Supplementary information

for

Persistent Petroleum Pollution: Microbial Responses in Bunger Hills, East Antarctica

Kellynn K. Y. Tan[1,2], Xabier Vázquez-Campos[1,2], Gwilym A. V. Price[3], Kasey Williams[3], Rebecca S. McWatters[3], Kristopher Abdullah[1,2], Náthali Machado de Lima[1,2], Thomas Laird[1], Angelique Ray[1,2], Jordan A. Vink[1,2], Dana Z. Tribbia[1,2], Daniel Wilkins[1,3], Tim Spedding[3], Belinda C. Ferrari[1,2*]

[1] *School of Biotechnology and Biomolecular Sciences, UNSW Sydney, NSW, Australia, 2052.*

[2] *Evolution and Ecology Research Centre, UNSW Sydney, Australia, 2052.*

[3] *Environmental Stewardship Program, Australian Antarctic Division, Department of Climate Change, Energy, the Environment and Water, 203 Channel Highway, Kingston, TAS, Australia, 7050.*

Supplementary notes

Supplementary Note 1. The processed dataset from amplicon sequencing included 2,587,999 prokaryotic and 3,242,740 eukaryotic gene sequences from 26 samples (Tables S21 and S22). The dataset included 7,770 Bacteria/Archaea ASVs classified to 35 bacterial phyla and four archaeal phyla, and 1,890 Eukarya ASVs belonging to 42 divisions (Tables S6 and S7).

Supplementary Note 2. NMDS ordination using environmental variables at 16S ASVs showed HT samples clustered tightly, reflecting similar community structures, while BT samples dispersed along the second axis, and HZ exhibited the greatest dispersion across both axes. For 18S ASVs, BT samples were distributed across both axes, while HT and HZ samples clustered closely (Fig. 2B).

Supplementary Note 3. It is worth noting that several contaminated samples were obtained from deeper subsurface soils, thus the effects of contamination and depth on microbial communities cannot be separated distinctly.

Supplementary Note 4. DESeq2 analysis showed fungal classes *Ascomycota* and *Basidiomycota*, many of which are effective degraders of petroleum hydrocarbons[1] and frequently detected in hydrocarbon-contaminated soils around Antarctica including Bunger Hills[2,3], were significantly (P adj. < 0.01) more abundant at HZ, alongside other predatory protists including *Variosea, Filosa-Granofilosea* and *Prostomatea*, and other unclassified *Fungi* (Extended Data Fig. 2B). The presence of a makeshift bamboo bridge at HZ prior to this study (duration unknown) and removed prior to sampling could play a role in the high abundance of fungi in these samples[4,5].Meanwhile, the predatory protist *Nebulidea*[6] and photosynthetic protists including *Bacillariophyceae*, *Trebouxiophyceae*, and *Chlorophyceae*, as well as other unclassified eukaryotes were significantly (P adj. < 0.01) more abundant in BT compared to HZ. *Nebulidea* and *Bacillariophyceae* were also more abundant in BT in comparison to HT, while *Euglenida* were more abundant in HT compared to HZ.

Supplementary Note 5. Spearman's correlation analysis of differentially abundant taxa demonstrated a strong positive relationship between post-SGC TRH concentration and the abundance of *Bdellovibrionota*, *Pseudomonadota*, *Ascomycota*, *Basidiomycota*, *Filosa-*

Granofilosea, *Bacillariophyceae*, and *Prostomatea*, and a negative correlation between these phyla abundances and pH and the rate of H_2 oxidation (Extended Data Figs. 3A and 3B). *Pseudomonadota* and *Bdellovibrionota* abundances negatively correlated with pH and the rate of H_2 oxidation.

Supplementary Note 6. Based on the MAG distribution across the different samples, two main clusters of MAGs were differentiated based on their detection (% of MAG with reads mapped) in each sample (Fig. S2). These two clusters separated MAGs associated with the contaminated (HZ) and uncontaminated (HT, BT) samples, with 169 and 131 MAGs respectively. With the exception of four MAGs with *prmA* and one predicted to harbour *alkB*, all the other MAGs containing hydrocarbon degradation gene annotations were characteristic of the contaminated samples (Fig. 3 and Table S13). This includes MAGs of novel taxa, whose parent is associated with hydrocarbon degradation potential such as *Nocardioides*[7], *Brevundimonas*[8], *Aquihabitans*[9], *Panacibacter*[10], and *Fimbriimonas*[11]. Similar to the trends observed in the whole metagenomes, genes associated with trace gas oxidation and dark carbon fixation were mainly present in the uncontaminated group of MAGs. This includes the high-affinity hydrogenases (except type 2a which were only detected in contaminated group of MAGs), RuBisCO form IE, and aerobic CO dehydrogenase (*coxL*), with the novel hydrogenase clade present in predominantly *Actinomycetota* and *Acidobacteriota* MAGs.

Supplementary Note 7. Prediction of optimal growth conditions with GenomeSPOT[12] also showed clear differences between the contaminated- and uncontaminated-associated MAGs. In general, MAGs associated with the contaminated samples showed lower optima pH (6.7 ± 0.6 vs 7.8 ± 0.5, $p = 0.00$), salinity (0.4 ± 0.8% vs 3.0 ± 1.6%, $p = 0.00$), and temperatures (26.1 ± 4.8 °C vs 33.0 ± 5.6 °C, $p = 0.00$) (Figs. S3–S5). A subcluster of 72 uncontaminated MAGs was associated with BH-23, the sample with high salinity, including 21 MAGs almost exclusively detected in that sample (Fig. S5). Analysis of the predicted optimal salinity for growth indicated indeed a preference for higher salt concentrations (3.3 ± 1.8%), statistically different to both the contaminated MAGs (0.4 ± 0.8%, P adj. = 0) and the other uncontaminated MAGs (2.8 ± 1.5%, P. adj. = 0.03). Two MAGs with growth preference for salinity were novel *Actinomycetota* taxa, BH-23_ACT12 (*Bungeriella frigidisoli* gen. nov., sp. nov.) and BH-23_ACT6 (*Cryoornithinimicrobium bungerii* gen. nov., sp. nov.). Six of the 23 MAGs with potential for atmospheric chemosynthesis were identified with growth preference for salinity, of which five were order *Euzebyales*[13] and one *Acidimicrobiales*.

Supplementary Note 8. Key aerobic respiration genes (Complex I-V) were detected across the metagenomes, with majority of MAGs possessing a partial electron transport chain (Figs. 4A and 4B and Tables S13–14). Low-affinity cytochrome oxidases (*coxA*/*ctaD*) were identified in 180 MAGs across all sampling locations, while high affinity type cytochrome oxidases (*ccoN*, *cydA*) were more prevalent in both metagenomes and MAGs from HZ, particularly in BH-11, indicating possible genetic adaptation to low oxygen (O_2) conditions due to O_2 requirements for hydrocarbon degradation[14]. Genes linked to nitrogen metabolism were detected in the metagenome. Nitrification genes (*pmoA*/*amoA*, *narG*/*nxrA*) were found in all samples, particularly in *Thermoproteota* MAGs, except BH-24 where *pmoA*/*amoA* were absent. MAGs that encoded denitrification genes (*nirK*, *norB*, *nosZ*) belonged to 14 phyla across sampling locations, while nitrogen fixation genes (*nifH*) were detected in HZ, particularly in BH-11, in *Pseudomonadota* MAGs but absent in BT and HT. Seven novel MAGs encoded genes for nitrogen metabolism including BH-09_BAC2 (*norB* and *nosZ*), BH-10_ACI3 (*norB*), BH-10_BAC2 (*nosZ*), BH-10_BAC5 (*nirK*), BH-10_CYA1 (*narG/nxrA* and *norB*), BH-24_ACT26 (*nirK*), and BH-24_ACT7 (*nirK*).

Supplementary Note 9. Alternate carbon fixation pathway markers, i.e., malonyl-CoA reductase/3-hydroxypropionate dehydrogenase ($NADP^+$) (*mcr*) for the 3-hydroxypropionate pathway, and ATP-citrate lyase (*aclB*) for the reverse TCA cycle were only detected in BH-11, and BH-23, respectively (Fig. 4B and Table S12). Neither marker was found associated with any MAG even despite the high LFGE values for *mcr* (-0.29).

Supplementary Note 10. Other chemoautotrophy-associated RuBisCO forms IC and ID were detected across metagenomes, with form IC being more abundant in BH-11, BH-20 and BH-24 (LFGE > 0), but near the detection limit in BH-10, and form ID being lower in abundance overall (LFGE < 0).

Taxonomic descriptions

Description of *Nitrosomicrobium* gen. nov.

Ni.tro.so.mi.cro’.bi.um. L. masc. adj. *nitrosus,* full of natron, here intended to mean nitrous; N.L. neut. n. *microbium,* a microbe; N.L. neut. n. *Nitrosomicrobium*, a nitrous microbe.

The type species is *Nitrosomicrobium frigidus*. The taxon is equivalent to placeholder genus g__TA-21 under the GTDB R220.

Description of *Nitrosomicrobium frigidus* sp. nov.

fri’gi.dus. L. neut. adj. *frigidus*, cold, frigid.

The type material is the metagenome assembled genome BH-18_THE1 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 2.0 Mbp in 231 contigs with an estimated completeness of 92.59% and 2.56% contamination, 16S (1469 bp), 23S (2048 bp), and 5S (120 bp) genes, and 43 tRNAs (20 unique: 19 standard plus tRNA-iMet). The GC content of this MAG is 35.8%. Predicted to be an ammonia oxidising archaeon.

This species is equivalent to the placeholder species s__TA-21 sp023251115 under the GTDB R220.

Description of *Nitrosobungeria* gen. nov.

Ni.tro.so.bun.ge’.ri.a. L. masc. adj. *nitrosus,* full of natron, here intended to mean nitrous; N.L. fem. n. *bungeria,* in reference to Bunger Hills, location in East Antarctica; *Nitrosobungeria*, a nitrous microbe from Bunger Hills.

The type species is *Nitrosobungeria shackeltonensis*. The taxon is equivalent to placeholder genus g__TH5893 under the GTDB R220.

Description of *Nitrosobungeria shackeltonensis* sp. nov.

shac.kle.ton.en’sis, N.L. fem. adj. *shackeltonensis,* inhabitant of the Shackelton ice shelf, East Antarctica, where Bunger Hills are located.

The type material is the metagenome assembled genome BH-18_THE2 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 3.6 Mbp in 204 contigs with an estimated completeness of 99.1% and 1.46% contamination, 16S (1472 bp), 23S (2054 bp), and 5S (120 bp) genes, and 46 tRNAs (21 unique: 20 standard plus tRNA-iMet). The GC content of this MAG is 37.8%. Predicted to be an ammonia oxidising archaeon.

Description of *Aquihabitans niveus* sp. nov.

ni've.us. L. masc. adj. *niveus*, snowy, snow-covered.

The type material is the metagenome assembled genome BH-10_ACT1 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 4.7 Mbp in 189 contigs with an estimated completeness of 97.03% and 1.79% contamination, 16S (1525 bp), partial 23S (457 bp), and 5S (117 bp) genes, and 46 tRNAs (21 unique: 20 standard plus tRNA-fMet). The GC content of this MAG is 71.6%.

Description of *Gelisolibacter* gen. nov.

Ge.li.so.li.bac'ter. L. neut. n. *gelu*, frost, cold; L. n. neut. *solum*, soil; N.L. masc. n. *bacter*, a rod; N.L. masc. n. *Gelisolibacter*, a rod from cold soil.

The type species is *Gelisolibacter meridionalis*. The taxon is equivalent to placeholder genus g__JACDBE01 under the GTDB R220.

Description of *Gelisolibacter meridionalis* sp. nov.

me.ri.dio.na'lis. L. fem. adj. *meridionalis*, southern; referring to the Southern hemisphere.

The type material is the metagenome assembled genome BH-24_ACT7 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 2.8 Mbp in 109 contigs with an estimated completeness of 98.4% and 2.8% contamination, 16S (1526 bp), 23S (3088 bp), and 5S (117 bp) genes, and 50 tRNAs (21 unique: 20 standard plus tRNA-fMet). The GC content of this MAG is 69.7%. Predicted to reduce nitrite (NO-forming nitrite reductase) and oxidise CO aerobically.

Description of *Cryoornithinimicrobium* gen. nov.

Cry.o.or.ni.thi.ni.mi.cro'bi.um. Gr. neut. n. *kryos*, icy cold; N.L. neut. n. *Ornithinimicrobium*, a bacterial genus; N.L. neut. n. *Cryoornithinimicrobium*, an *Ornithinimicrobium* from the cold.

The type species is *Cryoornithinimicrobium bungerii*. The taxon is equivalent to placeholder genus g__Ornithinimicrobium_A under the GTDB R220.

Description of *Cryoornithinimicrobium bungerii* sp. nov.

bun.ge.ri'i. N.L. gen. n., *bungerii*, referring to Bunger Hills, East Antarctica.

The type material is the metagenome assembled genome BH-23_ACT6 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 2.9 Mbp in 154 contigs with an estimated completeness of 90.24% and 0.47% contamination, 16S (1522 bp), 23S (3115 bp), and 5S (117 bp) genes, and 43 tRNAs (20 unique: 19 standard plus tRNA-fMet). The GC content of this MAG is 65.8%. Predicted to aerobically oxidise CO.

Description of *Nocardioides polaris* sp. nov.

po.la'ris. M.L. fem. adj. *polaris*, of, or pertaining to, a pole.

The type material is the metagenome assembled genome BH-09_ACT11 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 2.5 Mbp in 55 contigs with an estimated completeness of 94.69% and 0% contamination, 16S (1525 bp) and 5S (117 bp) genes, and 42 tRNAs (19 unique standard). The GC content of this MAG is 67%. Predicted to aerobically oxidise CO.

Description of *Actinosomnia* gen. nov.

Ac.ti.no.som'ni.a. Gr. fem. n. *aktis (gen. aktînos)*, ray, used to refer to actinomycetes; L. masc. n. *somnus*, sleep; N.L. fem. n. *Actinosomnia*, the dormant/sleepy actinomycete.

The type species is *Actinosomnia pattersoniae*. The taxon is equivalent to placeholder genus g__JACDCJ01 under the GTDB R220.

Description of *Actinosomnia pattersoniae* sp. nov.

pat.ter.so'ni.ae. N.L. gen. n. *pattersoniae*, in honour to Diana Patterson, first woman to lead an Australian Antarctic research station.

The type material is the metagenome assembled genome BH-20_ACT24 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 2.5 Mbp in 107 contigs with an estimated completeness of 93.99% and 0.16% contamination, 16S (1532 bp), 23S (4576 bp), and 5S (117 bp) genes, and 47 tRNAs (21 unique: 19 standard plus tRNA-fMet and tRNA-SeC). The GC content of this MAG is 68.3%. Predicted to consume H2 at atmospheric levels (high affinity [NiFe]-hydrogenase type 1m). Genome also harbours catechol 2,3-dioxygenase [EC:1.13.11.2].

This species is equivalent to the placeholder species s__JACDCJ01 sp013817655 under the GTDB R220.

Description of *Actinosomniaceae* fam. nov.

Ac.ti.no.som.ni.a'ce.ae. N.L. fem. n. *Actinosomnia*, the dormant/sleepy actinomycete; N.L. suff. -*aceae*, ending to denote a family; N.L. fem. pl. n. *Actinosomniaceae*, the *Actinosomnia* family.

The type genus is *Actinosomnia*. The taxon is equivalent to placeholder family f__UBA4738 under the GTDB R220.

Description of *Actinosomniales* ord. nov.

Ac.ti.no.som.ni.a'les. N.L. fem. n. *Actinosomnia*, the dormant/sleepy actinomycete; N.L. suff. -*ales*, ending to denote an order; N.L. fem. pl. n. *Actinosomniales*, the *Actinosomnia* order.

The type genus is *Actinosomnia*. The taxon is equivalent to placeholder order o__UBA4738 under the GTDB R220.

Description of *Actinopolaria* gen. nov.

Ac.ti.no.po.la'ri.a. Gr. fem. n. *aktis (gen. aktînos)*, ray, used to refer to actinomycetes; M.L. fem. adj. *polaris*, of, or pertaining to, a pole; N.L. fem. n. *Actinopolaria*, an actinomycete from the pole.

The type species is *Actinopolaria aerotropha*.

Description of *Actinopolaria aerotropha* sp. nov.

ae.ro.tro'pha. Gr. masc. n. *aêr (gen. aeros)*, air, gas; Gr. masc./fem. adj. *trophos*, feeder; N.L. fem. n. *aerotropha*, air-eater.

The type material is the metagenome assembled genome BH-24_ACT26 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 2.4 Mbp in 93 contigs with an estimated completeness of 98.12% and 0.45% contamination, 16S (1903 bp), 23S (1773 bp), and 5S (117 bp) genes, and 40 tRNAs (20 unique: 19 standard plus tRNA-SeC). The GC content of this MAG is 67.4%. Predicted to be able to reduce nitrate via NO-forming nitrite reductase and oxidise trace gases, i.e., CO and H_2, using the aerobic CO dehydrogenase and high affinity [NiFe]-hydrogenase type 1m.

Description of *Actinopolariaceae* fam. nov.

Ac.ti.no.po.la.ri.a'ce.ae. N.L. fem. n. *Actinopolaria*, an actinomycete from the pole; N.L. suff. -*aceae*, ending to denote a family; N.L. fem. pl. n. *Actinopolariaceae,* the *Actinopolaria* family.

The type genus is *Actinopolaria*. The taxon is equivalent to placeholder family f__WHSQ01 under the GTDB R220.

Description of *Actinopolariales* ord. nov.

Ac.ti.no.po.la.ri.a'les. N.L. fem. n. *Actinopolaria*, an actinomycete from the pole; N.L. suff. -*ales*, ending to denote an order; N.L. fem. pl. n. *Actinopolariales*, the *Actinopolaria* order.

The type genus is *Actinopolaria*. The taxon is equivalent to placeholder order o__CADDZG01 under the GTDB R220.

Description of *Bungeriella* gen. nov.

Bun.ger'i.el'la. N.L. fem. n., *Bungeriella,* a little thing from Bunger Hills, East Antarctica.

The type species is *Bungeriella frigidisoli*. The taxon is equivalent to placeholder genus g__JAJCYE01 under the GTDB R220.

Description of *Bungeriella frigidisoli* sp. nov.

fri.gi.di.so'li. L. masc. adj. *frigidus*, cold, frigid; L. neut. n. *solum*, soil; N.L. gen. n. *frigidisoli*, from cold soil.

The type material is the metagenome assembled genome BH-23_ACT12 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 2.4 Mbp in 140 contigs with an estimated completeness of 93.71% and 0.1% contamination, 16S (2399 bp) and 23S (2000 bp) genes, and 44 tRNAs (21 unique: 19 standard plus tRNA-fMet and tRNA-SeC). The GC content of this MAG is 64.7%. Predicted to fix carbon via Calvin–Benson–Bassham cycle using a RuBisCO type IE.

Description of *Bungeriellaceae* fam. nov.

Bun.ger'i.el.la'ce.ae. N.L. fem. n., *Bungeriella*, a little thing from Bunger Hills, East Antarctica; N.L. suff. *-aceae*, ending to denote a family; N.L. fem. pl. n. *Bungeriellaceae*, the *Bungeriella* family.

The type genus is *Bungeriella*. The taxon is equivalent to placeholder family f__JAHWKV01 under the GTDB R220.

Description of *Bungeriellales* ord. nov.

Bun.ger'i.el.la'les. N.L. fem. n., *Bungeriella*, a little thing from Bunger Hills, East Antarctica; N.L. suff. *-ales*, ending to denote an order; N.L. fem. pl. n. *Bungeriellales*, the *Bungeriella* order.

The type genus is *Bungeriella*. The taxon is equivalent to placeholder order o__JAHWKV01 under the GTDB R220.

Description of *Fimbriimonas antarctica* sp. nov.

an.tarc'ti.ca. L. neut. adj. *antarctica*, southern, pertaining to Antarctica.

The type material is the metagenome assembled genome BH-11_ARM2 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 4.0 Mbp in 138 contigs with an estimated completeness of 98.05% and 0.03% contamination, 16S (1507 bp), partial

23S (1480 bp), and 5S (117 bp) genes, and 47 tRNAs (22 unique: 20 standard plus tRNA-fMet and tRNA-SeC). The GC content of this MAG is 62.1%.

Description of *Panacibacter polaris* sp. nov.

po.la'ris. M.L. fem. adj. *polaris*, of, or pertaining to, a pole.

The type material is the metagenome assembled genome BH-10_BAC2 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 5.9 Mbp in 81 contigs with an estimated completeness of 100% and 0.07% contamination, 16S (1529 bp) and 23S (2541 bp) genes, and 48 tRNAs (21 unique: 20 standard plus tRNA-fMet). The GC content of this MAG is 38.5%. Predicted to aerobically oxidise CO, and reduce nitrite (NO-forming nitrite reductase). Genome also harbours a catechol 2,3-dioxygenase [EC:1.13.11.2] and a sensory rhodopsin.

Description of *Wilkeslandia* gen. nov.

Wil.kes.lan'di.a. N.L. fem. n. *Wilkeslandia*, a bacterium named after the Wilkes Land, district of Australian Antarctic Territory.

The type species is *Wilkeslandia alcanivorans*.

Description of *Wilkeslandia alcanivorans* sp. nov.

al.ca.ni'vo.rans. N.L. neut. n. *alcanum*, alkane, aliphatic hydrocarbon; L. part. adj. *vorans*, devourer; N.L. fem. adj. *alcanivorans*, devourer of alkanes.

The type material is the metagenome assembled genome BH-09_BAC2 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 2.6 Mbp in 16 contigs with an estimated completeness of 100% and 0.12% contamination, 16S (1527 bp), 23S (2941 bp), and 5S (117 bp) genes, and 35 tRNAs (21 unique: 20 standard plus tRNA-fMet). The GC content of this MAG is 37.5%. Predicted to degrade alkanes (alkane 1-monooxygenase) and at least capable of partial denitrification (predicted nitrous-oxide reductase [EC:1.7.2.4], and nitric oxide reductase subunit B [EC:1.7.2.5]).

Description of *Pseudopontibacter* gen. nov.

Pseu.do.pon.ti.bac’ter. Gr. neut. adj. *pseudes*, false; N.L. neut. n. *Pontibacter*, a bacterial genus; N.L. neut. n. *Pseudopontibacter*, a false *Pontibacter*.

The type species is *Pseudopontibacter australis*.

Description of *Pseudopontibacter australis* sp. nov.

aus.tra’lis. L. fem. adj. *australis*, southern.

The type material is the metagenome assembled genome BH-24_BAC1 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 5.6 Mbp in 1179 contigs with an estimated completeness of 98.42% and 2.26% contamination, 16S (1521 bp), 23S (212 bp), and 5S (112 bp) genes, and 28 tRNAs (17 unique standard). The GC content of this MAG is 54.5%.

Description of *Ventifactibacter* gen. nov.

Ven.ti.fact.i.bac’ter. NL. neut. n. *ventifactus*, a ventifact, rock that has been eroded by wind carrying sand or ice crystals; N.L. masc. n. *bacter*, a rod; N.L. masc. n. *Ventifactibacter*, the ventifact bacterium.

The type species is *Ventifactibacter hollidayae*.

Description of *Ventifactibacter hollidayae* sp. nov.

hol.li.day’ae. N.L. gen. n. *hollidayae*, in honour of Dr Louise Holliday, the first woman to winter in Antarctica for the Australian Antarctic Program serving as medical officer at Davis station.

The type material is the metagenome assembled genome BH-10_BAC5 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 3.4 Mbp in 46 contigs with an estimated completeness of 99.02% and 1.73% contamination, 16S (1518 bp), 23S (2939 bp), and 5S (115 bp) genes, and 42 tRNAs (21 unique: 20 standard plus tRNA-fMet). The GC content of this MAG is 35.2%. Predicted to be able to reduce nitrite (nitrite reductase (NO-forming) [EC:1.7.2.1]).

Description of *Ventifactibacteraceae* fam. nov.

Ven.ti.fact.i.bac.te.ra’ce.ae. N.L. masc. n. *Ventifactibacter*, the ventifact bacterium; N.L. suff. -*aceae*, ending to denote a family; N.L. fem. pl. n. *Ventifactibacteraceae*, the *Ventifactibacter* family.

The type genus is *Ventifactibacter*. The taxon is equivalent to placeholder family f__B-1AR under the GTDB R220.

Description of *Psychrobscuribacter* gen. nov.

Psy.chrobs.cu.ri.bac’ter. Gr. masc. adj. *psychros*, cold; N.L. masc. n. *Obscuribacter*, a bacterial genus; N.L. masc. n. *Psychrobscuribacter*, an *Obcuribacter* from the cold.

The type species is *Psychrobscuribacter pollutisoli*. The taxon is equivalent to placeholder genus g__PALSA-1081 under the GTDB R220.

Description of *Psychrobscuribacter pollutisoli* sp. nov.

pol.lu.ti.so’li. L. masc. perf. part. *pollutus*, polluted; L. n. neut. *solum*, soil; N.L. gen. n. *pollutisoli*, of polluted soil.

The type material is the metagenome assembled genome BH-10_CYA1 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 7.5 Mbp in 91 contigs with an estimated completeness of 100% and 4.31% contamination, 16S (1508 bp), 23S (2963 bp), and 5S (116 bp) genes, and 66 tRNAs (22 unique: 20 standard plus tRNA-fMet and tRNA-SeC). The GC content of this MAG is 49%. Predicted to be involved in nitrogen cycling (nitrate reductase, NarG; and nitric oxide reductase, NorB). The genome also harbours a sensory rhodopsin but no photosynthetic machinery.

Description of *Cryogemmata* gen. nov.

Cry.o.gem.ma’ta. Gr. neut. n. *kryos*, icy cold; L. inf. v. *gemmare*, to put forth buds, to bud, used to refer to *Gemmatimonadota*; N.L. fem. n. *Cryogemmata*, a *Gemmatimonadota* from the cold.

The type species is *Cryogemmata carboxiditropha*.

Description of *Cryogemmata carboxiditropha* sp. nov.

car.bo.xi.di.tro'pha. N.L. neut. adj. *carboxidum*, carbon monoxide; Gr. masc./fem. adj. *trophos*, feeder; N.L. neut. adj. *carboxiditropha*, carbon monoxide-eater.

The type material is the metagenome assembled genome BH-18_GEM1 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 2.5 Mbp in 116 contigs with an estimated completeness of 94.07% and 0.69% contamination, 16S (1572 bp) gene, and 46 tRNAs (20 unique: 18 standard plus tRNA-fMet and tRNA-SeC). The GC content of this MAG is 67.6%. Predicted to reduce nitrite (NO-forming nitrite reductase) and oxidise CO aerobically.

Description of *Cryogemmataceae* fam. nov.

Cry.o.gem.ma.ta'ce.ae. N.L. fem. n. *Cryogemmata*, a *Gemmatimonadota* from the cold; N.L. suff. -*aceae*, ending to denote a family; N.L. fem. pl. n. *Cryogemmataceae*, the *Cryogemmata* family.

The type genus is *Cryogemmata*. The taxon is equivalent to placeholder family f__JAHWKZ01 under the GTDB R220.

Description of *Cryogemmatales* ord. nov.

Cry.o.gem.ma.ta'les. N.L. fem. n. *Cryogemmata*, a *Gemmatimonadota* from the cold; N.L. suff. -*ales*, ending to denote an order; N.L. fem. pl. n. *Cryogemmatales*, the *Cryogemmata* order.

The type genus is *Cryogemmata*. The taxon is equivalent to placeholder order o__JACCXV01 under the GTDB R220.

Description of *Nanosomnibacter* gen. nov.

Na.no.som.ni.bac'ter. Gr. masc. n. *nânos*, a dwarf; L. masc. n. *somnus*, sleep; N.L. masc. n. *bacter*, a rod; N.L. masc. n. *Nanosomnibacter*, a sleepy/dormant small bacterium.

The type species is *Nanosomnibacter parvus*. The taxon is equivalent to placeholder genus g__CALBLZ01 under the GTDB R220.

Description of *Nanosomnibacter parvus* sp. nov.

par'vus. L. masc. adj. *parvus*, small.

The type material is the metagenome assembled genome BH-11_PAT4 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 0.9 Mbp in 17 contigs with an estimated completeness of 93.99% and 0.08% contamination, 16S (1475 bp), 23S (3502 bp), and 5S (117 bp) genes, and 42 tRNAs (21 unique: 20 standard plus tRNA-fMet). The GC content of this MAG is 51.3%.

Description of *Nanosomnibacteraceae* fam. nov.

Na.no.som.ni.bac.te.ra'ce.ae. N.L. masc. n. *Nanosomnibacter*, a sleepy/dormant small bacterium; N.L. suff. *-aceae*, ending to denote a family; N.L. fem. pl. n. *Nanosomnibacteraceae*, the *Nanosomnibacter* family.

The type genus is *Nanosomnibacter*. The taxon is equivalent to placeholder family f__CAILIB01 under the GTDB R220.

Description of *Nanosomnibacterales* ord. nov.

Na.no.som.ni.bac.te.ra'les. N.L. masc. n. *Nanosomnibacter*, a sleepy/dormant small bacterium; N.L. suff. *-ales*, ending to denote an order; N.L. fem. pl. n. *Nanosomnibacterales*, the *Nanosomnibacter* order.

The type genus is *Nanosomnibacter*. The taxon is equivalent to placeholder order o__CAILIB01 under the GTDB R220.

Description of *Nanosomnibacteria* clas. nov.

Na.no.som.ni.bac.te'ri.a. N.L. masc. n. *Nanosomnibacter*, a sleepy/dormant small bacterium; N.L. suff. *-ia*, ending to denote a class; N.L. fem. pl. n. *Nanosomnibacteria*, the *Nanosomnibacter* class.

The type genus is *Nanosomnibacter*. The taxon is equivalent to placeholder class c__UBA1384 under the GTDB R220.

Description of *Frigidisphaera* gen. nov.

Fri.gi.di.sphae’ra. L. masc. adj. *frigidus*, cold, frigid; Gr. fem. n. *sphaîra*, a sphere; N.L. fem. n. *Frigidisphaera*, a sphere from the cold.

The type species is *Frigidisphaera bungerii*. The taxon is equivalent to placeholder genus g__JACVCS01 under the GTDB R220.

Description of *Frigidisphaera bungerii* sp. nov.

bun.ge.ri’i. N.L. gen. n., *bungerii*, referring to Bunger Hills, East Antarctica.

The type material is the metagenome assembled genome BH-11_PLA1 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 3.2 Mbp in 39 contigs with an estimated completeness of 90.62% and 0.36% contamination, 16S (1484 bp), 23S (2744 bp), and 5S (107 bp) genes, and 47 tRNAs (21 unique: 20 standard plus tRNA-fMet). The GC content of this MAG is 67.5%.

Description of *Frigidisphaeraceae* fam. nov.

Fri.gi.di.sphae.ra’ce.ae. N.L. fem. n. *Frigidisphaera*, a sphere from the cold; N.L. suff. *-aceae*, ending to denote a family; N.L. fem. pl. n. *Frigidisphaeraceae*, the *Frigidisphaera* family.

The type genus is *Frigidisphaera*. The taxon is equivalent to placeholder family f__UBA1924 under the GTDB R220.

Description of *Cryolimnoglobus* gen. nov.

Cry.o Gr. neut. n. *kryos*, icy cold; N.L. masc. n. *Limnoglobus*, a bacterial genus; N.L. masc. n. *Cryolimnoglobus*, a *Limnoglobus* bacterium from the cold.

The type species is *Cryolimnoglobus antarcticus*.

Description of *Cryolimnoglobus antarcticus* sp. nov.

an.tarc’ti.cus. L. masc. adj. *antarcticus*, southern, pertaining to Antarctica.

The type material is the metagenome assembled genome BH-11_PLA2 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 6.6 Mbp in 220 contigs with an estimated completeness of 99.77% and 1.09% contamination, 16S (1504 bp), 23S (2758 bp), and 5S (109 bp) genes, and 57 tRNAs (22 unique: 20 standard plus tRNA-fMet and tRNA-SeC). The GC content of this MAG is 60.2%.

Description of *Brevundimonas antarctica* sp. nov.

an.tarc’ti.ca. L. fem. adj. *antarctica*, southern, pertaining to Antarctica.

The type material is the metagenome assembled genome BH-09_PSE1 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 3.2 Mbp in 138 contigs with an estimated completeness of 92.9% and 1.66% contamination, 16S (1463 bp), and 5S (115 bp) genes, and 42 tRNAs (19 unique: 18 standard plus tRNA-fMet). The GC content of this MAG is 67.4%. Predicted to oxidise sulfur (*soxB*, S-sulfosulfanyl-L-cysteine sulfohydrolase [EC:3.1.6.20]).

Description of *Frigidisolicola* gen. nov.

Fri.gi.di.so.li’co.la. L. masc. adj. *frigidus*, cold, frigid; L. n. neut. *solum*, soil; L. masc./fem. n. suff. -*cola*, inhabitant of, dweller; N.L. fem. n. *Frigidisolicola*, dweller of cold soils.

The type species is *Frigidisolicola castellviae*.

Description of *Frigidisolicola castellviae* sp. nov.

cas.tell.vi.ae. N.L. gen. n. *castellviae*, in honor of Josefina Castellví, the first Spanish woman to participate in and coordinate an international expedition to Antarctica.

The type material is the metagenome assembled genome BH-10_PSE17 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 4.2 Mbp in 112 contigs with an estimated completeness of 94.52% and 0.82% contamination, 16S (1531 bp), 23S (1860 bp), and 5S (114 bp) genes, and 45 tRNAs (21 unique: 19 standard plus tRNA-fMet and tRNA-SeC). The GC content of this MAG is 66.5%. Predicted to degrade hydrocarbons (long-chain alkane monooxygenase [EC:1.14.14.28], and phthalate 4,5-dioxygenase [EC:1.14.12.7]).

Description of *Dormimicrobium* gen. nov.

Dor.mi.mi.cro’bi.um. L. v. *dormio*, to sleep; N.L. neut. n. *microbium*, a microbe; N.L. neut. n. *Dormimicrobium*, a dormant microbe.

The type species is *Dormimicrobium murphyi*. The taxon is equivalent to placeholder genus g__JACDCP01 under the GTDB R220.

Description of *Dormimicrobium murphyi* sp. nov.

mur.phy’i. N.L. gen. n. *murphyi*, in honour of Herbet Dyce Murphy, Australian adventurer and explorer.

The type material is the metagenome assembled genome BH-24_PSE2 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 2.7 Mbp in 49 contigs with an estimated completeness of 100% and 0.1% contamination, 16S (1545 bp), 23S (4142 bp), and 5S (117 bp) genes, and 45 tRNAs (20 unique: 19 standard plus tRNA-fMet). The GC content of this MAG is 64.1%. Predicted to consume H_2 at atmospheric concentrations (high affinity [NiFe]-hydrogenase type 1l).

Description of *Dormimicrobiaceae* fam. nov.

Dor.mi.mi.cro.bi.a’ce.ae. N.L. neut. n. *Dormimicrobium*, a dormant microbe; N.L. suff. -*aceae*, ending to denote a family; N.L. fem. pl. n. *Dormimicrobiaceae*, the *Dormimicrobium* family.

The type genus is *Dormimicrobium*. The taxon is equivalent to placeholder family f__JACDCP01 under the GTDB R220.

Description of *Dormimicrobiales* ord. nov.

Dor.mi.mi.cro.bi.a’les. N.L. neut. n. *Dormimicrobium*, a dormant microbe; N.L. suff. -*ales*, ending to denote an order; N.L. fem. pl. n. *Dormimicrobiales*, the *Dormimicrobium* order.

The type genus is *Dormimicrobium*. The taxon is equivalent to placeholder order o__JACDCP01 under the GTDB R220.

Description of *Cryoterrimicrobium* gen. nov.

Cry.o.ter.ri.mi.cro’bi.um. Gr. neut. n. *kryos*, icy cold; N.L. neut. n. *Terrimicrobium*, a bacterial genus; N.L. neut. n. *Cryoterrimicrobium*, a *Terrimicrobium* from the cold.

The type species is *Cryoterrimicrobium chapmanii*.

Description of *Cryoterrimicrobium chapmanii* sp. nov.

chap.man’i.i. N.L. gen. n. *chapmanii*, in honour to limnologist Ann Chapman, first woman to lead an Antarctic expedition.

The type material is the metagenome assembled genome BH-09_VER1 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 6.1 Mbp in 69 contigs with an estimated completeness of 96.82% and 1.47% contamination, 16S (1526 bp), 23S (1995 bp), and 5S (116 bp) genes, and 58 tRNAs (21 unique: 20 standard plus tRNA-fMet). The GC content of this MAG is 58.7%.

Description of *Verrucomicrobium antarcticum* sp. nov.

an.tarc’ti.cum. L. neut. adj. *antarcticum*, southern, pertaining to Antarctica.

The type material is the metagenome assembled genome BH-11_VER1 (ACCNO) recovered from soil from Bunger Hills, East Antarctica. The MAG consists of 5.0 Mbp in 86 contigs with an estimated completeness of 99.66% and 0.58% contamination, 16S (1537 bp), 23S (2844 bp), and 5S (116 bp) genes, and 45 tRNAs (21 unique: 20 standard plus tRNA-fMet). The GC content of this MAG is 55%. Predicted to degrade phthalate (phthalate 4,5-dioxygenase [EC:1.14.12.7]).

Supplementary Figures

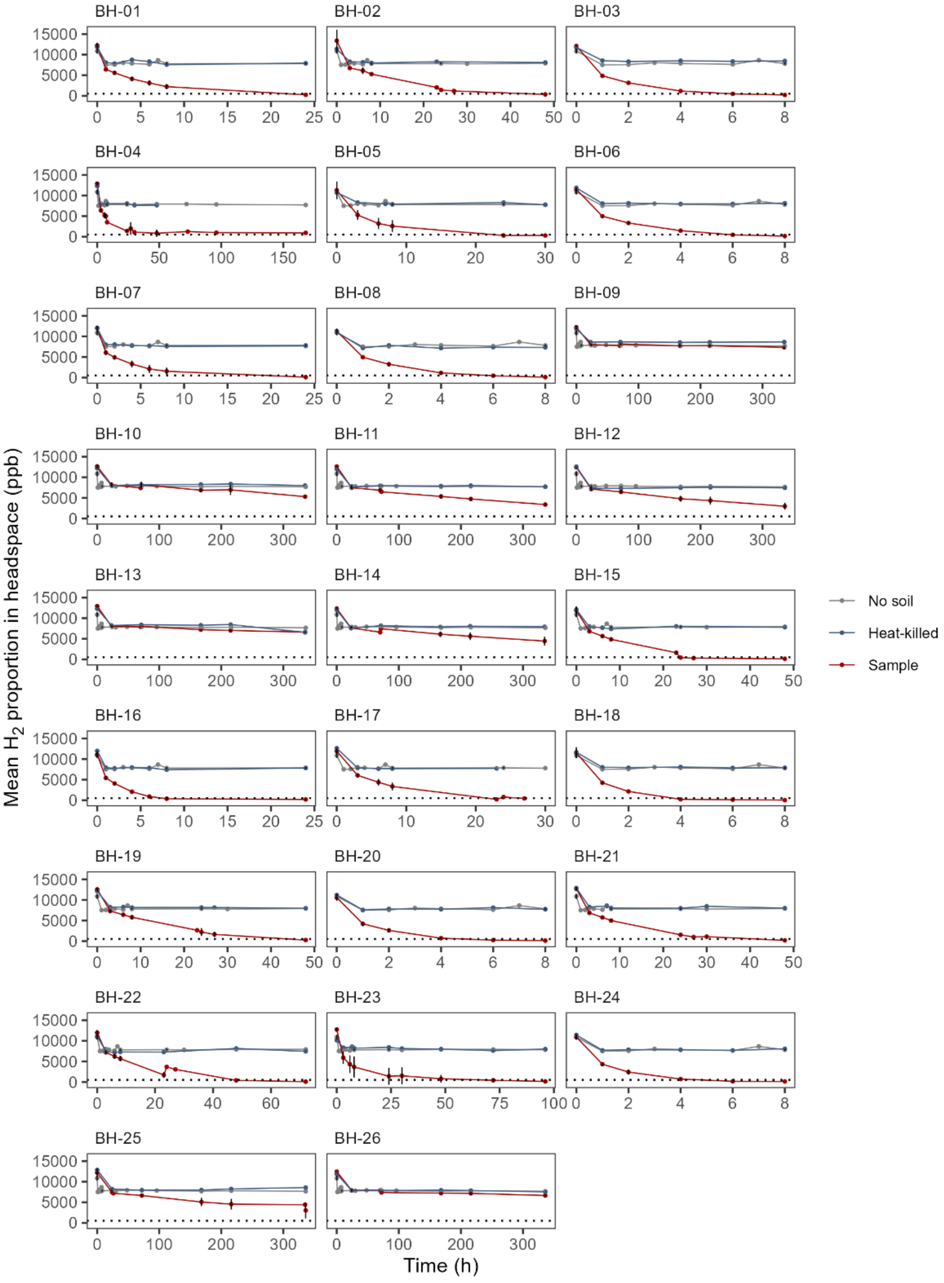

Figure S1. The oxidation of H_2 by Bunger Hills soil microcosms measured with gas chromatography. Shown are the H_2 oxidation activity of soils (1 g each), incubated at 10 °C in the dark. Sterile serum bottles, and heat-killed soils were used as controls. Points represent mean H_2 in the headspace for triplicates and vertical lines represent one standard deviation from the mean. The black dashed lines represent H_2 at atmospheric concentrations (~530 p.p.b.v.).

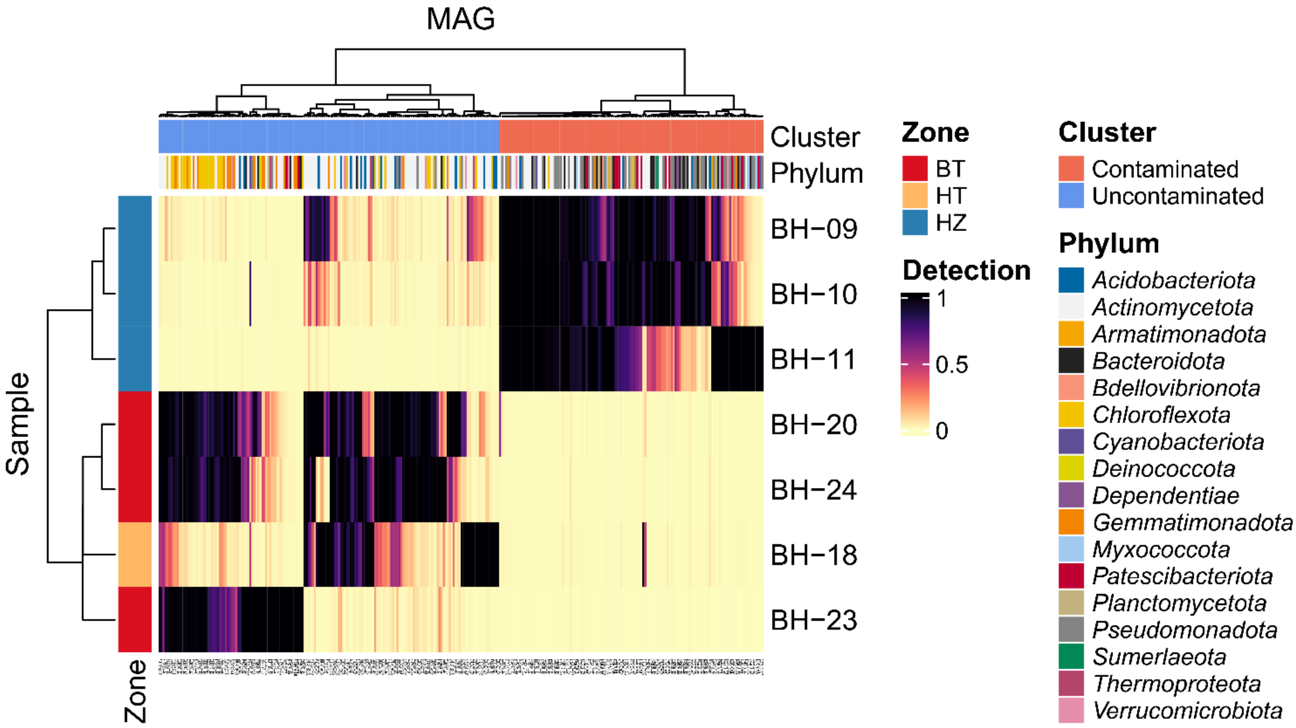

Figure S2. Clustering of MAG and metagenomic samples based on the detection of MAGs. MAG detection values were obtained with anvi'o. Clustering shows two major clusters separating the Helipad Zone/Contaminated area from the uncontaminated samples.

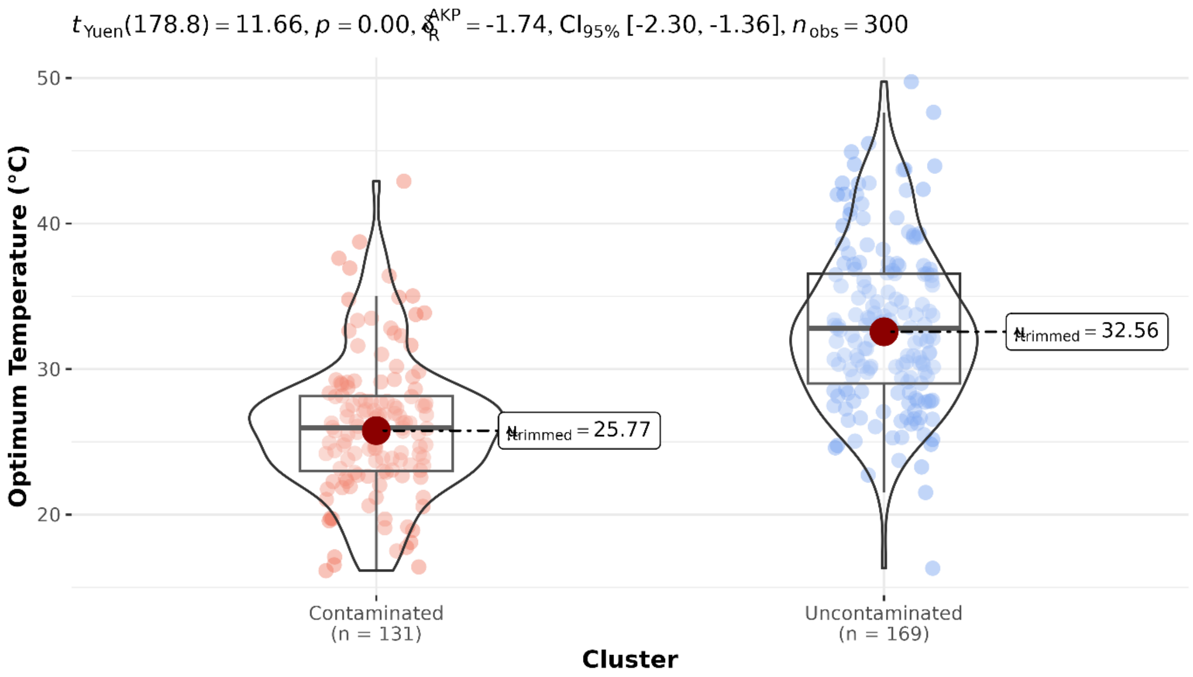

Figure S3. Optimum temperature predictions for growth for MAGs associated with samples from contaminated and uncontaminated zones. MAGs characteristic of the uncontaminated a significant higher predicted optimum temperature (Yuen's trimmed mean test).

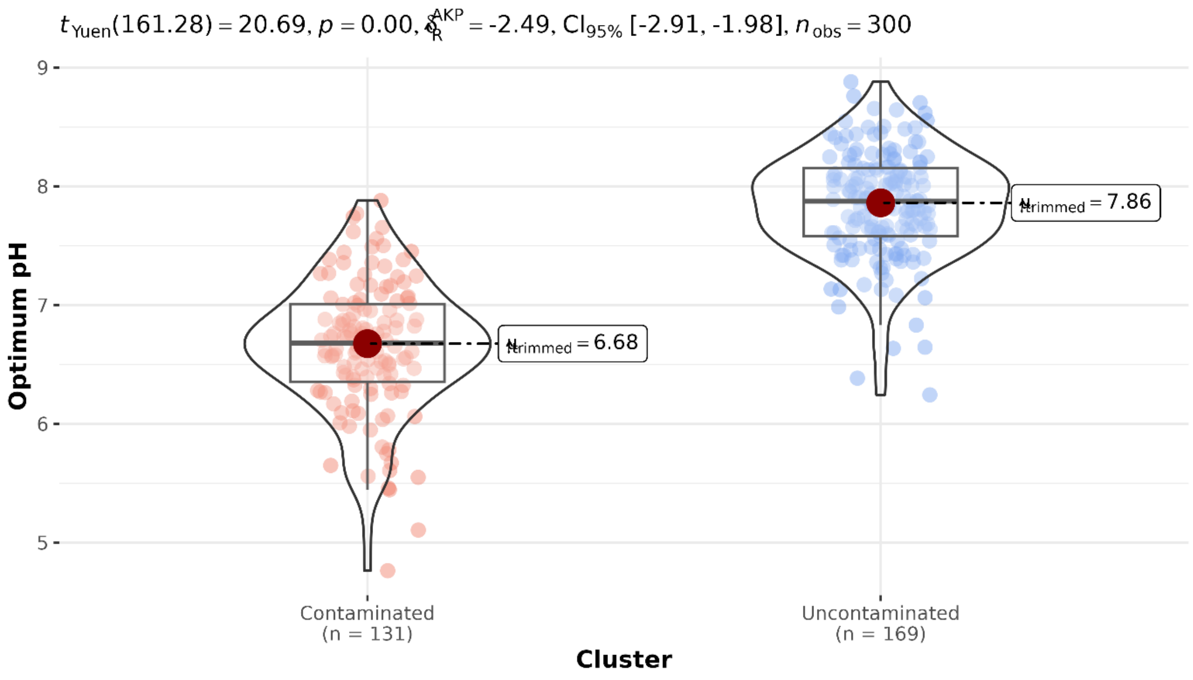

Figure S4. Optimum pH predictions for growth for MAGs associated with samples from contaminated and uncontaminated zones. MAGs characteristic of the uncontaminated a significant higher predicted optimum pH (Yuen's trimmed mean test).

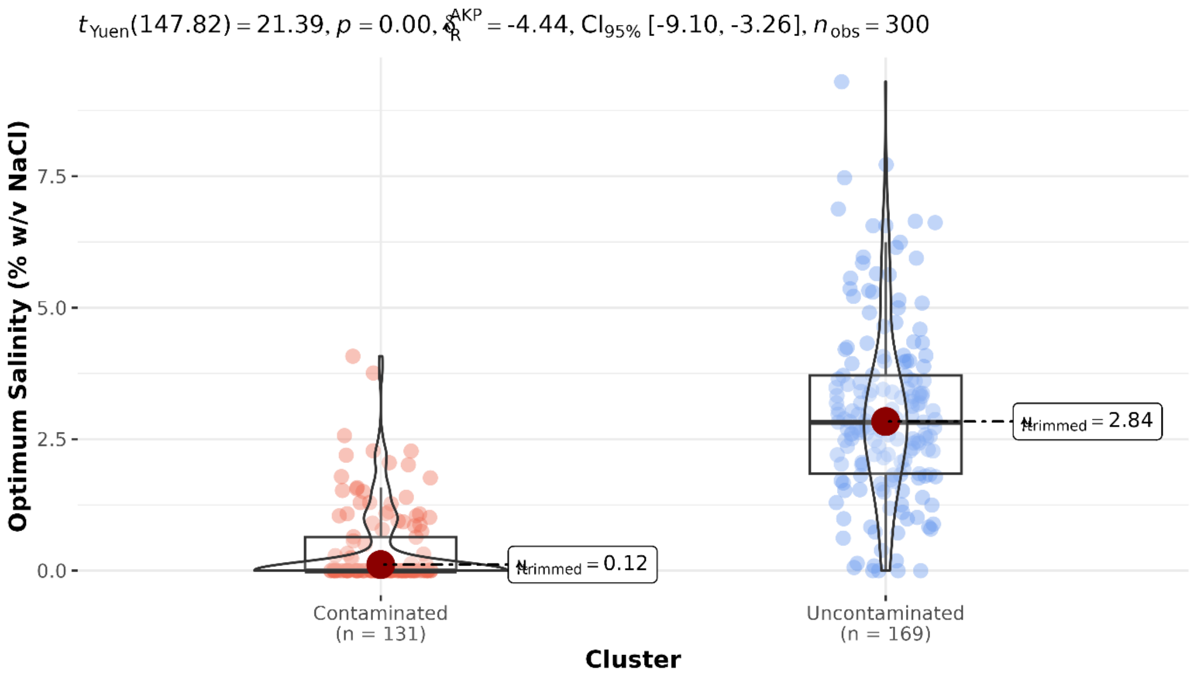

Figure S5. Optimum salinity predictions for growth for MAGs associated with samples from contaminated and uncontaminated zones. MAGs characteristic of the uncontaminated a significant higher predicted optimum salinity (Yuen’s trimmed mean test).

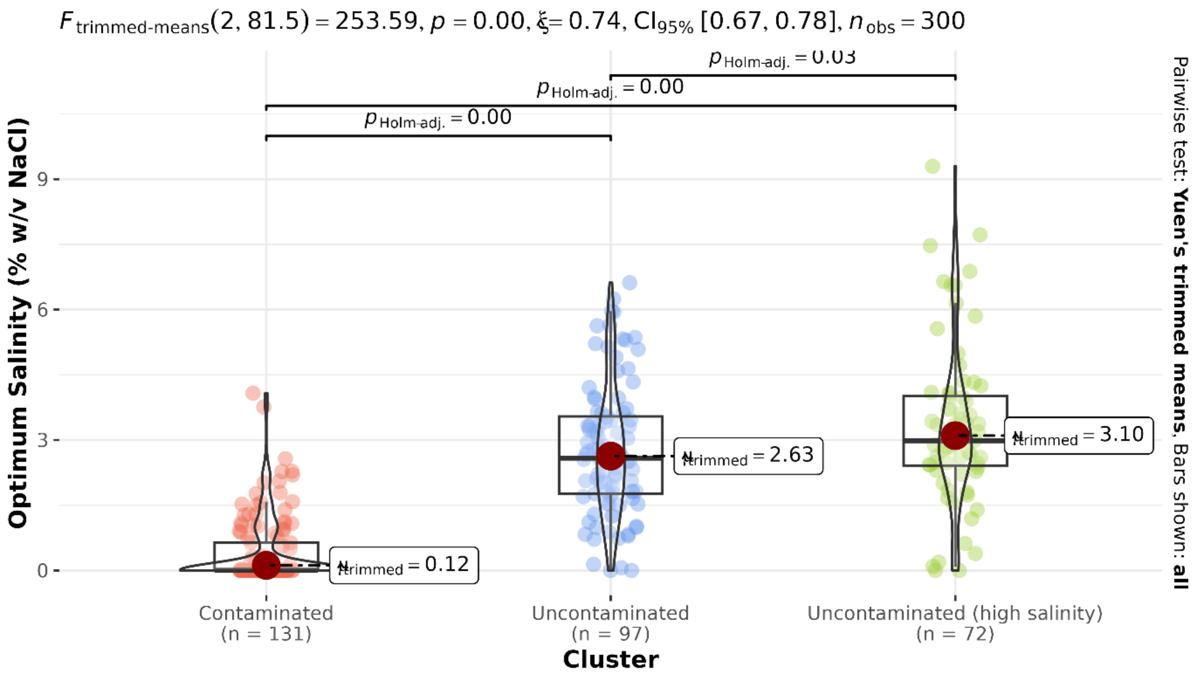

Figure S6. Optimum salinity predictions for growth for MAGs associated with samples from contaminated and uncontaminated zones. Here MAGs associated to the high salinity samples were segregated from the uncontaminated cluster. MAGs characteristic of the high-salinity uncontaminated sample (BH-23) show significantly higher predicted optimum salinity to both the contaminated and the rest of the uncontaminated (Yuen’s trimmed mean test).

References

1. Varjani, S. J. Microbial degradation of petroleum hydrocarbons. *Bioresour. Technol.* **223**, 277–286 (2017).

2. van Dorst, J. *et al.* Applying microbial indicators of hydrocarbon toxicity to contaminated sites undergoing bioremediation on subantarctic Macquarie Island. *Environ. Pollut.* **259**, 113780 (2020).

3. Kochkina, G. A. *et al.* Diversity of mycelial fungi in natural and human-affected Antarctic soils. *Polar Biol.* **42**, 47–64 (2019).

4. Kim, J.-J. *et al.* Fungi associated with bamboo and their decay capabilities. *Holzforschung* **65**, 271–275 (2010).

5. Dai, D. Q. *et al.* Bambusicolous fungi. *Fungal Divers.* **82**, 1–105 (2017).

6. Tikhonenkov, D. V. *et al.* Microbial predators form a new supergroup of eukaryotes. *Nature* **612**, 714–719 (2022).

7. Ma, Y. *et al.* *Nocardioides*: “Specialists” for Hard-to-Degrade Pollutants in the Environment. *Molecules* **28**, 7433 (2023).

8. Guermouche M’rassi, A., Bensalah, F., Gury, J. & Duran, R. Isolation and characterization of different bacterial strains for bioremediation of n-alkanes and polycyclic aromatic hydrocarbons. *Environ. Sci. Pollut. Res.* **22**, 15332–15346 (2015).

9. Zhen, M. *et al.* Combination of rhamnolipid and biochar in assisting phytoremediation of petroleum hydrocarbon contaminated soil using *Spartina anglica*. *SI Recent Adv. Environ. Sci.* **85**, 107–118 (2019).

10. Lladó, S. *et al.* Pyrosequencing reveals the effect of mobilizing agents and lignocellulosic substrate amendment on microbial community composition in a real industrial PAH-polluted soil. *J. Hazard. Mater.* **283**, 35–43 (2015).

11. Xiang, Y. *et al.* Neural network establishes co-occurrence links between transformation products of the contaminant and the soil microbiome. *Sci. Total Environ.* **924**, 171287 (2024).

12. Barnum, T. P. *et al.* Predicting microbial growth conditions from amino acid composition. Preprint at https://doi.org/10.1101/2024.03.22.586313 (2024).

13. Gonzalez-Pimentel, J. *et al.* The Marine Bacterial Genus *Euzebya* Is Distributed Worldwide in Terrestrial Environments: A Review. *Appl. Sci.* **13**, 9644 (2023).

14. Truskewycz, A. *et al.* Petroleum Hydrocarbon Contamination in Terrestrial Ecosystems—Fate and Microbial Responses. *Molecules* **24**, 3400 (2019).