



**HAL**  
open science

## A metagenomic insight into the microbiomes of geothermal springs in the Subantarctic Kerguelen Islands

Maxime Allieux, Stéven Yvenou, Alexander Merkel, Marc Cozannet, Johanne Aubé, Jolann Pommellec, Marc Le Romancer, Véronique Lavastre, Damien Guillaume, Karine Alain

► **To cite this version:**

Maxime Allieux, Stéven Yvenou, Alexander Merkel, Marc Cozannet, Johanne Aubé, et al.. A metagenomic insight into the microbiomes of geothermal springs in the Subantarctic Kerguelen Islands. *Scientific Reports*, 2022, 12, 10.1038/s41598-022-26299-4 . hal-03912182

**HAL Id: hal-03912182**

**<https://hal.science/hal-03912182>**

Submitted on 23 Dec 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



OPEN

## A metagenomic insight into the microbiomes of geothermal springs in the Subantarctic Kerguelen Islands

Maxime Alliou<sup>1</sup>, Stéven Yvenou<sup>1</sup>, Alexander Merkel<sup>2</sup>, Marc Cozannet<sup>1</sup>, Johanne Aubé<sup>1</sup>, Jolann Pommellec<sup>1</sup>, Marc Le Romancer<sup>3</sup>, Véronique Lavastre<sup>4</sup>, Damien Guillaume<sup>4</sup> & Karine Alain<sup>1</sup>✉

The Kerguelen Islands, located in the southern part of the Indian Ocean, are very isolated geographically. The microbial diversity and communities present on the island, especially associated to geothermal springs, have never been analyzed with high-throughput sequencing methods. In this article, we performed the first metagenomics analysis of microorganisms present in Kerguelen hot springs. From four hot springs, we assembled metagenomes and recovered 42 metagenome-assembled genomes, mostly associated with new putative taxa based on phylogenomic analyses and overall genome relatedness indices. The 42 MAGs were studied in detail and showed putative affiliations to 13 new genomic species and 6 new genera of *Bacteria* or *Archaea* according to GTDB. Functional potential of MAGs suggests the presence of thermophiles and hyperthermophiles, as well as heterotrophs and primary producers possibly involved in the sulfur cycle, notably in the oxidation of sulfur compounds. This paper focused on only four of the dozens of hot springs in the Kerguelen Islands and should be considered as a preliminary study of the microorganisms inhabiting the hot springs of these isolated islands. These results show that more efforts should be made towards characterization of Kerguelen Islands ecosystems, as they represent a reservoir of unknown microbial lineages.

Terrestrial hot springs are found all over the world, on all continents, and are abundant in areas of volcanic activity such as Iceland, Japan, Russia, Chile, Algeria or New Zealand<sup>1,2</sup>. With possible exception in the polyextreme Dallol area<sup>3</sup>, all studied geothermal environments harbor microbial cohorts. Microbial cohorts in terrestrial hot springs are often composed of *Bacteria* belonging to *Aquificae*, *Chloroflexi*, *Deinococcus-Thermus* and *Thermotogae*, and *Archaea* belonging to *Desulfurococcaceae*, *Thermoproteaceae* and *Thermococcaceae* (respectively referenced as follows in the Genome Taxonomy Database: *Aquificota*, *Chloroflexota*, *Deinococcota*, *Thermotogae*, *Desulfurococcaceae*/*Ignisphaeraceae*, *Thermoproteaceae*, and *Thermococcaceae*)<sup>4-6</sup>. Most of terrestrial hot springs, like those of Yellowstone (USA) or Kamchatka (Russia) areas, have been the subject of extensive microbiological investigations. These investigations include the study of the microbial community composition, the isolation and physiological characterization of microorganisms, the investigation of adaptive mechanisms of indigenous taxa, and the mining of extremophilic species for potential enzymes, activities or molecules of biotechnological interest (e.g.<sup>7-12</sup>). Studies of bacteria and archaea living in geothermal systems are essential for our knowledge of the history of life, as these environments are early Earth analogs and one of the possible cradles of life<sup>2,6,13,14</sup>. The microbial communities of hot springs in the polar regions are partly different in their composition. For example, the fumaroles of Deception Island (Antarctica) contain prokaryotic taxa belonging to *Verrucomicrobia*,

<sup>1</sup>Univ Brest, CNRS, IFREMER, IRP 1211 MicrobSea, Unité Biologie et Ecologie des Ecosystèmes Marins Profonds BEEP, IUEM, Rue Dumont d'Urville, 29280 Plouzané, France. <sup>2</sup>Research Center of Biotechnology of the Russian Academy of Sciences, Winogradsky Institute of Microbiology, Moscow, Russia. <sup>3</sup>UBO, UFR Sciences et Techniques, UR 7462, Laboratoire Géoarchitecture, Territoires, Urbanisation, Biodiversité, Environnement, Rennes, France. <sup>4</sup>UJM, CNRS, LGL-TPE UMR5276, 42023 Saint Etienne, France. ✉email: karine.alain@univ-brest.fr

*Proteobacteria*, *Planctomycetes*, ‘*Candidatus* Parcubacteria’, *Firmicutes*, *Chloroflexi*, *Calditrichaeota*, *Bacteroidetes*, *Thaumarchaeota*, *Nanoarchaeota*, *Euryarchaeota* and *Crenarchaeota* (respectively referenced as follows in the Genome Taxonomy Database: *Verrucomicrobiota*, *Proteobacteria*, *Planctomycetota*, *Patescibacteria*, *Firmicutes*, *Chloroflexota*, *Calditrichota*, *Bacteroidota*, *Thermoproteota*, *Nanoarchaeota*, *Methanobacteriota* and *Thermoproteota*)<sup>15–17</sup>. Due to their outstanding ecological and scientific values, specifically the presence of a wide diversity of rare endemic species or under-documented taxa, Deception Island and the other Antarctic geothermal ecosystems have been designated as Antarctic Special Protected Areas (ASPAs) by the Antarctic Treaty (management plan for Antarctic Specially protected area N° 175 and N° 140)<sup>18</sup>.

The volcanic Kerguelen Archipelago, which is part of the French Southern and Antarctic Lands, is situated in the southern part of the Indian Ocean (49° S, 69° E). Located at 3300 km from the nearest inhabited areas, it is amongst the most isolated islands from any continental landmass and contain a large part of the limited terrestrial habitats present at these latitudes. The Kerguelen Islands have the status of protected areas and are under the umbrella of international conventions supporting biodiversity protection (i.e., CITES, IUCN, Convention on Biological Diversity, IPBES). They represent an important UNESCO’s world heritage site and belongs to the national nature reserve (NNR) of the French Southern Lands. Such as the ASPAs, most of Kerguelen geothermal terrestrial micro-habitats are located within “strictly protected areas” (ZPI) set aside by the Nature Reserve, in order to protect biodiversity, and geological/geomorphological features. The Kerguelen Islands and the active volcanic Heard and MacDonal Islands are the only emerged entities among the vast Kerguelen oceanic Plateau. Kerguelen Archipelago is the third largest volcanic island complex in the world, after Iceland and Hawaii<sup>19</sup>. The last volcanic activity, dated  $26 \pm 3$  Ka, took place on the Rallier du Baty (RB) Peninsula in the south-western part of the Kerguelen Islands<sup>20</sup>. Current volcanic activity, due to the Kerguelen hotspot, is evidenced by fumaroles, mud pots, hydrothermal discharges and small hot springs that rise from sea level to at least 300 m altitude. The waters of these geothermal systems contain large amounts of dissolved minerals and have wide pH ranges, from acidic to alkaline, and temperature ranges, from 35 to over 100 °C. The geochemical properties of the most accessible parts of this system have been monitored more or less regularly over the last decades<sup>21,22</sup>.

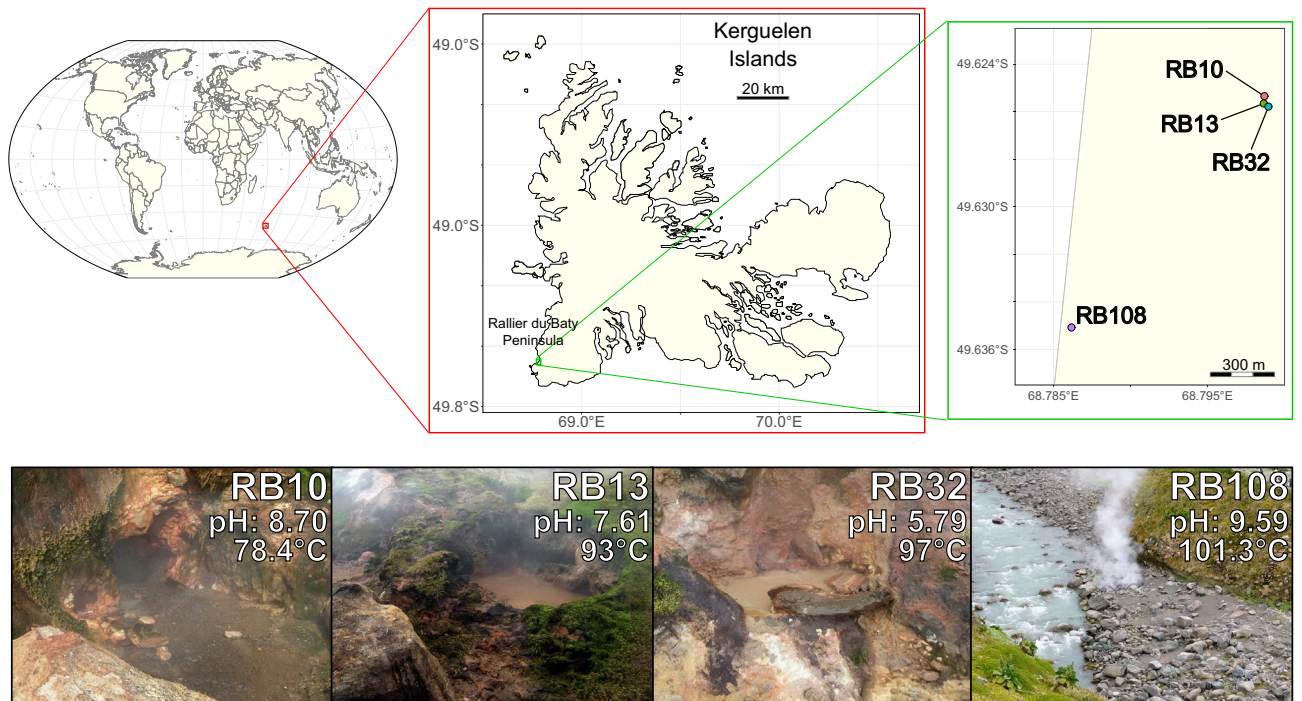
These geothermal habitats represent unique biodiversity sanctuaries in very insulated polar environments. Preliminary investigations based on 16S rRNA gene amplicon cloning and sequencing revealed a diverse collection of microbial community lineages composed of *Proteobacteria*, *Deinococcus-Thermus*, *Chloroflexi*, *Firmicutes*, *Actinobacteria* or *Aquificae*, as well as *Euryarchaeota*, *Crenarchaeota* (*Thermoproteales*, *Desulfurococcales*, *Acidilobales*, *Sulfolobales*) and *Thaumarchaeota* (respectively referenced as follows in the Genome Taxonomy Database: *Proteobacteria*, *Deinococota*, *Chloroflexota*, *Firmicutes*, *Actinobacteriota*, *Aquificota*, *Methanobacteriota*, *Thermoproteota* and *Thermoproteota*)<sup>23,24</sup>. Some of these lineages were also found at geothermal sites in Antarctica, and others represented common taxa in geochemical environments worldwide. A small number of new species have also been isolated, enriched or highlighted through molecular approaches from these regions<sup>23,25</sup>. Apart from these few studies, this area has not been subjected to any comprehensive microbiological investigation to date. The microbial diversity hosted in these hot springs remains largely unknown, as well as its functional potential (metabolism, physiology, adaptations). Microbial communities might be shaped by the biogeographic position and the physicochemical parameters of the hot springs (temperature, pH, in situ chemistry), that probably exert a strong selective pressure on indigenous communities<sup>17</sup>. Yet, these geothermal springs represent undoubtedly unique diversity, and reservoirs of new functions and innovation.

In this study, we focused on four small geothermal hot springs from the Kerguelen Islands whose microbial communities have never been studied before. We analyzed the metagenomes of the hot springs referenced as RB10, RB13 and RB32, located on the “plateau des Fumerolles” at an altitude of about 300 m on the west coast of the Rallier du Baty Peninsula, and of the ephemeral spring referenced as RB108 which flows slightly above sea level into the riverbed of the Infernet glacier (located at the base of the plateau des Fumerolles) (Fig. 1). This work was carried out from samples collected with the aim of cultivating thermophilic microorganisms, which had been stored at 4 °C for 2 years. Those storage conditions induce biases and limit the scientific discussion. Therefore, the generated data did not allow ecological or evolutionary interpretations. These samples were collected in a particularly isolated natural reserve that requires specific sampling permits, and are therefore very difficult to obtain, making them particularly valuable. This descriptive study provides the first sequenced and analyzed MAGs from these undisturbed areas and points out that more efforts must be made to characterize Kerguelen Island ecosystems. This insular area contains a reservoir of unknown microbial lineages, and thus possibly biological and genomic novelties, the study of which may prove useful in the future to discover new metabolic pathways or molecules and to understand the part of deterministic and stochastic processes underlying microbial assemblages.

## Results and discussion

**MAG binning and general features.** From the four hot springs, we assembled four associated metagenomes and then binned a total of 42 MAGs. We recovered 12 MAGs from RB10 hot spring, 13 from RB13, 14 from RB32 and 3 from RB108. Out of these 42 MAGs, 7 were of high-quality, 25 of nearly-high quality, 9 of medium quality and 1 of low quality (Table 1) based on metagenomic standards<sup>26</sup>. The GC% was quite variable, ranging from 25.76 to 70.35% among all MAGs and between 32.15 and 69.21% only among the high- and near high-quality MAGs. With the exception of RB108 from which we only recovered bacterial MAGs, we retrieved both bacterial and archaeal MAGs in the other hot springs. Two thirds of the MAGs (26/42) were assigned to the domain *Bacteria* and the rest to the domain *Archaea* (16/42) (Table 2).

**Taxonomic and phylogenomic analyses of MAGs.** The taxonomic affiliation of the MAGs was investigated in detail through the workflow classify of GTDB-Tk (v 2.1.0; GTDB reference tree 07-RS207) (Table 2)



**Figure 1.** Sampling locations at the plateau des Fumerolles in the Rallier du Baty Peninsula (Kerguelen Islands, French Southern and Antarctic Lands), and photographs of the four hot springs studied here with their associated temperature and pH conditions. The world map and details of the Kerguelen seashore were generated using the ‘cowplot’, ‘ggplot2’, ‘lwgeom’, ‘rnaturalearth’, ‘rnaturalearthdata’, ‘rworldmap’, ‘sf’ and ‘tidyverse’ libraries implemented on the RStudio software (v1.2.5001). A 2018 map of the Kerguelen islands was retrieved from the French national portal of maritime limits (<http://limitesmaritimes.gouv.fr/>) according to the 2017 Etalab open license (v2.0—<https://www.etalab.gouv.fr/>). Pictures depicting sampling sites were taken by Marc Le Romancer. Layout was carried out using the free open source vector image editor Inkscape (v1.0).

and through de novo phylogenomic analyses (Fig. S1a–i). We also tried to classify MAGs on the basis of overall genome relatedness indices (OGRI), which is detailed in supplementary material (Text S1, Table S2, Fig. S2).

De novo phylogenomic analyses globally confirmed the positioning of MAGs provided by GTDB-Tk, with high branching support. For *Bacteria*, GTDB-Tk analyses allowed us to place the MAGs in the following clades: six in the phylum *Aquificota* from the four different springs, comprising four MAGs belonging to the genus *Hydrogenivirga* (family *Aquificaceae*) (RB10-MAG07, RB13-MAG10, RB32-MAG07, RB108-MAG02), and two belonging to the family *Hydrogenobaculaceae* (RB10-MAG12, RB32-MAG11) (Table 2, Fig. S1a). Their closest cultured relatives originated either from hot springs or from deep-sea hydrothermal vents<sup>27</sup>. Three MAGs from three geothermal springs belonged to the phylum *Armatimonadota* (RB10-MAG03, RB13-MAG04, RB32-MAG03) and had no close cultured relatives. Seven MAGs have been classified into the phylum *Chloroflexota*: three MAGs belonging to the genus *Thermoflexus* from three different springs (RB10-MAG04, RB13-MAG05, RB32-MAG02), one affiliating with the genus *Thermomicrobium* (RB32-MAG08), one falling into the family *Ktedonobacteraceae* (RB108-MAG03), one belonging to the class *Dehalococcoidia* (RB32-MAG04) and another one whose phylogenetic position is more difficult to assert because it is a MAG of medium quality (RB32-MAG14). Six MAGs from four various hot springs belonged to the phylum *Deinococcota*, and to the genera *Thermus* (RB10-MAG08, RB10-MAG11, RB13-MAG09, RB32-MAG10, RB108-MAG01) and *Meiothermus* (RB13-MAG13). One MAG belonged to the family *Sulfurifustaceae* (RB13-MAG01), in the phylum *Proteobacteria* (*Gamma*-class). The MAG referenced as RB32-MAG13 was classified into the phylum *Patescibacteria*, in the class *Paceibacteria*, and was distantly related to MAGs originating from groundwater and from hot springs. Finally, two MAGs from two different springs belonged to the phylum WOR-3, in the *Candidatus* genus *Caldipriscus* (RB32-MAG12, RB10-MAG09).

For *Archaea*, almost all the MAGs reconstructed in this study, e.g. 15 of the 16 archaeal MAGs, belonged to the phylum *Thermoproteota*. Among them, four belonged to the genus *Ignisphaera* (RB10-MAG05, RB13-MAG08, RB13-MAG11, RB32-MAG05), three to the genus *Infirmifilum* (RB10-MAG06, RB13-MAG03, RB32-MAG09), two to the genus *Zestosphaera* (RB10-MAG02, RB13-MAG06), three to the family *Acidilobaceae* (RB10-MAG01, RB13-MAG02, RB32-MAG01) and two to the order *Geoarchaeales* (RB10-MAG10, RB32-MAG06). Additionally, one belonged to the family *Thermocladiaceae* (RB13-MAG07). Lastly, the MAG belonging to another phylum (RB13-MAG12) was affiliated with the *Aenigmataarchaeota*, class *Aenigmataarchaeia*, and was distantly related to MAGs from hot springs and from deep-sea hydrothermal vent sediments<sup>28,29</sup>.

Out of these 42 MAGs, at least 19 MAGs corresponded to different taxa at the taxonomic rank of species or higher according to GTDB (Table 2). Eighteen of them belonged to lineages with several cultivated representatives



Sample	MAG quality	Percent complet.	Percent contam.	Mean Coverage	GC content	Size (Mbp)	Contig number	N50 (kbp)	tRNA	rRNA sequences (5S-16S-23S)	Assembly accession number	Sample accession number
RB10-MAG01	NHQ	96.84	0.95	61.98	57.01	1.78	59	78.449	35	1-0-0	GCA_916099525	ERS7299774
RB10-MAG02	NHQ	98	0.63	40.11	38.88	1.49	212	15.625	33	1-0-0	GCA_916101595	ERS7299775
RB10-MAG03	NHQ	92.59	0.19	38.28	56.24	2.65	72	76.979	47	1/1/2000	GCA_916102575	ERS7299776
RB10-MAG04	HQ	92.73	1.82	125.94	65.49	2.5	119	38.887	45	1/1/2001	GCA_916102815	ERS7299777
RB10-MAG05	NHQ	94.34	0	20.12	33.97	1.6	309	8.06	17	1-0-0	GCA_916103375	ERS7299778
RB10-MAG06	HQ	97.79	0.74	74.51	53.43	1.72	322	11.699	19	1/1/2001	GCA_916103935	ERS7299779
RB10-MAG07	HQ	97.56	0.41	4822.7	53.78	1.44	153	17.997	43	1/1/2001	GCA_916103025	ERS7299780
RB10-MAG08	HQ	96.75	0.85	448.27	69.15	2.01	196	20.594	45	0-0-0	GCA_916101605	ERS7299781
RB10-MAG09	MQ	54.88	0	12.64	42.63	0.75	417	1.86	16	0-0-0	GCA_916098365	ERS7299782
RB10-MAG10	MQ	57.09	0	12.99	33.71	0.86	455	2	10	0-0-0	GCA_916100445	ERS7299783
RB10-MAG11	NHQ	91.61	0.14	118.61	65.66	1.78	159	22.734	33	0-0-0	GCA_916100255	ERS7299784
RB10-MAG12	MQ	54.21	0.81	13.87	31.69	0.75	413	1.839	12	1-0-0	GCA_916103405	ERS7299785
RB13-MAG01	NHQ	93.29	1.22	30.58	67.39	2.2	188	24.091	35	1-0-1	GCA_916101585	ERS7299786
RB13-MAG02	NHQ	96.52	0.63	1385.16	56.92	1.82	69	74.252	33	1-0-1	GCA_916101625	ERS7299787
RB13-MAG03	NHQ	97.79	0.74	189.7	53.64	1.67	246	22.516	47	1-0-0	GCA_916101685	ERS7299788
RB13-MAG04	NHQ	90.74	0.19	23.91	56.24	2.56	329	11.396	45	1-0-0	GCA_916101675	ERS7299789
RB13-MAG05	NHQ	93.03	1.01	31.18	65.49	2.35	243	14.935	17	1-0-1	GCA_916104055	ERS7299790
RB13-MAG06	NHQ	97.2	1.58	224.61	38.67	1.41	228	9.959	19	1-0-0	GCA_916104685	ERS7299791
RB13-MAG07	NHQ	92.65	0.74	36.1	52.52	1.24	149	14.053	43	1-0-0	GCA_916099355	ERS7299792
RB13-MAG08	NHQ	92.72	1.58	209.57	36.07	1.23	126	22.47	45	0-0-0	GCA_916104555	ERS7299793
RB13-MAG09	NHQ	100	0	118.59	65.33	2.01	179	35.402	16	0-0-0	GCA_916098735	ERS7299794
RB13-MAG10	HQ	97.43	0.41	118.59	53.62	1.41	147	14.581	10	1/1/2001	GCA_916103555	ERS7299795
RB13-MAG11	MQ	87.62	0.95	574.42	34.11	1.5	306	7.292	33	0-0-0	GCA_916103205	ERS7299796
RB13-MAG12	MQ	72.9	3.27	125.85	25.76	0.74	154	8.912	12	0-1-0	GCA_916101905	ERS7299797
RB13-MAG13	LQ	30.02	0.43	90.33	66.35	0.83	520	1.613	35	1-0-1	GCA_916101985	ERS7299798
RB32-MAG01	NHQ	97.47	0.95	29.75	56.8	1.75	57	88.979	33	0-0-0	GCA_916099235	ERS7299799
RB32-MAG02	HQ	94.55	1.82	421.42	65.38	2.54	148	29.008	47	1/1/2002	GCA_916099675	ERS7299800
RB32-MAG03	HQ	92.59	0.19	845.13	56.23	2.67	84	95.702	45	1/1/2001	GCA_916102075	ERS7299801
RB32-MAG04	NHQ	94.72	0	24.31	64.77	2.45	128	40.031	17	1-0-1	GCA_916098805	ERS7299802
RB32-MAG05	NHQ	98.73	0	28.06	34.36	1.76	238	16.008	19	1-0-1	GCA_916098685	ERS7299803
RB32-MAG06	NHQ	91.18	0	302.91	34.75	1.96	305	12.851	43	1-0-0	GCA_916102535	ERS7299804
RB32-MAG07	NHQ	97.83	0.41	1025.5	53.5	1.45	157	17.867	45	1-0-1	GCA_916109635	ERS7299805
RB32-MAG08	NHQ	94.95	0.93	1025.5	65.25	2.55	629	5.755	16	0-0-0	GCA_916109905	ERS7299806
RB32-MAG09	NHQ	94.85	1.47	94.92	53.78	1.57	334	8.549	10	1-0-0	GCA_916110065	ERS7299807
RB32-MAG10	NHQ	100	1.27	184.55	65.39	1.99	128	41.211	33	1-0-1	GCA_916105685	ERS7299808
RB32-MAG11	NHQ	94.7	2.28	20.34	32.15	1.3	295	6.491	12	0-0-0	GCA_916112325	ERS7299809
RB32-MAG12	MQ	80.35	0	11.95	42.38	1.2	362	20.577	35	1-0-0	GCA_916112475	ERS7299810
RB32-MAG13	MQ	70.82	0	129.49	27.82	0.55	176	19.96	33	0-0-1	GCA_916111365	ERS7299811
RB32-MAG14	MQ	55.12	0	11.32	70.35	1.47	834	1.812	47	0-0-0	GCA_916108235	ERS7299812
RB108-MAG01	NHQ	99.15	1.27	1120.09	69.21	2.04	126	30.509	45	0-0-0	GCA_916116595	ERS7299813
RB108-MAG02	NHQ	96.75	0	579.13	53.28	1.45	75	29.169	17	1-0-0	GCA_916109075	ERS7299814
RB108-MAG03	MQ	53.38	4.73	86.68	56.73	3.96	2105	1.944	19	0-0-0	GCA_916109065	ERS7299815

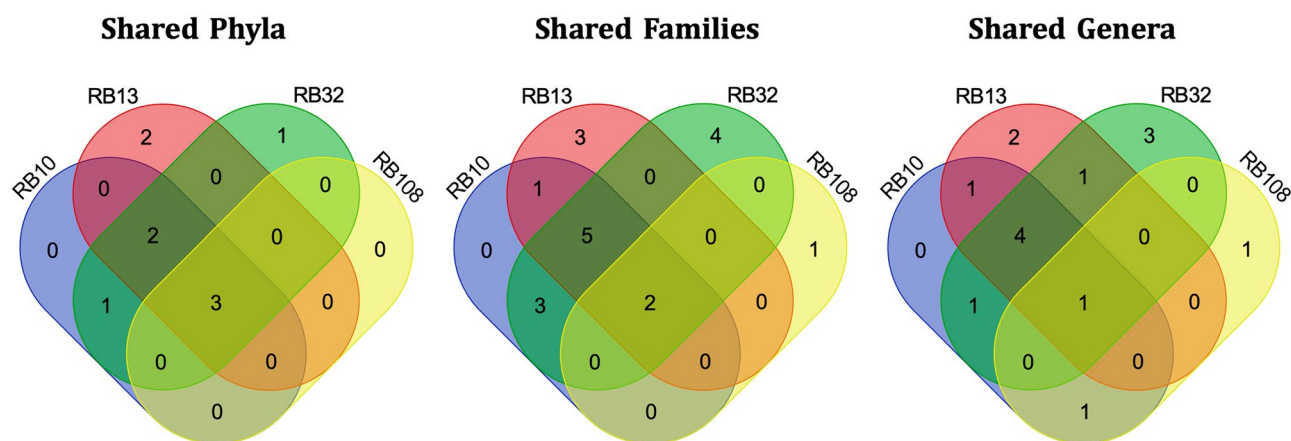
**Table 1.** General characteristics of the 42 MAGs obtained from RB10, RB13, RB32 and RB108 samples. In accordance with current standards, the bins were defined as high quality (HQ) (>90% completion, <5% contamination, presence of the 23S, 16S and 5S rRNA genes and at least 18 tRNAs), near high quality (NHQ) (>90% completion, <5% contamination, other criteria partially covered), medium quality (MQ) (≥50% completion, <10% contamination), and low quality (LQ) (<50% completion, <10% contamination) MAGs. Metagenomes and MAGs accession numbers are available on ENA (Study ID: PRJEB46766).

including the species *Thermus thermophilus*. 13 new genomic species within the GTDB genera *Hydrogenivirga*, *HRBIN17*, *Thermoflexus*, *SpSt-223*, *CADDYT01*, *Zestosphaera*, *Ignisphaera*, *Infirmifilum*, *Thermus*, *Thermus\_A*, *Meiothermus\_B*, *JAHLMO01* and *Caldipriscus*, and 6 putative new genomic genera belonging to the GTDB families *Hydrogenobaculaceae*, *Acidilobaceae*, *WAQG01*, *Thermocladaceae*, *Sulfurifustaceae* and *HR35* could be identified (Table 2). In addition, 9 MAGs belonged to lineages that are predominantly or exclusively known through environmental DNA sequences. Thus, these 42 MAGs comprised a broad phylogenetic range of *Bacteria* and *Archaea* at different levels of taxonomic organization, of which a large majority were not reported before.

The approaches implemented here were not intended to describe the microbial diversity present in these sources in an exhaustive way or to compare them in a fine way, and cannot allow it because of a 2-year storage at 4 °C. This long storage has probably led to changes in the microbial communities and to the selective loss or enrichment of some taxa. As a result, no analysis of abundance or absence of taxa can be conducted from these metagenomes and the results are discussed taking this bias into account. However, they do provide an overview of the microbial diversity effectively present. If we compare the phylogenetic diversity of the MAGs found in the four hot springs, we can observe that 3 shared phyla (*Deinococcota*, *Aquificota* and *Chloroflexota*: phyla names according to GTDB), 2 shared families (*Thermaceae* and *Aquificaceae*), and one shared genus (*Hydrogenivirga*) were found among the four sources (Fig. 2). In addition, hot springs RB10, RB13 and RB32, that are geographically close (<60 m), also share 2 other phyla (*Thermoproteota* and *Armatimonadota*) and 5 other families in

MAG ID	Domain	Phylum	Class	Order	Family	Genus	Species
RB10-MAG07	Bacteria	Aquificota	Aquificae	Aquificales	Aquificaceae	Hydrogenivirga	Not assigned
RB13-MAG10	Bacteria	Aquificota	Aquificae	Aquificales	Aquificaceae	Hydrogenivirga	Not assigned
RB32-MAG07	Bacteria	Aquificota	Aquificae	Aquificales	Aquificaceae	Hydrogenivirga	Not assigned
RB108-MAG02	Bacteria	Aquificota	Aquificae	Aquificales	Aquificaceae	Hydrogenivirga	Not assigned
RB10-MAG12	Bacteria	Aquificota	Aquificae	Aquificales	Hydrogenobaculaceae	Not assigned	Not assigned
RB32-MAG11	Bacteria	Aquificota	Aquificae	Aquificales	Hydrogenobaculaceae	Not assigned	Not assigned
RB10-MAG03	Bacteria	Armatimonadota	HRBIN17	HRBIN17	HRBIN17	HRBIN17	Not assigned
RB13-MAG04	Bacteria	Armatimonadota	HRBIN17	HRBIN17	HRBIN17	HRBIN17	Not assigned
RB32-MAG03	Bacteria	Armatimonadota	HRBIN17	HRBIN17	HRBIN17	HRBIN17	Not assigned
RB10-MAG04	Bacteria	Chloroflexota	Anaerolineae	Thermoflexales	Thermoflexaceae	Thermoflexus	Not assigned
RB13-MAG05	Bacteria	Chloroflexota	Anaerolineae	Thermoflexales	Thermoflexaceae	Thermoflexus	Not assigned
RB32-MAG02	Bacteria	Chloroflexota	Anaerolineae	Thermoflexales	Thermoflexaceae	Thermoflexus	Not assigned
RB32-MAG04	Bacteria	Chloroflexota	Dehalococcoidia	SpSt-223	SpSt-223	SpSt-223	Not assigned
RB32-MAG08	Bacteria	Chloroflexota	Chloroflexia	Thermomicrobiales	Thermomicrobiaceae	Thermomicrobium	Thermomicrobium sp002898975
RB32-MAG14	Bacteria	Chloroflexota	FW602-bin22	FW602-bin22	DSKJ01	DSKJ01	DSKJ01 sp011051445
RB108-MAG03	Bacteria	Chloroflexota	Ktedonobacteria	Ktedonobacterales	Ktedonobacteraceae	CADDYT01	Not assigned
RB10-MAG01	Archaea	Thermoproteota	Thermoproteia	Sulfolobales	Acidilobaceae	Not assigned	Not assigned
RB13-MAG02	Archaea	Thermoproteota	Thermoproteia	Sulfolobales	Acidilobaceae	Not assigned	Not assigned
RB32-MAG01	Archaea	Thermoproteota	Thermoproteia	Sulfolobales	Acidilobaceae	Not assigned	Not assigned
RB10-MAG02	Archaea	Thermoproteota	Thermoproteia	Sulfolobales	NBVN01	Zestosphera	Not assigned
RB13-MAG06	Archaea	Thermoproteota	Thermoproteia	Sulfolobales	NBVN01	Zestosphera	Not assigned
RB10-MAG05	Archaea	Thermoproteota	Thermoproteia	Sulfolobales	Ignisphaeraceae	Ignisphaera	Not assigned
RB13-MAG08	Archaea	Thermoproteota	Thermoproteia	Sulfolobales	Ignisphaeraceae	Ignisphaera	Not assigned
RB13-MAG11	Archaea	Thermoproteota	Thermoproteia	Sulfolobales	Ignisphaeraceae	Ignisphaera	Not assigned
RB32-MAG05	Archaea	Thermoproteota	Thermoproteia	Sulfolobales	Ignisphaeraceae	Ignisphaera	Not assigned
RB10-MAG06	Archaea	Thermoproteota	Thermoproteia	Thermofilales	Thermofilaceae	Infirmifilum	Not assigned
RB13-MAG03	Archaea	Thermoproteota	Thermoproteia	Thermofilales	Thermofilaceae	Infirmifilum	Not assigned
RB32-MAG09	Archaea	Thermoproteota	Thermoproteia	Thermofilales	Thermofilaceae	Infirmifilum	Not assigned
RB10-MAG10	Archaea	Thermoproteota	Thermoproteia	Geoarchaeales	WAQG01	Not assigned	Not assigned
RB32-MAG06	Archaea	Thermoproteota	Thermoproteia	Geoarchaeales	WAQG01	Not assigned	Not assigned
RB13-MAG07	Archaea	Thermoproteota	Thermoproteia	Thermoproteales	Thermocliadiaceae	Not assigned	Not assigned
RB10-MAG08	Bacteria	Deinococcota	Deinococci	Deinococcales	Thermaceae	Thermus	thermophilus
RB108-MAG01	Bacteria	Deinococcota	Deinococci	Deinococcales	Thermaceae	Thermus	thermophilus
RB10-MAG11	Bacteria	Deinococcota	Deinococci	Deinococcales	Thermaceae	Thermus	Not assigned
RB13-MAG09	Bacteria	Deinococcota	Deinococci	Deinococcales	Thermaceae	Thermus_A	Not assigned
RB32-MAG10	Bacteria	Deinococcota	Deinococci	Deinococcales	Thermaceae	Thermus_A	Not assigned
RB13-MAG13	Bacteria	Deinococcota	Deinococci	Deinococcales	Thermaceae	Meiothermus_B	Not assigned
RB13-MAG12	Archaea	Aenigmatarchaeta	Aenigmatarchaeta	CG10238-14	EX4484-224	JAHLMO01	Not assigned
RB13-MAG01	Bacteria	Proteobacteria	Gammaproteobacteria	Acidiferrobacterales	Sulfurifustaceae	Not assigned	Not assigned
RB32-MAG13	Bacteria	Patescibacteria	Patescibacteria	UBA6257	HR35	Not assigned	Not assigned
RB10-MAG09	Bacteria	WOR-3	Hydrothermia	LBFQ01	LBFQ01	Caldirpriscus	Not assigned
RB32-MAG12	Bacteria	WOR-3	Hydrothermia	LBFQ01	LBFQ01	Caldirpriscus	Not assigned

**Table 2.** Classification of the MAGs based on the taxonomic classification of GTDB-Tk software (v2.1.0) and the Genome Taxonomy Database (07-RS207 release).



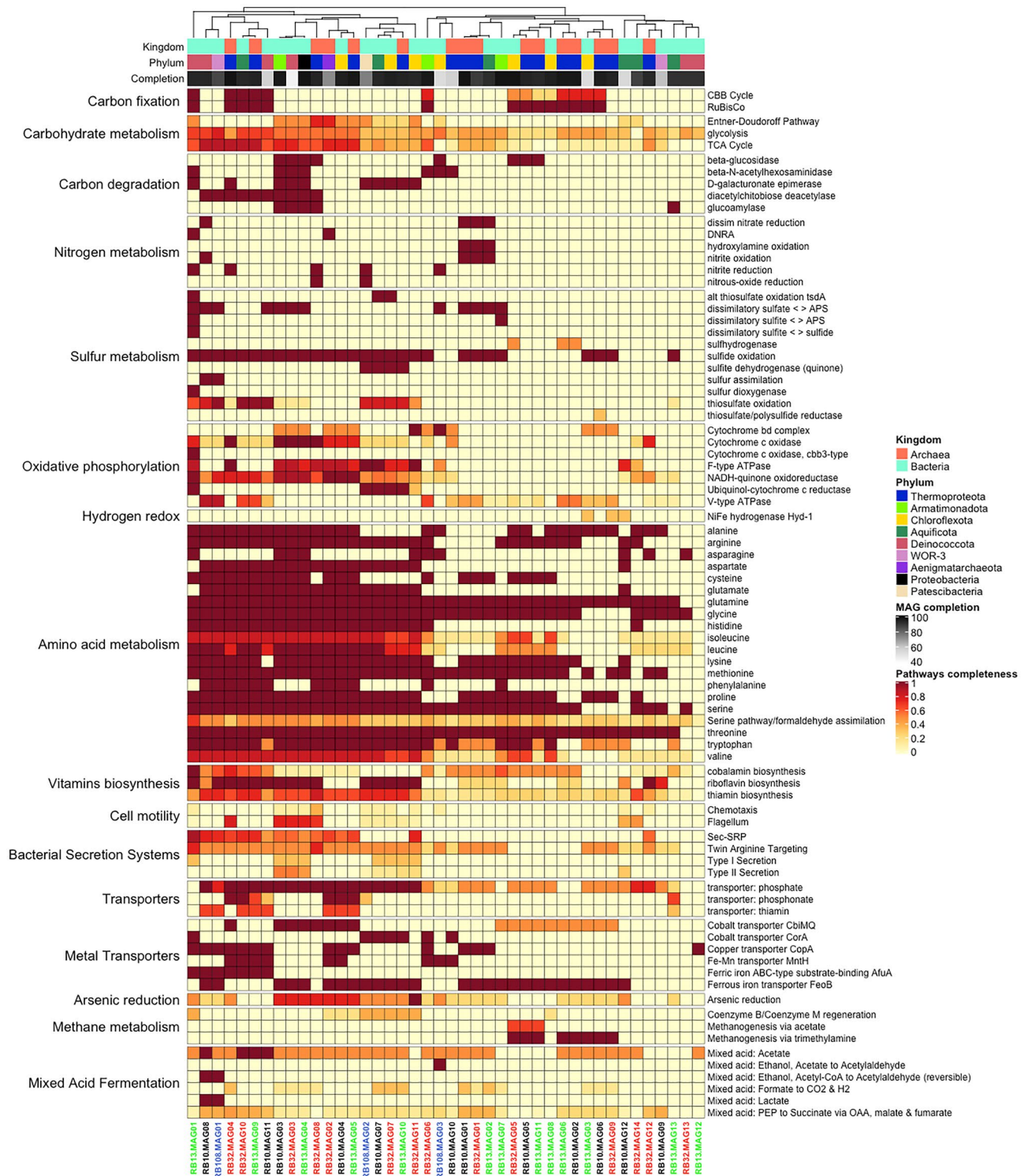
**Figure 2.** Venn diagrams showing the shared phyla, families and genera according to GTDB classification in the reconstructed MAGs from the hot springs RB10, RB13, RB32 and RB108.

common (*Acidilobaceae*, *Ignisphaeraceae*, *Thermofilaceae*, *Thermoflexaceae*, and HRBIN17) (Fig. 2). These phyla and families that are shared between sources are widespread lineages in terrestrial geothermal habitats (e.g.<sup>4–6,12</sup>). Phyla and families detected in the hot environments of Antarctica are also found here, such as *Patescibacteria*<sup>15</sup>. In summary, this metagenomic analysis highlighted the presence of bacterial and archaeal lineages commonly found in hot springs, and lineages found in hot habitats from polar areas (e.g.<sup>4–6,15,30</sup>). The microbial communities in these Kerguelen Islands hot springs samples were diverse, particularly in RB10, RB13, and RB32 hot springs. Within these lineages previously reported to occur in geothermal environments, a majority of the genomic taxa detected here were novel. Those results were obtained considering their taxonomic affiliation by GTDB-Tk, and their phylogenomic position with respect to closest relatives and the OGRI thresholds (16S rRNA gene sequence similarity, average nucleotide identity, and average amino acid identity) classically used to delineate different taxonomic ranks in cultured strains, used here as indicators of taxonomic differentiation (Table 2, Table S2).

**Functional potential of MAGs: putative metabolisms and adaptations.** A genomic characterization of the 42 MAGs has been performed to explore the possible metabolic pathways and adaptations of the microbial populations from which these MAGs originate. KEGG Decoder visualization highlighted various pathways associated with carbohydrate degradation, oxidative phosphorylation and sulfur, nitrogen, and amino acid metabolisms, among others (Fig. 3). To confirm these initial metabolic predictions, a further annotation was performed by combining data generated by Prokka with the MetaCyc database. Efforts have been directed at studying catabolic pathways, particularly those involving inorganic electron donors and acceptors. These results are not representative of the metabolic diversity of all the hot spring ecosystems studied, but they do reflect some of the microbial catabolism likely to be used in situ to produce energy and, by assumption, the most abundant ones. Metabolic predictions are presented hereafter, at different taxonomic ranks and have been compared to the known phenotypes of the closest cultivated taxa, and in some cases to genomic content of the closest relatives.

MAGs belonging to the genus *Thermoflexus* (RB10-MAG04, RB13-MAG05, RB32-MAG02) encode pathways for carbon monoxide oxidation (via aerobic carbon monoxide dehydrogenase), hydrogen oxidation and nitrate respiration; the only cultivated known representative of this genus is a heterotrophic bacterium<sup>31</sup>. The same pathways, except for the nitrate reduction pathway, are encoded in the complete genome of *Thermoflexus hugenholtzii* (NCBI: ASM1877156v1). In contrast, the genome of *T. hugenholtzii*, a strain isolated from a terrestrial hot spring in Nevada<sup>31</sup>, encodes the tetrathionate reduction and thiosulfate disproportionation pathways, which are not encoded in the three Kerguelen Island MAGs. The *Dehalococcoidia*'s MAG (RB32-MAG04) encodes only a carbon monoxide oxidation pathway; the cultivated members of this genus are strict anaerobic hydrogenotrophic, organohalide-respiring bacteria<sup>32</sup>. In the MAG associated with the genus *Thermomicrobium* (RB32-MAG08), we predicted pathways for dimethylsulfide degradation, thiosulfate disproportionation and carbon monoxide oxidation; only carboxydrotrophic growth has been reported in this genus and demonstrated by culture<sup>33</sup>. The same pathways are encoded in the complete genome of *Thermomicrobium roseum*, a strain isolated from a hot spring in Yellowstone National Park (NCBI: ASM2168v1)<sup>34</sup>. In the *Chloroflexota*'s MAG (RB32-MAG14) (belonging to the order *Chloroflexales*, Table S2), carbon monoxide oxidation and thiosulfate disproportionation pathways are present but no coding DNA sequence associated with phototrophy could be found, which may suggest a chemoorganotrophy mode of energy production<sup>35</sup>. The *Ktedonobacteraceae*'s MAG (RB108-MAG03) encodes enzymes for hydrogen oxidation (aerobic) pathways, carbon monoxide oxidation, dimethylsulfide degradation, selenate reduction, thiosulfate oxidation and disproportionation and finally tetrathionate oxidation; yet, the few taxa of this family isolated so far are mesophilic heterotrophic bacteria<sup>36</sup>. Within *Hydrogenobaculaceae* MAGs (RB10-MAG12, RB32-MAG11), we predicted a thiosulfate disproportionation pathway; most of the species within this family are capable of chemolithotrophic microaerophilic or anaerobic growth<sup>37</sup>. MAGs belonging to the genus *Hydrogenivirga* (RB10-MAG07, RB13-MAG10, RB32-MAG07, RB108-MAG02) possess genes encoding enzymes of aerobic respiration, thiosulfate oxidation, thiosulfate disproportionation, tetrathionate reduction, and hydrogen oxidation (aerobic and anaerobic); which is consistent with what is known about the genus (nitrate and oxygen respiration combined to hydrogen, sulfur, or thiosulfate oxidation)<sup>37</sup>. The same pathways, with the exception of the hydrogen oxidation pathway, are encoded in the genome of *Hydrogenivirga caldilitoris* (NCBI: ASM366400v1), a close relative isolated from a coastal hot spring in Japan<sup>27</sup>. In MAGs associated with the genus *Thermus* (RB10-MAG08, RB10-MAG11, RB13-MAG09, RB32-MAG10, RB108-MAG01), we predicted pathways for aerobic respiration, assimilatory sulfate reduction, hydrogen oxidation, selenate reduction, thiosulfate oxidation and thiosulfate disproportionation; cultivated species of this genus grow mainly chemoorganoheterotrophically by aerobic respiration, but some have genes coding for chemolithotrophic and anaerobic respiration enzymes<sup>38</sup>. The MAG belonging to the genus *Meiothermus* (RB13-MAG13) encodes pathways for carbon monoxide oxidation, hydrogen oxidation, thiosulfate oxidation and thiosulfate disproportionation; *Meiothermus* strains are known to grow chemoorganotrophically by oxygen or nitrate respiration<sup>39</sup>. For the RB13-MAG01 belonging to the *Sulfurifustaceae*, we predicted the genetic potential for aerobic respiration, ammonia oxidation, dissimilatory sulfate reduction, sulfite oxidation, sulfide oxidation (to sulfur globules), tetrathionate reduction, thiosulfate oxidation and thiosulfate disproportionation; *Sulfurifustaceae* (referenced as *Acidiferrobacteraceae* in the LPSN taxonomy) are known to be able to oxidize sulfur and iron, and the microorganism corresponding to this MAG may possess a larger panel of chemolithotrophic abilities<sup>40</sup>. For members of the *Armatimonadota* (RB10-MAG03, RB13-MAG04, RB32-MAG03), we predicted pathways for assimilatory sulfate reduction, carbon monoxide oxidation, selenate reduction and thiosulfate disproportionation; the members of the phylum have a phenotype of aerobic heterotrophs<sup>41</sup>. In *Zestosphaera*'s (RB10-MAG02, RB13-MAG06) and *Ignisphaera*'s (RB10-MAG05, RB13-MAG08, RB13-MAG11, RB32-MAG05) MAGs, we predicted sulfur and polysulfide reduction pathways; those MAGs could be classified as *Desulfurococcaceae* (LPSN taxonomy, Table S2) which are known as heterotrophs respiring sulfur species<sup>42,43</sup>. MAGs belonging to the class *Thermoproteia* (RB10-MAG10,





**Figure 3.** Metabolic pathway diagram of the 42 MAGs based on KEGG Decoder annotations, showing MAG classification according to GTDB-Tk and estimated genome completion.

RB13-MAG07, RB32-MAG06) encode dissimilatory sulfate reduction pathway; various catabolic pathways are described in this class<sup>44</sup>. In MAGs related to the genus *Caldipriscus* (RB10-MAG09, RB32-MAG12), phylum *Patescibacteria* (RB32-MAG13), family *Acidilobaceae* (RB10-MAG01, RB13-MAG02, RB32-MAG01), family *Thermofilaceae* (RB10-MAG06, RB13-MAG03, RB32-MAG09) and class *Aenigmataarchaeia* (RB13-MAG12), we did not predict any catabolic pathway of inorganic nutrients among those reported in the MetaCyc database. This could be explained by the low completion of the MAGs and/or the fact that only well-known pathways are documented in this database. However, all these MAGs have pathways associated with carbohydrate and



protein degradation. This may indicate that these taxa are chemoheterotrophs, which has already been reported in geothermal environments and already described for relatives of some of these taxa<sup>45,46</sup>.

Sulfide oxidation may be a possible energy production pathway for 28 MAGs based on KEGG Decoder (Fig. 3), since they code for a sulfide:quinone oxidoreductase (KEGG:K17218) and a flavoprotein chain of sulfide dehydrogenase (KEGG:K17229), but this hypothesis was not confirmed by MetaCyc except for RB13-MAG01. Due to high representations of sulfur metabolisms, genes encoded in MAGs were evaluated with DiSCo, which gave similar results to those obtained when analyzed with Pathway tools. DiSCo confirmed complete dissimilatory sulfate reduction pathways for two MAGs, predicted to be associated to sulfate reduction processes (RB13-MAG07) or sulfide oxidation processes by reverse sulfate reduction pathway (RB13-MAG01). The assimilatory sulfate reduction pathway is more represented in the overall dataset formed by all MAGs than the dissimilatory pathway, which is consistent with the low sulfate concentration measured in the four hot springs (Table S1). The thiosulfate disproportionation pathway predicted by MetaCyc in many MAGs simply refers to the detection of an enzyme, the rhodanese-type thiosulfate sulfurtransferase. However, in the current state of knowledge on the disproportionation pathways of inorganic sulfur compounds<sup>47,48</sup>, this enzyme alone does not allow the implementation of this catabolic pathway. If we consider all the genes present in these MAGs, nothing indicates that the microorganisms from which these MAGs originate can achieve the disproportionation of inorganic sulfur compounds.

Additionally, no enzymes clearly associated with photosystems I and II were found. Nevertheless, it cannot be ruled out that these energy production pathways are absent in microorganisms indigenous to these sources, due to sample storage bias and low completion of some MAGs. On the other hand, our results show that these sources host chemolithoautotrophic taxa involved in the carbon and sulfur cycle, and to a lesser extent in the hydrogen and nitrogen cycles. Several taxa are likely to be involved in the primary production of these sources through chemolithoautotrophy, but in addition, heterotrophs appear to be present and diverse in the collected samples. Additional studies will be required to better understand the metabolic diversity and trophic webs of these hot springs, in order to better understand the ecology of the microbial communities of the Kerguelen hot springs.

Regarding thermophily, we found that all MAGs encode heat shock proteins, mainly associated with the HSP20 family, with the exception of RB10-MAG12 and RB13-MAG12. The absence of Hsp encoding genes in these two MAGs is possibly due to the low genome completeness of these two MAGs. Under conditions of heat stress, it has been shown that the small heat shock proteins Hsp20, protect cellular proteins from aggregation and membrane lipids from destabilization, in some thermophilic archaea<sup>49</sup>. In taxa of these geothermal sources, these proteins could help the cells to counteract the deleterious effects of environmental stress and in particular of thermal stress. In addition, reverse gyrase coding sequences were found in 29 out of the 42 MAGs; these enzymes are known to be exclusive to hyperthermophiles and involved in DNA protection and repair at high temperatures<sup>50</sup>. Only MAGs RB10-MAG04, RB10-MAG09, RB10-MAG11, RB13-MAG01, RB13-MAG05, RB13-MAG09, RB13-MAG13, RB32-MAG02, RB32-MAG04, RB32-MAG08, RB32-MAG10, RB32-MAG12 and RB108-MAG03, belonging to the phylum *Chloroflexota*, the family *Sulfurifustaceae* or the genus *Caldipriscus* (GTDB taxonomy), do not encode any reverse gyrase gene. These results suggest the presence of numerous thermophilic and hyperthermophilic prokaryotes in these high temperature hot springs. Further cultural and physiological investigations from samples of these Kerguelen hot springs will be necessary to confirm these statements.

In conclusion, this first metagenomic overview of the microbial diversity of Kerguelen hot springs allowed the assembly of 42 MAGs, from four hot springs. Several MAGs correspond to putative new taxa, namely 13 new putative genomic species and 6 new putative genera affiliated to *Bacteria* and *Archaea* according to GTDB. Based on their genetic potential, these taxa appear to be chemolithoautotrophs and chemoheterotrophs and thus probably involved in the carbon, sulfur, hydrogen and nitrogen cycle. Many of these MAGs are likely to be derived from populations of thermophilic/hyperthermophilic bacteria and archaea. As geographically isolated sites, the Kerguelen Islands are reservoirs of diversity and taxa of novel microorganisms that should be interesting to study the evolution of microbial life and speciation processes. It has been difficult to fully assess the microbial metabolic diversity in these geothermal pools due to the inherent limitations of MAG reconstruction and the state of knowledge of microbial pathways that remains limited. However, these geothermal ecosystems could be reservoirs of biological and genomic novelty. The physiological properties and adaptive mechanisms of microorganisms inhabiting these unique environments will deserve to be examined in detail in the future by implementing large-scale metagenomics, metatranscriptomics and cultural analyses.

## Methods

**Sample collection and major elements analysis in water samples.** Water samples were collected from four hot springs during the 2016–2017 austral summer TALISKER field campaign (1st of December–11th of February) organized by the French Polar Institute Paul Emile Victor (<https://institut-polaire.fr/en/>). Water samples and water samples mixed with surficial sediments aliquots were collected. Water samples were collected in 250ML LDPE Nalgene bottles stored at 4 °C until ionic chromatography analysis. Mixed water and sediment samples were collected aseptically in sterile 50 mL Becton- Dickinson and Company-syringes, then stored anaerobically in sterile glass bottles at 4 °C. Field measurements of fluid parameters were performed using a HI9829 (Hanna instruments) multiparameter calibrated and equipped with sensors allowing the acquisition of pH, temperature (°C), alkalinity (mg/L), and electrical conductivity measurements (mS/cm) (Table S1). The major anions and cations were analyzed at LGL-TPE using ion-chromatography (Methrom ECO IC). A mixture of 3.2 mM Na<sub>2</sub>CO<sub>3</sub> and 1 mM NaHCO<sub>3</sub> was used as an eluent for analysis of anions and a chemical suppression module (MSM) was used to suppress the conductivity. For cations, 1.7 mM HNO<sub>3</sub> was used as an eluent. The anions and cations were separated using analytical columns, Metrosep A Supp5 Guard/4.0 and Metrosep C4 250/4.0, respectively (Table S1).

**DNA extraction and sequencing.** Hot spring's samples analyzed here were originally collected to grow thermophilic taxa. They were stored at 4 °C for 2 years before DNA was extracted. For each hot spring sample, three replicates of DNA extraction were conducted individually, and combined as a composition sample, before the sequencing. DNA was extracted with a standard PCI (Phenol:Chloroform:Isoamyl Alcohol (25:24:1)) protocol, as described elsewhere, from 10 g environmental matrix<sup>51</sup>, with the exception that 50 µM linear acrylamide were added to enhance nucleic acids precipitation (Invitrogen). One negative control was included and contained 10 mL of DNA-free sterile water. Elution of total DNA extracts was performed in 30–50 µL EB buffer (10 mM Tris–Cl, pH 8.5). Nucleic acid solution quality was determined using the NanoDrop 8000 (Thermo Scientific, Waltham, MA, USA) spectrophotometer. Double-strand DNA concentration was measured using the kit Quantifluor dsDNA system. DNA samples were sequenced by NovaSeq 6000 (2 × 150 bp, paired-end reads) technology by the Duke Center for Genomic and Computational Biology (GCB) (<https://genome.duke.edu/>).

**Sequence processing, metagenomic assembly and binning.** Metagenome sequences' quality were controlled by FastQC (v0.11.9—<https://github.com/s-andrews/FastQC>) and MultiQC (v1.9—<https://github.com/ewels/MultiQC>). Sequences were then processed with the snakemake of Anvi'o<sup>52–54</sup> (v7—<https://github.com/merenlab/anvio>), filtered by integrated minoch script (v2.8—<https://github.com/merenlab/illumina-utils>). Next, MetaSpades<sup>55</sup> (v3.14.1—<https://github.com/ablab/spades>) was used as genome assembler and Concoct<sup>56</sup> (v1.1.0—<https://github.com/BinPro/CONCOCT>) as genome binner with `anvi_cluster_contigs` function (“all against all” mode). Furthermore, MAGs were manually refined with `anvi-refine` function. Genome mapping was performed with bowtie2<sup>57</sup> (v2.4.2—<https://sourceforge.net/projects/bowtie-bio/files/bowtie2/2.4.2/>) and samtools<sup>58</sup> (v1.7—<https://samtools.github.io/>). MAGs' quality was estimated by Anvi'o and furthermore by CheckM<sup>59</sup> (v1.1.3—<https://ecogenomics.github.io/CheckM/>), both with default parameters. Total length, number of contigs, N50, and GC% contents were extracted with `anvi-summarize` function.

**Taxonomic and phylogenetic inference of metagenomic assemblies and MAGs.** According to the standards proposed elsewhere<sup>26</sup>, bins were defined as high-quality (HQ) MAGs (>90% completion, <5% contamination, presence of the 23S, 16S and 5S rRNA genes and at least 18 tRNAs), nearly high-quality (NHQ) MAGs (>90% completion, <5% contamination, other criteria partially covered), medium-quality (MQ) MAGs (≥50% completion, <10% contamination) and low-quality (LQ) MAGs (<50% completion, <10% contamination). MAGs often lack 16S rRNA genes due to their conserved and repetitive nature preventing their assembly<sup>60</sup>, so MAGs of near high-quality could be classified as high-quality MAGs by other authors taking this into consideration. The taxonomic affiliation of the MAGs was first investigated by placing the MAGs in a phylogenomic context. The phylogenetic reconstructions were based on 122 archaeal or 120 bacterial single copy conservative marker genes according to the Genome Taxonomy Database (07-RS207 release) and were constructed using de novo workflow implemented in GTDB-Tk (v1.4.1—<https://github.com/Ecogenomics/GTDBTk>)<sup>61,62</sup>. Visualization and trees analyses were done using ARB software<sup>63</sup>.

As the taxonomy proposed by GTDB is new and does not correspond exactly to the one recognized by the International Code of Nomenclature of prokaryotes (ICNP), we also analyzed data according to the rules of the Code and its nomenclature, because at the time of writing this article, SeqCode is being implemented<sup>64</sup>. For this purpose, we implemented a combination of genomic indices classically used for the delineation of the different taxonomic ranks, namely: 16S rRNA gene sequence similarity, average nucleotide identity score (ANI) and average amino-acid identity value (AAI). The approach followed and the results are given in supplementary material (Text S1, Table S2, Fig. S2).

**Metabolic profiling.** MAGs were processed with KEGG Decoder script (<https://github.com/bjtully/BioData/tree/master/KEGGDecoder>) from Anvi'o gene calls tables generated with `kegg_kofams` and then plotted with R packages (ComplexHeatmap, circlize, RColorBrewer, and dplyr) to get a general annotation with R (v.3.6.3)<sup>65,66</sup>. A more accurate annotation was performed with Prokka (v1.14.6—<https://github.com/tseemann/prokka>)<sup>67</sup> and associated outputs were analyzed by using the Pathway Tools software (v.24.5)<sup>68</sup> with the MetaCyc database (v.24.5)<sup>69</sup> to explore in details the putative metabolisms encoded in MAGs or public available genome assemblies. Regarding sulfur metabolisms, for `dsr` genes, the perl script DiSCo (v.1.0.0, <https://github.com/Genome-Evolution-and-Ecology-Group-GEEG/DiSCo>) was used on the Prokka protein sequences outputs of each MAG to highlight the specific genes<sup>70</sup>.

## Data availability

The metagenome bins generated and analyzed during the current study are available in the European Nucleotide Archive (ENA) (<https://www.ebi.ac.uk/ena/browser/home>), under the Project PRJEB46766 (Table 1).

Received: 5 April 2022; Accepted: 13 December 2022

Published online: 23 December 2022

## References

- Mehta, D. & Satyanarayana, T. Diversity of Hot Environments and Thermophilic Microbes. In *Thermophilic Microbes in Environmental and Industrial Biotechnology* 3–60 (Springer, 2013). [https://doi.org/10.1007/978-94-007-5899-5\\_1](https://doi.org/10.1007/978-94-007-5899-5_1).
- Des Marais, D. J. & Walter, M. R. Terrestrial hot spring systems: Introduction. *Astrobiology* **19**(12), 1419–1432. <https://doi.org/10.1089/ast.2018.1976> (2019).
- Belilla, J. *et al.* Hyperdiverse archaea near life limits at the polyextreme geothermal Dallol area. *Nat. Ecol. Evol.* **3**(11), 1552–1561. <https://doi.org/10.1038/s41559-019-1005-0> (2019).

4. Urbiet, M. *et al.* Copahue geothermal system: A volcanic environment with rich extreme prokaryotic biodiversity. *Microorganisms* **3**(3), 344–363. <https://doi.org/10.3390/microorganisms3030344> (2015).
5. Power, J. F. *et al.* Microbial biogeography of 925 geothermal springs in New Zealand. *Nat. Commun.* **9**, 1. <https://doi.org/10.1038/s41467-018-05020-y> (2018).
6. Lezcano, M. Á. *et al.* Biomarker profiling of microbial mats in the geothermal band of Cerro Caliente, Deception Island (Antarctica): Life at the edge of heat and cold. *Astrobiology* **19**(12), 1490–1504. <https://doi.org/10.1089/ast.2018.2004> (2019).
7. Brock, T. D., Brock, K. M., Belly, R. T. & Weiss, R. L. *Sulfolobus*: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch. Mikrobiol.* **84**(1), 54–68. <https://doi.org/10.1007/bf00408082> (1972).
8. Meyer-Dombard, D. R., Shock, E. L. & Amend, J. P. Archaeal and bacterial communities in geochemically diverse hot springs of Yellowstone National Park, USA. *Geobiology* **3**(3), 211–227. <https://doi.org/10.1111/j.1472-4669.2005.00052.x> (2005).
9. Kublanov, I. V. *et al.* Biodiversity of thermophilic prokaryotes with hydrolytic activities in hot springs of Uzon Caldera, Kamchatka (Russia). *Appl. Environ. Microbiol.* **75**(1), 286–291. <https://doi.org/10.1128/aem.00607-08> (2009).
10. Reigstad, L. J., Jørgensen, S. L. & Schleper, C. Diversity and abundance of Korarchaeota in terrestrial hot springs of Iceland and Kamchatka. *ISME J.* **4**(3), 346–356. <https://doi.org/10.1038/ismej.2009.126> (2010).
11. Wemheuer, B., Taube, R., Akyol, P., Wemheuer, F. & Daniel, R. Microbial diversity and biochemical potential encoded by thermal spring metagenomes derived from the Kamchatka Peninsula. *Archaea* **2013**, 1–13. <https://doi.org/10.1155/2013/136714> (2013).
12. Wilkins, L. G. E., Ettinger, C. L., Jospin, G. & Eisen, J. A. Metagenome-assembled genomes provide new insight into the microbial diversity of two thermal pools in Kamchatka, Russia. *Sci. Rep.* **9**, 1. <https://doi.org/10.1038/s41598-019-39576-6> (2019).
13. Deamer, D., Damer, B. & Kompanichenko, V. Hydrothermal chemistry and the origin of cellular life. *Astrobiology* **19**(12), 1523–1537. <https://doi.org/10.1089/ast.2018.1979> (2019).
14. Van Kranendonk, M. J., Deamer, D. W. & Djokic, T. Life springs. *Sci. Am.* **317**(2), 28–35. <https://doi.org/10.1038/scientificamerican0817-28> (2017).
15. Muñoz, P. A., Flores, P. A., Boehmwald, F. A. & Blamey, J. M. Thermophilic bacteria present in a sample from Fumarole Bay, Deception Island. *Antarct. Sci.* **23**(6), 549–555. <https://doi.org/10.1017/s0954102011000393> (2011).
16. Amenábar, M. J., Flores, P. A., Pugin, B., Boehmwald, F. A. & Blamey, J. M. Archaeal diversity from hydrothermal systems of Deception Island, Antarctica. *Polar Biol.* **36**(3), 373–380. <https://doi.org/10.1007/s00300-012-1267-3> (2013).
17. Bendia, A. G. *et al.* A mosaic of geothermal and marine features shapes microbial community structure on deception Island Volcano, Antarctica. *Front. Microbiol.* **9**, 20. <https://doi.org/10.3389/fmicb.2018.00899> (2018).
18. ATCM (Antarctic Treaty Consultative Meeting) (2012). Measure 8. Management plan for Antarctic Specially Protected Area No. 140. Parts of Deception Island, South Shetland Islands. Final report of the thirty-fifth ATCM, Hobart, Australia, June 11–20. [www.ats.aq](http://www.ats.aq).
19. Giret, A., Gregoire, M., Cottin, J. Y. & Michon, G. In *Kerguelen, a Third Type Of Oceanic Island? "The Antarctic Region: Geological Evolution and Processes"* (ed. Ricci, C. A.) 735–741 (Terra Antarctica Publication, 1997).
20. Gagnevin, D. *et al.* Open-system processes in the genesis of silica-oversaturated alkaline rocks of the Rallier-du-Baty Peninsula, Kerguelen Archipelago (Indian Ocean). *J. Volcanol. Geoth. Res.* **123**(3–4), 267–300. [https://doi.org/10.1016/s0377-0273\(02\)00509-7](https://doi.org/10.1016/s0377-0273(02)00509-7) (2003).
21. Parikka, K. J., Jacquet, S., Colombet, J., Guillaume, D. & Le Romancer, M. Abundance and observations of thermophilic microbial and viral communities in submarine and terrestrial hot fluid systems of the French Southern and Antarctic Lands. *Polar Biol.* **41**(7), 1335–1352. <https://doi.org/10.1007/s00300-018-2288-3> (2018).
22. Renac, C., Moine, B., Goudour, J. P., Le Romancer, M. & Perrache, C. Stable isotope study of rainfall, river drainage and hot springs of the Kerguelen archipelago, SW Indian Ocean. *Geothermics* **83**, 101726. <https://doi.org/10.1016/j.geothermics.2019.101726> (2020).
23. Postec, A., Ciobanu, M., Birrien, J. L., Prieur, D. & Le Romancer, M. (2009). Microbial biodiversity of thermophilic communities in remote geothermal springs of Saint-Paul and Kerguelen Islands. In *Proceedings of "Thermophiles 2009" Conference, Beijing, China*, August, 16–21.
24. Gramain, A., Brillet, F., Birrien, J. L., & Le Romancer, M. (2011). Novel hyperthermophilic archaea flying over the Kerguelen islands: A dissemination study. In *"Thermophiles 2011" Conference, Big Sky, Montana USA*, September 11–16.
25. Cozannet, M. *et al.* New insights into the ecology and physiology of *Methanomassiliicoccales* from terrestrial and aquatic environments. *Microorganisms* **9**(1), 30. <https://doi.org/10.3390/microorganisms9010030> (2021).
26. Bowers, R. M. *et al.* Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* **35**(8), 725–731. <https://doi.org/10.1038/nbt.3893> (2017).
27. Nakagawa, S. *et al.* *Hydrogenivirga caldilitoris* gen. nov., sp. nov., a novel extremely thermophilic, hydrogen- and sulfur-oxidizing bacterium from a coastal hydrothermal field. *Int. J. Syst. Evol. Microbiol.* **54**(6), 2079–2084. <https://doi.org/10.1099/ijs.0.03031-0> (2004).
28. Seitz, K. W. *et al.* Asgard archaea capable of anaerobic hydrocarbon cycling. *Nat. Commun.* **10**, 1. <https://doi.org/10.1038/s41467-019-09364-x> (2019).
29. Zhou, Z. *et al.* Genome- and community-level Interaction insights into carbon utilization and element cycling functions of *Hydrothermarchaeota* in hydrothermal sediment. *Systems* **5**, 1. <https://doi.org/10.1128/msystems.00795-19> (2020).
30. Kochetkova, T. V. *et al.* Hot in cold: Microbial life in the hottest springs in permafrost. *Microorganisms* **8**(9), 1308. <https://doi.org/10.3390/microorganisms8091308> (2020).
31. Dodsworth, J. A. *et al.* *Thermoflexus hugenholtzii* gen. nov., sp. nov., a thermophilic, microaerophilic, filamentous bacterium representing a novel class in the *Chloroflexi*, *Thermoflexia* classis nov., and description of *Thermoflexaceae* fam. nov. and *Thermoflexales* ord. nov. *Int. J. Syst. Evol. Microbiol.* **64**(Pt. 6), 2119–2127. <https://doi.org/10.1099/ijs.0.055855-0> (2014).
32. Löffler, F. E. *et al.* *Dehalococcoides mccartyi* gen. nov., sp. nov., obligately organohalide-respiring anaerobic bacteria relevant to halogen cycling and bioremediation, belong to a novel bacterial class, *Dehalococcoidia* classis nov., order *Dehalococcoidales* ord. nov. and family *Dehalococcoidaceae* fam. nov., within the phylum *Chloroflexi*. *Int. J. Syst. Evol. Microbiol.* **63**(2), 625–635. <https://doi.org/10.1099/ijs.0.034926-0> (2013).
33. King, C. E. & King, G. M. *Thermomicrobium carboxidum* sp. nov., and *Thermorudis peleae* gen. nov., sp. nov., carbon monoxide-oxidizing bacteria isolated from geothermally heated biofilms. *Int. J. Syst. Evol. Microbiol.* **64**(8), 2586–2592. <https://doi.org/10.1099/ijs.0.060327-0> (2014).
34. Wu, D. *et al.* Complete genome sequence of the aerobic CO-oxidizing thermophile *Thermomicrobium roseum*. *PLoS One* **4**(1), e4207. <https://doi.org/10.1371/journal.pone.0004207> (2009).
35. Hanada, S. The Phylum *Chloroflexi*, the family *Chloroflexaceae*, and the related phototrophic families *Oscillochloridaceae* and *Roseiflexaceae*. In *The Prokaryotes* 515–532 (Springer, 2014). [https://doi.org/10.1007/978-3-642-38954-2\\_165](https://doi.org/10.1007/978-3-642-38954-2_165).
36. Yabe, S. *et al.* *Dictyobacter aurantiacus* gen. nov., sp. nov., a member of the family *Ktedonobacteraceae*, isolated from soil, and emended description of the genus *Thermosporothrix*. *Int. J. Syst. Evol. Microbiol.* **67**(8), 2615–2621. <https://doi.org/10.1099/ijs.0.001985> (2017).
37. Gupta, R. S. The Phylum *Aquificae*. In *The Prokaryotes* 417–445 (Springer, 2014). [https://doi.org/10.1007/978-3-642-38954-2\\_119](https://doi.org/10.1007/978-3-642-38954-2_119).
38. Ming, H. *et al.* *Thermus thermamylovorans* sp. nov., isolated from a hot spring. *Int. J. Syst. Evol. Microbiol.* **70**(3), 1729–1737. <https://doi.org/10.1099/ijs.0.003965> (2020).

39. Albuquerque, L. & da Costa, M. S. The Family *Thermaceae*. In *The Prokaryotes* 955–987 (Springer, 2014). [https://doi.org/10.1007/978-3-642-38954-2\\_128](https://doi.org/10.1007/978-3-642-38954-2_128).
40. Issotta, F. *et al.* Insights into the biology of acidophilic members of the *Acidiferrobacteraceae* family derived from comparative genomic analyses. *Res. Microbiol.* **169**(10), 608–617. <https://doi.org/10.1016/j.resmic.2018.08.001> (2018).
41. Lee, K. C. Y., Dunfield, P. F. & Stott, M. B. The Phylum *Armatimonadetes*. In *The Prokaryotes* 447–458 (Springer, 2014). [https://doi.org/10.1007/978-3-642-38954-2\\_388](https://doi.org/10.1007/978-3-642-38954-2_388).
42. Niederberger, T. D., Götz, D. K., McDonald, I. R., Ronimus, R. S. & Morgan, H. W. *Ignisphaera aggregans* gen. nov., sp. nov., a novel hyperthermophilic crenarchaeote isolated from hot springs in Rotorua and Tokaanu, New Zealand. *Int. J. Syst. Evol. Microbiol.* **56**(5), 965–971. <https://doi.org/10.1099/ijs.0.63899-0> (2006).
43. St John, E. *et al.* A new symbiotic nanoarchaeote (*Candidatus* Nanoclepta minutus) and its host (*Zestosphaera tikerensis* gen. nov., sp. nov.) from a New Zealand hot spring. *Syst. Appl. Microbiol.* **42**(1), 94–106. <https://doi.org/10.1016/j.syapm.2018.08.005> (2019).
44. Itoh, T. The Family *Thermoproteaceae*. In *The Prokaryotes* 389–401 (Springer, 2014). [https://doi.org/10.1007/978-3-642-38954-2\\_330](https://doi.org/10.1007/978-3-642-38954-2_330).
45. Belkova, N. L., Tazaki, K., Zakharova, J. R. & Parfenova, V. V. Activity of bacteria in water of hot springs from Southern and Central Kamchatskaya geothermal provinces, Kamchatka Peninsula, Russia. *Microbiol. Res.* **162**(2), 99–107. <https://doi.org/10.1016/j.micres.2006.01.006> (2007).
46. Sahn, K. *et al.* High abundance of heterotrophic prokaryotes in hydrothermal springs of the Azores as revealed by a network of 16S rRNA gene-based methods. *Extremophiles* **17**(4), 649–662. <https://doi.org/10.1007/s00792-013-0548-2> (2013).
47. Slobodkin, A. I. & Slobodkina, G. B. Diversity of sulfur-disproportionating microorganisms. *Microbiology* **88**(5), 509–522. <https://doi.org/10.1134/s0026261719050138> (2019).
48. Allieux, M. *et al.* Genomic characterization and environmental distribution of a thermophilic anaerobe *Dissulfurirhabdus thermomarina* SH388<sup>T</sup> involved in disproportionation of sulfur compounds in shallow sea hydrothermal vents. *Microorganisms* **8**(8), 1132. <https://doi.org/10.3390/microorganisms8081132> (2020).
49. Roy, M., Gupta, S., Patranabis, S. & Ghosh, A. The oligomeric plasticity of Hsp20 of *Sulfolobus acidocaldarius* protects environment-induced protein aggregation and membrane destabilization. *Biochim. Biophys. Acta Biomembr.* **1860**(12), 2549–2565. <https://doi.org/10.1016/j.bbame.2018.09.005> (2018).
50. Lulchev, P. & Klostermeier, D. Reverse gyrase—recent advances and current mechanistic understanding of positive DNA supercoiling. *Nucleic Acids Res.* **42**(13), 8200–8213. <https://doi.org/10.1093/nar/gku589> (2014).
51. Charbonnier, F., Forterre, P., Erauso, G. & Prieur, D. Purification of plasmids from thermophilic and hyperthermophilic archaeobacteria. In *Archaea: A Laboratory Manual* (eds Robb, F. T. *et al.*) 87–90 (Cold Spring Harbor Laboratory Press, 1995).
52. Koster, J. & Rahmann, S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* **28**(19), 2520–2522. <https://doi.org/10.1093/bioinformatics/bts480> (2012).
53. Shaiber, A. *et al.* Functional and genetic markers of niche partitioning among enigmatic members of the human oral microbiome. *Genome Biol.* **21**, 1. <https://doi.org/10.1186/s13059-020-02195-w> (2020).
54. Eren, A. M. *et al.* Community-led, integrated, reproducible multi-omics with anvio. *Nat. Microbiol.* **6**(1), 3–6. <https://doi.org/10.1038/s41564-020-00834-3> (2021).
55. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: A new versatile metagenomic assembler. *Genome Res.* **27**(5), 824–834. <https://doi.org/10.1101/gr.213959.116> (2017).
56. Alneberg, J. *et al.* Binning metagenomic contigs by coverage and composition. *Nat. Methods* **11**(11), 1144–1146. <https://doi.org/10.1038/nmeth.3103> (2014).
57. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**(4), 357–359. <https://doi.org/10.1038/nmeth.1923> (2012).
58. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352> (2009).
59. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**(7), 1043–1055. <https://doi.org/10.1101/gr.186072.114> (2015).
60. Parks, D. H. *et al.* Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* **2**, 1533–1542. <https://doi.org/10.1038/s41564-017-0012-7> (2017).
61. Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: A toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* <https://doi.org/10.1093/bioinformatics/btz848> (2019).
62. Parks, D. H. *et al.* A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat. Biotechnol.* **38**(9), 1079–1086. <https://doi.org/10.1038/s41587-020-0501-8> (2020).
63. Ludwig, W. *et al.* ARB: A software environment for sequence data. *Nucleic Acids Res.* **32**(4), 1363–1371. <https://doi.org/10.1093/nar/gkh293> (2004).
64. Hedlund, B. P. *et al.* SeqCode: A nomenclatural code for prokaryotes described from sequence data. *Nat. Microbiol.* **7**, 1702–1708. <https://doi.org/10.1038/s41564-022-01214-9> (2022).
65. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize implements and enhances circular visualization in R. *Bioinformatics* **30**(19), 2811–2812. <https://doi.org/10.1093/bioinformatics/btu393> (2014).
66. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**(18), 2847–2849. <https://doi.org/10.1093/bioinformatics/btw313> (2016).
67. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**(14), 2068–2069. <https://doi.org/10.1093/bioinformatics/btu153> (2014).
68. Karp, P. D. *et al.* Pathway Tools version 23.0 update: Software for pathway/genome informatics and systems biology. *Brief. Bioinform.* **22**(1), 109–126. <https://doi.org/10.1093/bib/bbz104> (2021).
69. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes—a 2019 update. *Nucleic Acids Res.* **48**(D1), D445–D453. <https://doi.org/10.1093/nar/gkz862> (2020).
70. Neukirchen, S. & Sousa, F. L. DiSCo: A sequence-based type-specific predictor of Dsr-dependent dissimilatory sulphur metabolism in microbial data. *Microbial Genom.* **7**, 7. <https://doi.org/10.1099/mgen.0.000603> (2021).

## Acknowledgements

We would like to thank Guillaume Lannuzel for providing assistance with molecular biology analyses. We are grateful to the Territoire des Terres Australes et Antarctiques Françaises (TAAF), the IPEV (Institut Polaire Paul-Emile Victor) and its logistical staff for logistic support for transit and fieldwork in the Kerguelen archipelago. We thank M. Jebbar and Z. Shao, co-supervisors of M.A., for letting him do this study which was not in the scope of his thesis. We acknowledge the LABGeM (CEA/Genoscope & CNRS UMR8030), the France Génomique and French Bioinformatics Institute national infrastructures (funded as part of Investissements d’Avenir program managed by Agence Nationale pour la Recherche, contracts ANR-10-INBS-09 and ANR-11-INBS-0013) are



acknowledged for support within the MicroScope annotation platform. We are grateful for the financial support of IPEV provided by the TALISKER program N°1077 to D.G., the HOTVIR program N°408 to M.L.R. and the SEKMET program N°1231 to K.A. and M.L.R. Metagenomic analyses were funded by the Sino-French IRP 1211 MicrobSea to K.A., by the MERLIN “Abyss” program funded by Ifremer to K.A. and by the Scientific Council of IUEM to K.A. The study was supported by two grants from the French Ministry of Higher Education and Research, and from the Région Bretagne, to M.A. and S.Y.

### Author contributions

Conceptualization, M.A. and K.A.; formal analysis, M.A., S.Y., A.M., V.L., D.G., J.A., J.P. and K.A.; funding acquisition, D.G., M.L.R. and K.A.; investigation, M.A., S.Y., A.M., M.C., D.G., J.A., and K.A.; supervision, K.A.; writing—original draft, M.A., S.Y. and K.A.; review and editing, all coauthors. All authors have read and agreed to the published version of the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-26299-4>.

**Correspondence** and requests for materials should be addressed to K.A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022