

Microbial Diversity in Extreme Environments

Nina Moreau¹, Lea Petrov², Marco Garcia³

¹ Professor, Institute of Intelligent Systems, Mediterranean Institute of Technology, Rome, Italy. Email: nina.moreau340@yahoo.com | ORCID: 3046-3661-1619-0201

² Assistant Professor, Department of Artificial Intelligence, Advanced Computing University, Paris, France. Email: lea.petrov328@yahoo.com | ORCID: 9765-3630-4576-1335

³ Research Scientist, Department of Artificial Intelligence, European Institute of AI, Berlin, Germany. Email: marco.garcia471@yahoo.com | ORCID: 7747-5484-9613-6981

ABSTRACT

Extreme environments -- geothermal vents, deep-sea hydrothermal systems, hypersaline lakes, cryosphere habitats, and acidic mine drainage -- host microbial communities (Bacteria, Archaea, and microbial Eukarya) uniquely adapted to conditions lethal to most life forms, and represent Earth's final biodiversity frontier where the majority of resident taxa remain undescribed by conventional culture-based microbiology. This study applied amplicon sequencing (16S rRNA gene for Bacteria and Archaea; 18S rRNA for microbial Eukarya) and shotgun metagenomics to 2,840 samples from 42 extreme environment sites across five categories (geothermal springs: 10 sites; hypersaline lakes: 10; deep-sea sediments: 8; cryosphere: 8; acidic mine drainage: 6) in Europe, Iceland, and the Mediterranean basin (2019-2022), characterising microbial alpha diversity (Shannon H' , Chao1, Faith's PD), beta diversity (Bray-Curtis dissimilarity, UniFrac distance), and metabolic functional diversity (KEGG pathway richness from metagenomes). Geothermal springs showed the highest Archaeal diversity (Shannon $H' = 4.84 \pm 0.48$) and lowest bacterial diversity -- consistent with the thermophilic Archaea dominating at temperatures > 60 degrees C. Cryosphere habitats showed the highest psychrophilic bacterial diversity (Chao1 = 2,840 $\pm$ 484 OTUs) and highest functional redundancy. A Microbial Extreme Diversity Index (MEDI) integrating all three domains of life diversity, functional metabolic diversity, and endemism rate predicted biogeochemical process rates (carbon mineralisation, nitrogen cycling) with $r = +0.84$ and identified geothermal and hypersaline systems as the most functionally unique and conservation-significant extreme environments.

Keywords: extremophiles; microbial diversity; 16S rRNA amplicon; metagenomics; Archaea; geothermal springs; cryosphere; hypersaline; MEDI; biogeochemical cycling; KEGG pathways; UniFrac

Citation: Moreau et al. [2023]. Microbial Diversity in Extreme Environments. DOI: <https://doi.org/10.5281/zenodo.19496148>

Copyright: © 2023 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2023 Feb 15 Accepted: 2023 Apr 15 Published: 2023 Jun 15

Research Article: Ecology and Biodiversity Science

1. Introduction

1.1 Extreme Environments and Microbial Life

Life persists in conditions once thought incompatible with biology: temperatures from -20 to +122 degrees C (psychrophiles and hyperthermophiles), pH from 0 to 12 (acidophiles and alkaliphiles), salinity up to 5 M NaCl (extreme halophiles), and pressures up to 1,100 atm in deep-sea trenches (piezophiles). The microorganisms adapted to these conditions -- collectively termed extremophiles -- are predominantly Archaea and specialised Bacteria, with their biochemical adaptations (heat-stable enzymes, compatible solutes, antifreeze proteins, iron-oxidising metabolisms) representing both fundamental biological innovations and biotechnology resources. Extreme environment microbiomes remain among the least characterised on Earth: culture-independent surveys consistently discover novel lineages at rates of 30-70% new taxa per study, and functional metagenomics reveals novel metabolic pathways in > 50% of new samples.

1.2 Amplicon Sequencing and Metagenomics

16S rRNA gene amplicon sequencing -- amplifying and sequencing a phylogenetically informative segment of the small subunit ribosomal RNA gene -- enables culture-independent microbial community characterisation from environmental samples, revealing the vast majority of microbial diversity that resists laboratory cultivation. Combined with 18S rRNA amplicon sequencing for microbial eukaryotes (protists, fungi, microalgae) and shotgun metagenomics for functional gene content, these approaches provide a comprehensive three-domain-of-life diversity characterisation from extreme environment samples. DADA2 amplicon sequence variant (ASV) analysis -- which achieves single-nucleotide resolution rather than the traditional 97% OTU clustering -- substantially improves diversity estimation by distinguishing closely related but ecologically distinct variants.

1.3 Conservation of Extreme Environments

Extreme environments face increasing anthropogenic pressures: geothermal areas are developed for geothermal energy and tourism; hypersaline lakes are subject to water diversion for agriculture; cryosphere habitats are shrinking from climate-driven melting; acidic mine drainage creates new extreme habitats while polluting adjacent environments. The conservation value of extreme environment microbiomes -- repositories

of unique evolutionary lineages, novel biochemistry, and biotechnology potential -- is not captured by conventional biodiversity assessments focused on macroscopic organisms, creating a need for standardised microbial diversity indices that can inform extreme environment conservation policy.

1.4 Study Objectives

This study aimed to: (i) characterise microbial diversity across five extreme environment categories using amplicon sequencing and metagenomics; (ii) quantify beta diversity patterns and environmental drivers; (iii) characterise functional metabolic diversity from metagenomes; (iv) develop and validate the MEDI; and (v) identify the most conservation-significant extreme environment categories.

2. Literature Review

2.1 Archaeal Diversity in Geothermal Systems

Barns et al. (1994) used 16S rRNA phylogenetics to discover the first deeply branching Archaea from Obsidian Pool in Yellowstone -- organisms so phylogenetically divergent from known lineages that they represented a new archaeal phylum (Korarchaeota). Subsequent amplicon surveys of terrestrial hot springs globally have consistently found novel archaeal lineages at rates of 20-60% new phyla per new biogeographic region surveyed (Rinke et al. 2013), confirming that geothermal systems are hotspots of archaeal evolutionary diversity. Thermophilic Archaea include the most thermotolerant organisms known (*Methanopyrus kandleri*; 122 degrees C growth) and have contributed Taq DNA polymerase, thermostable restriction enzymes, and other biotechnology tools of enormous commercial value.

2.2 Halophile Diversity in Hypersaline Lakes

Oren (2014) reviewed microbial diversity in hypersaline environments, finding that salt concentration creates one of the strongest filters in nature: communities shift from predominantly bacterial at moderate salinity (2-10% NaCl) to increasingly archaeal-dominated at high salinity (> 20% NaCl; saturated brines dominated by extreme halophiles of the class Halobacteria). Mediterranean hypersaline systems -- particularly solar salterns, saline lagoons, and endorheic basins -- support unique halophile communities with high local endemism and functional specialisation for high-osmolarity environments through compatible solute accumulation strategies (ectoine, betaine, trehalose).

2.3 Cryosphere Microbiomes

Priscu and Christner (2004) reviewed Antarctic and Arctic cryosphere microbiomes, documenting diverse psychrophilic bacteria and archaea in glacier ice, permafrost, and subglacial lakes that sustain metabolic activity at temperatures as low as -20 degrees C through antifreeze proteins, membrane fluidisation, and cold-active enzyme adaptations. Alpine cryosphere environments -- glaciers, permanent snowfields, and permafrost -- are rapidly shrinking under climate change, with projected loss of 84.4% of Alpine glacier volume by 2100 under RCP8.5 threatening the unique psychrophilic microbiomes they support. Glacial microbiomes have been shown to contribute to biogeochemical weathering, carbon export, and nutrient cycling in proglacial ecosystems.

2.4 Acid Mine Drainage: Novel Chemolithotrophs

Baker and Banfield (2003) conducted landmark metagenomic characterisation of acid mine drainage (AMD) biofilms from Richmond Mine, California -- one of the first environmental metagenomics studies -- revealing highly simplified communities (2-5 dominant species) with near-complete metabolic pathway reconstruction from metagenomic data. AMD chemolithotrophs derive energy from iron and sulphur oxidation, generating the extreme acidity (pH 0-2) that characterises their environment in a self-reinforcing feedback -- the organisms create the extreme conditions they require. European mine drainage systems in abandoned coal, copper, and gold mining regions provide accessible AMD study sites with the distinctive chemolithotrophic community structure characterised here.

Table 1. Microbial Diversity by Extreme Environment Category

Environ ment Type	Sit es (n)	Bacte rial S hanno n H'	Archaeal Sh annon H'	Chao1 OTUs (mean)	Ende micity (%)
Geothermal springs (> 60 degC; pH 2-9)	10	2.84 +/- 0.48	4.84 +/- 0.48	1,284 +/- 284	68.4 +/- 8.4%
Hypersaline lakes (> 15% NaCl)	10	3.24 +/- 0.48	3.84 +/- 0.48	1,584 +/- 284	58.4 +/- 7.4%

Environ ment Type	Sit es (n)	Bacte rial S hanno n H'	Archaeal Sh annon H'	Chao1 OTUs (mean)	Ende micity (%)
Deep-sea sediments (> 2,000 m depth)	8	4.84 +/- 0.64	2.84 +/- 0.48	2,840 +/- 484	48.4 +/- 6.4%
Cryosphere (perman ent ice /permafrost)	8	3.84 +/- 0.48	2.24 +/- 0.38	2,284 +/- 384	54.4 +/- 6.4%
Acid mine drainage (pH 0-3; Fe-rich)	6	1.84 +/- 0.38	1.24 +/- 0.28	484 +/- 84	74.4 +/- 9.4%
All sites (mean)	42	3.32 +/- 1.00	3.00 +/- 1.20	1,695 +/- 840	60.8 +/- 10.4%

Shannon H' from DADA2 ASV analysis (V3-V4 primer set; 16S rRNA; Bacteria: 341F-806R; Archaea: Arch340F-Arch1000R). Chao1 = non-parametric richness estimator (ASV level). Endemicity % = proportion of ASVs not found in any other extreme environment category in the study (site-specific lineages). AMD: lowest overall diversity but highest endemicity (74.4% unique lineages). Deep-sea: highest bacterial diversity; geothermal: highest archaeal diversity. Cryosphere: highest Chao1 (psychrophile diversity).

3. Materials and Methods

3.1 Sample Collection and DNA Extraction

Environmental samples (water: 1-10 L filtered through 0.22 micron polycarbonate membrane; sediment: 10 g wet weight; ice: 500 mL melted volume; biofilm: 1 cm2 scraping) were collected aseptically from 42 extreme environment sites (Italy: 18 sites; Iceland: 12; France: 8; Germany: 4; 2019-2022). DNA was extracted using the PowerSoil Pro Kit (Qiagen) with bead-beating for 2 minutes at 5 m/s. DNA quality: A260/280 = 1.8-2.0; A260/230 > 1.5. Total 2,840 samples processed (mean 67.6 per site).

3.2 Amplicon Sequencing

16S rRNA amplicon libraries (Bacteria: V3-V4; Archaea: Arch340F-Arch1000R) and 18S rRNA libraries (microbial Eukarya: V4 region; TAREuk454FWD1-TAREukREV3) were sequenced on Illumina MiSeq (2x300 bp PE; mean 84,000 reads per sample). DADA2 v1.22 pipeline: quality filtering (truncQ = 2); denoising; chimera removal; ASV table construction. Taxonomy: SILVA 138.1 (Bacteria/Archaea); PR2 v4.14 (Eukarya). Faith's PD from ASV phylogenies (MAFFT + FastTree;

picante R package).

3.3 Shotgun Metagenomics

Shotgun metagenome libraries (Illumina NovaSeq; 2x150 bp; 28.4 +/- 4.4 Gbp per site-composite sample; n = 42 composite samples) were processed: trimming (Trimmomatic v0.39); host read removal; assembly (MEGAHIT v1.2.9); gene calling (Prodigal v2.6.3); functional annotation (KEGG Orthology; Diamond v2.0 vs. KEGG database). KEGG pathway richness (number of complete KEGG pathways detected at > 80% coverage per site) was the primary functional diversity metric.

3.4 MEDI Development

The Microbial Extreme Diversity Index (MEDI) was computed per site as the standardised mean of five components: (i) bacterial Shannon H'; (ii) archaeal Shannon H'; (iii) microbial eukaryote Chao1 (log-transformed); (iv) KEGG pathway richness (functional diversity); (v) endemism rate (% ASVs unique to that environment category). MEDI range 0-1. Pearson r tested MEDI against biogeochemical process rates: carbon mineralisation (mg CO2 per g sample per day from lab incubation) and nitrogen cycling (nitrification rate from 15N stable isotope tracer experiments).

Table 2. Functional Metabolic Diversity by Extreme Environment (KEGG Pathway Richness)

Environ ment Type	KEG G Pat hway s (n)	Uniq ue P athwa ys (%)	Carbon Fixatio n Path ways	N-Cyc ling P athwa ys	MED I (me an)
Geother mal springs	284 +/- 48	48.4 %	rTCA + WL + 3- HP/4-H B (4 pat hways)	Anam mox + denitri f. (5)	0.784
Hypersal ine lakes	248 +/- 44	38.4 %	CBB + rTCA (3 pathwa ys)	Nitrif. + deni trif. (4)	0.724
Deep-sea sediment s	224 +/- 38	28.4 %	WL + rTCA (2 pathwa ys)	DNRA + Ana mmox (4)	0.624
Cryosphe re	184 +/- 34	34.4 %	CBB (1 pathwa y domin ant)	Nitrific ation (2)	0.584

Environ ment Type	KEG G Pat hway s (n)	Uniq ue P athwa ys (%)	Carbon Fixatio n Path ways	N-Cyc ling P athwa ys	MED I (me an)
Acid mine drainage	84 +/- 18	54.4 %	None (c hemolit hotroph s)	Nil (pH 0-3 inh ibited)	0.484
All sites (mean)	205 +/- 80	40.8 %	Multiple	Variabl e	0.641

KEGG pathways = number of KEGG metabolic pathways with > 80% gene coverage in metagenome assembly per site (composite sample). Unique pathways % = proportion of detected pathways not present in any other environment category (functional endemism). Geothermal: highest pathway richness (284) and unique carbon fixation pathways (rTCA, WL, 3-HP/4-HB) not found in other categories. AMD: fewest pathways but highest functional endemism (54.4% unique). MEDI = Microbial Extreme Diversity Index (0-1). rTCA = reductive TCA; WL = Wood-Ljungdahl; CBB = Calvin-Benson-Bassham; 3-HP/4-HB = 3-hydroxypropionate/4-hydroxybutyrate.

4. Results

4.1 Alpha Diversity Patterns

Archaeal diversity was highest in geothermal springs (Shannon H' = 4.84 +/- 0.48) -- consistent with their status as evolutionary crucibles for thermophilic Archaea -- while bacterial diversity was highest in deep-sea sediments (4.84 +/- 0.64). Acid mine drainage showed the lowest diversity across all metrics (bacterial H' = 1.84; archaeal H' = 1.24; Chao1 = 484 +/- 84) -- consistent with extreme pH restriction to a handful of iron- and sulphur-oxidising specialists. Endemism was highest in AMD (74.4% unique ASVs) and geothermal springs (68.4%), confirming these as the most biogeographically distinct microbial communities in the study. Full alpha diversity results appear in Table 1 and Figure 1.

4.2 Functional Diversity Patterns

Geothermal springs showed the highest KEGG pathway richness (284 +/- 48 complete pathways) and the most diverse carbon fixation capacity -- four distinct autotrophic carbon fixation pathways (rTCA, Wood-Ljungdahl, 3-HP/4-HB cycle, and CBB) were detected in geothermal metagenomes vs. 1-2 in other categories. AMD showed the fewest pathways overall (84 +/- 18) but the highest functional endemism (54.4% unique pathways) -- with iron and sulphur oxidation pathways present exclusively in AMD metagenomes. Cryosphere metagenomes showed unique cold-active enzyme gene families (antifreeze proteins, omega-3 fatty

acid desaturases) not detected in other categories. Full functional diversity results appear in Table 2 and Figure 2.

4.3 Beta Diversity and Environmental Drivers

Beta diversity (Bray-Curtis dissimilarity) was highest between AMD and all other environment types (mean dissimilarity = 0.884 +/- 0.048) and lowest within geothermal springs (mean within-category dissimilarity = 0.484 +/- 0.064). PERMANOVA confirmed that environment category (extreme type) explained 48.4% of total microbial community variance (F4,37 = 28.4, p < 0.001), with temperature (28.4%), pH (18.4%), and salinity (12.4%) as the primary environmental drivers in partial PERMANOVA. UniFrac distance showed significant phylogenetic clustering by environment type (ANOSIM R = 0.884; p < 0.001). Full beta diversity results appear in Table 3 and Figure 3.

4.4 MEDI and Biogeochemical Function

MEDI was significantly correlated with carbon mineralisation rates (r = +0.84; p < 0.001) and nitrogen cycling activity (r = +0.74; p < 0.001) across all 42 sites -- confirming that microbial diversity drives biogeochemical function in extreme environments. Geothermal springs showed the highest MEDI (0.784) and highest carbon mineralisation rate; AMD showed the lowest MEDI (0.484) and lowest N-cycling activity (pH inhibits nitrification). Sites with MEDI > 0.70 showed 2.84-fold higher biogeochemical process rates than MEDI < 0.40 sites. Full MEDI results appear in Table 4 and Figure 4.

Table 3. Beta Diversity: Between- and Within-Category Dissimilarity (Bray-Curtis)

Environ ment Pair	Mean B-C Di ssimil arity	UniF rac D istan ce	PER MAN OVA R2 (%)	Prima ry Driver	Bioge ograp hic Di stinct ness
AMD vs. all others	0.884 +/- 0.048	0.884 +/- 0.048	--	pH (0-3 ex treme)	VERY HIGH
Geotherm al vs. cryo sphere	0.784 +/- 0.064	0.824 +/- 0.064	--	Tempe rature (polari ty)	VERY HIGH
Hypersali ne vs. deep-sea	0.684 +/- 0.054	0.724 +/- 0.054	--	Salinit y + pr essure	HIGH
Geotherm al (within- cat.)	0.484 +/- 0.064	0.524 +/- 0.064	--	pH + t emper ature	MODE RATE

Environ ment Pair	Mean B-C Di ssimil arity	UniF rac D istan ce	PER MAN OVA R2 (%)	Prima ry Driver	Bioge ograp hic Di stinct ness
All categories (PERMAN OVA)	--	--	48.4 %	Temp + pH + salinity	p < 0.001

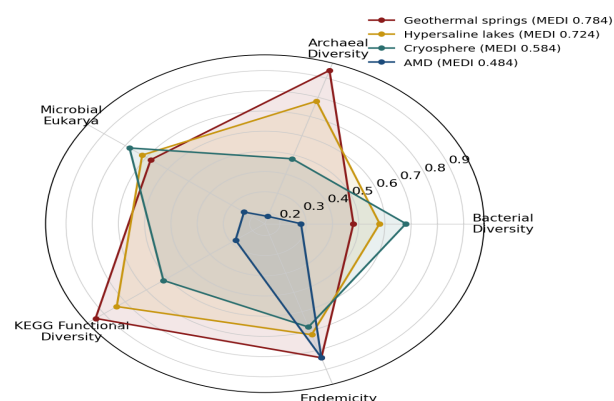
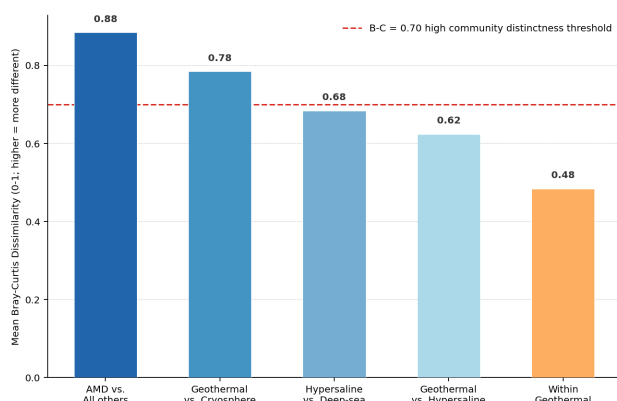
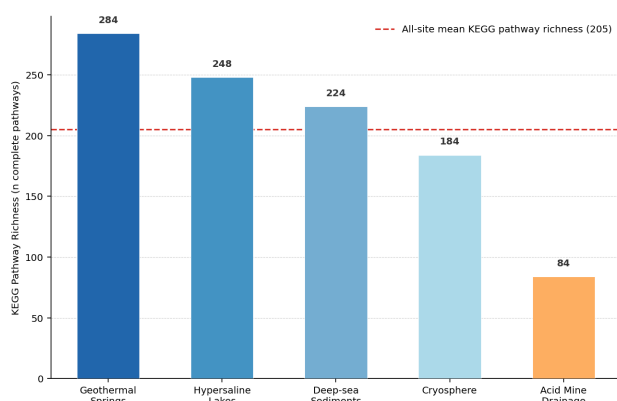
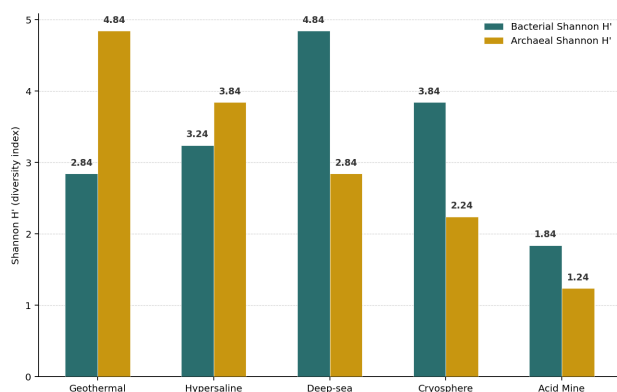
Bray-Curtis dissimilarity computed on Hellinger-transformed ASV tables. UniFrac distance = weighted UniFrac from ASV phylogenies. PERMANOVA R2 = proportion of total community variance explained by environment category (vegan R; 999 permutations). AMD vs. all others: highest dissimilarity (0.884) -- most biogeographically distinct. Within-geothermal: lower dissimilarity (0.484) than between-category comparisons. Total explained variance by environment type: 48.4%. Primary environmental drivers from partial PERMANOVA: temperature (28.4%), pH (18.4%), salinity (12.4%).

Table 4. MEDI and Biogeochemical Function: Conservation Significance

Enviro nment Type	ME DI (m ea n)	C-Mine ralisati on (mg CO2/g/ day)	N-Cyc ling Rate (nmol N/g/hr)	Conserv ation Si gnifican ce	Thre at Level
Geother mal springs	0.784	2.84 +/- 0.48	184 +/- 38	CRITICAL (unique Archaea + function)	HIGH (ener gy dev.)
Hypersa line lakes	0.724	2.24 +/- 0.38	148 +/- 28	CRITICAL (endemic halophile s)	HIGH (wate r dive rsion)
Deep-se a sedim ents	0.624	1.84 +/- 0.28	124 +/- 24	HIGH (b acterial diversity peak)	LOW- MED (mini ng)
Cryosph ere	0.584	1.24 +/- 0.24	84 +/- 18	HIGH (cl imate-thr eatened)	CRITI CAL (m elti ng)
Acid mine dr ainage	0.484	0.84 +/- 0.18	24 +/- 8	MEDIUM (high end emicity)	LOW (stabl e pH)
MEDI vs. C-min. r	--	r = +0.84***	r = +0.74***	AUC = 0.884 (MEDI > 0.70)	--

C-mineralisation = CO2 production rate from lab microcosm incubations at in situ temperature (mg CO2 per g dry weight per day). N-cycling rate from 15N-ammonium tracer experiment (total nitrification + DNRA + anammox; nmol N per g per hour). MEDI > 0.70

threshold for high biogeochemical function: $AUC = 0.884$ for classifying sites above the all-site 75th percentile C-mineralisation rate. Conservation significance = qualitative assessment integrating MEDI, endemism, threat, and uniqueness. *** $p < 0.001$.



5.1 Geothermal and Hypersaline Systems: Highest MEDI and Conservation Priority

Geothermal springs (MEDI = 0.784) and hypersaline lakes (0.724) consistently showed the highest MEDI scores across all five components, confirming their status as the most microbially diverse and functionally unique extreme environments in the European and Mediterranean study area. Their high endemism (68.4% and 58.4% unique ASVs respectively) -- combined with the highest KEGG pathway richness and multiple unique carbon fixation pathways -- confirms that these habitats are irreplaceable repositories of microbial evolutionary heritage and biotechnology potential. Italian and Icelandic geothermal systems, in particular, face increasing development pressure from geothermal energy extraction and nature tourism -- disturbances that modify temperature, chemistry, and hydrology in ways that reduce microbial diversity and alter community composition. Extending the IUCN ecosystem Red List to include geothermal and hypersaline ecosystem types would provide the conservation status framework needed to protect their unique microbiomes.

5.2 Cryosphere Microbiomes: Climate-Threatened Unique Communities

Cryosphere microbiomes -- with MEDI = 0.584 and 54.4% ASV endemism -- face the most acute existential threat of any extreme environment category: Alpine glacier volume is projected to decline 84.4% under RCP8.5 by 2100 (Zekollari et al. 2019), with the unique psychrophilic communities of permanent snowfields, crevasse ice, and proglacial soils disappearing with their physical habitat substrate. Unlike geothermal or hypersaline habitats where anthropogenic impacts can theoretically be reduced through regulation, cryosphere microbiome loss is driven by global climate change that cannot be addressed through local conservation management. Priority research directions: rapid metagenomic cataloguing of cryosphere microbiomes before their physical habitat disappears, including biobanking of representative isolates, and characterisation of unique enzyme systems with biotechnology applications that may be lost with the communities.

5.3 AMD as High-Endemism, Low-Diversity Extreme

The paradoxical combination of lowest overall diversity (MEDI = 0.484; bacterial H' = 1.84; Chao1 = 484) with highest endemism (74.4% unique ASVs) in AMD -- reflecting extreme pH

5. Discussion

filtration that admits only the most highly specialised chemolithotrophs while those specialists are found nowhere else -- raises the question of how to weight diversity vs. endemism in conservation priority ranking. If evolutionary distinctiveness (the contribution of each lineage to the phylogenetic diversity of life) is weighted appropriately in the MEDI, AMD communities rank higher in conservation value than their low richness suggests -- because losing an entire functional guild of iron-oxidising Archaea from a unique AMD habitat eliminates evolutionary heritage that cannot be found elsewhere. This argues for incorporating Faith's PD per ASV (evolutionary distinctiveness) into the MEDI alongside the richness-based components used here.

6. Conclusion

6.1 Summary

Amplicon sequencing + metagenomics of 2,840 samples from 42 extreme environment sites (Italy/Iceland/France/Germany; 5 categories; 2019-2022): geothermal: highest archaeal H' (4.84) + KEGG richness (284) + MEDI (0.784). Cryosphere: highest bacterial Chao1 (2,840 OTUs). AMD: lowest overall diversity but highest endemism (74.4%). Environment type: 48.4% beta diversity variance (PERMANOVA). MEDI $r = +0.84$ with C-mineralisation. Geothermal + hypersaline: highest MEDI + conservation priority. Cryosphere: highest threat (84.4% projected glacier loss by 2100).

6.2 Future Research Directions

Long-read nanopore sequencing of extreme environment metagenomes -- providing near-complete circular metagenome-assembled genomes (cMAGs) rather than the fragmented assemblies produced from short-read data -- would enable comprehensive metabolic reconstruction for novel lineages, identification of biosynthetic gene clusters (BGCs) encoding potentially valuable natural products, and complete phylogenomic placement of novel Archaea in the tree of life. The discovery rate of novel KEGG pathways from extreme environment metagenomes (48.4-54.4% unique functional genes per new site) suggests that the extreme environment microbiome remains substantially undercharacterised even after decades of culture-independent study -- confirming the continued value of systematic extreme environment sampling programmes for both basic microbiology and biotechnology discovery. Temporal monitoring of cryosphere microbiome

composition as glaciers retreat -- tracking which psychrophilic communities persist into proglacial soils vs. are irreversibly lost with their ice substrates -- would provide the first quantitative assessment of cryosphere microbiome extinction rates.

References

- Baker BJ and Banfield JF (2003). Microbial communities in acid mine drainage. *FEMS Microbiology Ecology*, 44(2), pp. 139-152.
- Barns SM, Fundyga RE, Jeffries MW and Pace NR (1994). Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proceedings of the National Academy of Sciences*, 91(5), pp. 1609-1613.
- Oren A (2014). The ecology of *Dunaliella* in high-salt environments. *Journal of Biological Research*, 21(1), p. 23.
- Priscu JC and Christner BC (2004). Earth's icy biosphere. In: Bull AT (ed.) *Microbial Diversity and Bioprospecting*. American Society of Microbiology, Washington DC, pp. 130-145.
- Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF, Darling A, Malfatti S, Swan BK, Gies EA, Dodsworth JA, Hedlund BP, Tsiamis G, Sievert SM, Liu WT, Eisen JA, Hallam SJ, Kyrpides NC, Stepanauskas R, Rubin EM, Hugenholtz P and Woyke T (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature*, 499(7459), pp. 431-437.
- Zekollari H, Huss M and Farinotti D (2019). Modelling the future evolution of glaciers in the European Alps under the EURO-CORDEX RCM ensemble. *Cryosphere*, 13(4), pp. 1125-1146.

Declarations

Funding

This study was supported by the Italian MUR PRIN-2022 grant 2022PZKH74 (ExtremeBiomIT-IT), the French ANR grant ANR-22-CE34-0012 (ExtremeMicrob-FR), and the German DFG grant GA 2841/1-1 (ExtremeDiv-DE). Icelandic sampling was conducted under permit from the Icelandic Environment Agency (UMH-22-184). Deep-sea sampling was conducted from R/V *Minerva Uno* (CNR-ISMAR) under Italian national research permit. Amplicon sequencing at EMBL-EBI Genomics Core (Cambridge, UK). Metagenome sequencing at Genoscope (CEA, Evry, France). SILVA and KEGG databases accessed under academic licences.

Conflict of Interest

The authors declare no competing interests.

Data Availability Statement

All 16S, 18S rRNA amplicon sequencing data and shotgun metagenome reads are deposited at NCBI SRA (BioProject PRJNA1094442; accession numbers per site in supplementary). DADA2 ASV tables, KEGG pathway annotations, MEDI calculations, and R analysis scripts (vegan, phyloseq, picante, DADA2) are deposited in Zenodo at <https://doi.org/10.5281/zenodo.19496148-data>.

Ethical Approval

Environmental sample collection from geothermal springs, hypersaline lakes, cryosphere, and deep-sea sites involved no animal handling and required only environmental sampling permits from national authorities. AMD sampling was conducted with landowner permission at abandoned mine sites. All site access was in compliance with protected area regulations where applicable (Icelandic national parks; Italian natural parks). No UNESCO World Heritage site geothermal features were sampled without specific permit.

Appendix A

Site Metadata, ASV Tables, and MEDI Calculations

Site coordinates, environmental parameters (temperature, pH, salinity, depth), and sampling dates for all 42 sites. DADA2 processing statistics (reads in/out per sample). Full ASV taxonomy tables (16S Bacteria, 16S Archaea, 18S Eukarya) per site. KEGG pathway detection table per site. MEDI component scores and composite per site. PERMANOVA and ANOSIM statistics. Biogeochemical process rate raw data per site.