. (1963) *Geology of the Clinchport Quadrangle Virginia*, Vol. **5**. Virginia, USA: Virginia Division of Mineral Resources Report of Investigations, p. 47.

Burgoyne, J., Crepeau, R., Jensen, J., Smith, H., Baker, G., and Leavitt, S.D. (2021) Lampenflora in a show cave in the Great Basin is distinct from communities on naturally lit rock surfaces in nearby wild caves. *Microorganisms* **9**: 1188.

Burki, F., and Keeling, p.J. (2014). Rhizaria. *Curr Biol* **24**: R103–107.

Cahoon, A.B., Huffman, A.G., Krager, M.M., and Crowell, R.M. (2018). A meta-barcoding survey of freshwater planktonic protists in Appalachia—Natural Tunnel State Park, Virginia, USA. *Metabarcoding and Metagenomics* **2**: e26939.

Calbet, A., and Saiz, E. (2005) The ciliate-copepod link in marine ecosystems. *Aquat Microb Ecol* **38**: 157–167.

Callahan, B.J., McMurdie, p.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016) DADA2: high resolution sample inference from Illumina amplicon data. *Nat Methods* **13**: 581–583.

Carty, S., and Parrow, M.W. (2015) Dinoflagellates. In *Freshwater Algae of North America*, 2nd ed. Wehr, J.D., Sheath, R.G., and Kociolek, J.P. (eds). Cambridge, MA: Academic Press.

Cavalier-Smith, T. (2014) Gregarine site-heterogeneous 18S rDNA trees, revision of gregarine higher classification, and the evolutionary diversification of Sporozoa. *Eur J Protistol* **50**: 472–495.

Chen, Y., Wu, L., Boden, R., Hillebrand, A., Kumaresan, D., Moussard, H., *et al*. (2009) Life without light: microbial diversity and evidence of sulfur- and ammonium-based chemolithoautotrophy in Movile Cave. *ISME J* **3**: 1093–1104.

Coppellotti, O., and Guidolin, L. (2003) Taxonomy and ecology of ciliate fauna (Protozoa, Ciliophora) from karst caves in North-East Italy. *Subterr Biol* **1**: 3–11.

Corliss, J.O. (2002) Biodiversity and biocomplexity of the protists and an overview of their significant roles in maintenance of our biosphere. *Acta Protozool* **41**: 199–219.

Culver, D.C., and Pipan, T. (2009) *The Biology of Caves and Other Subterranean Habitats*. Oxford: Oxford University Press, p. 256.

del Campo, J., Heger, T.J., Rodríquez-Martínez, R., Worden, A.Z., Richards, T.A., Massana, R., and Keeling, p.J. (2019) Assessing the diversity and distribution of apicomplexans in host and free-living environments using high-throughput amplicon data and a phylogenetically informed reference framework. *Front Microbiol* **10**: 2373.

del Campo, J., Kolisko, M., Boscaro, V., Santoferrara, L.F., Nenarokov, S., Massana, R., *et al*. (2018) EukRef: phylogenetic curation of ribosomal RNA to enhance understanding of eukaryotic diversity and distribution. *PLoS Biol* **16**: e2005849.

del Campo, J., Sierack, M.E., Molestina, R., Keeling, P., Massana, R., and Ruiz-Trillo, I. (2014) The others: our biased perspective of eukaryotic genomes. *Trend Ecol Evol* **29**: 252–259.

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

Engel, A.S. (2010) Microbial diversity of cave ecosystems. In *Geomicrobiology: Molecular and Environmental Perspective*. Barton, L.L., Mandl, M., and Loy, A. (eds). Netherlands, Dordrecht: Springer, pp. 219–238.

Engel, A.S., and Northup, D.E. (2008). Caves and karst as models for advancing microbial sciences. In *Frontiers of Karst Research: Proceedings and Recommendations of the Workshop Held May 3 through 5, 2007* J.B. Martin and W.B. White, (eds). San Antonio, TX: Leesburg: Karst Waters Institute, Inc. (pp. 37–48).

Falasco, E., Ector, L., Isaia, M., Wetzel, C.E., Hoffmann, L., and Bona, F. (2014) Diatom flora in subterranean ecosystems: a review. *Int J Speleol* **43**: 231–251.

Fawley, M.W., Fawley, K.P., and Cahoon, A.B. (2021) Finding needles in a haystack – extensive diversity in the Eustigmatophyceae revealed by community metabarcode analysis targeting the *rbcL* gene using lineage-directed primers. *J Phycol* **57**: 1636–1647.

Gómez, F. (2012) A quantitative review of the lifestyle, habitat and trophic diversity of dinoflagellates (Dinoflagellata, Alveolata). *Syst Biodivers* **10**: 267–275.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., *et al*. (2013) The protist ribosomal reference database (PR^2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res* **41**: D597–D604.

© 2022 The Authors. *Environmental Microbiology Reports* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

Hasegawa, M., Kishino, H., and Yano, T. (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* **22**: 160–174.

Hess, S. (2017) Hunting for agile prey: trophic specialization in leptophryid ameobae (Vampyrellida, Rhizaria) revealed by two novel predators of planktonic algae. *FEMS Microbiol Ecol* **93**: fix104.

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., and Vinh, L.S. (2018) UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* **35**: 518–522.

Huang, X., and Miller, W. (1991) A time-efficient, linear-space local similarity algorithm. *Adv Appl Math* **12**: 337–357.

Huggins, L.G., Koehler, A.V., Ng-Nguyen, D., Wilcox, S., Schunack, B., Inpankaew, T., and Traub, R. (2019) A novel metabarcoding diagnostic tool to explore protozoan haemoparasite diversity in mammals: a proof-of-concept study using canines from the tropics. *Sci Rep* **9**: 12644.

Hwang, S.-F., Strelkov, S.E., Feng, J., Gossen, B.D., and Howard, R.J. (2012) Plasmodiophora brassicae: a review of an emerging pathogen of the Canadian canola (Brassica napus) crop. *Mol Plant Pathol* **13**: 105–113.

Igreja, R.P. (2011) Infectious diseases associated with caves. *Wilderness Environ Med* **22**: 115–121.

Janouškovec, J., Paskerova, G.G., Miroliubova, T.S., Mikhailov, K.V., Birley, T., Aleoshin, V.V., and Simdyanov, T.G. (2019) Apicomplexan-like parasites are polyphyletic and widely but selectively dependent on cryptic plastid organelles. *Elife* **8**: e49662.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermiin, L.S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**: 587–589.

Kamikawa, R., Azuma, T., Ishii, K.-I., Matsuno, Y., and Miyashita, H. (2018). Diversity of organellar genomes in non-photosynthetic diatoms. *Protist*, **169**: 351–361.

Laiz, L., Groth, I., Gonzalez, I., and Saiz-Jimenez, C. (1999) Microbiological study of the dripping waters in Altamira cave (Santillana del Mar, Spain). *J Microbiol Methods* **36**: 129–138.

Lavoie, K.H., Winter, A.S., Read, K.J.H., Hughes, E.M., Spilde, M.N., and Northup, D.E. (2017) Comparison of bacterial communities from lava cave microbial mats to overlying surface soils from Lava Beds National Monument, USA. *PLoS One* **12**: e0169339.

Lynn, D.H. (2017) Ciliophora. In *Handbook of the Protists*, 2nd ed. Archibald, J.M., Simposon, A.G.B., and Slamovits, C.H. (eds). Cham, Switzerland: Springer Press, pp. 679–730.

Maurya, R., and Pandey, A.K. (2020) Importance of protozoa *Tetrahymena* in toxicological studies: a review. *Sci Total Environ* **741**: 140058.

Mazei, Y., Belyakova, O., Trulova, A., Guidolin, L., and Coppellotti, O. (2012) Testate amoebae communities from caves of some territories in European Russia and northeastern Italy. *Protist* **7**: 42–50.

Miller, R.L., and Brosge. W.P., (1954) Geology and Oil Resources of the Jonesville District, Lee County, Virginia: USGS Geological Survey Bulletin No. 99, 210.

Morard, R., Garet-Delmas, M.-J., Mahé, F., Romac, S., Poulain, J., Kucera, M., and de Vargas, C. (2018) Surface Ocean metabarcoding confirms limited diversity in planktonic foraminifera but reveals unknow hyperabundant lineages. *Sci Rep* **8**: 2539.

More, K., Simpson, A.G.B., and Hess, S. (2019) Two new marine species of Placopus (Vampyrellida, Rhizaria) that perforate the theca of Tetraselmis (Chlorodendrales, Viridiplantae). *J Eukaryot Microbiol* **66**: 560–573.

Morse, K.V., Richardson, D.R., Brown, T.L., VanGundy, R. D., and Cahoon, A.B. (2021) Longitudinal metabarcode analysis of karst bacterioplankton microbiomes provide evidence of epikarst to cave transport and community succession. *PeerJ* **9**: e10757.

Nicholls, K.H., and Wujek, D.E. (2015) Chrysophyceae and Phaeothamniophyceae. In *Freshwater Algae of North America*, 2nd ed. Wehr, J.D., Sheath, R.G., and Kociolek, J. P. (eds): Cambridge, MA: Academic Press, pp. 537–586.

Northup, D.E., and Lavoie, K.H. (2001) Geomicrobiology of caves: a review. *Geomicrobiol J* **18**: 199–222.

Ortiz, M., Legatzki, A., Neilson, J.W., Fryslie, B., Nelson, W. M., Wing, R.A., *et al*. (2014) Making a living while starving in the dark: metagenomic insights into the energy dynamics of a carbonate cave. *ISME J* **8**: 478–491.

Ortiz, M., Neilson, J.W., Nelson, W.M., Legatzki, A., Byrne, A., Yu, Y., *et al*. (2013) Profiling bacterial diversity and taxonomic composition on speleothem surfaces in Kartchner Caverns. *AZ Microb Ecol* **65**: 371–383.

Pfendler, S., Karimi, B., Maron, P.-A., Ciadamidaro, L., Valot, B., Bousta, F., *et al*. (2018) Biofilm biodiversity in French and Swiss show caves using the metabarcoding approach: fist data. *Sci Total Environ* **615**: 1207–1217.

Piano, E., Bona, F., Falasco, E., La Morgia, V., Badino, G., and Isaia, M. (2015) Environmental drivers of phototrophic biofilms in an Alpine show cave. *Sci Total Environ* **536**: 1007–1018.

Pierce, R.W., and Turner, J.T. (1992) Ecology of planktonic ciliates in marine food webs. *Rev Aquat Sci* **6**: 139–181.

Pinheiro, M.D., Power, M.E., Butler, B.J., Dayeh, V.R., Slawson, R., Lee, L.E., *et al*. (2007) Use of *Tetrahymena* thermophila to study the role of protozoa in inactivation of viruses in water. *Appl Environ Microbiol* **73**: 643–649.

Reboul, G., Moreira, D., Bertolino, P., Hellebrand-Voiculescu, A.M., and López-García, P. (2019) Microbial eukaryotes in the suboxic chemosynthetic ecosystem of Movile Cave, Romania. *Environ Microbiol Rep* **11**: 464–473.

Santoferrara, L.F., Alder, V.V., and McManus, G.B. (2017) Phylogeny, classification and diversity of Choreotrichia and Oligotrichia (Ciliophora, Spirotrichea). *Mol Phylogenet Evol* **112**: 12–22.

Sherwood, A.R., Carlile, A.L., Neumann, J.M., Kociolek, J. P., Johansen, J.R., Lowe, R.L., *et al*. (2014) The Hawaiian freshwater algae biodiversity survey (2009-2014): systematic and biogeographic trends with an emphasis on the macroalgae. *BMC Ecol* **14**: 28.

Sherwood, A.R., and Presting, G.G. (2007) Universal primers amplify a 23S rDNA plastid marker in eukaryotic algae and cyanobacteria. *J Phycol* **43**: 605–608.

Sigala-Regalado, I., Mayén-Estrada, R., and Morales-Malacara, J.B. (2011) Spatial and temporal distribution of protozoa at Ceuva De Los Riscos, Querétaro, México. *J Caves Karst Stud* **73**: 55–62.

© 2022 The Authors. *Environmental Microbiology Reports* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

Simdyanov, T., Guillou, L., Diakin, A.Y., Mikhailov, K.V., Schrével, J., and Aleoshin, V.V. (2017) A new view on the morphology and phylogeny of eugregarines suggested by the evidence from the gregarine *Ancora sagittate* (Leuckart, 1860) Labbé, 1899 (Apicomplexa: Eugregarinida). *PeerJ* **5**: e3354.

Simon, K.S., Benfield, E.F., and Macko, S.A. (2003) Food web structure and the role of epilithic biofilms in cave streams. *Ecology* **84**: 2395–2406.

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D.M., Breiner, H.-W., and Richards, T.A. (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol* **19**: 21–31.

Stoecker, D.K. (1999) Mixotrophy among dinoflagellates. *J Euk Microbiol* **46**: 397–401.

Sun, M., Li, Y., Cai, X., Liu, U.Y., Chen, Y., and Pan, X. (2021). Further insights into the phylogeny of peniculid ciliates (Ciliophora, Oligohymenophorea) based on multigene data. *Mol Phylogenet Evol* **154**: 107003.

Takahashi, K., Moestrup, Ø., Jordan, R.W., and Iwataki, M. (2015) Two new freshwater Woloszynskioids *Asulcocephalium miricentonis* gen. et sp. Nov. and *Leiocephalium pseudosanguineum* gen. et sp. Nov. (Suessiaceae, Dinophyceae) Lacking an apical furrow apparatus. *Protist* **166**: 638–658.

Tang, E.p.Y. (1996) Why do dinoflagellates have lower growth rates? *J Phycol* **32**: 80–84.

Thompson, B.A., Richardson, D.R., VanGundy, R.D., and Cahoon, A.B. (2019) Metabarcoding comparison of prokaryotic microbiomes from Appalachian karst caves to surface soils in Southwest Virginia, USA. *J Caves Karst Stud* **81**: 244–253.

Tomczyk-Żak, K., and Zielenkiewicz, U. (2016) Microbial diversity in caves. *Geomicrobiol J* **33**: 20–38.

Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., and Minh, B.Q. (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res* **44**: W232–W235.

Verni, V., and Gualtieri, P. (1997) Feeding behavior in ciliated protists. *Micron* **28**: 487–504.

Wisechart, A., Mhuantong, W., Tangphatsornruang, S., Chantasingh, D., and Pootanakit, K. (2019) Shotgun metagenomic sequencing from Manao-Pee cave, Thailand, reveals insight into the microbial community structure and its metabolic potential. *BMC Microbiol* **19**: 144.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Supplemental Fig. 1. Cave 18S rDNA ASVs Classified as Ciliophora. A Maximum Likelihood phylogeny was generated using IQTREE and the HKY base substitution model. All reference sequences were drawn from the phylum Ciliophora according to records archived in GenBank. Reference sequences are black, the cave ASVs red, ASVs collected from scooped water are marked with asterisks, and an outgroup is green. ASVs found in surface samples are enclosed by rectangles. Bootstrap values are listed at each node. The scale represents number of substitutions.

Supplemental Fig. 2. Cave 18S rDNA ASVs Classified as Rhizarians. Maximum Likelihood phylogeny generated using IQTREE and the HKY substitution model. Reference sequences are black, cave ASVs purple, and an outgroup sequence is green. Dashed lines represent unknown taxonomy within Rhizaria. Im. = phylum Imbricatae. Bootstrap values are listed at each node. The scale represents number of substitutions.

Supplemental Fig. 3. Stramenopiles identified by their similarities to archived 18S rDNA sequences. Maximum Likelihood phylogeny generated using IQTREE and HKY base substitution model. Bacillariophyta ASVs are tan and Chrysophytes red. ASVs collected from scooped water are marked with asterisks. ASVs found in surface samples are enclosed by rectangles. Bootstrap values are listed at each node. The scale represents number of substitutions.

Supplemental Fig. 4. Stramenopiles identified by their similarities to archived 23S rDNA sequences. This Maximum Likelihood phylogeny was generated using IQTREE and HKY base substitution model. ASVs found in surface samples are enclosed by rectangles. Bootstrap values are listed at each node. The scale represents number of substitutions.

Supplemental Fig. 5. Cave 18S rDNA ASVs Most Closely Related to Amoebozoan Reference Sequences. Maximum Likelihood phylogeny generated using IQTREE and the ModelFinder option. Reference sequences are black and cave ASVs orange. Bootstrap values are listed at each node. The scale represents number of substitutions.

Supplemental File 1. The list of GenBank files used to build the 18S rDNA reference trees.

Supplemental File 2. IQTREE output file for phylogenetic analysis of all nuclear 18S rDNA barcodes collected during this study.

Supplemental File 3. Eukaryotic taxonomic predictions for the cave nuclear 18S rDNA ASVs based on the PR2 database.

Supplemental File 4. IQTREE output file for Dinoflagellate nuclear 18S rDNA barcodes collected during this study.

Supplemental File 5. IQTREE output file for Ciliophora nuclear 18S rDNA barcodes collected during this study.

Supplemental File 6. IQTREE output file for Apicomplexan nuclear 18S rDNA barcodes collected during this study.

Supplemental File 7. IQTREE output file for Rhizaria nuclear 18S rDNA barcodes collected during this study.

Supplemental File 8. IQTREE output file for Stramenopile nuclear 18S rDNA barcodes collected during this study.

Supplemental File 9. IQTREE output file for plastid 23S rDNA barcodes collected during this study.

Supplemental File 10. IQTREE output file for Amoebozoan nuclear 18S rDNA barcodes collected during this study.

© 2022 The Authors. *Environmental Microbiology Reports* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,