

Annotating public fungal ITS sequences from the built environment according to the MlxS-Built Environment standard – a report from a May 23-24, 2016 workshop (Gothenburg, Sweden)

Kessy Abarenkov¹, Rachel I. Adams², Laszlo Irinyi^{3,4,5}, Ahto Agan^{6,*},
Elia Ambrosio^{1,6,7,*}, Alexandre Antonelli^{8,9,*}, Mohammad Bahram^{6,10,*},
Johan Bengtsson-Palme^{11,*}, Gunilla Bok^{12,*}, Patrik Cangren^{8,*},
Victor Coimbra^{13,*}, Claudia Coleine^{14,*}, Claes Gustafsson^{15,*}, Jinhong He^{16,*},
Tobias Hofmann^{8,*}, Erik Kristiansson^{17,*}, Ellen Larsson^{8,*}, Tomas Larsson^{18,*},
Yingkui Liu^{8,*}, Svante Martinsson^{8,*}, Wieland Meyer^{3,4,5,*}, Marina Panova^{19,*},
Nuttapon Pombubpa^{20,*}, Camila Ritter^{8,*}, Martin Ryberg^{10,*}, Sten Svantesson^{8,*},
Ruud Scharn^{21,*}, Ola Svensson^{8,*}, Mats Töpel^{18,*}, Martin Unterscher^{22,*},
Cobus Visagie^{23,24,*}, Christian Wurzbacher^{8,*}, Andy F. S. Taylor^{25,28},
Urmaz Kóljalg^{1,6}, Lynn Schriml^{26,27}, R. Henrik Nilsson⁸

1 University of Tartu, Natural History Museum, Tartu, Estonia **2** University of California, Plant & Microbial Biology, Berkeley, CA 94720, USA **3** Sydney Medical School-Westmead Hospital, Molecular Mycology Research Laboratory, Centre for Infectious Diseases and Microbiology, Sydney, Australia **4** University of Sydney, Marie Bashir Institute for Infectious Diseases and Biosecurity, Sydney, Australia **5** Westmead Institute for Medical Research, Westmead, Australia **6** Institute of Ecology and Earth Sciences, University of Tartu, Tartu, Estonia **7** Via Calamandrei 2, 53035 Monteriggioni, Siena, Italy **8** University of Gothenburg, Department of Biological and Environmental Sciences, Box 461, 405 30 Göteborg, Sweden **9** Gothenburg Botanical Garden, Carl Skottsbergs gata 22A, 413 19 Göteborg, Sweden **10** Department of Organismal Biology, Evolutionary Biology Centre, Uppsala University, SE 75236 Uppsala, Sweden **11** University of Gothenburg, Department of Infectious Diseases, Sahlgrenska Academy, Guldhedsgatan 10, 413 46 Göteborg, Sweden **12** SP Technical Research Institute of Sweden, Box 857, 501 15 Borås, Sweden **13** Universidade Federal de Pernambuco (UFPE), Departamento de Micologia, Centro de Ciências Biológicas (CCB), Av. Prof. Nelson Chaves, s/n, 50670-901 Recife, Pernambuco, Brazil **14** University of Tuscia, Department of Ecological and Biological Sciences, 01100, Viterbo, Italy **15** University of Gothenburg, Herbarium GB, Box 461, 405 30 Göteborg, Sweden **16** South China Botanical Garden, The Chinese Academy of Sciences, 723 Xingke Road, Tianhe District, Guangzhou 510650, China **17** Chalmers University of Technology, Department of Mathematical Sciences, 412 96 Göteborg, Sweden **18** University of Gothenburg, Department of Marine Sciences, Box 460, 405 30 Göteborg, Sweden **19** University of Gothenburg, Department of Marine Sciences - Tjärnö, 452 96, Strömstad, Sweden **20** University of California Riverside, Department of Plant Pathology and Microbiology and Institute of Integrative Genome Biology, Riverside, California 92521 USA **21** University of Gothenburg, Department of Earth

* Co-authors contributed equally.

Sciences, Box 460, 40530 Göteborg, Sweden **22** Ernst-Moritz-Arndt University Greifswald, Institute of Botany and Landscape Ecology, Soldmannstr. 15, 17487 Greifswald, Germany **23** Agriculture and Agri-Food Canada, Biodiversity (Mycology), 960 Carling Avenue, Ottawa, ON Canada K1A 0C6 **24** University of Ottawa, Department of Biology, 30 Marie-Curie, Ottawa, ON Canada K1N 6N5 **25** The James Hutton Institute, Craigiebuckler, Aberdeen AB15 8QH, Scotland UK **26** Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, MD, USA **27** Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD, USA **28** Institute of Biological and Environmental Sciences, Cruickshank Building, University of Aberdeen, Aberdeen AB24 3UU, UK

Corresponding author: Kessy Abarenkov (kessy.abarenkov@ut.ee)

Academic editor: A. Miller | Received 26 July 2016 | Accepted 8 September 2016 | Published 26 September 2016

Citation: Abarenkov K, Adams RI, Irinyi L, Agan A, Ambrosio E, Antonelli A, Bahram M, Bengtsson-Palme J, Bok G, Cangren P, Coimbra V, Coleine C, Gustafsson C, He J, Hofmann T, Kristiansson E, Larsson E, Larsson T, Liu Y, Martinsson S, Meyer W, Panova M, Pombubpa N, Ritter C, Ryberg M, Svantesson S, Scharn R, Svensson O, Töpel M, Unterseher M, Visagie C, Wurzbacher C, Taylor AFS, Kõljalg U, Schriml L, Nilsson RH (2016) Annotating public fungal ITS sequences from the built environment according to the MIxS-Built Environment standard – a report from a May 23–24, 2016 workshop (Gothenburg, Sweden). MycoKeys 16: 1–15. doi: 10.3897/mycokeys.16.10000

Abstract

Recent molecular studies have identified substantial fungal diversity in indoor environments. Fungi and fungal particles have been linked to a range of potentially unwanted effects in the built environment, including asthma, decay of building materials, and food spoilage. The study of the built mycobiome is hampered by a number of constraints, one of which is the poor state of the metadata annotation of fungal DNA sequences from the built environment in public databases. In order to enable precise interrogation of such data – for example, “retrieve all fungal sequences recovered from bathrooms” – a workshop was organized at the University of Gothenburg (May 23–24, 2016) to annotate public fungal barcode (ITS) sequences according to the MIxS-Built Environment annotation standard (<http://gensc.org/mixs/>). The 36 participants assembled a total of 45,488 data points from the published literature, including the addition of 8,430 instances of countries of collection from a total of 83 countries, 5,801 instances of building types, and 3,876 instances of surface-air contaminants. The results were implemented in the UNITE database for molecular identification of fungi (<http://unite.ut.ee>) and were shared with other online resources. Data obtained from human/animal pathogenic fungi will furthermore be verified on culture based metadata for subsequent inclusion in the ISHAM-ITS database (<http://its.mycologylab.org>).

Key words

Built environment, Indoor fungi, ITS, Annotation, Mycobiome

Introduction

Fungi are found throughout the biosphere, and the built environment is no exception. The taxonomic composition of indoor fungal communities tends to reflect the local outdoor communities, although the majority of fungal particles found indoors is

thought to represent spores, hyphal fragments, and other dormant and passively distributed stages (Seo et al. 2015). Although most of the fungi recovered from indoor environments would not be able to live in the built environment for any extended period of time, a minority of these species are able to cope with, and will even thrive in, the harsh conditions that the built environment presents (Hamada and Abe 2010; Nevalainen et al. 2015; Zupančič et al. 2016). These species are mainly saprotrophic, and their degree of active growth largely depends on water availability (Adams et al. 2013). They can be a serious cause of decay and other concerns in water-damaged buildings, but they are also found in buildings not subject to moisture issues – and even in buildings where very strict sanitization and filtration regimes are applied (e.g., La Duc et al. 2012; Chęcinska et al. 2015). Exposure to aerosolized fungal particles has been linked to asthma onset in humans and may furthermore play a role in eczema development and other issues in human health (Reijula et al. 2003; Knutsen et al. 2012). Indoor fungi may also contribute to other unwanted processes, such as food spoilage and wall staining (Varga et al. 2014). The built mycobiome is thus of interest to a range of scientific fields, including mycology, medicine, food biology, construction, and engineering.

Traditional, morphology-based studies of fungal spores and cultures derived from indoor sampling have recognized ca. 90 species of common indoor fungi (Flannigan et al. 2002). Efforts based on high-throughput DNA sequencing, in contrast, have revealed a vast and hitherto unknown diversity of indoor fungi. In a global study of indoor dust samples, Amend et al. (2010) using next-generation sequencing found ca. 4,500 fungal operational taxonomic units (OTUs; Blaxter et al. 2005) approximately at the species level. Similarly, another next-generation sequencing-powered study – Nonnenmann et al. (2012) – recovered 450 fungal species from 50 indoor dust samples in Yakima valley, WA (USA). Although precise species delimitation and species counts from next-generation sequencing data remain challenging (Nguyen et al. 2016), the taxonomic span of the fungal assemblages recovered in Amend et al. (2010) and Nonnenmann et al. (2012) is far larger than that occupied by the fungi traditionally thought of as common indoor fungi (cf. Flannigan et al. 2002). Thus, whereas these studies should not be used as estimates of the total number of indoor fungi, they do testify to the substantial diversity of fungi in the built environment. The lack of taxonomic reference sequences makes precise identification of many of these species problematic, and it is not unusual that a sizable proportion of the OTUs in environmental sequencing studies remain unassigned beyond the kingdom or phylum levels (e.g., Tedersoo et al. 2014; Grossart et al. 2016; Fouquier et al. 2016; Nilsson et al. 2016). There is clearly a need to generate reliable reference sequences, most notably from type material, to address this issue (cf. Schoch et al. 2014). However, the estimated number of extant species of fungi – 1.5–6 million (Hawksworth 2001; Taylor et al. 2014) – stands in stark contrast to the number of described species (~130,000 as of March 2016; www.speciesfungorum.org), and strongly suggests that molecular identification of fungi will remain challenging for the foreseeable future. In some cases, even reference barcode (nuclear

ribosomal internal transcribed spacer, ITS) sequences from type material will not be enough. Several fungal genera regularly recovered from built environment samples - such as *Aspergillus*, *Cladosporium*, *Fusarium*, and *Penicillium* - show little or no ITS variation across sets of two to several species (Bensch et al. 2012; Samson et al. 2014; Visagie et al. 2014; O'Donnell et al. 2015). Additional genetic markers are needed for robust species-level identification in these cases.

A second problem that compounds the scientific understanding of the built mycobiome has been the lack of a standardized vocabulary for sequence annotation. The International Nucleotide Sequence Database Collaboration (INSDC; Cochrane et al. 2016) holds more than 5,000 Sanger-derived fungal ITS (barcode) sequences from the built environment, but their level of metadata annotation differs widely. This unfortunately applies to most available fungal ITS sequences (cf. Nilsson et al. 2014); for example, a modest 43% are known to be annotated with something as simple and straightforward as country of collection (Tedersoo et al. 2011). In addition, where metadata exist they are not always provided in standardized and searchable formats, making precise queries difficult. There is, for instance, no straightforward way to download all fungal ITS sequences from bathrooms, or to target the substrate of gypsum board. It is reasonable to think that analysis of fungi recovered from bathrooms may prove a rewarding scientific enterprise, as indeed should be the case for fungi collected on specific building materials, under different moisture regimes, or any other particular parameter or setting. The full potential of such searches cannot presently be utilized due to the poor state of sequence annotation – primarily omitted by the original sequence authors – in the public sequence databases.

The new MIxS-Built Environment annotation standard (Glass et al. 2014; <http://gensc.org/mixs/>) addresses the need for a thorough, standardized vocabulary for microbiological analysis of the built microbiome. If all relevant fungal ITS sequences in the INSDC were annotated according to this standard, then this would open up the body of extant molecular data to detailed, precise scientific queries in the context of the built mycobiome. Going through and annotating large sequence sets is a daunting effort for any researcher, but fortunately such efforts are easy to split among a set of individual researchers. This paper presents the outcome of a sequence metadata annotation workshop (University of Gothenburg, May 23–24, 2016) to annotate the ~6,500 public fungal ITS sequences from the built environment according to the most relevant parts of the MIxS-Built Environment annotation standard. In recognition of the fact that fungi found indoors are typically found outdoors as well, the workshop also annotated closely related outdoor sequences according to basic geo-ecological parameters. The workshop was organized jointly with the UNITE and ISHAM databases (Köljalg et al. 2013; Irinyi et al. 2015). UNITE is a general-purpose sequence management environment seeking to reconcile molecular ecology and taxonomy of fungi and fungal communities. The ISHAM database centers on identification of human and animal pathogenic fungi to guide antifungal treatment choices. Both databases focus, at least for the time being, on the ITS region and share views on the importance of openness, free accessibility, and community participation.

Materials and methods

The workshop comprised 20 physical participants, mainly local Ph.D. students and post-docs – but also other researchers – in systematics and ecology. In addition, another 16 researchers participated remotely through Skype, Google Docs, and email. The participants focused on the public fungal ITS sequences of the INSDC as mirrored in the UNITE and ISHAM databases. To single out INSDC sequences associated with the built environment, we used a set of 24 keywords such as “dust”, “gypsum”, and “floor” (Suppl. material 1). Keyword matches were made to the title of the underlying publication (the INSDC field “title”), the INSDC fields “source” and “tissue type”, and the UNITE field “sequence source”. We refer to this set of sequences as the *built mycobiome set* (BMS). To single out outdoor sequences with a direct relation to the BMS, we extracted all UNITE species hypotheses with at least one BMS sequence. We then built the *outdoor mycobiome set* (OMS) from all sequences that did not match any of our keywords but that were found in the same species hypothesis as at least one BMS sequence. Sequences that initially were assigned to the BMS set, but that on closer inspection turned out not to qualify as the built mycobiome (“collected outside hospital”, for example), were transferred to the OMS set.

For each BMS sequence we tried to locate any underlying publication through the INSDC fields TITLE, JOURNAL, and PUBMED. If these were not informative, we resorted to ISI Thompson, Google/Google Scholar, and ResearchGate searches. We examined the publications for the nine items of the MIXS-Built Environment annotation standard that we felt were the most relevant and the most likely to be covered by the studies: building occupancy type, indoor space, indoor surface, surface material, surface-air contaminant, space typical state, substructure type, ventilation type, and filter type (<http://gensc.org/mixs/>). In addition we also targeted the country and host of collection and the nature of the fungus-host association (e.g., “plant: wood”, “plant: leaf”, and “human/animal: skin”), as applicable, for all sequences. We only targeted metadata and information that was clearly and unequivocally specified in the paper. A research professional (G. Bok) from a building-related technical institute was present to assist with technical, analytical, and construction-related questions in the context of the built environment. For the OMS we similarly retrieved the underlying publications and annotated the sequences to country and host of collection plus host association (as applicable, and if and when these data were missing). All results were entered into an Excel sheet for upload into UNITE and ISHAM (after culture-based verification in the case of the latter), and for sharing with other online resources.

Results

A total of 6,526 BMS and 11,574 OMS sequences from a total of 255 separate studies were annotated with at least one metadata item. A total of 45,488 annotations were made during the workshop. For example, “building occupancy type” was established for 5,801 sequences, and “ventilation type” was established for 2,235 sequences (Table 1;

Table 1. Results of the annotation workshop, specified for the built mycobiome sequence set (BMS) and the outdoor mycobiome sequence set (OMS). Countries and hosts of collections plus host association were assembled for both of these. The number of sequences processed, plus the number of underlying published and unpublished scientific studies, are also provided. For the BMS, the nine MIXS-Built Environment annotation standard items targeted at the workshop are specified in separate columns. The sequence numbers shown in the table refer to the number of sequences annotated for each data item.

	Number of sequences (annotated)	Number of studies	Country of collection	Different countries	Host of collection	Different hosts
BMS	6550 (6526)	144	2447	29	881	15
OMS	16766 (11574)	128	5983	83	5632	859
Total	23316 (18100)	255 unique	8430	83 unique	6513	865 unique
	Host association	Comments	Building occupancy type	Indoor space	Indoor surface	Surface material
BMS	764	2348	5801	1223	1207	1318
OMS	2892	1293	N/A	N/A	N/A	N/A
Total	3656	3641	5801	1223	1207	1318
	Surface-air contaminant	Space typical state	Substructure type	Ventilation type	Filter type	
BMS	3876	5618	96	2235	1874	
OMS	N/A	N/A	N/A	N/A	N/A	
Total	3876	5618	96	2235	1874	

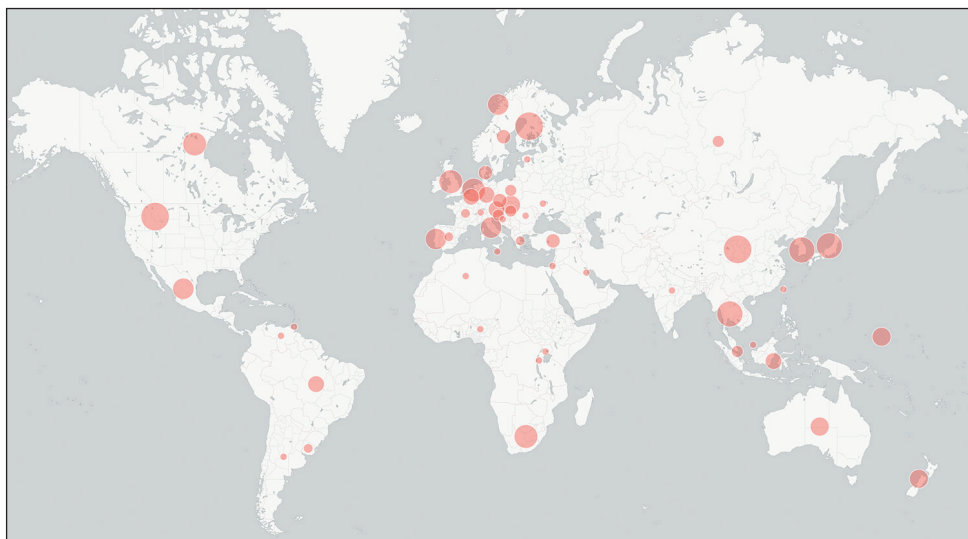


Figure 1. Analysis of the BMS sequences for country of collection. Country centroids marked with bubbles of different size on the global map indicate the number of BMS sequences originating from these countries (54 distinct countries, sequence count ranging from 1 to 2,914). For an additional 2.9% of the sequences, country information could not be restored during the workshop. The figure includes pre-existing data plus the data added during the workshop, such that these charts indicate the scientific state of ITS-based Sanger-derived sequencing of the built mycobiome as of spring 2016. Sequences that were not annotated with a single built environment-related term in the INSDC were not included in this effort, and are not represented in these charts.

Figures 1–3). The results were uploaded into UNITE via its data management system PlutoF (<https://plutof.ut.ee>; Abarenkov et al. 2010) for open query by the scientific community and was shared with the INSDC as an Excel sheet (Suppl. material 2).

Discussion

The workshop compiled a total of 45,488 metadata items, making them available for scientific query through UNITE and other venues. These metadata, although typically “published” and thus “available”, were previously not open for direct query. This highlights the wealth of relevant scientific information that lies buried in the last few decades’ worth of scientific publications – formally available, yet only available to those who know where to look, and reachable only to those with access to that literature. Fortunately, we live in a digital age where the infrastructure for recovering and sharing such information is falling into place (Martin and Martin 2010). Furthermore, there is a growing awareness of the need to annotate newly generated sequences beyond the barest minimum when these are first deposited into public sequence databases (Hyde et al. 2013; Schoch et al. 2014). Such annotations unlock significant scientific potential of those molecular data, increase the citability of the underlying scientific studies, and

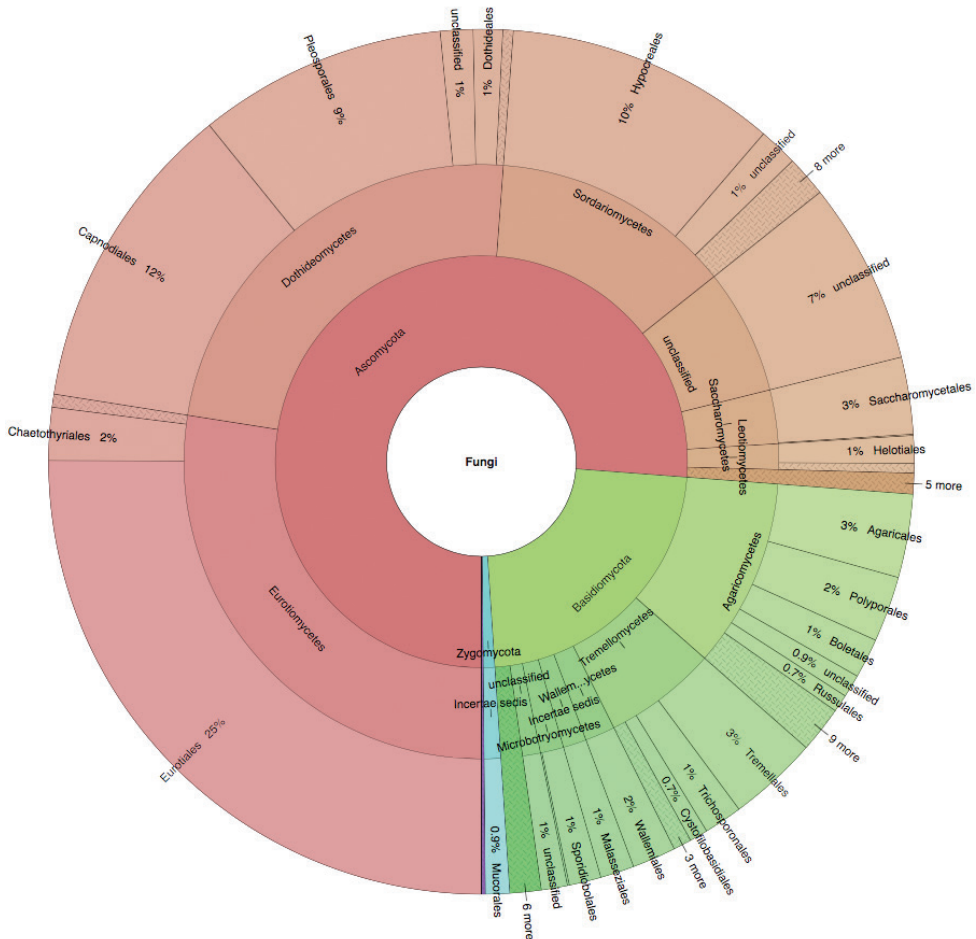


Figure 2. Krona chart of the taxonomic affiliation of the BMS sequences down to order level. The Krona chart lists all annotated BMS sequences except those classified as *Fungi* sp. (36.4%) and those of non-fungal origin (0.9%). An interactive version of the Krona chart is provided as Suppl. material 3. The figure includes pre-existing data plus the data added during the workshop, such that these charts indicate the scientific state of ITS-based Sanger-derived sequencing of the built mycobiome as of spring 2016. Sequences that were not annotated with a single built environment-related term in the INSDC were not included in this effort, and are not represented in these charts.

fulfill funding agencies' demands for openness and maximum scientific use of research funding. We certainly hope that the mycological community will be quick to embrace a more integrative approach to sequence annotation. The public sequence databases can similarly make it even easier and faster to provide such metadata upon sequence submission. We speculate that excessive time consumption is the primary reason why some sequence depositors do not annotate their sequences as well as they could have.

We managed to process nearly all BMS sequences – for which we could retrieve the underlying publication(s) – for at least one metadata item. A total of 4,985 sequences

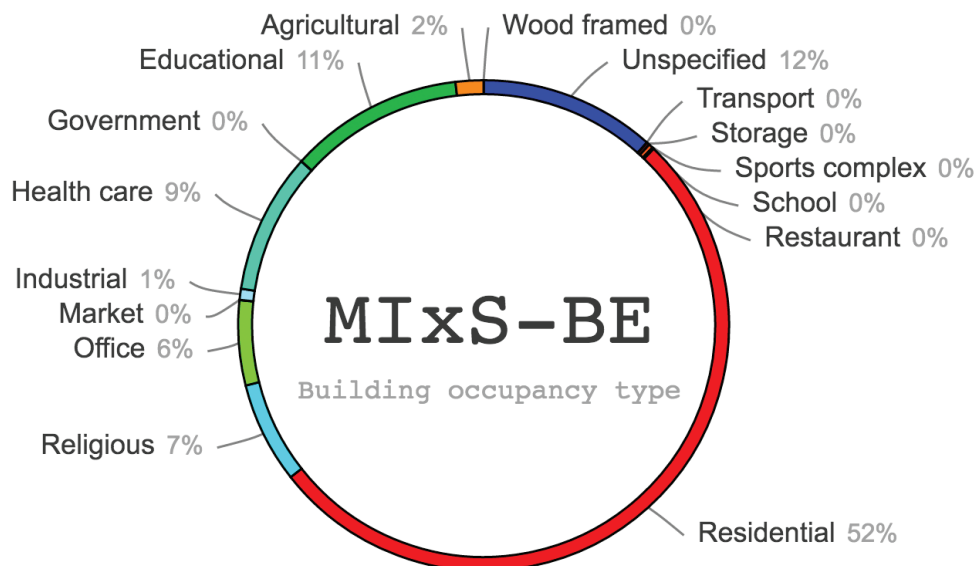


Figure 3. Analysis of the MIxS-BE “building occupancy type” (type of building where the underlying sample was taken).

were false positives – our keywords indicated them to belong to the BMS whereas in reality they did not. A sequence could stem from “outside city hospital” (keyword “hospital”), for instance. These sequences were annotated for country and host of sampling, plus the nature of the relation to the host, whenever the underlying scientific study could be retrieved and interpreted. It is reasonable to assume that our initiative suffered from a fair number of false negatives as well – sequences that should have been a part of the BMS, but that were not. Although we used no fewer than 24 keywords in our efforts to capture the built environment, we presumably missed one or more important terms in the field. We similarly missed out on all built-environment sequences that featured no relevant annotation whatsoever – perhaps just a species name and the country of origin were available. Thus, whereas we managed to do at least something about nearly all BMS sequences we recovered, we do not claim to have annotated all public fungal ITS sequences from the built environment.

The workshop identified several potential venues for amendments to the MIxS-BE standard. For example, “floor” was found to be a common place for sampling of, e.g., dust, yet the data point of “floor” could not easily be fitted into any extant MIxS-BE category. Similarly, “air” could not be represented in a straightforward way in the MIxS-BE standard (but rather applied to other packages of the MIxS standard). We also felt the need for a “laboratory” flag to indicate that a sequence stemmed from sampling in a laboratory. In addition, we were surprised by the number of fungal sequences generated from environments that must be considered to qualify as “built” or at least altered by man, but that nevertheless were difficult to fit into the present MIxS-BE categories. The examples included tombs, crypts, and mummies (Šimonovičová et

al. 2015), tumuli and other prehistoric remains (e.g., Kiyuna et al. 2011; Fernandez-Cortes et al. 2011), spacecraft (Satoh et al. 2011), and indoor historical paintings or artifacts such as the Turin shroud (López-Miras et al. 2013; Barcaccia et al. 2015). In these cases, we tried to capture the essence of the underlying sequence entries to the extent that the MIXS-BE standard allowed. We used our free-text field “Comment” to provide additional information that we felt was important with respect to future queries of these entries. These potential venues for improvements of the MIXS-BE standard have been communicated to MIXS-BE representatives from the Genomic Standards Consortium’s MIXS Compliance and Implementation working group (<http://gensc.org/mixs/mixs-compliance-and-implementation/>).

Conclusions

The present study used a workshop-style approach to accomplish a task that would have taken several months for a single researcher to accomplish. Costs were kept low by recruiting many of the participants among local Ph.D. students and postdocs in systematics and ecology, and workshop participation was made attractive by providing the opportunity to contribute to this workshop report. We can recommend this model when tackling projects of a similar kind, such as data assembly and analysis in molecular ecology and systematics. As an added benefit, the more junior participants obtain experience in scientific collaboration and communication as well as in carrying out scientific projects (cf. Ryberg et al. 2016). The workshop was funded by an Alfred P. Sloan foundation grant to improve the support for the built mycobiome in UNITE and elsewhere. Other events include a forthcoming (2017) taxonomic sequence annotation workshop and the generation and public release of sequences from type material. We invite feedback and participation in these events, and we welcome any other idea to take molecular identification of the built mycobiome to the next level.

Acknowledgements

We gratefully acknowledge financial support from the Alfred P. Sloan Foundation and the Swedish Research Council of Environment, Agricultural Sciences, and Spatial Planning (FORMAS, 215-2011-498). VRMC thanks CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico (SWE 232695/2014-8) for providing a shared Ph.D. scholarship. A.A. is funded by the Swedish Research Council (B0569601), the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013, ERC Grant Agreement n. 331024), and a Wallenberg Academy Fellowship.

References

- Abarenkov K, Tedersoo L, Nilsson RH et al. (2010) PlutoF - a web-based workbench for ecological and taxonomical research, with an online implementation for fungal ITS sequences. *Evolutionary Bioinformatics* 6: 189–196. doi: 10.4137/EBO.S6271
- Adams RI, Miletto M, Taylor JW, Bruns TD (2013) Dispersal in microbes: fungi in indoor air are dominated by outdoor air and show dispersal limitation at short distances. *The ISME Journal* 7(7): 1262–1273. doi: 10.1038/ismej.2013.28
- Amend AS, Seifert KA, Samson R, Bruns TD (2010) Indoor fungal composition is geographically patterned and more diverse in temperate zones than in the tropics. *Proceedings of the National Academy of Sciences* 107(31): 13748–3753. doi: 10.1073/pnas.1000454107
- Barcaccia G, Galla G, Achilli A, Olivieri A, Torroni A (2015) DNA analysis of dust particles sampled from the Turin Shroud. *MATEC Web of Conferences* 36: 03001. doi: 10.1051/mateconf/20153603001
- Bensch K, Braun U, Groenewald JZ, Crous PW (2012) The genus *Cladosporium*. *Studies in Mycology* 72: 1–401. doi: 10.3114/sim0003
- Blaxter M, Mann J, Chapman T, Thomas F, Whitton C, Floyd R, Abebe E (2005) Defining operational taxonomic units using DNA barcode data. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 360(1462): 1935–1943. doi: 10.1098/rstb.2005.1725
- Checinska A, Probst AJ, Vaishampayan P, White JR, Kumar D, Stepanov VG, Fox GE, Nilsson RH, Pierson DL, Perry J, Venkateswaran K (2015) Microbiomes of the dust particles collected from the International Space Station and Spacecraft Assembly Facilities. *Microbiome* 3(1): 1. doi: 10.1186/s40168-015-0116-3
- Cochrane G, Karsch-Mizrachi I, Takagi T, International Nucleotide Sequence Database Collaboration (2016) The International Nucleotide Sequence Database Collaboration. *Nucleic Acids Research* 44(D1): D48–D50. doi: 10.1093/nar/gkv1323
- Fernandez-Cortes A, Cuezva S, Sanchez-Moral S, Cañaveras JC, Porca E, Jurado V, Martin-Sanchez PM, Saiz-Jimenez C (2011) Detection of human-induced environmental disturbances in a show cave. *Environmental Science and Pollution Research* 18(6): 1037–1045. doi: 10.1007/s11356-011-0513-5
- Flannigan B, Samson RA, Miller JD (2002) *Microorganisms in home and indoor work environments*, 2nd edition. CRC Press.
- Fouquier J, Schwartz T, Kelley ST (2016) Rapid assemblage of diverse environmental fungal communities on public restroom floors. *Indoor air*: In press. doi: 10.1111/ina.12279
- Glass EM, Dribinsky Y, Yilmaz P et al. (2014) MlXs-BE: a MlXs extension defining a minimum information standard for sequence data from the built environment. *ISME Journal* 8(1): 1–3. doi: 10.1038/ismej.2013.176
- Grossart HP, Wurzbacher C, James TY, Kagami M (2016) Discovery of dark matter fungi in aquatic ecosystems demands a reappraisal of the phylogeny and ecology of zoospore fungi. *Fungal Ecology* 19: 28–38. doi: 10.1016/j.funeco.2015.06.004
- Hamada N, Abe N (2010) Comparison of fungi found in bathrooms and sinks. *Biocontrol Science* 15(2): 51–56. doi: 10.4265/bio.15.51

- Hawksworth DL (2001) The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycological Research* 105(12): 1422–1432. doi: 10.1017/S0953756201004725
- Hyde KD, Udayanga D, Manamgoda DS et al. (2013) Incorporating molecular data in fungal systematics: a guide for aspiring researchers. *Current Research in Environmental and Applied Mycology* 3(1): 1–32. doi: 10.5943/cream/3/1/1
- Irinyi L, Serena C, Garcia-Hermoso D et al. (2015) International Society of Human and Animal Mycology (ISHAM)-ITS reference DNA barcoding database -the quality controlled standard tool for routine identification of human and animal pathogenic fungi. *Medical Mycology* 53(4): 313–337. doi: 10.1093/mmy/myv008
- Kiyuna T, An KD, Kigawa R, Sano C, Miura S, Sugiyama J (2011) Molecular assessment of fungi in “black spots” that deface murals in the Takamatsuzuka and Kitora Tumuli in Japan: *Acremonium* sect. *Gliomastix* including *Acremonium tumulicola* sp. nov. and *Acremonium felinum* comb. nov. *Mycoscience* 52(1): 1–17. doi: 10.1007/s10267-010-0063-6
- Knutsen AP, Bush RK, Demain JG et al. (2012) Fungi and allergic lower respiratory tract diseases. *Journal of Allergy and Clinical Immunology* 129(2): 280–291. doi: 10.1016/j.jaci.2011.12.970
- Kóljalg U, Nilsson RH, Abarenkov K et al. (2013) Towards a unified paradigm for sequence-based identification of Fungi. *Molecular Ecology* 22(21): 5271–5277. doi: 10.1111/mec.12481
- La Duc MT, Vaishampayan P, Nilsson RH, Torok T, Venkateswaran K (2012) Pyrosequencing-derived bacterial, archaeal, and fungal diversity of spacecraft hardware destined for Mars. *Applied and Environmental Microbiology* 78(16): 5912–5922. doi: 10.1128/AEM.01435-12
- López-Miras M, Piñar G, Romero-Noguera J, Bolívar-Galiano FC, Ettenauer J, Sterflinger K, Martín-Sánchez I (2013) Microbial communities adhering to the obverse and reverse sides of an oil painting on canvas: identification and evaluation of their biodegradative potential. *Aerobiologia* 29(2): 301–314. doi: 10.1007/s10453-012-9281-z
- Martin NF, Martin F (2010) From galactic archeology to soil metagenomics—surfing on massive data streams. *New Phytologist* 185(2): 343–347. doi: 10.1111/j.1469-8137.2009.03138.x
- Nevalainen A, Täubel M, Hyvärinen A (2015) Indoor fungi: companions and contaminants. *Indoor Air* 25(2): 125–156. doi:10.1111/ina.12182
- Nguyen NP, Warnow T, Pop M, White B (2016) A perspective on 16S rRNA operational taxonomic unit clustering using sequence similarity. *npj Biofilms and Microbiomes* 2: 16004. doi: 10.1038/npjbiofilms.2016.4
- Nilsson RH, Hyde KD, Pawlowska J et al. (2014) Improving ITS sequence data for identification of plant pathogenic fungi. *Fungal Diversity* 67(1): 11–19. doi: 10.1007/s13225-014-0291-8
- Nilsson RH, Wurzbacher C, Bahram M et al. (2016) Top 50 most wanted fungi. *MycoKeys* 12: 29–40. doi: 10.3897/mycokeys.12.7553
- Nonnenmann MW, Coronado G, Thompson B, Griffith WC, Hanson JD, Vesper S, Faustman EM (2012) Utilizing pyrosequencing and quantitative PCR to characterize fungal populations among house dust samples. *Journal of Environmental Monitoring* 14(8): 2038–2043. doi: 10.1039/c2em30229b
- O'Donnell K, Ward TJ, Robert VARG, Crous PW, Geiser DM, Kang S (2015) DNA sequence-based identification of *Fusarium*: current status and future directions. *Phytoparasitica* 43(5): 583–595. doi: 10.1007/s12600-015-0484-z

- Reijula K, Leino M, Mussalo-Rauhamaa H, Nikulin M et al. (2003) IgE-mediated allergy to fungal allergens in Finland with special reference to *Alternaria alternata* and *Cladosporium herbarum*. *Annals of Allergy, Asthma & Immunology* 91(3): 280–287. doi: 10.1016/S1081-1206(10)63531-4
- Ryberg M, Kristiansson E, Wurzbacher C, Nilsson RH (2016) New Ph.D. students in the empirical sciences should be recruited into ongoing scientific studies right from the start. *Journal of Brief Ideas*. doi 10.5281/zenodo.48158
- Samson RA, Visagie CM, Houbaken J et al. (2014) Phylogeny, identification and nomenclature of the genus *Aspergillus*. *Studies in Mycology* 78: 141–173. doi: 10.1016/j.simyco.2014.07.004
- Satoh K, Nishiyama Y, Yamazaki T et al. (2011) Microbe-I: fungal biota analyses of the Japanese experimental module KIBO of the International Space Station before launch and after being in orbit for about 460 days. *Microbiology and Immunology* 55(12): 823–829. doi: 10.1111/j.1348-0421.2011.00386.x
- Schoch CL, Robbertse B, Robert V et al. (2014) Finding needles in haystacks: linking scientific names, reference specimens and molecular data for Fungi. Database, 2014, bau061. doi: 10.1093/database/bau061
- Seo S, Ji YG, Yoo Y, Kwon MH, Choung JT (2015) Submicron fungal fragments as another indoor biocontaminant in elementary schools. *Environmental Science: Processes & Impacts* 17(6): 1164–1172. doi: 10.1039/C4EM00702F
- Šimonovičová A, Kraková L, Pangallo D, Majorošová M, Piecková E, Bodoriková S, Dörnhöferová M (2015) Fungi on mummified human remains and in the indoor air in the Kuffner family crypt in Sládkovičovo (Slovakia). *International Biodeterioration & Biodegradation* 99: 157–164. doi: 10.1016/j.ibiod.2014.12.011
- Taylor DL, Hollingsworth TN, McFarland JW, Lennon NJ, Nusbaum C, Ruess RW (2014) A first comprehensive census of fungi in soil reveals both hyperdiversity and fine-scale niche partitioning. *Ecological Monographs* 84(1): 3–20. doi: 10.1890/12-1693.1
- Tedersoo L, Abarenkov K, Nilsson RH et al. (2011) Tidying up International Nucleotide Sequence Databases: ecological, geographical, and sequence quality annotation of ITS sequences of mycorrhizal fungi. *PLoS ONE* 6(9): e24940. doi: 10.1371/journal.pone.0024940
- Tedersoo L, Bahram M, Pölme S et al. (2014) Global diversity and geography of soil fungi. *Science* 346(6213): 1256688. doi: 10.1126/science.1256688
- Varga J, Kocsubé S, Szigeti G, Baranyi N et al. (2014) Occurrence of black *Aspergilli* in indoor environments of six countries. *Archives of Industrial Hygiene and Toxicology* 65(2): 219–223. doi: 10.2478/10004-1254-65-2014-2450
- Visagie CM, Houbaken J, Frisvad JC, Hong SB, Klaassen CHW, Perrone G, Seifert KA, Varga J, Samson RA (2014) Identification and nomenclature of the genus *Penicillium*. *Studies in Mycology* 78: 343–371. doi: 10.1016/j.simyco.2014.09.001
- Zupančič J, Babič MN, Zalar P, Gunde-Cimerman N (2016) The black yeast *Exophiala dermatitidis* and other selected opportunistic human fungal pathogens spread from dishwashers to kitchens. *PLoS ONE* 11(2): e0148166. doi: 10.1371/journal.pone.0148166

Supplementary material 1

Keywords used to identify fungal sequences from the built environment in the INSDC

Authors: Kessy Abarenkov Rachel I. Adams Irinyi Laszlo Ahto Agan, Elia Ambrosio, Alexandre Antonelli, Mohammad Bahram, Johan Bengtsson-Palme, Gunilla Bok, Patrik Cangren, Victor Coimbra, Claudia Coleine, Claes Gustafsson, Jinhong He, Tobias Hofmann, Erik Kristiansson, Ellen Larsson, Tomas Larsson, Yingkui Liu, Svante Martinsson, Wieland Meyer, Marina Panova, Nuttapon Pombubpa, Camila Ritter, Martin Ryberg, Sten Svantesson, Ruud Scharn, Ola Svensson, Mats Töpel, Martin Unterseher, Cobus Visagie, Christian Wurzbacher, Andy F.S. Taylor Urmas Köljalg Lynn Schriml R. Henrik Nilsson

Data type: text

Explanation note: Keywords used to identify fungal sequences from the built environment in the INSDC.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 2

Annotations made during the workshop

Authors: Kessy Abarenkov Rachel I. Adams Irinyi Laszlo Ahto Agan, Elia Ambrosio, Alexandre Antonelli, Mohammad Bahram, Johan Bengtsson-Palme, Gunilla Bok, Patrik Cangren, Victor Coimbra, Claudia Coleine, Claes Gustafsson, Jinhong He, Tobias Hofmann, Erik Kristiansson, Ellen Larsson, Tomas Larsson, Yingkui Liu, Svante Martinsson, Wieland Meyer, Marina Panova, Nuttapon Pombubpa, Camila Ritter, Martin Ryberg, Sten Svantesson, Ruud Scharn, Ola Svensson, Mats Töpel, Martin Unterseher, Cobus Visagie, Christian Wurzbacher, Andy F.S. Taylor Urmas Köljalg Lynn Schriml R. Henrik Nilsson

Data type: metadata

Explanation note: The annotations made during the workshop shown with original INSDC data. For the BMS, we targeted nine MIXS-BE items plus country of collection, host of collection, host association, and a general “Comment” field. For the OMS, we targeted country of collection, host of collection, host association, and a general “Comment” field.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 3

Krona chart

Authors: Kessy Abarenkov Rachel I. Adams Irinyi Laszlo Ahto Agan, Elia Ambrosio, Alexandre Antonelli, Mohammad Bahram, Johan Bengtsson-Palme, Gunilla Bok, Patrik Cangren, Victor Coimbra, Claudia Coleine, Claes Gustafsson, Jinhong He, Tobias Hofmann, Erik Kristiansson, Ellen Larsson, Tomas Larsson, Yingkui Liu, Svante Martinsson, Wieland Meyer, Marina Panova, Nuttapon Pombubpa, Camila Ritter, Martin Ryberg, Sten Svantesson, Ruud Scharn, Ola Svensson, Mats Töpel, Martin Unterseher, Cobus Visagie, Christian Wurzbacher, Andy F.S. Taylor Urmas Köljalg Lynn Schriml R. Henrik Nilsson

Data type: html

Explanation note: Interactive Krona chart for visualizing the taxonomic distribution of annotated BMS sequences down to order level. Sequences classified as Fungi sp. (36.4%) or non-fungal (0.9%) were excluded from this dataset.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

High habitat-specificity in fungal communities in oligo-mesotrophic, temperate Lake Stechlin (North-East Germany)

Christian Wurzbacher^{1,2,3,4}, Norman Warthmann^{1,2,5}, Elizabeth C. Bourne^{1,2},
Katrin Attermeyer^{3,6}, Martin Allgaier^{2,3}, Jeff R. Powell⁷, Harald Detering^{1,2,8},
Susan Mbedi^{2,9}, Hans-Peter Grossart^{3,10}, Michael T. Monaghan^{1,2}

1 Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Department of Ecosystem Research, Berlin, Germany **2** Berlin Center for Genomics in Biodiversity Research, Berlin, Germany **3** Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Department of Experimental Limnology, Stechlin, Germany **4** University of Gothenburg, Department of Biological and Environmental Sciences, Gothenburg, Sweden **5** The Australian National University, Research School of Biology, Division of Plant Sciences, Canberra, ACT, Australia **6** Uppsala University, Department of Ecology and Genetics, Limnology, Uppsala, Sweden **7** Hawkesbury Institute for the Environment, Western Sydney University, Penrith, NSW, Australia **8** University of Vigo, Department of Biochemistry, Genetics and Immunology, Vigo, Spain **9** Museum für Naturkunde Leibniz-Institute for Evolution and Biodiversity Science, Berlin, Germany **10** Potsdam University, Institute for Biochemistry and Biology, Potsdam, Germany

Corresponding author: Christian Wurzbacher (christian@wurzbacher.cc)

Academic editor: K. Voigt | Received 22 June 2016 | Accepted 7 October 2016 | Published 21 October 2016

Citation: Wurzbacher C, Warthmann N, Bourne EC, Attermeyer K, Allgaier M, Powell JR, Detering H, Mbedi S, Grossart H-P, Monaghan MT (2016) High habitat-specificity in fungal communities in oligo-mesotrophic, temperate Lake Stechlin (North-East Germany). MycoKeys 16: 17–44. doi: 10.3897/mycokeys.16.9646

Abstract

Freshwater fungi are a poorly studied ecological group that includes a high taxonomic diversity. Most studies on aquatic fungal diversity have focused on single habitats, thus the linkage between habitat heterogeneity and fungal diversity remains largely unexplored. We took 216 samples from 54 locations representing eight different habitats in the meso-oligotrophic, temperate Lake Stechlin in North-East Germany. These included the pelagic and littoral water column, sediments, and biotic substrates. We performed high throughput sequencing using the Roche 454 platform, employing a universal eukaryotic marker region within the large ribosomal subunit (LSU) to compare fungal diversity, community structure, and species turnover among habitats. Our analysis recovered 1027 fungal OTUs (97% sequence similarity). Richness estimates were highest in the sediment, biofilms, and benthic samples (189–231 OTUs), intermediate in water samples (42–85 OTUs), and lowest in plankton samples (8 OTUs). NMDS grouped the eight studied habitats into six clusters, indicating that community composition was strongly influenced

by turnover among habitats. Fungal communities exhibited changes at the phylum and order levels along three different substrate categories from littoral to pelagic habitats. The large majority of OTUs (> 75%) could not be classified below the order level due to the lack of aquatic fungal entries in public sequence databases. Our study provides a first estimate of lake-wide fungal diversity and highlights the important contribution of habitat heterogeneity to overall diversity and community composition. Habitat diversity should be considered in any sampling strategy aiming to assess the fungal diversity of a water body.

Key words

Freshwater fungi, aquatic fungi, metabarcoding, LSU, GMYC, habitat specificity, Chytridiomycota, Cryptomycota, Rozellomycota, community ecology, lake ecosystem, biofilm, sediment, plankton, water sample, benthos, reed, fungal diversity

Introduction

Aquatic fungi play an important role in the cycling of carbon and nutrients in ecosystems (Gleason et al. 2008; Wurzbacher et al. 2010; Jobard et al. 2010; Grossart and Rojas-Jimenez 2016). Fungi may be involved in many stages of nutrient cycling, but can also be quite specific in their ecological functions. The degradation of recalcitrant plant, algal and animal residues may be carried out by a number of poorly known groups within the phyla Chytridiomycota and Rozellomycota (Corsaro et al. 2014, syn. Cryptomycota; Jones et al. 2011), and by ecological groups of aquatic hyphomycetes and yeasts (reviewed by Wurzbacher et al. 2010; Jobard et al. 2010). Parasitism by Chytridiomycota species facilitates the trophic transfer of nutrients from otherwise inedible phytoplankton to filter-feeding zooplankton (termed the “mycoloop”; Kagami et al. 2007, 2014). Aquatic fungi also form symbiotic relationships, such as endophytic or mycorrhiza-forming fungi (Kohout et al. 2012) or Chytridiomycota symbioses with algae (Picard et al. 2013). Despite their important functional role in lakes, the biodiversity of freshwater fungi remains poorly known.

Estimates of total fungal diversity currently range from 1.5–3 M species worldwide (Hawksworth 2012). Of these, roughly 100,000 species are described, with only ca. 3000 of these from aquatic habitats (Shearer et al. 2007; Tsui et al. 2016). The low diversity of aquatic compared to terrestrial (e.g., soil) fungi partly results from the fact that mycological studies in aquatic systems remain rare. Apart from a few well studied lotic ecosystems and wetlands (Wong et al. 1998; Shearer et al. 2007; Gulis et al. 2009; Krauss et al. 2011), the total diversity of aquatic fungi has not been linked to habitat heterogeneity. Most studies in freshwaters have focussed on marshlands (reviewed in Kuehn 2008) and examined the open water, leaf litter or emergent macrophytes (e.g., *Typha*, *Phragmites*). Studies in lakes have often concentrated on seasonal patterns in the water column (e.g., van Donk and Ringelberg 1983; Holfeld 1998; Lefèvre et al. 2012; Rasconi et al. 2012) or have compared different lakes (e.g. Zhao et al. 2011; Lefèvre et al. 2012; Taib et al. 2013). Several studies have found evidence for vertical and horizontal structuring of fungal communities in the water column (Lefèvre et al. 2007; Chen et al. 2008; Lepère et al. 2010), suggesting that there is an important spatial component of diversity. A recent

meta-analysis of global diversity found that aquatic fungi clustered in habitat-specific biomes, with freshwater biomes having the highest diversity at the phylum level (Panzer et al. 2015). The authors attributed this to the high substrate diversity and temporal dynamics of environmental parameters in freshwater ecosystems.

Considering the multitude of available niches and fungal lifestyles in aquatic habitats (Karling et al. 1977; Wurzbacher et al. 2010), the actual species number of aquatic fungi is likely to be much higher than what is currently recognized. Freshwater systems contain a great diversity of habitats including the boundaries that connect them to terrestrial and groundwater ecosystems (Vadeboncoeur et al. 2002; Schindler and Scheuerell 2002). Temperate, stratified lakes encompass horizontal gradients from shallow (littoral zone) to open water (pelagic zone) habitats, as well as vertical gradients from the surface associated epilimnion (often photic, light) to the deeper hypolimnion (often aphotic, dark) and the sediment. Shore regions are transition zones between terrestrial and aquatic habitats, and include biogeochemical gradients and macrostructures such as aquatic macrophytes, animals, plant debris and biofilms. These shore regions may thus be “hot spots” of aquatic, amphibious and terrestrial fungal diversity (Wurzbacher et al. 2010). In contrast, pelagic habitats have little or no macrostructure, and pelagic fungi may be limited to planktonic substrates such as dissolved organic matter (DOM), phytoplankton and zooplankton (living or dead). In particular, accompanying the change of substrate from coarse particulate organic matter (CPOM) near the edges of the lake to fine particulate organic matter (FPOM) in the open water, filamentous Dikarya are expected to be replaced by less abundant single celled yeasts and flagellated Chytridiomycota (Wurzbacher et al. 2010). We hypothesize that such a change in “fungal morphotypes” to unicellular fungi is linked to a change in the abundance and size of substrates present in the various lake habitats.

We examined the fungal diversity of a temperate lake in North-East Germany (Lake Stechlin) using a high throughput sequencing and metabarcoding approach. Our first aim was to examine the effect of habitat specificity on the fungal community by measuring the extent to which different habitat types contained similar communities, or whether there was a pronounced taxa turnover among habitats. Our second aim was to test the morphotype hypothesis, specifically whether fungal groups present were related to the availability of major types of particulate organic matter (POM). We expected that the broad diversity of substrate size and structures sampled (e.g. plankton, macrophytes) would reveal a more heterogeneous fungal community than previously detected by traditional lake sampling strategies.

Methods

Sampling site

Lake Stechlin is a deep (maximum depth: 69.5 m), oligo-mesotrophic, dimictic hard-water lake in North-East Germany (53°10'N; 13°02'E). It has a surface area of 4.25 km² and is divided into three distinct basins (Figure 1). The lake has a littoral reed

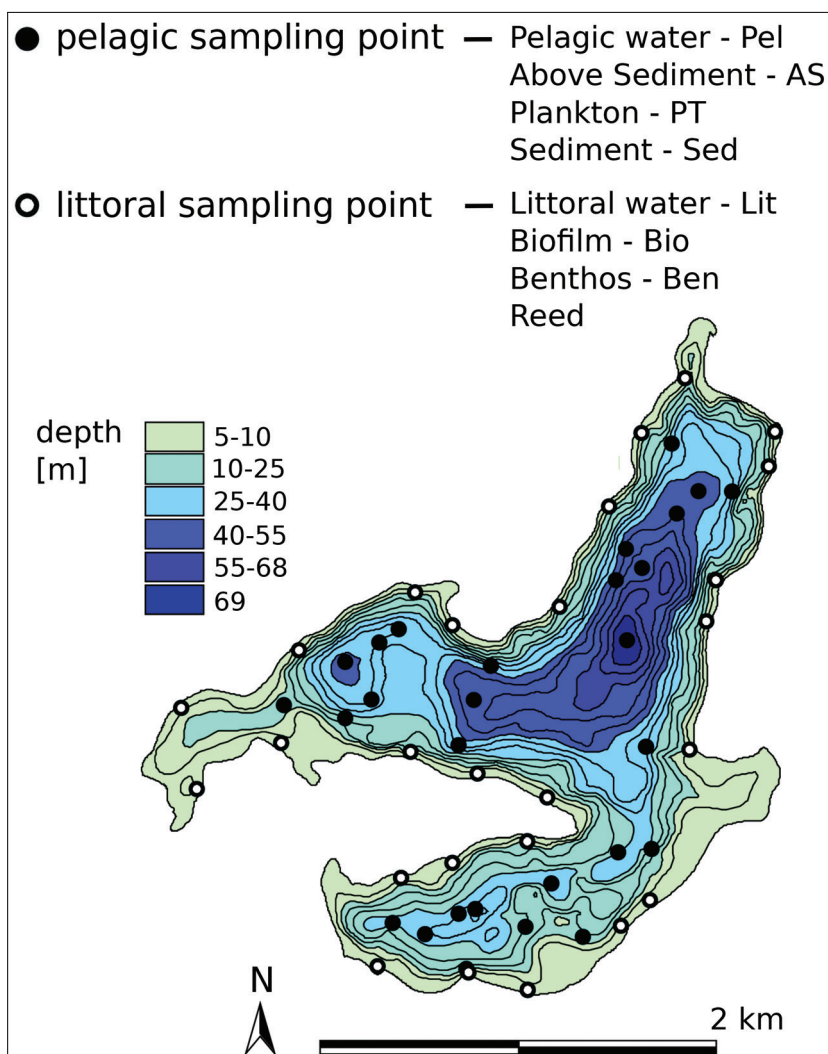


Figure 1. Sampling sites in Lake Stechlin. Integrated water samples, above-sediment water, plankton (> 55 μm), and sediment were taken from pelagic locations. Surface water samples, reed plants (*Phragmites australis*), biofilm samples (from stone, wood and macrophytes) and benthic samples (detritus, macrozoobenthos) were taken from littoral locations.

belt of *Phragmites australis* that is interspersed with areas of underwater macrophytes (mainly *Characea*). It is surrounded by mixed forest dominated by *Pinus sylvestris* and *Fagus sylvatica*. Lake Stechlin is part of the global lake ecological observatory network (GLEON) and has been monitored since 1959 (Casper 1985). Of the many publications from Lake Stechlin, few have examined the fungi (Casper 1965; Luo et al. 2011; Wurzbacher et al. 2014). This study thus represents the first attempt to characterize Lake Stechlin's mycobiota. During the course of our field sampling (April–June 2010),

Table 1. Overview of the total abundance of eukaryotic sequences and OTUs (97% similarity clustering) recovered for each lake habitat. Fungal contribution (reads, OTUs) was calculated as median percentage with standard deviation. Habitat abbreviations used in Figs 2 and 3 are indicated to the right of habitat names. POM mainly consisted of three types: fine (FPOM), coarse particulate organic matter (CPOM), or a mixture of both (MIX). Reported values were obtained from analysis of 3 samples in each habitat, except for Reed habitat which had only 1 sample. Each of these samples contained pooled DNA from three time points (April – June 2010). The shared Chao OTU richness is given as a range between a conservative and a non-conservative estimate (see Method section).

	Habitat	POM type	n	Total reads/ total OTU	Fungal reads % \pm SD	Fungal OTU % \pm SD	Shared Chao (total)	Shared Chao (fungal)
<i>Pelagic</i>	Pelagic water - Pel	FPOM	3	19795/470	2.4 \pm 0.4	5.8 \pm 0.6	303–711	42–107
	Above Sediment -AS	FPOM	3	18701/592	2.2 \pm 0.3	7.7 \pm 2.5	467–1005	85–148
	Plankton - PT	FPOM	3	17902/165	0.4 \pm 0.2	3.7 \pm 0.5	151–294	8–8*
	Sediment - Sed	MIX	3	79190/1759	5.1 \pm 1.2	23.5 \pm 0.9	612–3219	189–853
<i>Littoral</i>	Littoral water - Lit	FPOM	3	29128/521	1.5 \pm 0.2	6.8 \pm 1.0	363–883	54–116
	Biofilm - Bio	MIX	3	26302/1053	11.0 \pm 6.8	25.1 \pm 2.6	541–1971	225–640
	Benthos - Ben	CPOM	3	59169/1070	63.1 \pm 13.1	30.9 \pm 8.4	580–2280	231–1323
	Reed	CPOM	1	4277/179	88.7	35.8	n.a.	n.a.

n.a. (non applicable)

* Chao estimate may be not reliable for small sample sizes

the phytoplankton community was dominated by diatoms and by filamentous cyanobacteria (*Dolichospermum flos-aquae*). The nutrient status of the lake during the sampling period is detailed in the Suppl. material 1.

Sampling

We sampled eight different habitat types (Table 1) at three time points encompassing spring and early summer 2010 (8–9 April; 11–12 May; and 9–10 June). Sampling was carried out relatively early in the year to avoid an over-representation of wood-degrading Basidiomycetes that are introduced as airborne spores from the surrounding forest between July to November (personal observation). Our sampling scheme was designed to cover for both pelagic (defined here as areas with > 20 m depth and > 100 m from shore) and littoral (< 10 m from shore) habitats (Figure 1). Habitats were defined as follows: “Pelagic” samples consisted of a 1 litre water sample integrated from three depths: 1 m below the surface, at mid-depth, and at 2–3 m above the sediment. These were collected using a Niskin-type water sampler (Hydro-Bios, Germany); “Plankton” was obtained from an integrated sample (surface to 2–3 m above sediment) from a plankton net (55 μ m mesh; Hydro-Bios, Germany); “Above Sediment” was a water sample from 0–20 cm above the sediment that was retrieved together with “Sediment”, which itself comprised 1 ml of the uppermost cm of the core, using a sediment corer (6 cm diameter; Uwitec, Austria). “Littoral” samples consisted of a 1 litre water sample taken from 0.5–1 m depth in the

littoral zone; “Reed” samples were taken from aerial, submerged, and rhizosphere parts of reed plants, following the physical removal of biofilm; “Biofilm” samples were taken from stones, woody debris, and reed stems (removed using a scalpel); and “Benthos” consisted of detritus and zoobenthos sampled from the littoral zone using a sediment grabber (Ekman-Birge bottom sampler, Hydrobios, Germany). Each of the eight habitat types was sampled at 3 locations in each of the 3 basins at each of the 3 time points ($n = 27$), for a total of 216 samples. Samples were pooled by combining one sample from each of the three basins, resulting in 3 representative samples of each habitat per time point. These were further pooled for sequencing analysis (see below). Water samples were filtered on a 0.22 μm Sterivex filter (Millipore, USA), plankton-net samples were filtered onto a 12 μm cellulose acetate filter (Sartorius AG, Germany), and 1 ml of sediment was transferred to a cryotube for storage. All samples and filters were stored at $-80\text{ }^{\circ}\text{C}$ until further processing. We categorized the typically predominant POM type for each habitat. Water samples were FPOM dominated and Reed and Benthos habitats were CPOM dominated, while Sediment and Biofilm was classified as a mixture of both POM types (Table 1).

DNA extraction

Total DNA was extracted using the Power Soil kit (MoBio Laboratories, Carlsbad, USA) for Sediment samples; the Qiagen Plant kit (Qiagen, Hilden, Germany) for Reed, Biofilm, and Benthos samples; and the Qiagen Blood & Tissue kit for Littoral, Pelagic, Above Sediment and Plankton samples. Manufacturers’ instructions were followed with the following modifications: Reed and Benthos samples were homogenized with a mill (Pulverisette 9, rpm = self-optimize speed, 20 sec, Fritsch, Germany) and all other samples were subjected to a bead-beating step prior to extraction (MMX400, $2 \times 2\text{ min}$, $f = 30\text{ sec}^{-1}$, Retsch, Germany). We added 20 μl Proteinase K (Qiagen, Netherlands) to the lysis buffer for Sediment, Reed, Biofilm, and Benthos samples, and incubated these for 1 h at $56\text{ }^{\circ}\text{C}$. DNA concentrations were measured using a PicoGreen assay (Invitrogen, USA). Approximately 20 ng of DNA was used as template for PCR.

Library preparation for pyrosequencing

DNA metabarcoding was carried out on all samples using the D1/D2 variable region of the ribosomal LSU with the eukaryotic primers NLF184cw (TACCCGCT-GAAYTTAAGCATAT; modified from Van der Auwera et al. 1994) and Euk573rev (AGACTCCTTGGTCCRTGT; modified from NLR818, Van der Auwera et al. 1994). After *in silico* tests using TestPrime (Klindworth et al. 2012) we found the primer pair covered 84% of all eukaryotes deposited in the SILVA database (LSU r123 version) when allowing for two mismatches, neither of which was in the last 3 bp of the 3’ region. The primer pair potentially excludes single eukaryotic lineages within Amoebozoa, Excavata, Cercozoa. Within fungi it covers 93.4% of deposited sequences in all phyla, except Microsporidia. Oomycetes were covered at 76%. Among the fungal phyla, the

lowest coverage was 85% for Basidiomycota, followed by Zygomycota with 93%. Primers were modified with 5' sequencing adaptors (extended primer list in Suppl. material 2), consisting of barcodes recommended by Roche and Lennon et al. (2010) and Lib-L adapters (Roche). PCR was conducted with AccuPrime Taq Polymerase High Fidelity (Invitrogen, USA) in a 40 μ l reaction with the following conditions: initial denaturation for 3 min at 98 °C followed by 32 cycles of 1 min denaturation at 94 °C and 2 min annealing/elongation at 60 °C. The quality and intensity of the amplicons were checked on an agarose gel to ensure semi-quantitative assumptions (Lindahl et al. 2013). PCR amplicons were purified using AMPure XP Beads (Beckman Coulter) and quality was verified by microfluidics electrophoresis (Bioanalyzer, Agilent). The 9 PCR products per habitat (3 replicates per sampling time) were then pooled into three final replicates for sequencing, each of which contained all 3 time points. As a result, the sequencing triplicates were representative for the habitat biota within the sampled timespan of April-June. Pooling also helps to ameliorate PCR bias and template stochasticity. We sequenced only one of the triplicates of the reed habitat. Afterwards all amplicons were pooled equimolar for emulsion PCR and subjected to pyrosequencing library preparation and sequencing following the manufacturer's recommendations (Lib-L, FLX titanium chemistry, Roche, Switzerland). The sequence data was deposited at ENA (<http://www.ebi.ac.uk/ena>) under following accession number: PRJEB14236.

Sequence data processing

Sequences were processed as briefly outlined in Suppl. material 3. Raw 454 sequencing data were transformed by coding any nucleotide with a Phred score < 11 as N. We removed all reads shorter than 300 nt and trimmed reads with trailing Ns. The D1 region is highly variable and has a pronounced length polymorphism, which renders an accurate alignment difficult. We therefore defined an end position to serve as an alignment anchor by screening the SILVA reference database (v123) for a conservative eukaryotic region located within our amplicon. We identified a conserved 42-nt sequence (GAG-NCCGATAGNNNACAAGTANNGNGANNGAAAGWTGNAAAG) located after the D1 region as being suitable to serve as a stable 3' end for the alignment by using the probe design tool of ARB (Ludwig et al. 2004). We subsequently clipped all filtered reads (fastq format) after the last nucleotide using Shore oligo-match: a sequence context-aware clipping tool (Ossowski et al. 2008). This normalized the length of the reads to a fixed position in a global alignment (average read length: 360 ± 13 , $n = 596k$). We allowed for mismatches by scoring each match with 3, mismatches with -1, and gaps with -4. The threshold for clipping was set to $\text{score}_{\text{MAX}} > 0.5$ and the effect on the size-frequency distribution can be found in Suppl. material 3. Unclipped sequences were rejected and analysed separately (Suppl. material 3).

Clipped reads were processed in Mothur following 4-5-4 SOP (Schloss et al. 2009, accessed in August 2012). Quality filtering was achieved by using the sliding-window option (quality threshold of 25). For the alignment-based procedure, we constructed a reference dataset with long, high-quality reads processed with pyrotag-

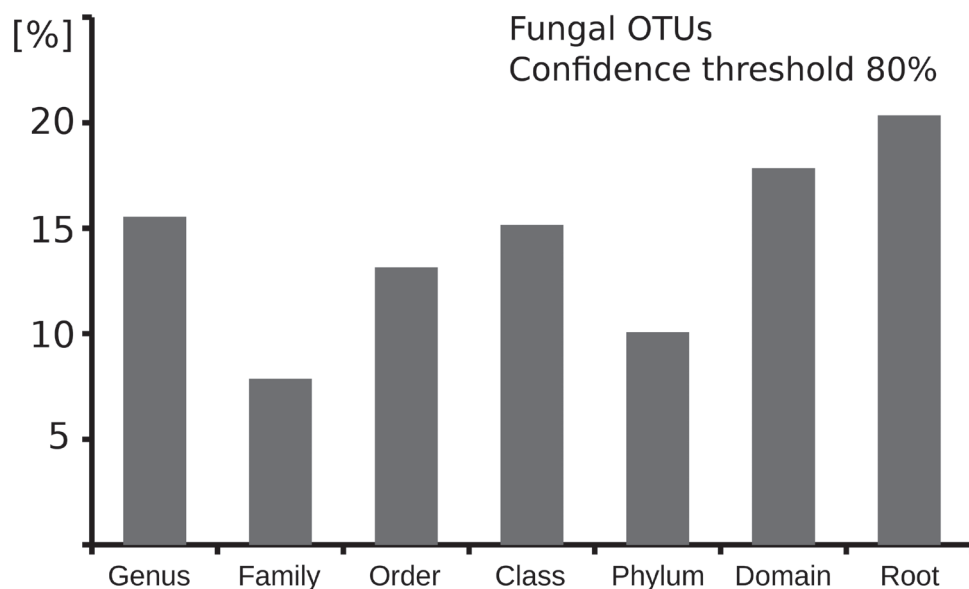


Figure 2. Classification of fungal OTUs by the RDP classifier to the different taxonomic levels as a percentage. Classification cutoff was a confidence level of 0.8.

ger (<http://pyrotagger.jgi-psf.org>) using a cutoff at > 500 nt. These were aligned to the eukaryotic backbone provided by the SILVA database LSURef (version 111; www.arb-silva.de) using the SINA aligner (Pruesse et al. 2012). This reference alignment was used to align our reads in Mothur. After clustering at 97% sequence similarity (average neighbour algorithm), the OTU abundance matrix was imported into R (www.r-project.org, version 3.3.1.) for further analysis (see below). As a comparison to a fixed 97% sequence similarity cutoff, we employed a coalescent-based clustering analysis as implemented by the gmyc model (Powell et al. 2011; Fujisawa and Barraclough 2013) with an UPGMA tree for the early diverging lineages (456 OTUs). OTUs were classified by the RDP classifier using the RDP fungal LSU training dataset with a confidence level of 80% (version 11; Liu et al. 2011; Figure 2; see Suppl. material 4 for classifications).

Habitat richness and statistics

OTU count was positively correlated with read count (Pearson's $r = 0.90$), thus we avoided single sample based richness estimates. For richness estimates of the habitats we applied a shared corrected Chao index (Chiu et al. 2014). More specifically, we provided a range for the Chao estimates based on a lower conservative OTU filtering and an upper overestimate based on unfiltered OTUs. In the former case all OTUs that occurred only in one sample of the dataset (independent of the absolute OTU

frequency) were removed while in the latter case all OTUs including singletons were kept. Supporting rarefaction curves displaying the sampling effort for Eukarya and Fungi based on a singleton filtered OTU matrix are provided in the Suppl. material 5).

POM and habitat types were compared by employing parametric statistics (ANOVA and Tukey honest significant difference) on “logit” transformed proportional read (read counts) or OTU (OTU counts) data. Plankton samples were excluded due to their skewing effect on the distribution caused by their low fungal proportion (including Plankton samples will still lead to a significant Kruskal-Wallis test, $p < 0.001$, but renders PostHoc tests difficult to apply). Normality and homogeneity of variances were confirmed by Shapiro-Wilk tests and Levene’s Tests, respectively.

Multivariate analyses

For all subsequent β -diversity analyses, an OTU table without singletons was used to account for noise in the data (e.g., Reeder and Knight 2009). Differences among habitats within the fungal sub-community were examined with a non-metric multidimensional scaling (NMDS) ordination plot based on the Cao distance (Cao et al. 1997), which accounts for variable sampling intensity. Ellipses correspond to the standard deviation around the habitat group centroids. Stress values below 0.1 can be considered as a very good fit. We additionally tested for significance when separating habitats (excluding Reed) and POM type and using a PERMANOVA (1000 permutations) on the distance matrices. The robustness of the results were evaluated by comparing them with a presence/absence transformed OTU matrix using Jaccard distances, as well as with a classified taxonomy abundance table generated with SILVA NGS (see below) using Cao distance. Both additional analyses resulted in similar outcomes (Suppl. material 6).

Alternative sequence data processing

As an alternative to the OTU-based RDP classification of our sequences, we performed two additional analyses with the aim to gain resolution for the taxonomic classifications of our sequences (see also Suppl. material 3). First, clipped sequences were demultiplexed and quality trimmed in Mothur as described above and then submitted to SILVA NGS (www.arb-silva.de/ngs/) (Quast et al. 2013) for classification at the minimum similarity level of 85% against the LSU reference database (version 123). This resulted in 57 fungal taxonomic paths (unique taxonomic names, hierarchical, see Quast et al. 2013; Suppl. material 4). Second, we performed an analysis, in which we pooled all clipped sequences of one habitat and then subjected these to a blast search (Blast+) against the nt database (GenBank, accessed January 2015) for eukaryotes. Sequences were then classified using the LCA classifier implemented in Megan5 (Huson et al. 2011) using the following parameters: Min. Score = 100,

Max. Expected = 0.01, Top % = 5.0, Min. Support % = 0.01, Min. Support = 2, LCA = 75 %, Min. Complexity = 0. Habitats were compared based on square root normalization (Suppl. material 7).

Phylogenetic inference

We recovered several Neocallimastigales (rumen fungi) sequences, as classified by SILVA NGS and also by the RDP classifier with a low probability (< 55%). We took 26 representative sequences and constructed a phylogenetic tree with an extension of the reference dataset from James et al. (2006) in order to confirm or reject this potential classification. A matrix was aligned using the SINA aligner (Pruesse et al. 2012) followed by manual inspection. For tree reconstruction, we used MrBayes (v3.2.6; Ronquist et al. 2012) with 10 million generations and an “invgamma” model.

Results

A total of 54 sampling stations, representing eight habitat types, were sampled at three time points in spring of 2010 and analysed using pyrosequencing of the large ribosomal subunit (LSU) as a universal eukaryotic marker. Across all habitats, the total number of eukaryotic OTUs was 3695, as estimated using alignment-based clustering at 97%, 47% of which were singletons. The lower limit of shared estimated OTUs (shared corrected Chao index) varied considerably among habitats, with the highest values found in Sediment, Benthos and Biofilm habitats, and the lowest in Plankton and water samples (Table 1). Of the total OTUs, 1027 (27%) were classified as fungi by RDP (48% of which were singletons). The gmyc method of OTU delimitation for the non-Dikarya taxa (mainly aquatic lineages that comprised 52% of the fungal OTUs in our data) resulted in 65% of OTUs with more than one occurrence, compared with 68% of units defined by gmyc clustering. The ability of each type of taxonomic unit to predict habitat (97%: adj. $r^2 = 0.59$; gmyc: adj. $r^2 = 0.61$) was very similar and we thus decided, to hereafter use the more conservative OTUs based on the 97% criterion. The shared lower estimated OTUs were following similar trends as for all eukaryotes with Benthos and Biofilm ranking highest (231 and 225 estimated OTUs, respectively) and markedly lower ranks for the water samples (42–85 estimated OTUs) and only 8 estimated OTUs for Plankton samples (Table 1). Both, the fungal proportions (fungal reads) as well as the proportional fungal diversity (fungal OTUs) were significantly different for each POM type (reads: ANOVA, $F = 104.4$, $p = 6.6 \cdot 10^{-10}$, Tukey Post Hoc Test $p < 0.001$ (Figure 3); OTUs: ANOVA, $F = 132.9$, $p = 1.08 \cdot 10^{-10}$, Tukey Post Hoc Test $p < 0.01$).

Fungal community composition was significantly structured into different habitats according to the NMDS clustering of OTUs (Figure 4, stress = 0.08; PERMANOVA, $r^2 = 0.71$, $p < 0.001$) and POM types (PERMANOVA, $r^2 = 0.36$, $p < 0.001$). The three wa-

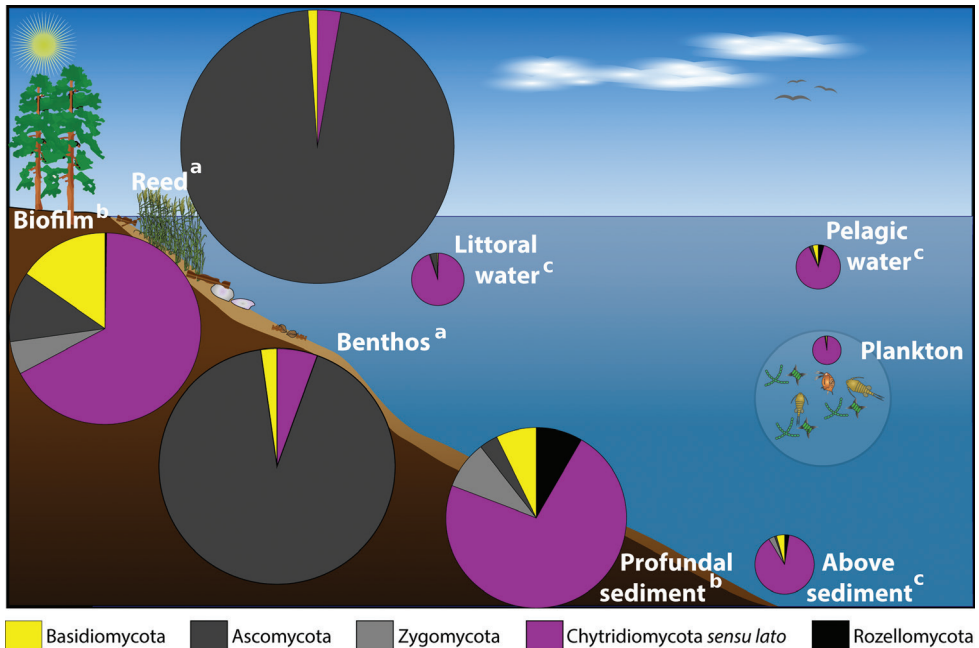


Figure 3. Contribution of fungal phyla within each of the lakes habitat. Pie chart diameters are scaled by the contribution of fungal diversity to the total eukaryotic diversity in each habitat. Lower case letters group significantly different POM types in terms of OTU diversity and fungal reads (Tukey Post-Hoc tests, $p < 0.01$).

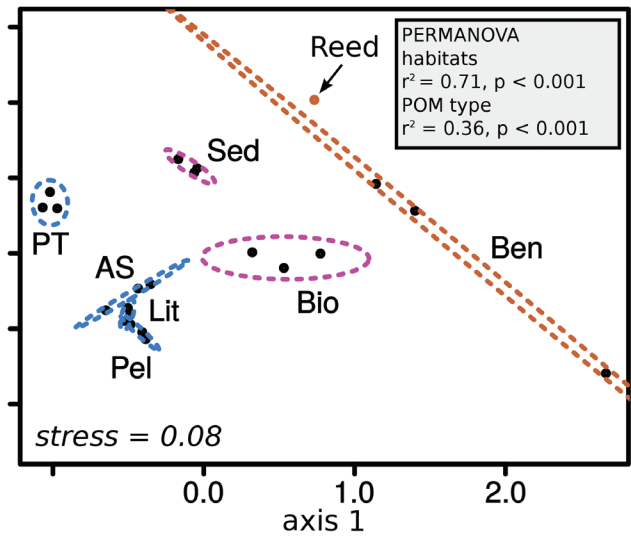


Figure 4. Habitat specificity of the fungal community in Lake Stechlin. NMDS plots based on the fungal OTU matrix (1027 OTUs). Ordination is based on Cao distances (Cao et al. 1997), which are insensitive to differences in sampling effort. Ellipses are based on standard deviations around habitat centroids (based on a confidence level of 0.95) and are coloured according to their POM type: FPOM (blue), MIX (magenta), CPOM (brown). Habitat codes and POM categories are taken from Table 1.

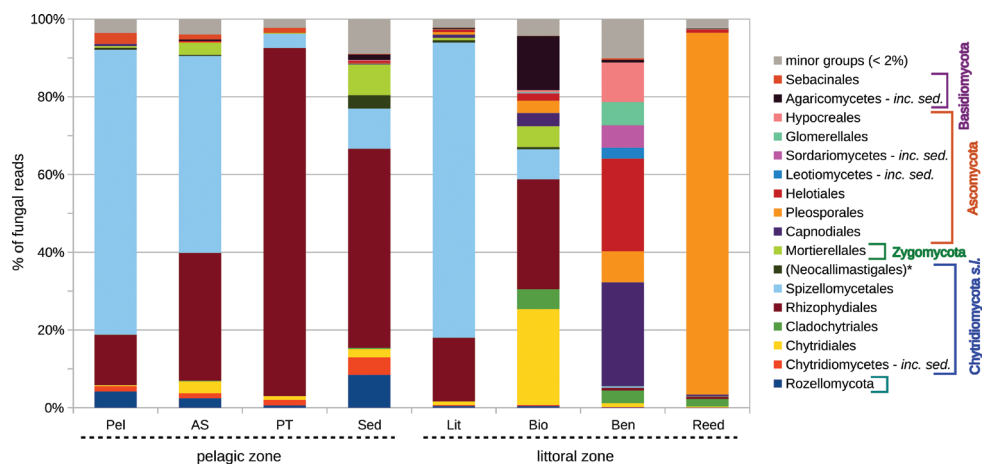


Figure 5. Distribution of fungal orders in Lake Stechlin habitats based on the mean percentages as determined with SILVA NGS. Fungal phyla are highlighted with brackets. Habitat codes are taken from Table 1. (*) marks the false classified Neocallimastigales as clarified by phylogenetic inference (see discussion paragraph on sediments).

ter samples (Pelagic, Littoral, and Above Sediment) appeared to be very similar, whereas all other habitats were distinct (Figure 4). Comparable results were found by NMDS clustering of the presence/absence OTU matrix or by clustering the fungal taxonomic paths generated by SILVA NGS, indicating that the habitat clustering was robust and took place at even higher taxonomic levels (phylum to order level; Suppl. material 6).

Only 23% of the fungal OTUs could be classified to the family or genus level, and around 20% of the sequences could not be assigned to the kingdom level at 0.8 confidence threshold (Figure 2). RDP classifier seem to provide limited classification success when certain early diverging lineages are targeted (e.g., see “Pel” sample in Suppl. material 4). Thus we decided to also evaluate other alternatives. By using Blast against the nucleotide database of NCBI most fungal sequences could only be classified as “fungi” or “environmental samples” (mean: 70.2% of sequences; range: 36.5–89.7% of sequences in a given habitat; see Suppl. material 7). Then by processing the sequences by SILVA NGS we obtained a classification on the order level. Hence, we will use RDP to discuss the fine scale resolution and the SILVA classification for overall comparisons on the order or phyla level (Figure 3, 5). The orders Spizellomycetales and Rhizophydiales (both Chytridiomycetes) comprised the majority of fungal sequences in the four pelagic habitats and the Littoral water sample, with a greater proportion of Spizellomycetales in the three types of water samples compared to more Rhizophydiales in the Plankton and Sediment habitats (Figure 5). In contrast, the Biofilm habitat harboured a good representation of all major fungal phyla (Figure 3) with Chytridiales, Rhizophydiales (both Chytridiomycetes), and Agaricomycetes (Basidiomycota) forming the most prominent orders (Figure 4). Capnodiales and Helotiales (both Ascomycota) were the most prominent orders in the Benthic habitat whereas the Reed habitat was dominated by Pleosporales (Ascomycota) (Figure 4).

Only a small proportion of fungal sequences (0–6%) could be assigned to what we assume are forest taxa (Agaricales, Auriculariales, Boletales, Cantharellales, Gleophyllales, Hymenochaetales, Polyporales, Russulales; Suppl. material 1). This proportion was significantly different among habitats (Kruskal-Wallis test, $df = 5$, $p = 0.023$), being highest in the Sediment samples (mean = 4.9%, SD = 0.9). From all habitats, we recovered sequences from oomycetes (i.e. *Albugo*, *Aphanomyces*, *Phytophthora*, *Pythium*, *Saprolegnia*), a group that was formerly related to aquatic fungi and that occupied similar ecological niches (Sparrow 1960). They have a Chytridiomycota-like life cycle and serve as parasites (e.g., agent of the European cray-fish plague) and saprophytes in aquatic systems. These sequences were 1–3 orders of magnitude lower in abundance compared to the fungal sequences, with maxima in Benthos and Sediment samples (Suppl. material 1).

Discussion

In the following discussion, we first address methodological considerations and then discuss fungal diversity separately for the major habitats.

Methodological considerations

The occurrence of early diverging fungal lineages as well as members of the Dikarya renders a comprehensive assessment of the aquatic mycobiota challenging. This is due to the difficulty in finding a universally suitable marker (i.e., “DNA barcode”) with both sufficient coverage of evolutionary distant groups and meaningful resolution within any of the individual groups. We employed the D1 region of the LSU as a marker because of its high variability while still being conservative enough to amplify across the fungal kingdom (Porter and Golding 2012). Both D1 and D2 regions were formerly used as molecular markers for fungi, especially yeast (Kurtzman and Robnett 1997) and perform almost as well as the commonly used ITS region in discriminating fungal groups (Schoch et al. 2012). The LSU is an established phylogenetic marker for Chytridiomycota (Letcher et al. 2006) and, unlike the ITS region, it can be used to delimit distant aquatic fungal lineages (Lefèvre et al. 2012; Wurzbacher et al. 2014). The small ribosomal subunit (SSU) is also well established for early diverging lineages (e.g., Jobard et al. 2012; Ishii et al. 2015); however, it is less suitable for fungal groups within Dikarya (Lindahl et al. 2013; Tedersoo et al. 2015) and would fail to generate meaningful OTUs for a broad spectrum of simultaneously occurring fungal phyla, such as in our biofilm habitat. Currently, the major disadvantage of using LSU regions as taxonomic markers for aquatic fungi is the lack of reference sequences that allow assignment. Although the RDP classifier identified at least two aquatic hyphomycetes genera in our dataset (*Spirosphaera* and *Tetracladium*) most previous work on this ecological group was done with ITS (Duarte et al. 2014). The ITS region may pose a

better solution for those lake habitats that were dominated by Dikarya, however, as of 2014 only 26% of described aquatic hyphomycetes species had an ITS database record (Duarte et al. 2014). The RDP database had also problems with classifying Chytridiomycota (and Rozellomycota) from the water samples and similar problems may arise with freshwater Chytridiomycota for ITS data (currently there are 1121 Chytridiomycota sequences in UNITE version 7.0, excluding Batrachochytrium sequences) pointing to larger gaps in the reference datasets for aquatic species. The UNITE species hypothesis concept introduced by Kõljalg et al. (2013) might be a good interim solution for dealing with undocumented species; however, the LSU offered alternatives here, namely (i) it was possible to employ species delimitation methods for clustering (GMYC) and classification (see below) and (ii) it allowed the multivariate analyses for using evolutionary (or trait based) based diversity indexes such as UniFrac, which is frequently used in microbial ecology (Lozupone et al. 2011).

Water and large plankton

All three water habitats (Pelagic, Littoral, and Above Sediment) had a low proportion of fungal sequences. Each were characterised by a predominance of Chytridiomycota. This was also the case for the habitats directly connected to processes in the open water (Plankton and Sediment). The proportion of fungal sequences in all samples classified as FPOM was significantly lower (Figure 3). Previous results from Lake Stechlin also reported a low proportion of fungi in water samples (Luo et al. 2011), which may be related to the fact that we did not enrich for fungi by prefiltration or by primer selection. Lefèvre et al. (2012) provided a summary of fungal and chytrid percentages ranging from 1–50% in water samples, relating these observations to prefiltration and primer pair used. Like Monchy et al. (2011), we observed similar communities in all of the water samples (Littoral, Pelagic, and Above Sediment). The greater number of OTUs per read abundance may originate from a rare fungal parasite community (Mangot et al. 2013), whereby parasitic chytrids can recruit for temporally variable infection opportunities such as may occur over time scales of a few weeks (e.g., Ibelings et al. 2004; Alster and Zohary 2007).

There was a limited number of fungal taxa associated with water borne zoo- and phytoplankton samples (Plankton; > 55 µm), which presumably should represent attached or infective stages of fungi. 84–93% of these fungi belonged to Rhizophydiales, a group of well described phytoplankton parasites. By contrast, Rhizophydiales accounted for only 14–17% in the pelagic (open) water samples. This is insofar important because most microscopic studies on chytrids refer to infected algae of approximate the size of 50 µm or larger (e.g., Hohlfeld 1998; Ibelings et al. 2004; Rasconi et al. 2012). However, in an unbiased water sample (i.e. not fractionated by filtration or enriched by a plankton net), they were replaced as dominant group by the order Spizellomycetales, which are common saprotrophs in soil and may underline to the importance of saprotrophic chytrids in aquatic environments (Wurzbacher et al. 2014).

This may establish the mycoloop (Kagami et al. 2014) as a trophic link during times with low prevalence of algal infections, based on the mineralization of detritus (cf. Gleason et al. 2008). Finally, there was a low proportion of Rozellomycota in the large plankton. Rozellomycota are discussed as *inter alia* attached algal parasites (Jones et al. 2011); however, their under-representation in Plankton samples indicates that they were not relevant parasites of the larger plankton in Lake Stechlin, where Chytridiomycota occupied this niche. Due to the small size of Rozellomycota, they may rather have a specialization towards smaller hosts, which do not provide enough resources for Chytridiomycota to complete their life cycle.

Sediment

The profundal sediment temperature in Lake Stechlin remains *ca.* 4 °C year-round while the upper sediment surface (~ 5 mm) is usually oxic. The sediment has a high water content (> 95%) and high organic matter content at the sampled sites because it receives sinking matter from pelagic organisms. Thus the sediments serve as a fungal spore bank. We therefore expected to observe elevated proportions of forest fungi that had probably blown in or been washed in as spores. The dominant fungal group was the Rhizophydiales (Chytridiomycota), as also found in the large plankton samples. Similar to their hosts, parasitic chytrids develop thick-walled resting spores (cysts), which can be found in sediment, while other parasitic species can actively infect algal resting stages in sediments (Canter 1948, Canter 1968). The few studies that have investigated lake or pond sediments reported Chytridiomycota and Rozellomycota (at that time referred to as LKM11 & LKM15) to be the dominant fungal phyla (Luo et al. 2005; Slapeta et al. 2005). Rozellomycota species appear to occur in the hypolimnion of lakes (Lepère et al. 2010) and also in anoxic habitats (Jones et al. 2011), but their ecological function remains unclear (Grossart et al. 2016). Similarly enigmatic was the appearance of Zygomycota (Mortierella) at the sediment surface in our study. Some of them can grow at low temperatures under oxic conditions, e.g. under snow packs in sub-alpine regions (Schmidt et al. 2008). Very surprising was the appearance of Neocallimastigomycota, which are by definition obligate, mutualistic, anaerobic rumen fungi. They are exceptional in that they break down a broad variety of plant polymers under anaerobic conditions (Solomon et al. 2016). These fungi must have had an environmental ancestor and it is possible that anoxic sediments may represent such an ancestral habitat. However, the sequences were only approx. 90% similar to *Orpinomyces* and the RDP classifier assigned a low probability to this classification (< 0.55). Lefèvre et al. (2012) also described sequences from lake plankton samples that may support such a new environmental lineage of “rumen fungi”. Our test of whether those sequences clustered within the Neocallimastigomycota in a phylogenetic tree (Suppl. material 8) found no support for this. The sequences resembled unknown fungal lineages or belonged to Rozellomycota or Zygomycota lineages with a moderate probability.

Biofilm (Periphyton)

Biofilm samples appeared to represent an intermediate fungal habitat between sediments and benthic samples by including a high diversity of early diverging lineages as well as elevated proportions of Dikarya (16–36%, Figure 3). Fungi formed a significant proportion of the overall eukaryotic biofilm community recovered (28% of the OTUs recorded in the habitat were identified as fungi), dominated by biofilm-forming algae (the ratio of fungi to periphyton/epilithic algae was roughly 1:5, see Suppl. material 4, SILVA NGS). Biofilms represent a complex environment (exhibiting the highest eukaryotic taxon richness of all eight habitats) and this is also reflected by a broad range of fungal groups and taxa. Along with Rhizophydiales and Spizellomycetales, we found other chytrids of the orders Chytridiales and Cladochytridiales. The spatial proximity of host cells in periphyton could be ideal for chytrid species, facilitating a high encounter rate with potential hosts and substrates. In contrast to the water samples, the RDP classifier was able to classify more Chytridiomycota to genus level: *Nowakowskiella*, *Chyridium* and *Betamycetes*. The RDP classifier also identified the earlier mentioned aquatic hyphomycetes. However a large part of sequences was only classified to phylum level (Suppl. material 4). These autotrophic lake biofilms seem to be a rich source of fungal biodiversity and pose promising target habitats for future studies. Biofilms (in our case mainly littoral periphyton and epilithic biofilms) have been rarely examined for fungi, and only a few studies on stream ecosystems have investigated the fungal occurrence (measured as ergosterol) on substrates other than leaves (Tank and Dodds 2003; Artigas et al. 2004; Aguilera et al. 2007; Frossard et al. 2012). In lakes and streams, periphyton can contribute substantially to the primary production of the whole ecosystem (Lalonde et al. 1991; Vadeboncoeur et al. 2007 and references therein; Vis et al. 2007) and can be the primary food source for macrozoobenthic grazers (Cattaneo and Mousseau 1995). Our findings suggest that it is not only a rich source of widely divergent fungal lineages, but that fungi might play an important ecological role in periphyton, turning over a significant amount of algal carbon and thus total carbon in the lake.

Benthic and reed samples (CPOM)

In contrast to water samples, fungal sequences were dominant in CPOM (Benthos, Reed) samples and their relative proportions were significantly elevated in this POM type. Samples consisted mainly of submerged plant residues in addition to algae and benthic animals. Mitosporic ascomycetes lineages were predominant, followed by a small percentage of chytrids (mainly Cladochytridiales) and very few Basidiomycota. This appears congruent with our initial “morphotype hypothesis”. Mitosporic ascomycetes are effective plant decomposers in freshwater systems (Gessner et al. 2007), where they are ecologically grouped together as aquatic hyphomycetes (e.g. *Spirosphaera*, a potential aero-aquatic hyphomycete increased to 12% in Benthos samples). The Ben-

thos habitat had a high proportion of fungal OTUs and fungal reads and is probably home to those fungi responsible for the breakdown of submerged plant remains. Interestingly, the importance of aquatic hyphomycetes for plant litter breakdown has thus far only been demonstrated in lotic environments, with lakes not yet investigated in detail (see Chauvet et al. 2016). In contrast to the benthic samples, the Reed sample exhibited a dominance of the fungal order of Pleosporales (93%). Early molecular work has already established the high diversity of reed endophytes (Neubert et al. 2005, Angelini et al. 2012) and we could confirm their presence. The reed sample can be seen as an outgroup in our study, as it comprised the emergent parts of plants. Sequences of the order Pleosporales were largely restricted to the reed and, to a lesser extent, benthic samples.

Conclusions

Fungi play an important role in the cycling of carbon and nutrients in a wide range of freshwater habitats (Bärlocher and Boddy 2016). While much of our understanding of their diversity and ecological roles stems from research in the terrestrial realm, there is increasing interest in their taxonomic and functional diversity in freshwater systems (Grossart et al. 2016; Grossart and Rojas-Jimenez 2016). We examined fungal community composition in eight different habitats of a single lake, in contrast to most studies which have compared water samples among seasons or lakes (e.g., Monchy et al. 2011; Taib et al. 2013). We found pronounced differences in diversity and community composition among the sampled habitat types, and conclude that the habitat heterogeneity within a single lake offers a wide range of fungal niches. The results extend previous research of fungal diversity and distribution in freshwaters and clearly indicate that lake biofilms can be hotspots for aquatic fungi. Most of the fungi from the water samples were rather homogeneous in their community composition, with a clear dominance of Chytridiomycota. This may be due to the predominance of FPOM in the sampled habitats. Our study highlights the importance of habitat heterogeneity and we hope will stimulate further research on under-sampled lake habitats, such as sediments, biofilms, and submerged macrophytes. A more holistic approach in evaluating fungal diversity, using a more comprehensive inclusion of habitat types and taxonomic markers, should provide deeper insights into the multiple ecological roles of fungi in diverse freshwater environments.

Acknowledgments

Research was partially supported by the Leibniz SAW/Pakt for Research and Innovation project “Mycolink” and by an IGB Fellowship in Freshwater Science awarded to NW, HPG, and MTM. CW acknowledges a Marie Skłodowska-Curie post doc grant (660122, CRYPTTRANS). This is publication 39 of the Berlin Center for Genomics in Biodiversity Research.

References

- Aguilera A, Souza-Egipsy V, Gómez F, Amils R (2007) Development and structure of eukaryotic biofilms in an extreme acidic environment, Río Tinto (SW, Spain). *Microbial Ecology* 53(2): 294–305. doi: 10.1007/s00248-006-9092-2
- Alster A, Zohary T (2007) Interactions between the bloom-forming dinoflagellate *Peridinium gatunense* and the chytrid fungus *Phlyctochytrium*. *Hydrobiologia* 578(1): 131–139. doi: 10.1007/s10750-006-0439-y
- Angelini P, Rubini A, Gigante D, Reale L, Pagiotti R, Venanzoni R (2012) The endophytic fungal communities associated with the leaves and roots of the common reed (*Phragmites australis*) in Lake Trasimeno (Perugia, Italy) in declining and healthy stands. *Fungal Ecology* 5(6): 683–693. doi: 10.1016/j.funeco.2012.03.001
- Artigas J, Romaní AM, Sabater S (2004) Organic matter decomposition by fungi in a Mediterranean forested stream: contribution of streambed substrata. *Annales de Limnologie-International Journal of Limnology (EDP Sciences)* 40: 269–277. doi: 10.1051/limn/2004025
- Bärlocher F, Boddy L (2016) Aquatic fungal ecology—How does it differ from terrestrial? *Fungal Ecology* 19: 5–13. doi: 10.1016/j.funeco.2015.09.001
- Canter HM (1948) Studies on British chytrids: VII. On *Phlyctochytrium mucronatum* n.sp. *Transactions of the British Mycological Society* 32(3–4): 236–240. doi: 10.1016/S0007-1536(49)80011-8
- Canter HM (1968) Studies on British chytrids XXVIII. *Rhizophydium nobile* sp. nov., parasitic on the resting spore of *Ceratium hirundinella* O.F.Müll. from the plankton. *Biological Journal of the Linnean Society* 179(2): 197–201. doi: 10.1111/j.1095-8312.1968.tb00977.x
- Cao Y, Williams WP, Bark AW (1997) Similarity measure bias in river benthic Aufwuchs community analysis. *Water Environment Research* 69(1): 95–106. doi: 10.2175/10614309-7X125227
- Casper SJ (1965) Hyphomyceten-Studien. I. Die Suesswasser Hyphomyceten des Stechlinsee-Gebietes. *Limnologica* 3: 257–270.
- Casper SJ (1985) *Lake Stechlin: A temperate oligotrophic lake*. Springer, Dodrecht. doi: 10.1007/978-94-009-5506-6
- Cattaneo A, Mousseau B (1995) Empirical analysis of the removal rate of periphyton by grazers. *Oecologia*, 103(2): 249–254. doi: 10.1007/BF00329087
- Chauvet E, Cornut J, Sridhar KR, Selosse MA, Bärlocher F (2016) Beyond the water column: aquatic hyphomycetes outside their preferred habitat. *Fungal Ecology* 19: 112–127. doi: 10.1016/j.funeco.2015.05.014
- Chen M, Chen F, Yu Y, Ji J, Kong F (2008) Genetic diversity of eukaryotic microorganisms in Lake Taihu, a large shallow subtropical lake in China. *Microbial Ecology* 56(3): 572–583. doi: 10.1007/s00248-008-9377-8
- Chiu CH, Wang YT, Walther BA, Chao A (2014) Improved nonparametric lower bound of species richness via a modified Good-Turing frequency formula. *Biometrics* 70(3): 671–682. doi: 10.1111/biom.12200

- Corsaro D, Walochnik J, Venditti D, Steinmann J, Müller KD, Michel R (2014) Microsporidia-like parasites of amoebae belong to the early fungal lineage Rozellomycota. *Parasitology research* 113(5): 1909–1918. doi: 10.1007/s00436-014-3838-4
- Duarte S, Bärlocher F, Cássio F, Pascoal C (2014) Current status of DNA barcoding of aquatic hyphomycetes. *Sydowia* 66(2): 191–202. doi: 10.12905/0380.sydowia66(2)2014-0191
- Frossard A, Gerull L, Mutz M, Gessner MO (2012) Fungal importance extends beyond litter decomposition in experimental early-successional streams. *Environmental Microbiology* 14(11): 2971–2983. doi: 10.1111/j.1462-2920.2012.02865.x
- Fujisawa T, Barraclough TG (2013) Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent (GMYC) approach: a revised method and evaluation on simulated datasets. *Systematic Biology* 62: 707–724. doi: 10.1093/sysbio/syt033
- Gessner MO, Gulis V, Kuehn KA, Chauvet E, Suberkropp K (2007) Fungal decomposers of plant litter in aquatic ecosystems. In: Kubicek CP, Druzhinina IS (Eds) *The Mycota IV: Environmental and Microbial Relationships* (2nd edn). Springer, Berlin, 301–324.
- Gleason FH, Kagami M, Lefevre E, Sime-Ngando T (2008) The ecology of chytrids in aquatic ecosystems: roles in food web dynamics. *Fungal Biology Reviews* 22(1): 17–25. doi: 10.1016/j.fbr.2008.02.001
- Grossart HP, Rojas-Jimenez K (2016) Aquatic fungi: targeting the forgotten in microbial ecology. *Current Opinion in Microbiology* 31: 140–145. doi: 10.1016/j.mib.2016.03.016
- Grossart HP, Wurzbacher C, James TY, Kagami M (2016) Discovery of dark matter fungi in aquatic ecosystems demands a reappraisal of the phylogeny and ecology of zoosporic fungi. *Fungal Ecology* 19: 28–38. doi: 10.1016/j.funeco.2015.06.004
- Gulis V, Kuehn KA, Suberkropp K (2009) Fungi. In: Likens EG (Ed.) *Encyclopedia of inland waters*. Academic Press, Oxford, 233–243. doi: 10.1016/B978-012370626-3.00129-0
- Hawksworth DL (2012) Global species numbers of fungi: are tropical studies and molecular approaches contributing to a more robust approach? *Biodiversity and Conservation* 21(9): 2245–2433. doi: 10.1007/s10531-012-0335-x
- Holfeld H (1998) Fungal infections of the phytoplankton: seasonality, minimal host density, and specificity in a mesotrophic lake. *New Phytologist* 138(3): 507–517. doi: 10.1046/j.1469-8137.1998.00126.x
- Huson DH, Mitra S, Ruscheweyh HJ, Weber N, Schuster SC (2011) Integrative analysis of environmental sequences using MEGAN4. *Genome Research* 21(9): 1552–1560. doi: 10.1101/gr.120618.111
- Ibelings BW, De Bruin A, Kagami M, Rijkeboer M, Brehm M, Van Donk E (2004) Host parasite interactions between freshwater phytoplankton and chytrid fungi (Chytridiomycota). *Journal of Phycology* 40(3): 437–453. doi: 10.1111/j.1529-8817.2004.03117.x
- Ishii N, Ishida S, Kagami M (2015) PCR primers for assessing community structure of aquatic fungi including Chytridiomycota and Cryptomycota. *Fungal Ecology* 13: 33–43. doi: 10.1016/j.funeco.2014.08.004
- James TY, Letcher PM, Longcore JE, Mozley-Standridge SE, Porter D, Powell MJ, Griffith GW, Vilgalys R (2006) A molecular phylogeny of the flagellated fungi (Chytridiomycota) and description of a new phylum (Blastocladiomycota). *Mycologia* 98(6): 860–871. doi: 10.3852/mycologia.98.6.860

- Jobard M, Rasconi S, Sime-Ngando T (2010) Diversity and functions of microscopic fungi: a missing component in pelagic food webs. *Aquatic Sciences* 72(3): 255–268. doi: 10.1007/s00027-010-0133-z
- Jobard M, Rasconi S, Solinhac L, Cauchie HM, Sime-Ngando T (2012) Molecular and morphological diversity of fungi and the associated functions in three European nearby lakes. *Environmental Microbiology* 14(9): 2480–2494. doi: 10.1111/j.1462-2920.2012.02771.x
- Jones MD, Forn I, Gadelha C, Egan MJ, Bass D, Massana R, Richards TA (2011) Discovery of novel intermediate forms redefines the fungal tree of life. *Nature* 474(7350): 200–203. doi: 10.1038/nature09984
- Kagami M, de Bruin A, Ibelings BW, Van Donk E (2007) Parasitic chytrids: their effects on phytoplankton communities and food-web dynamics. *Hydrobiologia* 578(1): 113–129. doi: 10.1007/s10750-006-0438-z
- Kagami M, Miki T, Takimoto G (2014) Mycoloop: chytrids in aquatic food webs. *Frontiers in Microbiology* 5: 166. doi: 10.3389/fmicb.2014.00166
- Karling JS (1977) *Chytridiomycetorum Iconographia*. J. Cramer, Montecello, NY, 414 pp.
- Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO (2012) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research*. doi: 10.1093/nar/gks808
- Kõljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AE, Bahram M, Bates ST, Bruns TD, Bengtsson-Palme J, Callaghan TM, Douglas B, Drenkhan T, Eberhardt U, Dueñas M, Grebenc T, Griffith GW, Hartmann M, Kirk PM, Kohout P, Larsson E, Lindahl BD, Lücking R, Martín MP, Matheny PB, Nguyen NH, Niskanen T, Oja J, Peay KG, Peintner U, Peterson M, Pöldmaa K, Saag L, Saar I, Schüßler A, Scott JA, Senés C, Smith ME, Suija A, Taylor DL, Telleria MT, Weiss M, Larsson KH (2013) Towards a unified paradigm for sequence-based identification of fungi. *Molecular ecology* 22(21): 5271–5277. doi: 10.1111/mec.12481
- Kohout P, Sýkorová Z, Čtvrtlíková M, Rydlova J, Suda J, Vohník M, Sudova R (2012) Surprising spectra of root-associated fungi in submerged aquatic plants. *FEMS Microbiology Ecology* 80(1): 216–235. doi: 10.1111/j.1574-6941.2011.01291.x
- Krauss GJ, Solé M, Krauss G, Schlosser D, Wesenberg D, Bärlocher F (2011) Fungi in freshwaters: ecology, physiology and biochemical potential. *FEMS Microbiology Reviews* 35(4): 620–651. doi: 10.1111/j.1574-6976.2011.00266.x
- Kuehn KA (2008) The role of fungi in the decomposition of emergent wetland plants. *Novel Techniques and Ideas in Mycology* 20: 19–41.
- Kurtzman CP, Robnett CJ (1997) Identification of clinically important ascomycetous yeasts based on nucleotide divergence in the 5' end of the large-subunit (26S) ribosomal DNA gene. *Journal of Clinical Microbiology* 35(5): 1216–1223.
- Lalonde S, Downing JA (1991) Epiphyton biomass is related to lake trophic status, depth, and macrophyte architecture. *Canadian Journal of Fisheries and Aquatic Sciences* 48(11): 2285–2291. doi: 10.1139/f91-268
- Lefèvre E, Bardot C, Noël C, Carrias JF, Viscogliosi E, Amblard C, Sime-Ngando T (2007) Unveiling fungal zooflagellates as members of freshwater picoeukaryotes: evidence from a molecular diversity study in a deep meromictic lake. *Environmental Microbiology* 9(1): 61–71. doi: 10.1111/j.1462-2920.2006.01111.x

- Lefèvre E, Letcher PM, Powell MJ (2012) Temporal variation of the small eukaryotic community in two freshwater lakes: emphasis on zoosporic fungi. *Aquatic Microbial Ecology* 67(2): 91–105. doi: 10.3354/ame01592
- Lepère C, Masquelier S, Mangot JF, Debroas D, Domaizon I (2010) Vertical structure of small eukaryotes in three lakes that differ by their trophic status: a quantitative approach. *The ISME Journal* 4(12): 1509–1519. doi: 10.1038/ismej.2010.83
- Lennon NJ, Lintner RE, Anderson S, Alvarez P, Barry A, Brockman W, Daza R, Erlich RL, Giannoukos G, Green L, Hollinger A, Hoover CA, Jaffe DB, Juhn F, McCarthy D, Perrin D, Ponchner K, Powers TL, Rizzolo K, Robbins D, Ryan E, Russ C, Sparrow T, Stalker J, Steelman S, Weiland M, Zimmer A, Henn MR, Nussbaum C, Nicol R (2010) A scalable, fully automated process for construction of sequence-ready barcoded libraries for 454. *Genome Biology* 11(2): R15. doi: 10.1186/gb-2010-11-2-r15
- Letcher PM, Powell MJ, Churchill PF, Chambers JG (2006) Ultrastructural and molecular phylogenetic delineation of a new order, the Rhizophydiales (Chytridiomycota). *Mycological Research* 110(8): 898–915. doi: 10.1016/j.mycres.2006.06.011
- Lindahl BD, Nilsson RH, Tedersoo L, Abarenkov K, Carlsen T, Kjøller R, Kõljalg U, Pennanen T, Rosendahl S, Stenlid J, Kausrud H (2013) Fungal community analysis by high-throughput sequencing of amplified markers—a user's guide. *New Phytologist* 199(1): 288–299. doi: 10.1111/nph.12243
- Liu KL, Porras-Alfaro A, Kuske CR, Eichorst SA, Xie G (2012) Accurate, rapid taxonomic classification of fungal large-subunit rRNA genes. *Applied and environmental microbiology* 78(5): 1523–1533. doi: 10.1128/AEM.06826-11
- Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R (2011) UniFrac: an effective distance metric for microbial community comparison. *The ISME journal* 5(2): 169–172. doi: 10.1038/ismej.2010.133
- Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, Buchner A, Lai T, Steppi S, Jobb G, Förster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, König A, Liss T, Lüssmann R, May M, Nonhoff B, Reichel B, Strehlow R, Stamatakis A, Stuckmann N, Vilbig A, Lenke M, Ludwig T, Bode A, Schleifer KH (2004) ARB: a software environment for sequence data. *Nucleic Acids Research* 32(4): 1363–1371. doi: 10.1093/nar/gkh293
- Luo Q, Krumholz LR, Najar FZ, Peacock AD, Roe BA, White DC, Elshahed MS (2005) Diversity of the microeukaryotic community in sulfide-rich Zodletone Spring (Oklahoma). *Applied and Environmental Microbiology* 71(10): 6175–6184. doi: 10.1128/AEM.71.10.6175-6184.2005
- Luo W, Bock C, Li HR, Padisák J, Krienitz L (2011) Molecular and microscopic diversity of planktonic eukaryotes in the oligotrophic Lake Stechlin (Germany). *Hydrobiologia* 661(1): 133–143. doi: 10.1007/s10750-010-0510-6
- Mangot JF, Domaizon I, Taib N, Marouni N, Duffaud E, Bronner G, Debroas D (2013) Short-term dynamics of diversity patterns: evidence of continual reassembly within lacustrine small eukaryotes. *Environmental Microbiology* 15(6): 1745–1758. doi: 10.1111/1462-2920.12065
- Monchy S, Sancier G, Jobard M, Rasconi S, Gerphagnon M, Chabé M, Cian A, Meloni D, Niquil N, Christaki U, Viscogliosi E, Sime-Ngando T (2011) Exploring and quantifying

- fungal diversity in freshwater lake ecosystems using rDNA cloning/sequencing and SSU tag pyrosequencing. *Environmental Microbiology* 13(6): 1433–1453. doi: 10.1111/j.1462-2920.2011.02444.x
- Neubert K, Mendgen K, Brinkmann H, Wirsal S (2006) Only a few fungal species dominate highly diverse mycofloras associated with the common reed. *Applied and Environmental Microbiology* 72(2): 1118–1128. doi: 10.1128/AEM.72.2.1118-1128.2006
- Ossowski S, Schneeberger K, Clark RM, Lanz C, Warthmann N, Weigel D (2008) Sequencing of natural strains of *Arabidopsis thaliana* with short reads. *Genome Research* 18(12): 2024–2033. doi: 10.1101/gr.080200.108
- Panzer K, Yilmaz P, Weiß M, Reich L, Richter M, Wiese J, Schmaljohann R, Labes A, Imhoff JF, Glöckner FO, Reich M (2015) Identification of habitat-specific biomes of aquatic fungal communities using a comprehensive nearly full-length 18S rRNA dataset enriched with contextual data. *PLoS ONE* 10: e0134377. doi: 10.1371/journal.pone.0134377
- Picard KT, Letcher PM, Powell MJ (2013) Evidence for a facultative mutualist nutritional relationship between the green coccoid alga *Bracteacoccus* sp. (Chlorophyceae) and the zoospore fungus *Rhizidium phycophilum* (Chytridiomycota). *Fungal biology* 117(5): 319–328. doi: 10.1016/j.funbio.2013.03.003
- Porter TM, Golding GB (2012) Factors that affect large subunit ribosomal DNA amplicon sequencing studies of fungal communities: classification method, primer choice, and error. *PLoS ONE* 7(4): e35749. doi: 10.1371/journal.pone.0035749
- Powell JR, Monaghan MT, Oepik M, Rillig MC (2011) Evolutionary criteria outperform operational approaches in producing ecologically relevant fungal species inventories. *Molecular Ecology* 20(3): 655–666. doi: 10.1111/j.1365-294X.2010.04964.x
- Pruesse E, Peplies J, Glöckner FO (2012) SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28(14): 1823–1829. doi: 10.1093/bioinformatics/bts252
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41(1): 590–596. doi: 10.1093/nar/gks1219
- Rasconi S, Niquil N, Sime-Ngando T (2012) Phytoplankton chytridiomycosis: community structure and infectivity of fungal parasites in aquatic ecosystems. *Environmental Microbiology* 14(8): 2151–2170. doi: 10.1111/j.1462-2920.2011.02690.x
- Reeder J, Knight R (2009) The ‘rare biosphere’: a reality check. *Nature Methods* 6(9): 636–637. doi: 10.1038/nmeth0909-636
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology* 61(3): 539–542. doi: 10.1093/sysbio/sys029
- Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences* 109(16): 6241–6246. doi: 10.1073/pnas.1117018109
- Schindler DE, Scheuerell MD (2002) Habitat coupling in lake ecosystems. *Oikos* 98(2): 177–189. doi: 10.1034/j.1600-0706.2002.980201.x

- Schmidt SK, Wilson KL, Meyer AF, Gebauer MM, King AJ (2008) Phylogeny and ecophysiology of opportunistic “snow molds” from a subalpine forest ecosystem. *Microbial Ecology* 56(4): 681–687. doi: 10.1007/s00248-008-9387-6
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* 75(23): 7537–7541. doi: 10.1128/AEM.01541-09
- Shearer CA, Descals E, Kohlmeyer B, Kohlmeyer J, Marvanová L, Padgett D, Porter D, Raja HA, Schmit JP, Thorton HA, Voglymayr H (2007) Fungal biodiversity in aquatic habitats. *Biodiversity and Conservation* 16(1): 49–67. doi: 10.1007/s10531-006-9120-z
- Slapeta J, Moreira D, López-García P (2005) The extent of protist diversity: insights from molecular ecology of freshwater eukaryotes. *Proceedings of the Royal Society of London B: Biological Sciences* 272(1576): 2073–2208. doi: 10.1098/rspb.2005.3195
- Solomon KV, Haitjema CH, Henske JK, Gilmore SP, Borges-Rivera D, Lipzen A, Brewer HM, Purvine SO, Wright AT, Theodorou MK, Grigoriev IV (2016) Early-branching gut fungi possess a large, comprehensive array of biomass-degrading enzymes. *Science* 351(6278): 1192–1195. doi: 10.1126/science.aad1431
- Sparrow F (1960) *Aquatic phycomycetes* (2nd edn). University of Michigan Press, Ann Arbor. doi: 10.5962/bhl.title.5685
- Taib N, Mangot JF, Domaizon I, Bronner G, Debroas D (2013) Phylogenetic affiliation of SSU rRNA genes generated by massively parallel sequencing: new insights into the freshwater protist diversity. *PLoS ONE* 8(3): e58950. doi: 10.1371/journal.pone.0058950
- Tank JL, Dodds WK (2003) Nutrient limitation of epilithic and epixylic biofilms in ten North American streams. *Freshwater Biology* 48(6): 1031–1049. doi: 10.1046/j.1365-2427.2003.01067.x
- Tedersoo L, Anslan S, Bahram M, Pölme S, Riit T, Liiv I, Kõljalg U, Kisand V, Nilsson RH, Hildebrand F, Bork P, Abarenkov K (2015) Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. *Mycology* 10: 1–43. doi: 10.3897/mycokeys.10.4852
- Tsui CK, Baschien C, Goh TK (2016) *Biology and Ecology of Freshwater Fungi*. In: Li DW (Ed.) *Biology of Microfungi*. Springer International Publishing, 285–313. doi: 10.1007/978-3-319-29137-6_13
- Vadeboncoeur Y, Vander Zanden MJ, Lodge DM (2002) Putting the Lake Back Together: Reintegrating Benthic Pathways into Lake Food Web Models: Lake ecologists tend to focus their research on pelagic energy pathways, but, from algae to fish, benthic organisms form an integral part of lake food webs. *Bioscience* 52(1): 44–54. doi: 10.1641/0006-3568(2002)052[0044:PTLBTR]2.0.CO;2
- Vadeboncoeur Y, Peterson G, Vander Zanden MJ, Kalff J (2007) Benthic algal production across lake size gradients: interactions among morphometry, nutrients, and light. *Ecology* 89(9): 2542–2552. doi: 10.1890/07-1058.1
- van der Auwera G, Chapelle S, De Wächter R (1994) Structure of the large ribosomal subunit RNA of *Phytophthora megasperma*, and phylogeny of the oomycetes. *FEBS Letters* 338(2): 133–136. doi: 10.1016/0014-5793(94)80350-1

- van Donk E, Ringelberg J (1983) The effect of fungal parasitism on the succession of diatoms in Lake Maarsseveen I (The Netherlands). *Freshwater Biology* 13(3): 241–251. doi: 10.1111/j.1365-2427.1983.tb00674.x
- Vis C, Hudon C, Carignan R, Gagnon P (2007) Spatial analysis of production by macrophytes, phytoplankton and epiphyton in a large river system under different water-level conditions. *Ecosystems* 10(2): 293–310. doi: 10.1007/s10021-007-9021-3
- Wong MK, Goh TK, Hodgkiss IJ, Hyde KD, Ranghoo VM, Tsui CK, Ho WH, Wong WS, Yuen TK (1998) Role of fungi in freshwater ecosystems. *Biodiversity and Conservation* 7(9): 1187–1206. doi: 10.1023/A:1008883716975
- Wurzbacher C, Bärlocher F, Grossart HP (2010) Fungi in lake ecosystems. *Aquatic Microbial Ecology* 59(2): 125–149. doi: 10.3354/ame01385
- Wurzbacher C, Rösler S, Rychla A, Grossart HP (2014) Importance of saprotrophic freshwater fungi for pollen degradation. *PLoS ONE* 9(4): e94643. doi: 10.1371/journal.pone.0094643
- Zhao B, Chen M, Sun Y, Yang J, Chen F (2011) Genetic diversity of picoeukaryotes in eight lakes differing in trophic status. *Canadian Journal of Microbiology* 57(2): 115–126. doi: 10.1139/W10-107

Supplementary material I

Accompanying data tables

Authors: Christian Wurzbacher, Norman Warthmann, Elizabeth Bourne, Katrin Attermeyer, Martin Allgaier, Jeff R. Powell, Harald Detering, Susan Mbedi, Hans-Peter Grossart, Michael T. Monaghan

Data type: PDF file

Explanation note: Two additional tables stating 1) the physicochemistry of the lake during the sampled time span, 2) the proportions of fungal-like organism and terrestrial fungi (“forest fungi”, see main text).

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 2

List of pyrosequencing primers.

Authors: Christian Wurzbacher, Norman Warthmann, Elizabeth Bourne, Katrin Attermeyer, Martin Allgaier, Jeff R. Powell, Harald Detering, Susan Mbedi, Hans-Peter Grossart, Michael T. Monaghan

Data type: PDF file

Explanation note: List of employed pyrosequencing fusion primers.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 3

Detailed results of the clipping procedure

Authors: Christian Wurzbacher, Norman Warthmann, Elizabeth Bourne, Katrin Attermeyer, Martin Allgaier, Jeff R. Powell, Harald Detering, Susan Mbedi, Hans-Peter Grossart, Michael T. Monaghan

Data type: ZIP file

Explanation note: File archive in “zip” format. Decompress with e.g., “unzip” for inspecting individual files. Files describing the bioinformatic processing and the effects of the clipping procedure.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 4

Krona charts of classified OTUs (RDP) and sequences (SILVA NGS).

Authors: Christian Wurzbacher, Norman Warthmann, Elizabeth Bourne, Katrin Attermeyer, Martin Allgaier, Jeff R. Powell, Harald Detering, Susan Mbedi, Hans-Peter Grossart, Michael T. Monaghan

Data type: ZIP file

Explanation note: File archive in “zip” format. Decompress with e.g., “unzip” for inspecting individual html files. Open html files with an internet browser (network connection must be enabled). Krona charts presenting the classified sequences based on a) OTUs classified by the RDP classifier, and b) based clipped reads processed by the SILVA NGS pipeline.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 5

Rarefaction analysis

Authors: Christian Wurzbacher, Norman Warthmann, Elizabeth Bourne, Katrin Attermeyer, Martin Allgaier, Jeff R. Powell, Harald Detering, Susan Mbedi, Hans-Peter Grossart, Michael T. Monaghan

Data type: PDF file

Explanation note: Rarefaction curves based on the OTU table without OTU singletons for all eukaryotes (left panel) and for fungal OTUs (right panel).

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 6

Alternative ordinations based on binary data and taxonomy

Authors: Christian Wurzbacher, Norman Warthmann, Elizabeth Bourne, Katrin Attermeyer, Martin Allgaier, Jeff R. Powell, Harald Detering, Susan Mbedi, Hans-Peter Grossart, Michael T. Monaghan

Data type: PDF file

Explanation note: Two alternative NMDS accompanied by corresponding PERMANOVA statistics based on a) the presence/absence of OTUs (left panel) and the Jaccard index, and b) the taxonomic classification of SILVA NGS and the Cao index. The ellipses (coloured to POM type) are based on standard deviations around centroids with a confidence level of 0.95.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 7

Comprehensive Blast analysis

Authors: Christian Wurzbacher, Norman Warthmann, Elizabeth Bourne, Katrin Attermeyer, Martin Allgaier, Jeff R. Powell, Harald Detering, Susan Mbedi, Hans-Peter Grossart, Michael T. Monaghan

Data type: PDF file

Explanation note: Blast results against the NCBI nt database (June 2015) for all eukaryotes presented in a Megan5 plot.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 8

Phylogenetic tree of *Neocallimastigales* sequences

Authors: Christian Wurzbacher, Norman Warthmann, Elizabeth Bourne, Katrin Attermeyer, Martin Allgaier, Jeff R. Powell, Harald Detering, Susan Mbedi, Hans-Peter Grossart, Michael T. Monaghan

Data type: PDF file

Explanation note: Phylogenetic bayesian tree based on 10 million generations. Sequences that were classified as *Neocallimastigales* in the manuscript are coloured in red with the frequency of sequences in the dataset in black. The name corresponds to the representative read identifier in the dataset (accessible at ENA nucleotide archive: PRJEB14236). Node labels are posterior probabilities above 0.8. Statistically supported branches have an probability of 0.95-1.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.