

UC Merced

UC Merced Previously Published Works

Title

Fungal diversity and function in metagenomes sequenced from extreme environments

Permalink

<https://escholarship.org/uc/item/7j85j7c9>

Authors

de Mesquita, Clifton P Bueno

Vimercati, Lara

Wu, Dongying

et al.

Publication Date

2024-12-01

DOI

10.1016/j.funeco.2024.101383

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



Fungal diversity and function in metagenomes sequenced from extreme environments

Clifton P. Bueno de Mesquita^a, Lara Vimercati^b, Dongying Wu^a, Mary K. Childress^b, August Danz^b, Arthur C. Grupe^{b,c}, Danny Haelewaters^{b,d}, Natalie M. Hyde^{b,e}, Thiago Kossmann^b, Charles Oliver^b, Candice Perrotta^b, Benjamin D. Young^b, Steven K. Schmidt^b, Susannah G. Tringe^{a,f}, C. Alisha Quandt^{b,*}

^a DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA

^b Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO, 80309, USA

^c Biology Department, University of Wisconsin-La Crosse, La Crosse, WI, 54601, USA

^d Biology Centre of the Czech Academy of Sciences, Institute of Entomology, 370 05, České Budějovice, Czech Republic

^e Institute of Microbiology of the Czech Academy of Sciences, 14220 Praha 4, Czech Republic

^f Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA

ARTICLE INFO

Handling Editor: Prof. Benjamin E. Wolfe

Keywords:

Community structure
Extremophiles
Fungi
Adaptation

ABSTRACT

Fungi are increasingly recognized as key players in various extreme environments. Here we present an analysis of publicly-sourced metagenomes from global extreme environments, focusing on fungal taxonomy and function. The majority of 855 selected metagenomes contained scaffolds assigned to fungi. Relative abundance of fungi was as high as 10% of protein-coding genes with taxonomic annotation, with up to 289 fungal genera per sample. Despite taxonomic clustering by environment, fungal communities were more dissimilar than archaeal and bacterial communities, both for within- and between-environment comparisons. Relatively abundant fungal classes in extreme environments included Dothideomycetes, Eurotiomycetes, Leotiomycetes, Pezizomycetes, Saccharomycetes, and Sordariomycetes. Broad generalists and prolific aerial spore formers were the most relatively abundant fungal genera detected in most of the extreme environments, bringing up the question of whether they are actively growing in those environments or just surviving as spores. More specialized fungi were common in some environments, such as zoospore taxa in cryosphere water and hot springs. Relative abundances of genes involved in adaptation to general, thermal, oxidative, and osmotic stress were greatest in soda lake, acid mine drainage, and cryosphere water samples.

1. Introduction

Microbes have a greater ability to grow in extreme environments compared to most animals and plants (Rothschild and Mancinelli, 2001). While considerable research in recent decades has gone towards understanding bacterial and archaeal diversity in extreme environments (Narasimharao et al., 2012; Shu and Huang, 2022), less attention has been focused on documenting fungal extremophiles despite knowledge that fungi are some of the most recalcitrant species able to survive extreme conditions (Bridge and Spooner, 2012; Onofri et al., 2015; Vimercati et al., 2016).

Many studies of fungal extremophiles have focused on individual species used as model organisms to understand specific adaptations,

growth optima, and genetic underpinnings of extremophilic living (Coleine et al., 2022). We still have a limited understanding of fungal communities in different types of extreme environments, and if there are consistent taxa that thrive in specific extreme conditions. Likewise, it is not known whether certain functional guilds or lineages are over or underrepresented in these environments.

In this study we focus on the following nine extreme environments: acid mine drainage, cryosphere soil, cryosphere water, deserts, glacial forefields, hot springs, hydrothermal vents, hypersaline environments, and soda lakes. Within each of these environments there are one to many key stressors that organisms must tolerate in order to survive (Table 1). Acid mine drainage sites, including mine rock dumps and tailings impoundments (Akcil and Koldas, 2006), are characterized by high levels

* Corresponding author.

E-mail address: alisha.quandt@colorado.edu (C.A. Quandt).

<https://doi.org/10.1016/j.funeco.2024.101383>

Received 17 April 2024; Received in revised form 3 August 2024; Accepted 6 August 2024

Available online 29 August 2024

1754-5048/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Table 1

Summary of extreme environment types included in this study, along with subtypes, major stressors, and adaptations that organisms have to cope with those stressors.

Environment	Subtypes	Major Stressors	Adaptations
Acid mine drainage	Coal, copper, iron, silver, gold, zinc, quartz, pyrite	Low pH, high metal	Metal sequestration or efflux
Cryosphere - soil	Cryoconite hole soils, supraglacial sediment, Antarctic soils	Low temperature, low nutrient	Modified membrane lipids, cold-shock proteins, dormancy, melanin
Cryosphere - water	Cryoconite hole water, ice, glacial meltwater	Low temperature, low nutrient	Modified membrane lipids, cold-shock proteins, dormancy, melanin
Desert	Hot desert soils, biological soil crust, saline desert soils	Low water activity, high salt	Melanin, heat-shock proteins, compatible solutes
Glacial forefield	Arctic glacial forefields from Greenland, Sweden, Norway	Low temperature, low nutrient	Modified membrane lipids, cold-shock proteins, dormancy, melanin
Hot spring	Microbial mats, sediments	High temperature	Heat-shock proteins, high GC content, modified cell membrane, DNA supercoils
Hypersaline	Solar salterns, hypersaline lakes	High salt, low water activity	Compatible solutes, melanin
Hydrothermal vent	Black smoker, chimney, sediment, plume, microbial mats, vent fluid	High temperature	Heat-shock proteins, high GC content, modified cell membrane, DNA supercoils
Soda lake	Soda lakes, haloalkaline lakes	High pH, high salt, low water activity	Compatible solutes, alkaline active enzymes

of toxic metals such as chromium, copper, zinc, cadmium, and lead, and low pH, typically <4 (Saria et al., 2006). The cryosphere refers to perennially cold snow and ice masses that occur at high latitudes and high elevations, including both terrestrial ice and sea ice (Marshall, 2011), and is characterized by low nutrient concentrations and nutrient cycling rates, although there can be significant variation in nutrient levels based on geographic location and factors such as distance to the sea (Schmidt et al., 2022). The cryosphere includes habitats called cryoconite holes, formed when sediment on top of ice heats up and forms a melt pool; these habitats contain soil, water, and ice (Wharton et al., 1985). Adjacent to the cryosphere, in some cases, are glacial forefields, which are newly exposed soils or proto-soils high in sand content and low in organic matter and nutrient concentrations (Schmidt et al., 2008), often with phosphorus as the limiting nutrient rather than nitrogen (Bueno de Mesquita et al., 2020; Darcy et al., 2018; King et al., 2008; Knelman et al., 2021).

Desert biomes include both cold deserts and hot deserts that are most effectively defined as areas where the potential for evapotranspiration is more than precipitation (Quinn, 2008). Microbes in deserts are forced to endure low water activity conditions (Lebre et al., 2017). Hot springs are habitats formed over geothermally active hotspots within the Earth and can be defined as habitats with water temperatures greater than the core human body temperature of 36.7 °C (Pentecost et al., 2003) and ranging up to 99 °C, with different communities dominating above and below 75 °C (Guo et al., 2020; Poddar and Das, 2018). Hydrothermal vents, including “black smoker” (~350 °C) and “lost city” (~50–90 °C) types, occur on the ocean floor, normally near tectonic boundaries, where geological activity releases hot water as high as 373 °C (Martin et al., 2008). Hypersaline environments are defined as environments with salt concentrations greater than that of seawater (35 ppt) (Mcgenity and Oren, 2012), which include man-made saltern ponds for salt production,

and natural lakes and inland seas with low freshwater inputs. Soda lakes are best characterized by their elevated pH, typically greater than 10 and reaching up to 12. Factors contributing to the elevated pH in these environments typically include high evaporative rates as compared to inflow, high salt concentrations, and low levels of magnesium and calcium ions (Duckworth et al., 1996).

Several adaptations exist at the molecular level for organisms to live in conditions with high temperature, low temperature, low water activity, high pH, low pH, and/or high metal concentrations (Shu and Huang, 2022). These include modified cell membranes, melanized cells, heat-shock and cold-shock proteins, high GC content, DNA supercoils, increased unsaturated fatty acids, dormancy, accumulation of KCl in the cytoplasm (salt-in strategy), accumulation of compatible osmotic solutes to exclude salt from the cytoplasm (salt-out strategy), impermeable cell membranes, cytoplasmic buffering, use of iron rivets, metal binding, and the formation of sulfate-metal complexes (Gostinčar et al., 2009; Konings et al., 2002; Shu and Huang, 2022). Differences exist among the three domains of life in strategies to cope with the various stressors of extreme environments. For example, the salt-in strategy is mostly found in archaea, and while only eukaryotes use glycerol as a primary compatible solute, archaea and bacteria use other molecules such as trehalose, glycine betaine, and ectoine (Gunde-Cimerman et al., 2018). Extremophilic fungi are often tolerant to multiple stressors (i.e., poly-extremophilic) (Gostinčar et al., 2015), slow-growing (Gostinčar et al., 2022), and opportunistic (Gostinčar et al., 2018).

Here we utilized 855 publicly available metagenomes from extreme environments worldwide to examine fungal community structure and function. Our first goal was to describe the diversity of fungi found in these environments. More specifically, we tested the hypothesis that fungal communities are taxonomically distinct in each extreme environment, similar to trends in archaea and bacteria. Furthermore, we hypothesized that particular taxonomic and functional groups previously linked to specific stressors would be overrepresented in environments corresponding to those stressors.

2. Materials and methods

Metagenomic datasets from nine different extreme environment types were acquired from the U.S. Department of Energy Joint Genome Institute’s (JGI) IMG/M online database (Chen et al., 2021). The environments included acid mine drainage, cryosphere (water/ice and soil), deserts, glacial forefields, hot springs, hypersaline environments, hydrothermal vents, and soda lakes (Fig. 1a). Complete metadata, including original classification of environments according to GOLD categories (Mukherjee et al., 2021), are available in Table S1. IMG/M was searched on June 24, 2022 both by using the environment names as keywords, and by browsing metagenomes by environment type. The dataset was filtered to contain only assembled metagenomes (no metatranscriptomes) that were publicly available, were from 2012 to the time of the search (June 24, 2022), were sequenced with Illumina rather than 454 technology, contained scaffolds assigned to genus level (of any taxa), and that were confirmed to be from the target environment types. Glacial chronosequence samples other than those listed as closest to the glacier forefront were removed from the glacial forefield list, as were any other samples that were misclassified after examining the detailed sample names and other metadata. The available glacial forefields included samples from Greenland, Norway, and Sweden but none from Antarctica or more temperate latitudes. Samples that were collected from the target environments but then enriched or manipulated in the laboratory before sequencing were also excluded. Nine additional samples from the cryosphere and glacial forefields that became available by August 2023 were also included. This resulted in 855 metagenomes spanning all regions of the globe, including polar, temperate, and tropical latitudes (Fig. 1A).

Key sample metadata (such as sampling date, geographic coordinates, sequencing and assembly methods, sequencing results, and

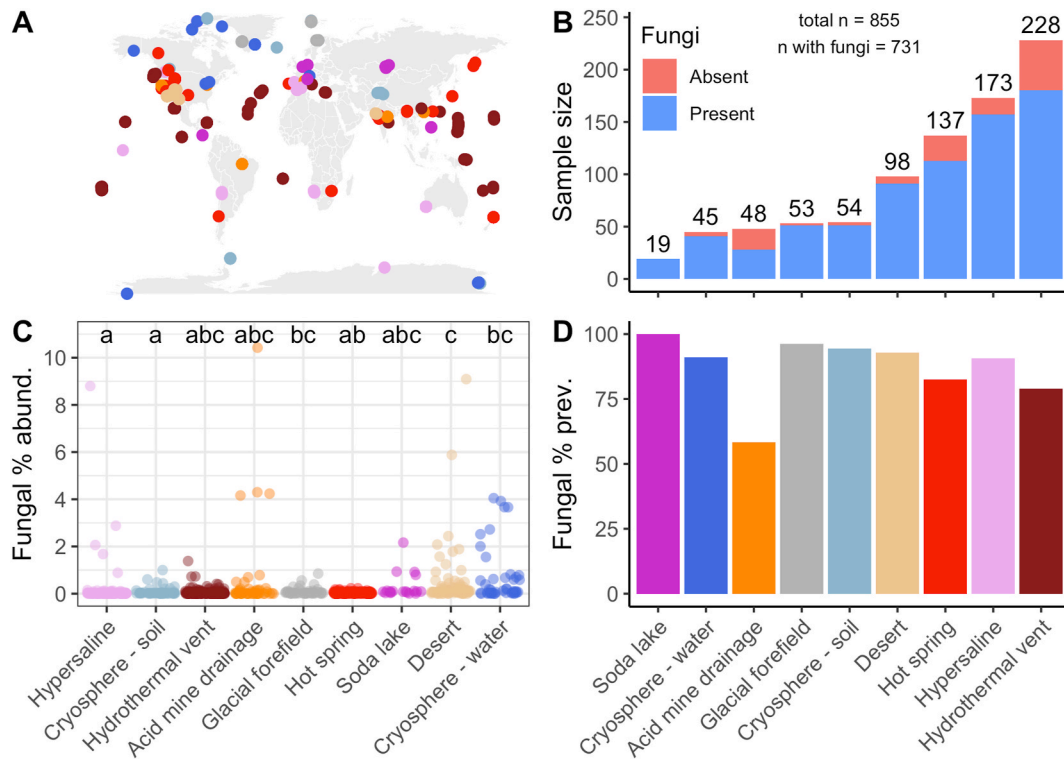


Fig. 1. Sample map (A), sample size (B), fungal percent relative abundance (C), and fungal prevalence (D) for the 855 analyzed metagenomes from various extreme environments. Different letters in (C) represent significant pairwise differences among environments (zero-inflated beta regression, $p < 0.05$). Note the order of the x-axis differs between panels; it is sorted by sample size in panels B and D and by median fungal relative abundance in panel C. Colors in panel A match those in panels C and D.

whether data were available for public use) were downloaded from IMG/M by selecting the appropriate columns (Table S1). The environment type given to each sample was curated such that Antarctica desert samples were classified as cryosphere soils, cryoconite samples were classified as “cryosphere - water” unless that sample explicitly stated that it was the sediment that was sequenced in which case it was classified as “cryosphere - soil”, and samples from solar salterns that had basic pH were classified as hypersaline instead of soda lake.

Taxonomic relative abundance profiles at the genus level were downloaded from IMG/M using the “Gene counts” download option in the Statistical Analysis tool (Chen et al., 2021). Coverage information was not available for all the metagenomes, so the “Gene copies” method was not used. The IMG/M annotation pipeline annotates protein coding genes with taxonomy with the IMG-NR reference database and lastal 983 for alignment. The version number of the IMG/M pipeline when available is given in Table S1. Each new version of the IMG/M pipeline uses an updated version of the IMG-NR reference database, which currently contains 446 fungal isolate genomes (Table S2). Scaffolds are assigned the lineage of the lowest taxonomic rank with at least 50% of the best hits (minimum 30% identity, minimum 80% alignment length) of the protein coding sequences on that scaffold, and then all protein coding genes on the scaffold receive the same taxonomic assignment. This is IMG/M’s equivalent to the “high confidence BLAST hit” used by the J. Craig Venter Institute (Tanenbaum et al., 2010). The resulting taxonomy table was examined for consistency, and any genera that had differing higher level taxonomy were all given the same most recent higher-level taxonomy. Furthermore, taxa that had been renamed since their annotation were renamed according to the most recent version of the IMG-NR reference database (Table S2). Taxa that were annotated by a previous database version but were no longer in the current database that had not been renamed were deleted from the analysis, as they were not complete genomes and this was rightfully updated by the IMG/M staff. Sequences that were unknown at the domain level, or classified as

viruses or plasmids, were filtered from the dataset. The number of protein coding sequences with a taxonomic annotation in the resulting table averaged $361,620.7 \pm 195,12.19$ standard error. The total number of sequences with an assigned genus was strongly correlated with the assembled metagenome size (number of base pairs) (linear regression $R^2 = 0.86$), while the number of sequences assigned to fungal genera was only weakly correlated with assembled metagenome size (linear regression $R^2 = 0.13$). Fungal functional relative abundance profiles according to KEGG Orthology (KO) (Kanehisa et al., 2016) were acquired by first using a custom Python script (Wu, 2023) to retain only scaffolds with fungal taxonomic assignment, and then summing assigned KOs by K number for each sample to acquire KO counts. KO counts were normalized using the *DESeq2* R package (Love et al., 2014). Fungal KO counts from metagenomes on IMG/M suffer from biases because eukaryotic gene callers were not used in the annotation pipeline. For example, some protein families are inflated due to gene fragmentation, while others fail to be detected. Biases are stronger for filamentous fungi due to intron-exon structure, and less for yeast. Thus, KO absences and relative abundances in fungi should be interpreted with caution; however, we still make comparisons among the environment types, assuming if a KO is identified with IMG/M methods in one sample it should be identified in all samples in which it is truly present.

Zero-inflated beta regression, implemented with the *gamlss* R package (Rigby and Stasinopoulos, 2005), followed by Tukey posthoc tests, implemented with the *emmeans* R Package (Lenth, 2021), was used to test for the effect of environment type on total fungal relative abundances (calculated as the counts of taxa assigned to a fungal phylum divided by the total counts with assigned taxonomy), which are constrained between 0 and 1 and did not follow a normal distribution. Kruskal-Wallis tests followed by Nemenyi posthoc tests implemented in the R package *PMCMRplus* (Pohlert, 2022) were used to test for the effect of environment type on fungal alpha-diversity (genus richness and Shannon diversity) as well as taxonomic relative abundances. For each

environment, we examined the most relatively abundant fungal taxa according to counts per million assembled metagenomic base pairs (CPM) at the class level among studies, using the four studies with the highest sample size if there were more than four studies of an environment type. Community composition at each taxonomic level (genus through phylum) was compared across environments by calculating Bray-Curtis and Jaccard dissimilarity matrices from genus-level CPM data for samples containing fungi, and testing for the effect of environment, year, assembler, and latitude with PERMANOVA, implemented in the *vegan* R package (Oksanen et al., 2022). Differences in within-environment variation were tested with PERMDISP, also implemented in *vegan*. Relationships between community dissimilarity and geographic distance were tested with Mantel tests in *vegan*. Indicator species analysis was performed using the *indicspecies* R package with the “r.g” association function and a significance level of 0.05 (De Cáceres and Legendre, 2009). Overall composition of KOs (DESeq2 normalized) was assessed in a similar fashion, but only on samples with a minimum of 750 KOs from the fungal portion of the metagenome. This cutoff removed metagenomes with very few annotated fungal KOs to avoid statistical artifacts, while retaining samples from eight out of nine environment types. A curated list of KOs involved in various forms of stress tolerance was extracted from the literature (Baeza et al., 2021; Furness, 2020; Liu et al., 2020), and mean relative abundances per environment were plotted as heatmaps with hierarchical clustering. Heatmaps were made with the *heatmap* R package (Kolde, 2019), while all other figures were generated with the *mctoolsr* (Leff, 2022), *ggplot2* (Wickham, 2016) and/or *cowplot* (Wilke, 2020) R packages. All analyses were performed with R version 4.2.3 (R Core Team, 2023). Data and analyses are available on GitHub (<https://doi.org/10.5281/zenodo.10642140>).

3. Results

Of the 855 samples that met requirements for inclusion in our analysis, 731 contained fungi (Fig. 1b). Fungal prevalence (presence/absence) among environment types ranged from 58% (acid mine drainage) to 100% (soda lakes) (Fig. 1d). Relative abundance of fungal sequences among assembled contigs across all of the metagenomes ranged from 0 to 10.4% (mean 0.18% \pm 0.03% SE). Fungi had significantly higher relative abundances in cryosphere waters, deserts, and glacial forefields, compared to cryosphere soils and hypersaline environments (Fig. 1c). Assembled metagenome sizes ranged from a minimum of 380,429 base pairs to over 1 billion base pairs in some samples

(Fig. S1).

The mean number of fungal-assigned protein coding sequences per sample was 555.6 ± 63.5 standard error. There were 293 known fungal genera identified across the whole dataset. The number of fungal genera identified per sample with the methodology used ranged from 0 to 289 (mean = 47 ± 2 SE). Hypersaline and acid mine drainage environments had the lowest number of genera on average (Fig. 2a). Fungal Shannon diversity calculated at the genus level ranged from 0 to 5 and followed a similar trend as genus richness, with lowest diversity in hypersaline and acid mine drainage samples (Fig. 2b).

Environment significantly affected fungal taxonomic composition at all taxonomic levels (PERMANOVA, $p = 0.001$). At the genus level, environment explained 15% of the variation (Fig. 3a; PERMANOVA pseudo- $F = 17.9$, $R^2 = 0.15$, $p = 0.001$). However, there was a considerable amount of variation within environment types as well, related to differences in geography and conditions as well as methodology. Sequencing year and metagenomic assembly method also had significant effects on composition and explained a substantial amount of variation (PERMANOVA $R^2 = 0.11$ and 0.06 , respectively), but these effects were separate from the environment effect. There was also a significant effect of latitude on composition (PERMANOVA pseudo- $F = 6.6$, $R^2 = 0.01$, $p = 0.001$). Within environments, composition was significantly affected by sampling location. For example, there were four different glacial forefields sampled, and location significantly affected glacial forefield fungal community composition (PERMANOVA, pseudo- $F = 2.9$, $R^2 = 0.32$, $p = 0.001$). The amount of variation within environments was not homogeneous (PERMDISP, $F = 13.6$, $p < 0.001$). The mean distance between each sample and the group centroid for each environment was greatest in hypersaline environments and lowest in cryosphere soils (Table 2). Fungal communities were more dissimilar between and within environment types compared to both bacteria and archaea (Fig. 3b). There was slightly more clustering among archaeal and bacterial communities as well, but still a high degree of within-environment variation, with mean Bray-Curtis dissimilarities above 0.4 even for within-environment pairwise comparisons (Fig. 3b, Fig. S2). Fungal Bray-Curtis dissimilarity was significantly positively correlated with geographic distance (Mantel test, $r = 0.12$, $p < 0.001$), more so than archaeal ($r = 0.08$, $p < 0.001$) and bacterial ($r = 0.02$, $p = 0.04$) communities.

Ascomycota was the dominant fungal phylum in all environment types, accounting for 0.1 to 7 CPM on average. Other phyla—Basidiomycota, Blastocladiomycota, Chytridiomycota, Microsporidia, Mucoromycota, and Zoopagomycota—were present in most

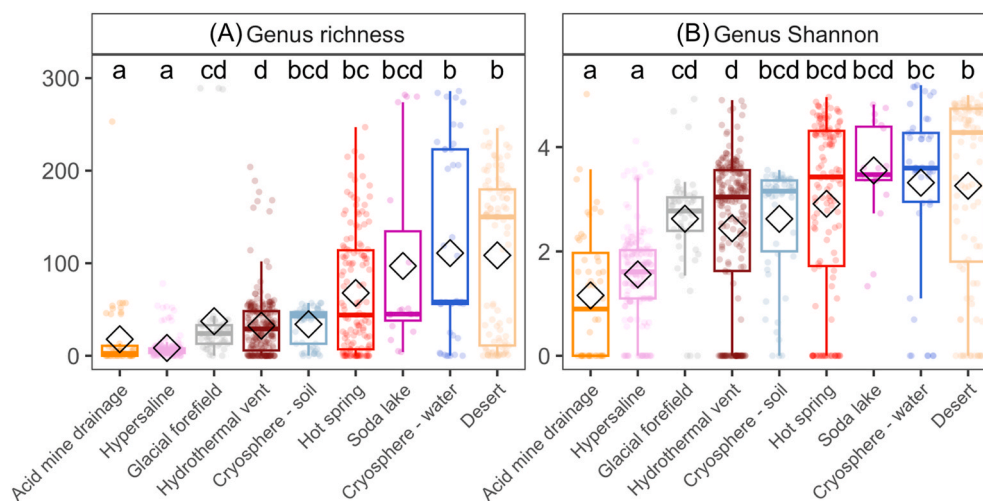


Fig. 2. Fungal alpha-diversity showing genus richness (a) and genus Shannon diversity (b). Black diamonds represent means. Different letters represent significant differences (Nemenyi posthoc test, $p < 0.05$). Note the difference in the y-axis between panels. The x-axis order is the same in both panels and is ordered by increasing median genus richness.

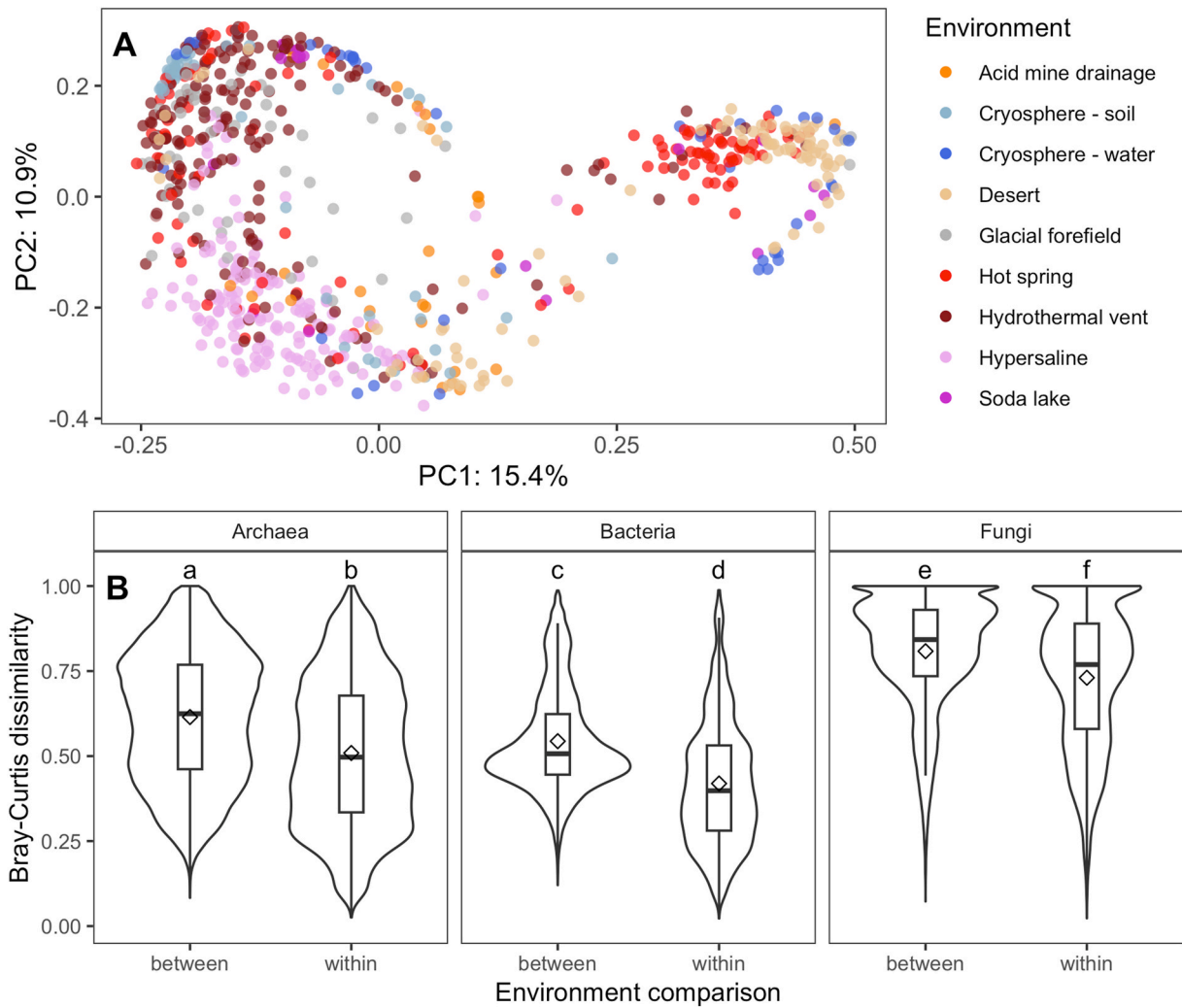


Fig. 3. Principal coordinates analysis of fungal genus Bray-Curtis dissimilarity (A) and Bray-Curtis dissimilarity of archaeal, bacterial, and fungal communities at the genus level for pairwise comparisons between and within environments (B). In panel B, the datasets were filtered to 718 samples containing archaea, bacteria, and fungi (257403 pairwise comparisons). In panel B, diamonds represent the means, and different letters represent significant differences in means (Tukey posthoc, $p < 0.05$).

Table 2

PERMANOVA and PERMDISP results for fungal genus-level taxonomic composition (Bray-Curtis dissimilarity). Df = degrees of freedom, SS = sum of squares, GD = mean group distance to centroid.

Test	Variable	Df	SS	R ²	F	Pr(>F)	Symbol
PERMANOVA	Environment	8	35.86	0.15	18.8	0.001	***
	Year	10	26.30	0.11	11.0	0.001	***
	Assembler	13	13.83	0.06	4.5	0.001	***
	Latitude	1	1.64	0.01	6.9	0.001	***
	Residual	696	166.36	0.68			
	Total	728	244.00	1			
PERMDISP	Environment	8	2.91	0.36	13.6	<0.001	***
	Residuals	722	19.35	0.03			
	Environment	GD					
PERMDISP	Hypersaline	0.61					
	Acid mine drainage	0.59					
	Hot spring	0.50					
	Cryosphere - water	0.50					
	Desert	0.50					
	Soda lake	0.47					
	Glacial forefield	0.47					
	Hydrothermal vent	0.46					
	Cryosphere - soil	0.41					

environment types but in lower relative abundances (Fig. S3a). Relative abundances of all seven phyla differed significantly across environments (Table S3). The most relatively abundant 15 classes across the whole dataset were Agaricomycetes, Basidiobolomycetes, Blastocladiomycetes, Chytridiomycetes, Dothideomycetes, Eurotiomycetes, Glomeromycetes, Leotiomyces, Mortierellomycetes, Mucoromycetes, Neocallimastigomycetes, Pezizomycetes, Saccharomycetes, Sordariomycetes, and Xylonomycetes. These 15 classes made up the vast

majority of the fungal community on average in each environment type (~90–99%) (Fig. S3b). Within Ascomycota, at the class level there were differences in the most relative abundant taxa among environment types (Fig. S3b). Relative abundances of all 39 classes differed significantly across environments (Table S4).

To examine general trends in taxonomic groups associated with different extreme environments across sites, we examined class-level fungal relative abundances in metagenomes from specific studies with

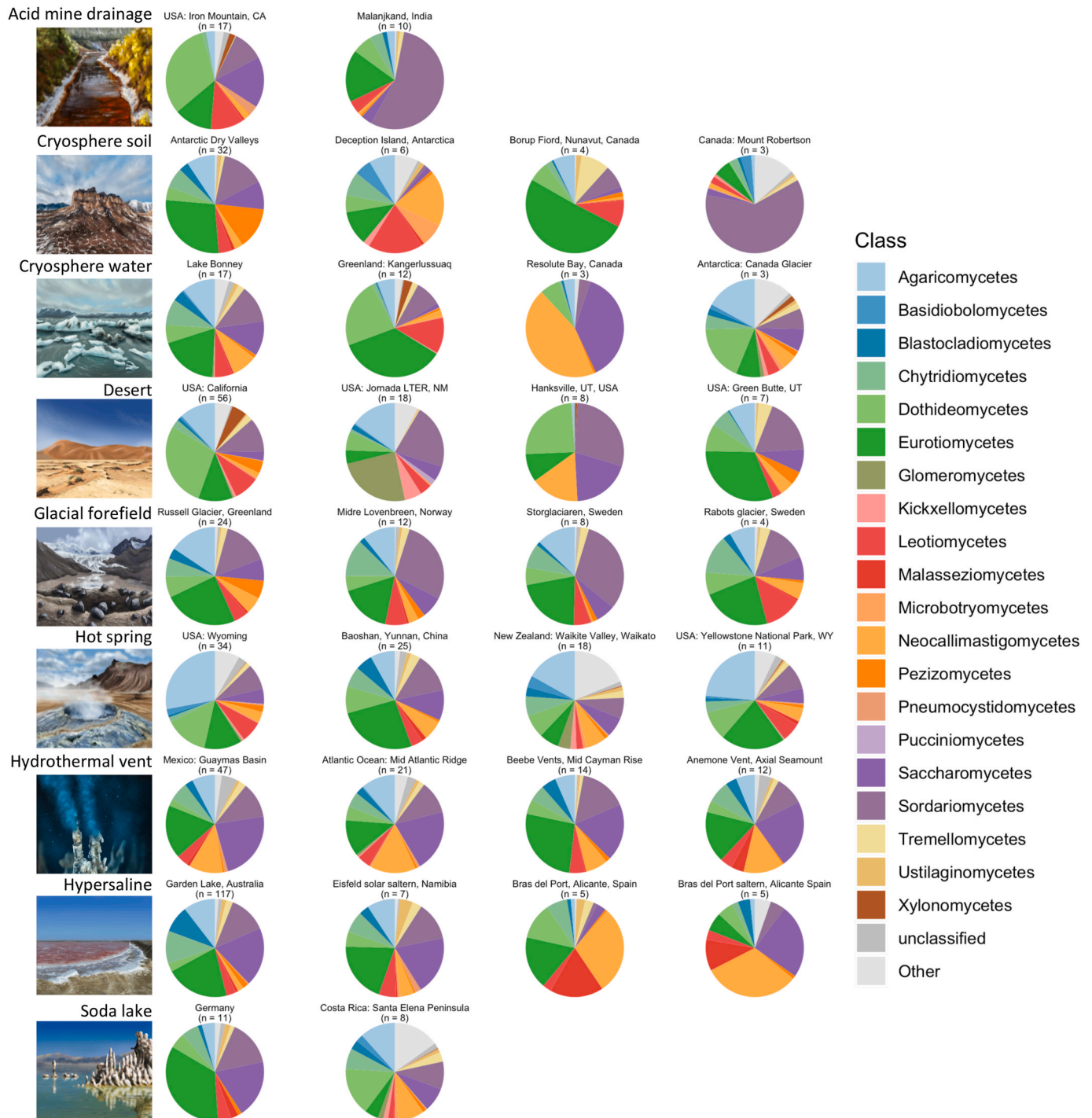


Fig. 4. Relative abundances of fungal classes among different studies in each environment. Each pie shows the mean relative abundance of classes in an individual study, with the location and number of samples in that study stated above the pie. For glacial forefields, the four pies come from the same study but are separated by location. Of the 37 fungal classes (including “unclassified”) present in the data, 20 with at least 25% relative abundance of fungi in at least one individual sample in the whole dataset are shown here, with all others aggregated into the “Other” category. For environments with more than 4 studies, the 4 with the greatest number of samples were chosen. Studies with only one sample are not included here.

the greatest number of samples in each environment type. There were significant differences in the most relatively abundant fungal classes across environments, within each environment among different studies, and between geographic locations in each environment (Fig. 4). Top fungal classes in acid mine drainage sites varied from Dothideomycetes in California (USA) to Sordariomycetes in Malanjand (India). Top classes in cryosphere soils ranged from Eurotiomycetes in Nunavut (Canada) to Sordariomycetes in Mount Robertson (Canada), to more evenly mixed communities in Antarctica. Cryosphere water samples from Kangerlussuaq (Greenland) were dominated by Eurotiomycetes and Dothideomycetes and those from Resolute Bay (Canada) were dominated by Saccharomycetes and Neocallimastigomycetes, while samples from Lake Bonney in Antarctica had diverse and even communities where all of these were represented. Deserts were dominated by Eurotiomycetes, Sordariomycetes, and Dothideomycetes, as well as Glomeromycetes in New Mexico (USA). Sordariomycetes were also key components of glacial forefields, particularly at Midre Lovénbreen (Norway), and Storglaciären (Sweden), while Eurotiomycetes was the most relatively abundant fungal class at the Russell Glacier in Kangerlussuaq (Greenland) (similar to the cryosphere water samples there) and Rabots Glacier (Sweden). Common dominant fungal taxa in hot springs were Agaricomycetes and Eurotiomycetes, as well as several other taxa that differed among the sites. Saccharomycetes were key components of the top four most sampled hydrothermal vent sites, along with Eurotiomycetes, Sordariomycetes, and Malasseziomycetes. Saccharomycetes were also a dominant fungal class in hypersaline environments, except in one of the studies of salterns in Bras del Port (Spain), which were dominated by Malasseziomycetes and Pezizomycetes. Lastly, among soda lakes, sites in Germany were dominated by Eurotiomycetes and those in Costa Rica had an even mix of 20 classes (Fig. 4, Table S5).

Several fungal families were identified as indicators of specific environments, especially soda lakes (62 indicators), followed by cryosphere water samples (22 indicators) with both including a phylogenetically diverse range of families (Fig. S4). Other indicator taxa included Chaetomiaceae, Cordycipitaceae, Nectriaceae, Plectosphaerellaceae, Podosporaceae, and Pyriculariaceae in acid mine drainage

sites, Ascobolaceae, Diademaceae, and Morchellaceae in deserts, and Psathyrellaceae in glacial forefields (Fig. S4).

The total number of KOs annotated in fungal-assigned scaffolds was 6160. The number of fungal KOs per sample ranged from 1 to 2,441, with a mean of 146 ± 14 SE. The number of fungal KOs per metagenome was closely and positively related to the number of fungal-assigned protein coding sequences per metagenome (second order polynomial regression, $p < 0.001$, $R^2 = 0.95$, Fig. S5), suggesting many of the metagenomes were sequenced at a depth too low for KO predictions to accurately represent the community. Because of this, samples with at least 750 fungal KOs ($n = 40$, containing all environments but hot springs) were chosen to examine further. Environmental differences in functional composition (Fig. 5, PERMANOVA $R^2 = 0.34$, $p = 0.001$ and $R^2 = 0.31$, $p = 0.001$ for Bray-Curtis and Jaccard, respectively) were less pronounced than differences in taxonomic composition when considering the same subset of 40 samples (Fig. S6, PERMANOVA $R^2 = 0.53$, $p = 0.001$ and $R^2 = 0.59$, $p = 0.001$ for Bray-Curtis and Jaccard, respectively; note that this differs from Fig. 3 which used the whole dataset), and were mostly driven by presence versus absence of KOs rather than by differences in relative abundance (Fig. 5). Clustering was similar when using Bray-Curtis and Jaccard metrics, as was the variation explained by the first two axes and by environment type (PERMANOVA $R^2 = 0.34$ and 0.31 , respectively).

KOs involved in fungal responses to a variety of stresses (Table S6) tended to be more relatively abundant, on average, in soda lake, cryosphere water, and acid mine drainage samples compared to samples from the other six environments (Fig. 6). Key exceptions to this trend were relatively abundant *otsB* trehalose-6-phosphate phosphatases in deserts, relatively abundant *SHO1* osmosensors in cryosphere soils, and relatively abundant *HSPA1* heat shock proteins and *ACE2* in hydrothermal vents (Fig. 6).

4. Discussion

In addition to archaea and bacteria, extreme environments are home to eukaryotes including fungi, which are increasingly recognized as key

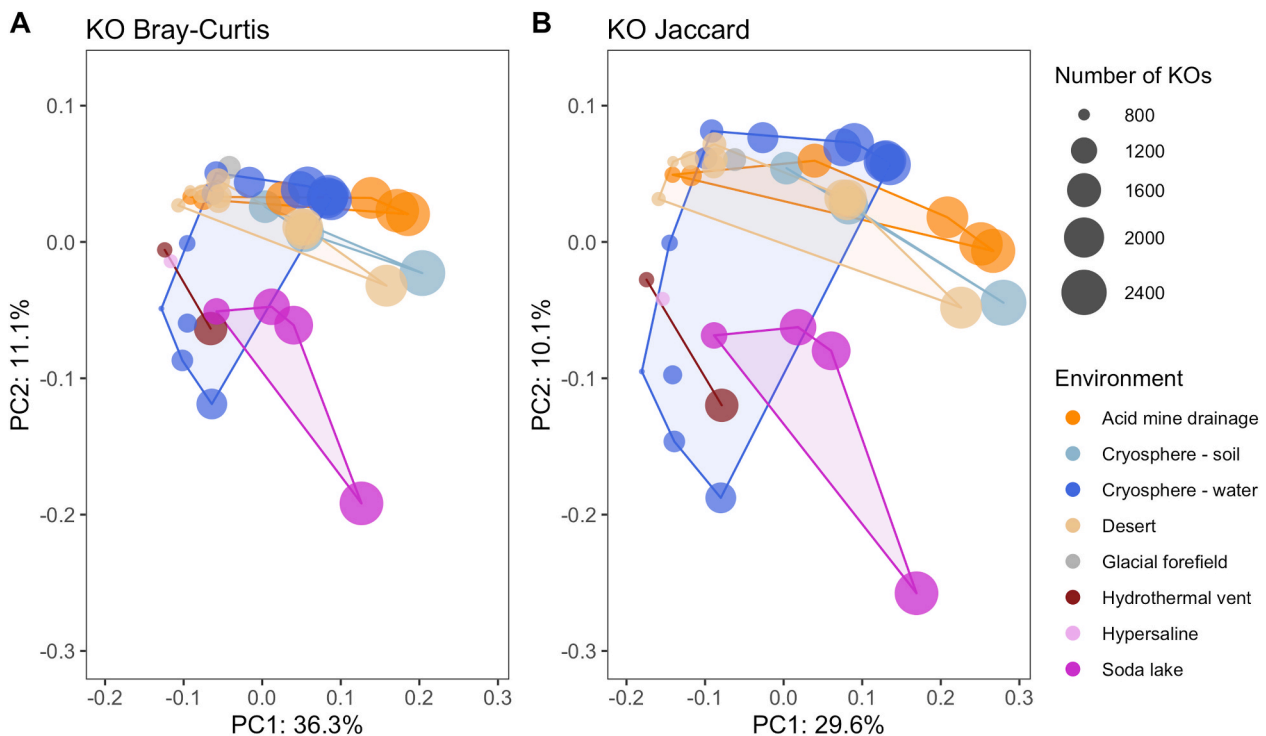


Fig. 5. Principal coordinates analysis of fungal KO Bray-Curtis dissimilarity (A) and Jaccard dissimilarity (B). Only samples with >750 KOs were analyzed ($n = 40$). Point sizes correspond to the number of KOs. Axes have the same scale in both panels.

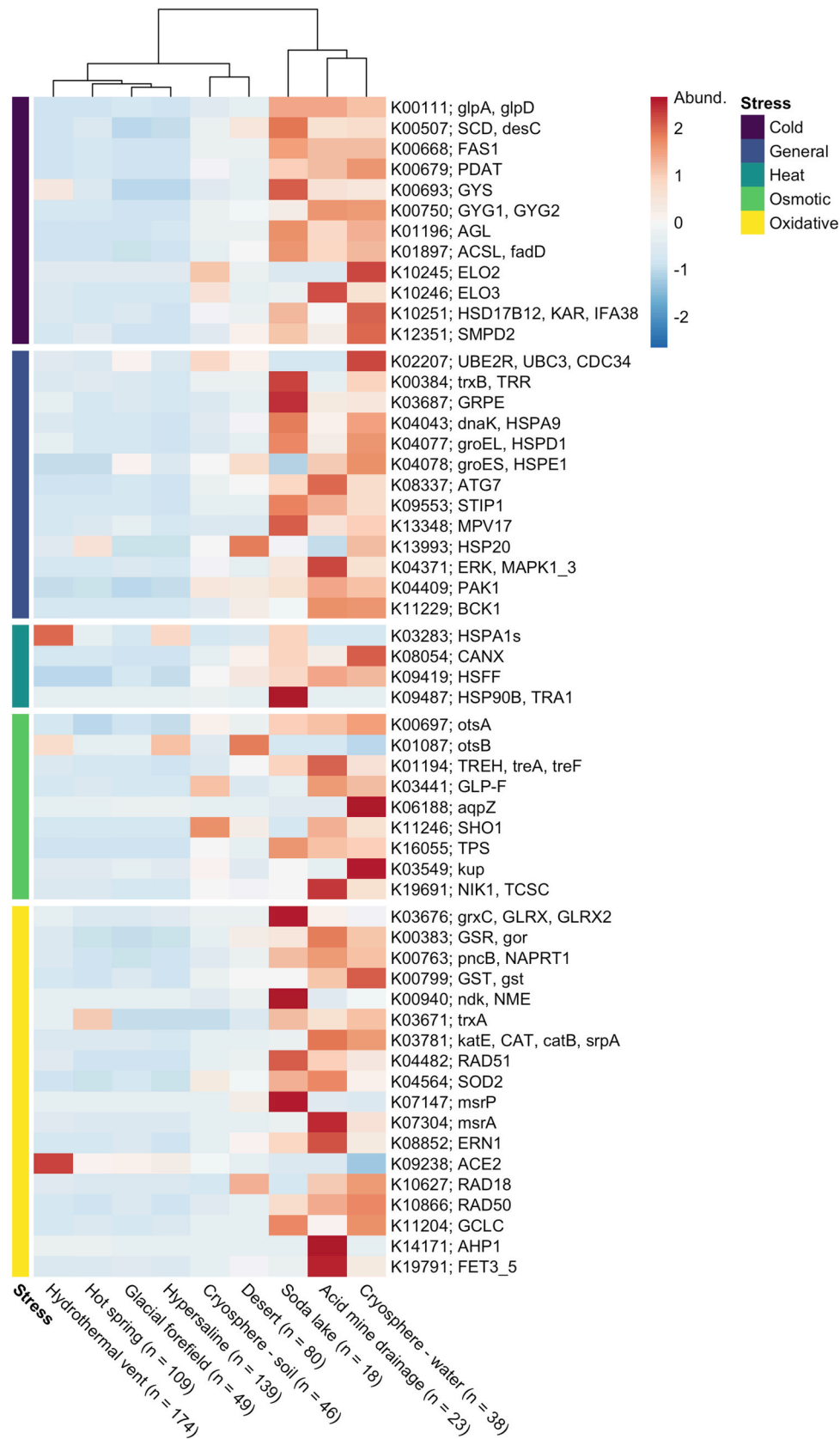


Fig. 6. Heatmap of KO relative abundances for KOs involved in stress tolerance for 676 metagenomes with fungal KOs. Relative abundances are z-scores calculated from DESeq-normalized counts. The annotation row color bar shows the type of stress. Columns are clustered according to Ward's D² hierarchical clustering method.

players in those environments (Coleine et al., 2022; Gostinčar et al., 2023). The diversity of fungal genera found in these environments was surprisingly high and is similar to more benign environments. For example, developed soils in alpine tundra on Niwot Ridge, Colorado, USA, a seasonally cold environment but not as extreme as the cryosphere soils or glacial forefield soils analyzed in our study, have similar genus richness based on ITS sequences, in the range of a couple hundred genera (Porazinska et al., 2018). The diversity of dominant taxa in extreme environments is also supported by the indicator taxa results. Although there were some instances of indicators of just one environment type, within environments such as soda lakes and cryosphere soils, there were many different indicator taxa from diverse lineages. Historically, the study of extremophilic fungi has focused on model organisms (Onofri et al., 2020; Zajc et al., 2014). Given the diversity found in our survey and the low relative abundance of model organisms, some of these taxa may not be as relevant for a comprehensive understanding of extremophilic fungal diversity and functioning. However, model organisms remain useful for experimental manipulations and detailed comparative genomic and transcriptomic studies. It is also important to note that not all extremophilic fungal taxa (Coleine et al., 2022) have genomes that are publicly available and present in the IMG-NR reference database (Table S2).

Our analysis enables some inferences to be made about the relative roles of deterministic versus neutral processes in fungal community assembly. It is hypothesized deterministic processes should dominate in extreme environments, while neutral processes should be relatively more important in more benign environments (Lemoine et al., 2023). Fungal communities were more dissimilar within environment type than archaeal and bacterial communities. This pattern is surprising given the specific habitat preferences found for bacteria (Caporaso et al., 2011; Delgado-Baquerizo et al., 2018) and protists (Oliverio et al., 2020). Major fungal classes differed significantly among best sampled study sites within a given environment (Fig. 4). This result suggests that even within the realm of “extreme environments”, there is still a high degree of environmental variability within environment types, although we did not quantify such differences due to a lack of shared measured variables across studies. Consequently, such variability could lead to the observed dominance of different taxa depending on local site conditions even within the same environment type. Alternatively, neutral processes could also generate such patterns. When considering these results with the statistically significant correlation between fungal community dissimilarity and geographical distance, there could be a higher degree of endemism or perhaps dispersal limitation in fungi compared to bacteria and archaea. Previous studies have shown that fungal communities in extreme polar sites are both habitat- and geographically-patterned (Zhang et al., 2020, 2021). Dispersal was also found to dominate bacterial community assembly in the extreme environment of the Shackleton Glacier in Antarctica (Lemoine et al., 2023). There is still some debate as to whether fungi or bacteria are more dispersal limited. It is hypothesized that fungi are more dispersal limited due to their larger spore sizes, but there is only mixed support for this hypothesis (Bahram et al., 2018; J. Chen et al., 2020; Xiao et al., 2018), likely due to variation in dispersal ability among bacterial taxa and among fungal taxa, which would prevent such a broad generalization from being made.

Fungi that thrive in extreme conditions can be divided into 1) ubiquitous and polyextremotolerant generalists and 2) rarely isolated specialists with narrow ecological amplitudes (Gostinčar et al., 2022). Some generalists also have a high degree of recombination and dispersal ability, resulting in homogeneous populations with no obvious specialization for any of the diverse habitats where they are found (Gostinčar et al., 2019). In our analyses, there was a lot of taxonomic overlap, more so than in archaea and bacteria, yet fungi were still more dissimilar at the entire community level. Some of the same fungal taxa were found in multiple extreme environments (Table S7), suggesting many generalists withstanding a broad spectrum of environmental stress are thriving. For example, Agaricomycetes were in the top 10 most relatively abundant

fungal classes (of 36 total classes identified) in all 9 environment types; this class was previously shown to dominate soils of most biomes globally (Tedersoo et al., 2014). Agaricomycetes is also the most speciose fungal class, with over 40,000 described species.

At the genus level, we found that generalists were often the most relatively abundant fungal taxa in extreme environments. *Aspergillus* and *Talaromyces* were among the 10 most relatively abundant genera detected in all or almost all of the environments analyzed (Table S5, Table S7), making them the most versatile genera in terms of adaptation to several different extreme environments, or the most easily dispersed genera and represent dormant spores in these environments. This demonstrates a potential high degree of plasticity and/or dormancy in these two genera, which have been found in space stations (Makimura et al., 2001; Romsdahl et al., 2018), hydrothermal vents, and hot springs (Boruta, 2018), and in the case of *Aspergillus* spores, can withstand extreme amounts of X-ray and UV radiation (Cortezão et al., 2020). This may explain why they are found in so many environments, but more work is needed to determine if they are capable of active growth in all the extreme environments in which they are detected with DNA-based methods. These two genera are members of a lineage known for their prolific production of wind-dispersed, asexual spores which are nearly ubiquitous in air and could be overrepresented in these extreme metagenomes for that reason. However, some species of *Aspergillus* and *Talaromyces* possess extremophilic adaptations and might be capable of growth in those environments. *Aspergillus penicillioides* has been shown to have adaptations to low water environments through regulating intracellular glycerol and the MAPK (mitogen-activated protein kinase) pathway, while *Aspergillus sydowii* has adaptations for saline environments that include unsaturated phospholipids, ergosterol synthesis, and compatible solute production (Coleine et al., 2022). Furthermore, *Aspergillus fumigatus*, *Talaromyces emersonii*, and *Talaromyces marnefei* each express multiple heat shock proteins when exposed to heat (Chen and Chen, 2004; Pongpom and Vanittanakom, 2016), supporting the finding that *Aspergillus* and *Talaromyces* were some of the most relatively abundant fungal genera in hot springs.

Fusarium is another generalist taxon that was among the most relatively abundant fungal genera detected in five of these extreme environments; in particular, it was the most relatively abundant fungal genus identified in acid mine drainages, and glacial forefields. *Fusarium* has been reported from Arctic and Antarctic soils, cryoconite holes, desert sands and rocks, hypersaline soils, and acid mines using culture-based and amplicon sequencing methods (Ameen et al., 2022; Bridge and Spooner, 2012; Gonçalves et al., 2016; Gross and Robbins, 2000; Mandeel, 2006; Murgia et al., 2019; Rathore et al., 2022; Ye et al., 2020; Zhang et al., 2021). *Fusarium* (as well as *Aspergillus*) isolates have been shown to grow in culture at pH 2 (Wheeler et al., 1991). Some *Fusarium* species have been shown to contain proton pumps and hydrogen transporting ATPases that are important adaptations for acidic environments (Brandão et al., 1992; García-Martínez et al., 2015). Since many of these samples are likely low in fungal biomass, it remains to be seen if the metagenomic *Fusarium* sequences represent functioning populations or dormant cells deposited from the atmosphere (Schmidt et al., 2017). Research on extremophilic fungi has so far focused on isolated specialists that withstand narrow ecological amplitudes, giving less attention to the broader generalists (such as *Aspergillus*, *Talaromyces*, and *Fusarium*) that our work suggests often make up most of the fungal sequences recorded in extreme environments.

Research in some of the most extreme high-elevation soil habitats on Earth has shown that a generalist polyextremophilic basidiomycete yeast (*Naganishia friedmannii llullensis*) has adapted using an opportunistic strategy in which cells lay dormant for long periods of time and then grow rapidly during brief periods of water availability by using a broad range of organic compounds and tolerating freeze-thaw cycles and UV radiation (Schmidt et al., 2017; Solon et al., 2018; Vimercati et al., 2016). It remains to be seen if *Aspergillus*, *Talaromyces*, and *Fusarium* are capable of employing a similar opportunistic strategy in the extreme

environments studied here, or if they are just dormant transients from the atmosphere. Notably, it has been suggested that *Talaromyces marnefei* can activate pathways to use alternative carbon sources other than glucose, which could be useful in oligotrophic environments (Pongpom and Vanittanakom, 2016). This deserves further investigation, especially because these genera are often not found to be the most relatively abundant fungal genera in marker gene (e.g., 18S rRNA gene, ITS) surveys of environments such as glacial forefields. For example, research at the forefield of the receding Marr Ice Piedmont in the Antarctic peninsula using 18S rRNA gene sequencing found that *Alatospora* (Leotiomycetes), *Protomyces* (Taphrinomycetes), and *Mrakia* (Tremellomycetes) were the most relatively abundant fungal genera early in the chronosequence, while *Aspergillus*, *Talaromyces*, and *Fusarium* were not detected (Vimercati et al., 2022). A recent ITS survey of a glacial forefield in Svalbard did not detect *Talaromyces* or *Fusarium*, but did detect a very low relative abundance of *Aspergillus* (Trejos-Espeleta et al., 2024). While shotgun metagenomic/metatranscriptomic and marker gene methods can sometimes be in agreement (Rivera Pérez et al., 2022), discrepancies between our metagenomic results and marker gene survey results could include differences in the taxonomic database and method of taxonomic assignment, the resolution of taxonomic assignments (assignments from marker genes are often not to genus level), primer biases (Tedersoo et al., 2015; Tedersoo and Lindahl, 2016), as well as other differences in the samples even if they were taken at the same site. Although the ITS gene is among the best methods for classifying fungal taxa to the genus level, a substantial number (24%) of amplicon sequence variants lacked genus-level taxonomy in a recent survey of rhizosphere soils (Bueno de Mesquita et al., 2024), and the relative abundance of unassigned taxa even at higher levels such as the order level can be as high as 50% (Pellitier et al., 2019). Advances in long-read sequencing technology are making improvements in fungal taxonomic assignment (Tedersoo et al., 2020).

As far as more specialized taxa worth highlighting, two of the top five genera from soda lakes are amphibian-associated taxa (*Basidiobolus* and *Batrachochytrium*), which may represent influx from the surrounding environment or amphibians that died in the lakes, as soda lakes are not known to house amphibians due their extreme salinity and high pH (Matagi, 2004). *Malassezia* was the most relatively abundant fungal genus in hydrothermal vents. While yeasts in this genus were originally described from mammalian skin, *Malassezia* spp. are now understood to occupy a diversity of habitats especially in marine ecosystems (Amend, 2014). Our results are in line with recent research showing that *Malassezia* is widespread in various marine environments (Steinbach et al., 2023). Cryosphere water and hot springs had multiple zoosporic taxa in their top five genera, including *Spizellomyces* and the anaerobic, herbivore-associated *Neocallimastix*. Other work in high alpine ecosystems that receive winter snowfall (including periglacial soils) (Freeman et al., 2009; Schmidt et al., 2012), as well as aqueous environments (Fuller and Jaworski, 1987) have shown dominance among zoosporic fungi and their ability to tolerate multiple extreme conditions (Gleason et al., 2010). Other high-throughput sequencing studies have reported Neocallimastigomycota fungi from arctic sea ice, sediments (Hassett and Gradinger, 2016) and lakes (Marchetta, 2022). It is also worth noting that (unlike *Aspergillus*, *Beauveria*, *Fusarium*, and *Talaromyces*) these zoosporic fungal sequences are unlikely to be the result of wind-deposited spores (Ingold, 1971) suggesting they may represent unstudied parts of these extreme environments' communities. Alternatively, they may simply represent influx from the excrement of migratory, herbivorous birds, which have been documented to harbor Neocallimastigomycota (Mahtab et al., 2024). Either way, future research should clarify how these and other environmentally isolated sequences are related to sequenced gut Neocallimastigomycota to see if they actually represent a novel, free-living lineage (Picard, 2017).

We used a focused approach to examine relative abundance trends in a small subset of key genes related to general, cold, heat, osmotic, and oxidative stress (Table S6). The functional gene analysis highlighted

some key adaptations that fungi have in these environments and also demonstrated a potential trend of generally greater relative abundances of extreme-adaptive KOs in soda lakes, acid mine drainage sites, and cryosphere water samples (Fig. 6). There were a few exceptions to this general pattern. In terms of general stress, hydrothermal vents had the highest relative abundance of HSPA1S. This protein is a chaperone that ensures proper protein folding and degradation of improperly folded proteins (Rosenzweig et al., 2019).

In terms of oxidative stress, hydrothermal vents had the highest relative abundance of ACE2. ACE2 is an activator of metallothionein expression (Butler and Thiele, 1991), whose role is linked to metals and oxidative stress (Ragasa et al., 2021). The presence of many trace metals in hydrothermal vents may explain the high relative abundance of ACE2. Soda lakes, on the other hand, had the highest relative abundance of nucleoside-diphosphate kinase (NDPK, K00940), another key player in signal transduction pathways involved in oxidative stress (Otero, 2000). Studies have shown that members of the NDPK family reduce genome instability caused by oxidative damage (C.-W. Chen et al., 2020), suggesting that oxidative stress in soda lakes may trigger genome instability in fungi.

Among genes associated with osmotic stress, aquaporin Z was most relatively abundant in cryosphere water, which might be vital for survival and adaptability of polyextremotolerant species in water-challenged environments (Gostiñcar et al., 2014). Aquaporins from fungal species that inhabit various extreme habitats have not been studied and might have either high efficiency or unique mechanisms of regulation that are still awaiting discovery (Xu et al., 2013). Similarly, KUP potassium transporter was also most relatively abundant in cryosphere water samples; studies suggest a positive correlation between capacity for incremented K^+ levels and survival advantages (Rajagopal et al., 2023), which may be especially relevant in environments exposed to frequent environmental fluctuations. otsB, which encodes for a gene involved in trehalose metabolism, was most relatively abundant in desert samples. Trehalose plays an important role in the adaptation of desert-derived fungi to environmental challenges including high temperature, freezing, dehydration, desiccation, and high osmotic pressure (Argüelles et al., 2017; Jiang et al., 2018). Lastly, the osmosensor SHO1 was most relatively abundant in cryosphere soils. Low temperature often imposes osmotic stress on soil microorganisms, providing a potential explanation as to why SHO1 was most relatively abundant in cryosphere soils.

4.1. Caveats

While we assembled a global dataset with a high sample size for this synthesis, there are several important caveats that are worth discussion. Despite an overall sample size of 855, this was dominated by hydrothermal vents and hypersaline environments, while soda lakes were only represented by 19 samples. In addition, geographic distribution within an environment was sometimes also limited, as was the case for glacial forefields, which are only represented in our study by northern hemisphere sites in Sweden, Norway, and Greenland. This hindered our ability to make generalizations about species common to particular stressors/environments (Fig. 4), where examining diversity from few or geographically limited studies may not be as informative.

We analyzed DNA-based data (metagenomes) rather than RNA-based data (metatranscriptomes), which measures the entire pool of DNA, including dead or dormant cells or spores, rather than the exclusively active community. The DNA from dead cells, referred to as "relic DNA" can constitute a major proportion of reads in some cases and can inflate estimates of diversity, especially in low biomass samples (Carini et al., 2016). It remains a challenge to extract enough high-quality RNA from many extreme environments, and there are not yet enough metatranscriptomes available to perform a synthesis of active communities in these environments. We expect this to change over time and would welcome a metatranscriptomic synthesis of the active fungal community

in the future. Furthermore, we only analyzed the assembled portion of the metagenome, which doesn't take into account unassembled reads.

Lastly, while IMG/M is one of the most comprehensive and standardized resources available for comparative metagenomics, functional profiles on IMG/M are biased against eukaryotes including fungi and the IMG-NR taxonomic database does not include reference genomes from all extremophilic fungi model organisms (Coleine et al., 2022). Notable omissions include the Strigulaceae (Dothideomycetes) family and the known extremophilic black yeast family Herpotriciellaceae (Eurotiomycetes) (Selbmann et al., 2020). Therefore, taxonomic results, particularly at the genus level, should be interpreted cautiously. ITS-based surveys are advantageous for classification at this level, and we welcome future syntheses of ITS-based surveys in extreme environments. A eukaryotic gene finder that can be scaled up to analyze metagenomes remains lacking, but improvements are currently being made (Bazant et al., 2023; Levy Karin et al., 2020; Pronk and Medema, 2022). Furthermore, it remains challenging to recover eukaryotic metagenome assembled genomes (or MAGs), which would enable further genomic investigation of fungi in extreme environments and elsewhere (Saraiva et al., 2023; West et al., 2018).

4.2. Conclusions

Despite nearly a century of research documenting the presence of certain fungi from extreme environments, our understanding of what fungal communities in these various extremes tend to look like and how they compare to one another is still limited. Our work suggests that extreme environments are inhabited by diverse groups of fungal taxa, including generalists that have typically not been the focus of previous research on extreme environments. Documenting the high degree of within-environment variation seen in these 855 metagenomic fungal communities suggests other factors may be playing a role in community structure (e.g., endemism, dispersal limitation, relic DNA), or that these environment types are too broad and variable to detect consistently dominant fungal taxa. More work is needed to explore the functional signatures of fungi in extreme environments, and a focus on determining which fungi are active and what they are doing transcriptionally should be the target of future research.

CRedit authorship contribution statement

Clifton P. Bueno de Mesquita: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Lara Vimercati:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Don-gyung Wu:** Software, Methodology, Investigation, Formal analysis, Data curation. **Mary K. Childress:** Writing – review & editing, Investigation, Data curation, Conceptualization. **August Danz:** Writing – review & editing, Investigation, Data curation, Conceptualization. **Arthur C. Grupe:** Writing – review & editing, Investigation, Data curation, Conceptualization. **Danny Haelwaters:** Writing – review & editing, Investigation, Data curation, Conceptualization. **Natalie M. Hyde:** Writing – review & editing, Investigation, Data curation, Conceptualization. **Thiago Kossmann:** Writing – review & editing, Investigation, Data curation, Conceptualization. **Charles Oliver:** Writing – review & editing, Investigation, Data curation, Conceptualization. **Candice Perrotta:** Writing – review & editing, Investigation, Data curation, Conceptualization. **Benjamin D. Young:** Writing – review & editing, Visualization, Investigation. **Steven K. Schmidt:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Susannah G. Tringe:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **C. Alisha Quandt:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation,

Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Natalia Ivanova and Elle Barnes for feedback on our analyses and manuscript. We thank Ylenia Vimercati Molano for illustrating the environment types in Fig. 4. This work was supported by funding from the National Science Foundation (ANT-2133684 to C.A.Q.). D.H. received support from the Czech Science Foundation (grant 21-06446S). The work conducted by the U.S. Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.funeco.2024.101383>.

References

- Akcil, A., Koldas, S., 2006. Acid mine drainage (AMD): causes, treatment and case studies. *J. Clean. Prod., improving environmental, economic and ethical performance in the mining industry. Part 2. Life cycle and process analysis and technical issues* 14, 1139–1145. <https://doi.org/10.1016/j.jclepro.2004.09.006>.
- Ameen, F., AlNadhari, S., Yassin, M.A., Al-Sabri, A., Almansob, A., Alqahtani, N., Stephenson, S.L., 2022. Desert soil fungi isolated from Saudi Arabia: cultivable fungal community and biochemical production. *Saudi J. Biol. Sci.* 29, 2409–2420. <https://doi.org/10.1016/j.sjbs.2021.12.011>.
- Amend, A., 2014. From dandruff to deep-sea vents: *malassezia*-like fungi are ecologically hyper-diverse. *PLoS Pathog.* 10, e1004277 <https://doi.org/10.1371/journal.ppat.1004277>.
- Argüelles, J.-C., Guirao-Abad, J.P., Sánchez-Fresneda, R., 2017. Trehalose: a crucial molecule in the physiology of fungi. In: *Reference Module in Life Sciences*. Elsevier. <https://doi.org/10.1016/B978-0-12-809633-8.12084-9>. B9780128096338120849.
- Baeza, M., Zúñiga, S., Peragallo, V., Barahona, S., Alcaino, J., Cifuentes, V., 2021. Identification of stress-related genes and a comparative analysis of the amino acid compositions of translated coding sequences based on draft genome sequences of antarctic yeasts. *Front. Microbiol.* 12, 623171 <https://doi.org/10.3389/fmicb.2021.623171>.
- Bahram, M., Hildebrand, F., Forslund, S.K., Anderson, J.L., Soudzilovskaia, N.A., Bodegom, P.M., Bengtsson-Palme, J., Anslan, S., Coelho, L.P., Harend, H., Huerta-Cepas, J., Medema, M.H., Maltz, M.R., Mundra, S., Olsson, P.A., Pent, M., Pölme, S., Sunagawa, S., Ryberg, M., Tedersoo, L., Bork, P., 2018. Structure and function of the global topsoil microbiome. *Nature* 560, 233–237. <https://doi.org/10.1038/s41586-018-0386-6>.
- Bazant, W., Blevins, A.S., Crouch, K., Beiting, D.P., 2023. Improved eukaryotic detection compatible with large-scale automated analysis of metagenomes. *Microbiome* 11, 72. <https://doi.org/10.1186/s40168-023-01505-1>.
- Boruta, T., 2018. Uncovering the repertoire of fungal secondary metabolites: from Fleming's laboratory to the International Space Station. *Bioengineered* 9, 12–16. <https://doi.org/10.1080/21655979.2017.1341022>.
- Brandão, R.L., Castro, I.M., Passos, J.B., Nicoli, J.R., Thevelein, J.M., 1992. Glucose-induced activation of the plasma membrane H⁺-ATPase in *Fusarium oxysporum*. *Microbiology* 138, 1579–1586. <https://doi.org/10.1099/00221287-138-8-1579>.
- Bridge, P.D., Spooner, B.M., 2012. Non-lichenized Antarctic fungi: transient visitors or members of a cryptic ecosystem? *Fungal Ecol. Fungi in Extreme Environments* 5, 381–394. <https://doi.org/10.1016/j.funeco.2012.01.007>.
- Bueno de Mesquita, C., Walsh, C., Tarble, Z., Kane, N., Koehler, B., Hulke, B., 2024. Environment, plant genetics, and their interaction all shape sunflower rhizosphere microbial communities. *Appl. Environ. Microbiol.* <https://doi.org/10.1101/2024.08.09.607089>. In review.
- Bueno de Mesquita, C.P., Brigham, L.M., Sommers, P., Porazinska, D.L., Farrer, E.C., Darcy, J.L., Suding, K.N., Schmidt, S.K., 2020. Evidence for phosphorus limitation in high-elevation unvegetated soils, Niwot Ridge, Colorado. *Biogeochemistry* 147, 1–13. <https://doi.org/10.1007/s10533-019-00624-y>.
- Butler, G., Thiele, D.J., 1991. ACE2, an activator of yeast metallothionein expression which is homologous to SWI5. *Mol. Cell Biol.* 11, 476–485. <https://doi.org/10.1128/mcb.11.1.476-485.1991>.
- Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., Knight, R., 2011. Global patterns of 16S rRNA diversity

- growing optimally at salinities above 15% NaCl. *Appl. Environ. Microbiol.* 80, 247–256. <https://doi.org/10.1128/AEM.02702-13>.
- Zhang, T., Wang, N., Yu, L., 2020. Soil fungal community composition differs significantly among the Antarctic, Arctic, and Tibetan Plateau. *Extremophiles* 24, 821–829. <https://doi.org/10.1007/s00792-020-01197-7>.
- Zhang, T., Wang, N.-F., Yu, L.-Y., 2021. Geographic distance and habitat type influence fungal communities in the arctic and antarctic sites. *Microb. Ecol.* 82, 224–232. <https://doi.org/10.1007/s00248-021-01742-7>.