Phenylobacterium ferrooxidans sp. nov., isolated from a sub-surface geothermal aquifer in Iceland

Eva Pouder, Erwann Vince, Karen Jacquot, Maimouna batoma Traoré, Ashley Grosche, Maria Ludwig, Mohamed Jebbar, Lois Maignien, Karine Alain and Sophie Mieszkin*

Author affiliations:

Univ Brest, CNRS, IFREMER, EMR 6002 BIOMEX, Unité Biologie et Écologie des Écosystèmes marins Profonds BEEP, F-29280 Plouzané, France

*** Corresponding author.**

E-mail address: Sophie.Mieszkin@univ-brest.fr (S. Mieszkin)

Abstract

A novel bacterial strain, HK31-G^T, was isolated from a subsurface geothermal aquifer (Hellisheidi, SW-Iceland) and was characterized using a polyphasic taxonomic approach. Phylogenetic analysis of 16S rRNA gene along with phylogenomic position indicated that the novel strain belongs to the genus *Phenylobacterium*. Cells are motile Gram-negative bacilli. Physiological characterization showed that strain HK31-G^T is a mesophilic bacterium able to grow from 10 to 30 °C, at pH values between 2 and 12 and at NaCl concentrations between 0 and 0.5%. Optimal growth was observed without sodium chloride at 25 °C and pH 6. Strain HK31-G^T is chemoorganoheterotroph and its major saturated fatty acids are $C_{18:1}\omega7c$, $C_{16:1}\omega6c$ and $C_{16:0}$, the predominant quinone is Q-10 and the major polar lipid is phosphatidylglycerol. The new strain also possesses the capacity to use ferrous iron (Fe(II)) as the sole energy source and can also be considered as a chemolithoautotrophic microorganism. The overall genome of strain HK31-G^T was estimated to be 4.46 Mbp in size with a DNA G+C content of 67.95%. Genes involved in iron metabolism were identified, but no genes typically involved in Fe(II) oxidation were found. Average nucleotide identity and digital DNA-DNA hybridization values between the genome of strain HK31-G^T and the genomes of its closest relatives are below the species delineation threshold. Therefore, given the polyphasic approach used, strain HK31-G^T represents a novel species of the genus *Phenylobacterium*, for which the name *Phenylobacterium ferrooxidans* sp. nov. is proposed. The type strain is HK31-G^T (DSM 116432^T = UBOCC-M-3429^T= LMG 33376^T).

Keywords: subsurface geothermal aquifer, *Phenylobacterium* sp., bacterial species, physiology, iron oxidation, nitrate reduction

Introduction

Carbon Capture Storage (CCS) technologies offer a way to sequester anthropogenic CO_2 emissions from the atmosphere in the subsurface, lowering environmental consequences such as greenhouse-gas emissions (Snæbjörnsdóttir et al., 2020). Through injections of CO_2 dissolved in water into deep basalts or peridotites rich in calcareous and ferromagnesian silicates, CO_2 will react and precipitate as solid Ca-Mg-Fe-carbonate minerals over a two-year timespan (Snæbjörnsdóttir et al., 2020). The carbonates formed will be stable for thousands of years. This deep mineral carbon storage in basalt was developed in the framework of the Carbfix project (https://www.carbfix.com) at the CarbFix-1 site near the Hellisheidi geothermal power plant (SW-Iceland) (Matter et al., 2011). Following the injection and dissolution of CO_2 into the surface, protons are released, resulting in acidic and oxidative conditions along with an increase of inorganic carbon source promoting growth of chemolithoautotrophic microbial communities (Mu et al. 2015; O'Mullan et al., 2015). At the CarbFix-1 site, it was shown that after CO_2 was injected into basalt, microbial richness decreased but lithoautotrophic Fe(II)-oxidizing *Betaproteobacteria* and aromatic compound degraders became dominant (Trias et al., 2017).

During a study of bacterial diversity associated with CCS gas injection at the CarbFix-1 site, but on different wells from the study by Trias et al. (2017) (Fig. S1), an attempt was also made to enrich and isolate neutrophilic Fe(II)-oxidizing bacteria (FeOB). The bacterial strain HK31-G^T was finally isolated and appears to belong to the genus *Phenylobacterium*. At the time of writing, this genus encompasses 17 validly published species, according to the List of Prokaryotic names with Standing in Nomenclature (LPSN), from a variety of environments, but mainly soil and water (Thomas et al., 2022). The genus *Phenylobacterium* belongs to the class *Alphaproteobacteria* and the family *Caulobacteraceae,* and the type species is *P. immobile* (Lingens et al., 1985). To date,

species of the genus *Phenylobacterium* are characterized as aerobic or facultatively anaerobic, Gram-negative, motile or non-motile, straight to slightly curved rods, coccobacilli or cocci, occurring either singly, in pairs or in short chains. The major respiratory quinone is Q-10 while $C_{18:1}\omega7c$ and/or $C_{18:1}\omega6c$ are the major fatty acids, and phosphatidylglycerol the main polar lipid. DNA G+C contents range from 64 to 72.3 % (Thomas et al., 2022).

In the present study, we performed a polyphasic taxonomic characterization of strain HK31-G^T and provided phenotypic, chemotaxonomic, phylogenomic and genomic evidence that it meets criteria for the delineation of a new species of the genus *Phenylobacterium*. Interestingly, we also reported the capacity of this strain to oxidize ferrous iron (Fe(II)) under microaerophilic as well as anaerobic and neutrophilic conditions highlighting the versatility of its metabolism. This is in line with the changing environmental conditions that could be encountered during CCS gas injection.

Materials and Methods

Isolation, ecology and deposit in public culture collections

Strain HK31-G^T was isolated from a deep subsurface hydrothermal aquifer (well HK-31) at the Carbfix-1 injection site of an adjacent geothermal power plant at Hellisheidi (SW-Iceland), which is the third largest geothermal power plant in the world (64° 02’ 14’’ N, 21° 24’ 03’’ W) (Matter et al., 2011). The Carbfix-1 was used for pilot scale injections of CO_2 in 2012 where 175 tons of pure commercial CO_2 was injected, while industrial scale injection is today carried out at the Carbfix-2 injection site (Aradóttir et al., 2015). At the Carbfix-1 site, the CO_2 injection well (HN-2, located at a depth of 2000 m and into which CO_2 was injected at a depth between 400 and 800 m) is located approximately at 1,750 m from the well HK-31, from which the new strain HK31-G^T is originated (Fig. S1) (Aradóttir et al., 2011). Water, for CO_2 injection, was pumped from the upstream well

HN-1, then re-injected with CO_2 into HN-2. Overall, the Carbfix-1 storage formation is composed of fresh basaltic lavas interbedded with hyaloclastites. The ground waters are located between 400 and 800 m depth and are characterized by poor dioxygen concentrations and by temperatures and pH ranging respectively from 18 to 33 °C and from 8.4 to 9.4 (Alfredsson et al., 2008; Aradóttir et al., 2011). Prior to sampling in July 2019, the well was pumped continuously for 24 hours and water samples were collected using sterile 50 mL Falcon tubes and then stored at 4 °C. Before storage, the pH of the aqueous samples was measured and was established at 8.78. The temperature was recorded at 24.1 °C and the oxido-reduction potential (EH) measured was 268 uS/sec.

In order to enrich and isolate neutrophilic Fe(II)-oxidizing bacteria (FeOB), semi-solid Fe(II)/O_2 gradient tubes under microaerophilic conditions were aseptically inoculated in the laboratory immediately after sample collection, as described elsewhere (Emerson and Merrill Floyd, 2005). After enrichment culture at 20 °C, a serial dilution-to-extinction method in semi-solid Fe(II)/O_2 gradient tubes was applied for isolation at the same temperature. A co-culture was finally obtained and the two strains were then isolated by plating this co-culture on Reasoner's 2A (R2A) agar medium at 20 °C and pH 7.2 (Khan et al., 2018; Reasoner et al., 1985), which enabled distinct colonies to be isolated. The strain was then purified by three successive quadrant streaks on R2A agar medium at 20 °C and pH 7.2 to obtain a pure culture. Strain HK31-G^T was then grown routinely on R2A broth or agar adjusted to pH 7.2 over respectively, 3 or 5 days at 20 °C, under agitation (250 r.p.m). Its purity was routinely confirmed by microscopic observations (Olympus BX60 and CX40) and by sequencing its 16S rRNA gene. Stock cultures were stored at −80 °C in R2A broth medium supplemented with 5% (v/v) dimethylsulfoxide (DMSO). *Phenylobacterium ferrooxidans* type strain HK31-G^T (DSM 116432^T ; UBOCC-M-3429^T ; LMG 33376^T) is available in three public culture collections, namely Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ;

https://www.dsmz.de/collection), UBO Culture Collection (UBOCC; https://www.univ-brest.fr/ubocc) and Belgian Coordinated Collections of Microorganisms (BCCM; https://bccm.belspo.be/).

Morphological, physiological and metabolic features

Colony morphology of strain HK31-G^T was observed on R2A agar adjusted to pH 6 for 5 days at 25 °C. Cell morphology and motility were determined by light microscopy (Olympus BX60 and CX40) and transmission electron microscopy (TEM, Jeol JEM 1400). Motility was also confirmed by using mannitol–motility–nitrate (MMN) agar medium (composed of 10 g·L^{-1} tryptic hydrolysate of casein, 1 g potassium nitrate, 7.5 g·L^{-1} mannitol, 40 mg·L^{-1} phenol red and 3.5 g·L^{-1} agar), which was also used to evidence mannitol fermentation and nitrate reductase activity. Gram-staining was determined using standard procedures and confirmed with a KOH (3%) test. Catalase and cytochrome oxidase activities were respectively evaluated using H_2O_2 and strips of *N*,*N*,*N*′,*N*′-tetramethyl-*p*-phenylenediamine dihydrochloride (Bio-Rad).

Physiological characterization of the novel strain HK31-G^T was carried out aerobically, in triplicates, on R2A (agar or broth) adjusted to pH 6 at 25 °C and under agitation (250 r.p.m.). Determination of the temperature range for growth and salt tolerance were respectively tested over the range 5–55 °C, at 5 °C intervals and 0–5% NaCl (w/v), at 0.5% intervals, both for 12 days on R2A agar at pH 6. The pH range for growth was tested from pH 3 to 12 (at 25 °C), with increments of 1 unit in R2A broth for 12 days. Cells were routinely enumerated by direct cell counting using a modified Thoma chamber (Preciss; surface: 0.0025 mm^2, depth: 10 µm). The following buffers (each at 20 mM, Sigma-Aldrich) were used to adjust the required pH: pH 4 and 5 with HOMOPIPES buffer, pH 6 with MES buffer, pH 7 with PIPES buffer, pH 8 with HEPES

buffer, and pH 9 and 10 with CAPSO buffer. For pH 3, and greater or equal to 11, no buffer was used. The growth kinetics of HK31-G^T under optimal conditions was then studied in triplicate, on R2A broth adjusted to pH 6 at 25 °C with no NaCl salt, under agitation (250 r.p.m.) for 13 days. Cell growth was monitored by direct cell counting, using a modified Thoma chamber, in order to determine the growth rate and doubling time of the strain under optimal culture conditions.

Metabolic features were estimated for the novel isolate HK31-G^T and closely related strains previously described and available in public culture collections: *P. haematophilum* LMG 11050^T (=DSM 21793^T=CCUG 26751^T), *P. conjunctum* FWC 21^T (LMG 24262^T) (Abraham et al., 2008), and *P. aquaticum* W2-3-4^T (KACC 18306^T) (Jo et al., 2016). Utilization of organic substrates as the sole carbon source was investigated using the mineral basis of R2A medium containing 0.3 g·L^{-1} dipotassium phosphate and 0.05 g·L^{-1} magnesium sulfate and adjusted to pH 6. Each substrate (xylose, glycerol, malonate, hydroxybutyrate, lactate, succinate, aspartate, acetate, glutamate, propionate, proline, alanine and phenylalanine) was supplied at a final concentration of 20 mM. Each strain was cultivated at its own optimal growth parameters (T°, pH and NaCl). Before inoculation, cells, at the end of their exponential growth phase, were harvested and washed three times with distilled water. The bacterial suspension was then adjusted to obtain a McFarland index value of 1. A medium consisting of R2A mineral base adjusted to pH 6, with no carbon source, was used as negative control for each carbon utilization bioassay and a positive control was performed using R2A broth adjusted to pH 6. Additional tests were carried out to determine: (i) mannitol fermentation and presence of nitrate reductase (MMN agar medium); (ii) fermentative pathways (mixed acids or butane-2,3-diol pathways; Clark and Lubs liquid medium); (iii) glucose and lactose fermentation, gas and H_2S production (Kligler–Hajna agar

medium); and (iv) citrate utilization (Simmons citrate agar medium). For these assays, cell washes and preparation of bacterial suspensions from cultures of HK31-G^T under optimal conditions on R2A medium were performed as described above and elsewhere (Mieszkin et al., 2021). Oxidative and fermentative utilization of carbohydrates as well as enzymatic activities for the strain HK31-G^T and closely related strains were also evaluated by using the API®20NE kit, API®20E kit and API®ZYM kit (BioMérieux) according to the manufacturer's instructions with slight modifications (McFarland index = 1; 3 days of incubation at 20°C).

Chemotaxonomic analyses

Characterization of respiratory quinones, polar lipids and fatty acids of cells of strain HK31-G^T was carried out by the identification Service of the DSMZ (Braunschweig, Germany) as described by Tindall (1990a, 1990b) and Kuykendall et al. (1988). To carry out these analyses, cells were grown in R2A broth, pH 6 at 25 °C under agitation (250 r.p.m) for 5 days and were then harvested by centrifugation (800 g; 10 min) at the end of their exponential phase of growth.

Iron oxidation metabolism

To evaluate the ability of strain HK31-G^T to oxidize iron under microaerophilic and anaerobic conditions (with nitrate as the terminal electron acceptor), its growth was evaluated using a modified DSMZ 730 medium (Emerson and Merrill Floyd, 2005), adapted to non-marine strains. The mineral basis was composed of NaCl (0.50 $g \cdot L^{-1}$), $MgCl_2.6H_2O$ (0.50 $g \cdot L^{-1}$), $CaCl_2.2H_2O$ (0.10 $g \cdot L^{-1}$), KCl (0.34 $g \cdot L^{-1}$), K_2HPO_4 (0.14 $g \cdot L^{-1}$), NH_4Cl (0.24·$g.L^{-1}$) and $FeCl_2.4H_2O$ at final concentration of 20 mM, added as

191 the sole electron donor. The medium was then adjusted to pH 6.5 and placed under an
192 anaerobic atmosphere (N_2 (100%; 1 bar)) before autoclaving. Solutions of Na_2S (0.30
193 $g \cdot L^{-1}$) under N_2 (100%) and Na_2CO_3 (1.50 $g \cdot L^{-1}$) under N_2/CO_2 (80/20) have been
194 prepared and autoclaved separately and then added to the medium after sterilization.
195 In addition, autoclaved solutions of vitamins (1 mL; DSMZ 141, 10×) and trace elements
196 (10 mL, DSMZ 141, 1×) were also added to the medium that was finally distributed (10
197 mL) in 50 mL penicillin flasks. Flasks were placed under N_2/CO_2 (80/20) atmosphere.
198 To be under microaerophilic conditions, 4 mL of 0.22 µm filtered ambient air were
199 injected into the flask to obtain a 2% final dioxygen concentration, while to be under
200 anaerobic conditions, 200 µL of KNO_3 solution (1 $mol \cdot L^{-1}$) were added to obtain a final
201 concentration of 20 $mmol \cdot L^{-1}$. Before inoculations that were performed in triplicate, cells
202 of strain HK31-G^T were previously grown in R2A broth, were washed three times with
203 the mineral basis described above, in order to remove all traces of carbon source. Then,
204 250 µL of the washed cells suspension were used to inoculate the Fe(II)-oxidizing
205 culture medium. Incubations were performed at 25 °C without agitation for 5 and 6 days
206 (end of exponential growth phase) under microaerophilic and anaerobic conditions,
207 respectively. Negative controls for each condition were added to the experiment (culture
208 medium without inoculation of cell suspension). For the third consecutive cultures in
209 these conditions, cell growth was estimated by direct cell count using epifluorescence
210 microscopy (Olympus BX60) at T0 day and T5 (microaerophilic condition) or T6
211 (anaerobic condition) days of incubation. For each condition, the experiments were
212 repeated twice and in an independent way. Each time, similar trends were obtained.
213 In parallel to each cell count, Fe(II) concentrations were estimated
214 spectrophotometrically using the Ferrozine method with slight modifications (Rouxel et
215 al., 2018; Viollier et al., 2000). At each measuring point, 150 µL of HCl 0.2% was added

to 150 µL of bacterial culture to stop the Fe(II) oxidation kinetics. To determine the Fe(II) concentrations, 100 µL of HCl at 6 mol.L^{-1}, were added, and then supplemented with the ferrozine solution (100 µL ; 10 mmol·L^{-1}) and the analytical buffer (ammonium acetate (200µL ; 10 mol·L^{-1}). The absorbance of the mixture was then measured spectrophotometrically at 562 nm. In order to link the OD_{562nm} obtained for the cultures with Fe(II) concentrations, a standard curve (OD_{562nm} *versus* Fe(II) concentrations (µmol.L^{-1})) was made using the culture medium (from 0 to 100 µmol.L^{-1}), as previously described (Fig. S2). To evaluate if the Fe(II) concentrations were significantly different between T0 and T5 or T6, the Mann-Whitney test (package rstatix version 0.7.2) was applied using RStudio (Version 4.2.1).

Transmission electron microscopy (TEM, Jeol JEM 1400) was finally performed to visualize cell-mineral interactions under both conditions.

Phylogenetic analysis

The complete double-strand 16S rRNA coding gene sequence of strain HK31-G^T was generated from an isolated colony, as described elsewhere (Alain et al., 2002). Pairwise 16S rRNA gene sequence similarity, of strains having validated published names was calculated using a global alignment algorithm implemented at the EzTaxon-e server (http://eztaxon-e.ezbiocloud.net/; (Yoon et al., 2017). Phylogenetic trees were reconstructed using the neighbor-joining (NJ), maximum parsimony (MP) and maximum likelihood (ML) methods using the software Seaview version 5.0.5 (Gouy et al., 2010). The evolutionary distances for the NJ, MP and ML methods were respectively calculated using the Kimura two-parameters, the Dnapars and the GTR models (Kimura, 1983; Saitou and Nei, 1987; Dereeper et al., 2008; Gouy et al., 2010; Guindon et al., 2010).

The robustness of the inferred topology was assessed by bootstrap analyses based on 1000 replications.

Genome sequencing, assembly and annotation

Genomic DNA of the novel strain HK31-G^T was extracted according to a standard PCI (Phenol-Chloroform-Isoamyl alcohol 25:24:1) protocol (Charbonnier *et al.*, 1995). Whole genome sequencing was performed by the Fasteris company (Plan-les-Ouates, Switzerland) using the Illumina MiSeq technology (2 × 150 bp paired reads; Micro Nano V2 chemistry). Reads quality control and genome assembly were performed on Galaxy France (https://usegalaxy.fr/) using, respectively, FastQC (Galaxy v0.73) (Andrews, 2010) and SPAdes assemblers (Galaxy Version 3.15.4) (Gangiredla *et al.*, 2021). Quality of the genome assembly was then obtained with Quast (Galaxy Version 5.2.0) (Guervich *et al.*, 2013). Genome completeness and potential redundancy were estimated with CheckM (Parks *et al.*, 2015) on the MicroScope Microbial Genome Annotation and Analysis Platform (MaGe; https://mage.genoscope.cns.fr) (Vallenet et al., 2020). The average nucleotide identity scores (ANI; OrthoANI values) and digital DNA-DNA hybridization (dDDH) scores between the genome of strain HK31-G^T and genomes of closely related type strains were respectively obtained using the ANI calculator tool from the EzBioCloud web server (https://www.ezbiocloud.net/tools/ani) and by the genome-to-genome distance calculator (GGDC 2.1), using formula 2 (Meiner-Kolthoff et al., 2013; Yoon et al., 2017). Genome assembly of the new strain HK31-G^T was annotated *via* the MaGe platform using KEGG and BioCyc databases but also using Prokka (Seemann, 2014) (Galaxy Version 1.14.6) on the Galaxy platform (Wee and Yap, 2021) and the NCBI prokaryotic genome annotation pipeline (PGAP) (Tatusova et al., 2016). In addition, EggNOG database was used to classify coding DNA sequences (CDS) to clusters of orthologous groups (COG) (Hernández-Plaza et al.,

2023). To identify the closest Metagenomes Assembled Genomes (MAGs) to the genome of strain HK31-G^T, the ANI scores between the genome of HK31-G^T and the 262 sequences of available MAGs affiliated to *Phenylobacterium* sp. on NCBI database, were calculated using FastANI tool on Galaxy France platform (Galaxy Version 1.3) (Jain et al., 2018).

The phylogenomic analysis focused on 20 references genomes representatives of the *Brevundimonas*, *Caulobacter* and *Phenylobacterium* genera and phylogenetically closely related to strain HK31-G^T (*Brevundimonas nasdae* Au29^T (GCF_019395145.1), *Brevundimonas vancanneytii* NCTC9239 (GCF_901421975.1), *Caulobacter mirabilis* FWC 38^T (GCF_002749615.1), *Caulobacter henricii* CB4 (GCF_001414055.1), *Caulobacter rhizosphaerae* KCTC 52215^T (GCF_010977555.1), *Caulobacter soli* Ji-3-8^T (GCF_011045195.1) *Caulobacter hibisci* KACC 18849^T (GCF_016135805.1), *Caulobacter flavus* RHGG3^T (GCF_003722335.1), *Caulobacter segnis* ATCC 21756^T (GCF_000092285.1), *Caulobacter vibrioides* CB15^T (GCF_000006905.1), *Phenylobacterium aquaticum* KACC 18306^T (GCF_022695515.1), *Phenylobacterium haematophilum* DSM 21793^T (GCF_014196295.1), *Phenylobacterium glaciei* 20 BVR1^T (GCF_016772415.1), *Phenylobacterium parvum* HYN0004^T (GCF_003150835.1), *Phenylobacterium immobile* ATCC 35973^T (GCF_001375595.1), *Phenylobacterium hankyongense* HKS-05^T (GCF_003254505.1), *Phenylobacterium deserti* YIM 73061^T (GCF_003254705.1), *Phenylobacterium soli* LX32^T (GCF_003254475.1), *Phenylobacterium kunshanense* BUT-10^T (GCF_003254525.1) and *Phenylobacterium zucineum* HLK1^T (GCF_000017265.1), the target genome HK31-G^T and its six closely related publicly available MAGs from groundwater (classified as the species *Phenylobacterium* sp030693625 on the Genome Taxonomy Database (GTDB, https://gtdb.ecogenomic.org), with accession numbers GCA_030693625, GCA_030645635, GCA_030704785, GCA_030696765, GCA_030683775 and

GCA_030652015). The phylogenomic tree was performed with the Anvi'o software (version 7.1) following the workflow published in 2017 (https://merenlab.org/2017/06/07/phylogenomics/) (Eren et al., 2021). Briefly, genomes were recovered from NCBI database and an Anvi'o contigs database was generated for each of them using the *anvi-gen-contigs-database* command. To compare genomes, the *anvi-get-sequences-for-hmms-hits* command was used to annotate single copy core genes listed in the Bacteria_71 HMMs profile collection, as well as ribosomal RNAs profiles ('Ribosomal_RNA_16S' (3 models), 'Ribosomal_RNA_23S' (2 models), and 'Ribosomal_RNA_5S' (5 models)). To build the tree, amino acid sequences were extracted, aligned and concatenated using the *anvi-get-sequences-for-gene-cluster* command. Then, to get the newick-formatted phylogeny artifact for these genomes and MAGs, the *anvi-gen-phylogenomic-tree* command was used with FastTree and the maximum-likelihood method (Price et al., 2009). The phylogenomic tree obtained was visualized using the anvi'o interactive interface thanks to the *anvi-interactive* command. The tree was rooted using the genome *Sphingomonas paucimobilis* ZJSH1 (GCA_016919545.1) as an outgroup. Finally, the single-copy-core genes corresponding to each of the genomes from the phylogenomic tree were used to calculate ANI values between each of them using the FastANI tool on Galaxy France platform (Galaxy Version 1.3) (Jain et al., 2018). Genes involved in iron metabolism were identified in the genome of strain HK31-G^T and those of its closely related MAGs using FeGenie in command line (version 1.2) according to Garber et al. (2020). Sequences homologous to the *cyc2* gene were also subjected to an HMM profile search using the HMMER 3.4 software package (Eddy, 1992).

Results and discussion

Morphological and physiological properties

Under optimal growth conditions, colonies of strain HK31-G^T are circular with regular edges and smooth surface and with a light creamish color. Cells are Gram-negative bacilli dividing by binary fission. They occur mainly singly (Fig. 1a) but can also form aggregates. Cell sizes range from 0.36 to 0.71 µm wide (mean 0.53 µm; *n*=42*)* and from 1.00 to 3.59 µm long (mean 1.87 µm; *n*=42). For many cells observed by TEM, refringent intracellular granules, which could be polyhydroxyalkanoates (PHA) storage granules, were also observed (Fig. 1b). Strain HK31-G^T is motile, as confirmed by observations on MMN agar medium and the observation of a single polar flagellum per cell by TEM (Fig. 1c, 1d). Interestingly, some cells of strain HK31-G^T harbor a stalk (Fig. 1a). Such cellular appendices were only previously observed for *P. conjunctum* FWC 21^T cells. However, contrary to the study of Abraham *et al.* (2008), no 'rosette' of prosthecate cells was observed. Like all closely related type strains, HK31-G^T is positive for oxidase activity but negative for catalase.

Growth was observed from 10 to 30 °C with optimal growth between 25 to 30 °C which is consistent with the optimal growth temperature of the other related type strains (Table 1). With respect to salt tolerance, the strain grew only at 0.5% NaCl and showed optimal growth without sodium chloride. Only *P. haematophilum* LMG 11050^T and *P. conjunctum* FWC 21^T expressed greater salt resistance with growth up to 2% NaCl (Table 1). Concerning the pH, growth of strain HK31-G^T was observed from pH 2 to 12 with optimal growth at pH 6. This large pH tolerance was not observed for the closest type strains that showed a narrower pH range but with an optimum that is also close to neutrality (Table 1). All these results are consistent with the environment where the strain was isolated. Indeed, at the time of sampling, a temperature of 24.1 °C and a pH of 8.78 were recorded, meaning that strain HK31-G^T is well adapted to these

environmental conditions. Under optimal growth conditions on R2A broth, the growth rate and the generation time of strain HK31-G^T are, respectively, 0.031 h^{-1} and 32 h 32 min (Fig. S3).

The novel isolate is chemoorganoheterotroph and grows by aerobic respiration. It catabolizes L-alanine, L-proline, phenylalanine, malonate and to a lesser extent, (D,L)-β-Hydroxybutyrate. *P. haematophilum* LMG 11050^T has a similar profile, except that it does not use malonate and (D,L)-β-hydroxybutyrate. The other closely related strains are not capable of using all the substrates tested, but only *P. aquaticum* W2-3-4^T is able to catabolize malonate as the sole carbon source (Table 1). Regarding the enzymatic activities that can differentiate the new strain HK31-G^T from the other closely related strains, the α- and ®-glucosidase as well as the α-mannosidase are only active for strain HK31-G^T as shown by the API®Zym kit. It should be noted that ®-galactosidase activity was also demonstrated for *P. haematophilum* LMG 11050^T when using API®20E kit with PNPG as substrate. As with *P. glaciei* 20VBR1^T and *P. aquaticum* W2-3-4^T, the strain HK31-G^T was able to use nitrate as a terminal electron acceptor, revealing nitrate reductase activity (Table 1). This activity was also experimentally demonstrated with the MMN medium for strain HK31-G^T and *P. aquaticum* W2-3-4^T.

Table 1. Differential physiological characteristics of strain HK31-G^T and related type strains of the genus *Phenylobacterium*. 1. HK31-G^T, 2. *P. glaciei* 20VBR1^T (DSM 111428^T) (Thomas et al., 2022), 3. *P. aquaticum* W2-3-4^T (KACC 18306^T) (Jo et al., 2016) , 4. *P. haematophilum* LMG 11050^T (Abraham et al., 2008), 5. *P. conjunctum* FWC 21^T (LMG 24262^T) (Abraham et al., 2008).

All data concerning nitrate reduction to nitrite, carbon sources utilization and API®20 NE, 20 E and ZYM come from this study. Otherwise, for other characteristics concerning related type strains, data come from the references cited above. Note that for *P. glaciei* 20VBR1^T, the strain was not available from the two culture collections where it was deposited and the results presented here were taken from (Thomas et al., 2022).

All the strains were positive for oxidase activity but negative for catalase. For carbon sources utilization, all the strains were negative for xylose, glycerol, aspartate, acetate, lactate, succinate, glutamate, propionate and glucose through butanediol pathway. None are able of citrate fermentation. In API®20 NE and API®20 E, all the strains were negative for β-galactosidase and arginine dihydrolase, assimilation of D-glucose, L-arabinose, D-mannose, N-acetyl-glucosamine, D-maltose, citrate, gluconate, caprate, adipate, malate and phenyl-acetate, production of H_2S, indole and acetoin and acid production from D-glucose, D-mannitol, inositol, D-sorbitol, L-rhamnose, D-saccharose, D-melibiose, amygdalin and L-arabinose. In API®ZYM kits, all the strains were positive for alkaline phosphatase, esterase, esterase lipase, leucine arylamidase, acid phosphatase and naphthol-AS-BI-phosphohydrolase, but negative for lipase, cystine arylamidase, chymotrypsine, β-glucuronidase and α-fucosidase.

Characteristics	**1**	**2**	**3**	**4**	**5**
Colony color	Light creamish	Light creamish	Light brown	Light creamish	Light brown
Cell size (µm)	0.4-0.7 × 1.0-3.6	0.4-0.5 × 1.0-2.0	0.4-0.6 × 2.5-4.0	0.3-0.4 × 0.9-2.5	0.5-0.7 × 1.2-1.6
Mobility	+	–	–	+	–
Prostheca	+	–	–	–	+
Temperature range (optimum) (°C)	10-30 (25-30)	10-35 (20)	18-40 (25-30)	10-40 (37)	20-40 (25-30)
pH growth range (optimum)	2-12 (6)	6.5-8 (7)	6.0-8.0 (7)	6.0-8.5	5.5-8
NaCl tolerance (optimum) (%, w/v)	0-0.5 (0)	0-2.5 (0.5)	0-0.5	0-2	0-2
Nitrate reduction to nitrite	+	+	+	–	–
Carbon sources utilization:					
L-Alanine	+	–	–	w	–
L-Proline	+	–	–	+	–
Phenylalanine	+	–	–	w	–
(D,L)-®-Hydroxybutyrate	w	–	–	–	–
Malonate	+	–	+	–	–
Mannitol (MMN)	–	–	+	–	–
Glucose (Kligler-Hajna)	–	–	–	+	–
API® 20NE and 20E results					
Lysine decarboxylase	–	NA	+	–	–
Ornithine decarboxylase	–	NA	+	–	–
Urease	–	–	+	–	+
Tryptophan deaminase	+	NA	+	–	–
Gelatinase	+	NA	+	–	+
Arginine dihydrolase	–	–	+	–	+

β -Glucosidase (aesculin hydrolysis)	+	NA	–	+	–
Gelatin hydrolysis	–	–	+	–	+
β -Galactosidase (PNPG)	+	NA	–	+	–
API® ZYM results					
Valine arylamidase	–	+	–	+	+
α-Galactosidase	+	–	–	–	–
β -Galactosidase	+	–	–	–	–
α-Glucosidase	–	+	–	+	+
β -Glucosidase	–	+	–	–	–
N-acetyl-β-glucosaminidase	+	–	–	+	+
α-mannosidase	+	–	–	–	–
DNA G+C content (mol%)	67.95	67.86	68.87	67.9	67.0
Isolation source	Deep subsurface geothermal aquifer	Glacier snout ice	Reservoir of water purifier	Human blood	Water biofilm

+, positive; –, negative; w, weakly; w, weak; NA, Not Available

Chemotaxonomic characteristics

The major respiratory quinone of strain HK31-G^T is Q-10 (98.79%) as described for other members of the family *Caulobacteraceae* (Thomas et al., 2022). In addition, a low amount of Q-9 (1.21%) was also detected. The polar lipids profile indicated the presence of phosphatidylglycerol (PG), which is in line with other species of *Phenylobacterium*, and by two unknown glycolipids (GL), an unknown glycophospholipid (GPL), two unknown phospholipids (PL) and four unidentified polar lipids (L) (Fig. S4). This profile is slightly different from that of the closest species *P. glaciei* 20VBR1^T and *P. aquaticum* W2-3-4^T. Indeed, in addition to PG, *P. glaciei* 20VBR1^T was characterized by an unknown PG, an unknown GL and four unknown lipids, while *P. aquaticum* W2-3-4^T was characterized by an unknown PL, four unknown GL and three unknown lipids (Thomas et al., 2022; Jo et al., 2016). This latter profile was similar to *P. conjunctum* FWC 21^T (LMG 24262^T) but quite different from that of *P. haematophilum* LMG 11050^T (Abraham et al., 2008). Polar lipids can therefore be used as biomarkers, in addition to other criteria, to distinguish *Phenylobacterium* species (Abraham et al., 2008).

403

404 The predominant cellular fatty acids (>10% of the total fatty acids) of strain HK31-G^T
405 were $C_{18:1}\omega7c$ (also in summed feature eight comprising $C_{18:1}\omega7c$ and/or $C_{18:1}\omega6c$;
406 44.54%), followed by $C_{16:1}\omega6c$ (also in summed feature three comprising $C_{16:1}\omega6c$
407 and/or $C_{16:1}\omega7c$; 19.30%) and the saturated fatty acid $C_{16:0}$ (11.62%) (Table 2). Cells
408 of strain HK31-G^T and its closely related genera are characterized by high amounts
409 (>10.2%) of summed feature eight and $C_{16:0}$ as already observed for members of the
410 genus *Phenylobacterium* (Jo et al., 2016; Stackebrandt et al., 2006). Summed feature
411 three seems more abundant and in similar amounts in cells of strain HK31-G^T and *P.*
412 *glaciei* 20VBR1^T (19.30% and 19.18%, respectively) compared to the closely related
413 species (ranging from 1.88% to 8.9%). Conversely, the methyl ester fatty acid
414 $C_{18:1}\omega5c$11-methyl, present in similar amount in cells of strain HK31-G^T and *P. glaciei*
415 20VBR1^T (2.54% and 6.55%, respectively), was less abundant compared to the closely
416 related genera (ranging from 10.8% to 28%). Some minor quantitative differences can
417 also be observed between the new strain HK31-G^T and its closest relatives. Indeed, low
418 amount of summed features 1 (comprising $C_{13:0}$3-OH and/or iso-$C_{15:1}$; 0.26%) and 9
419 (comprising $C_{17:1}$iso$\omega9c$ and/or $C_{16:0}$-methyl; 0.41%) and iso-$C_{15:0}$ (0.32%) were only
420 detected in cells of strain HK31-G^T (Table 2). Overall, these whole-cell fatty acid profiles
421 demonstrate that strain HK31-G^T shares typical characteristics of species of the genus
422 *Phenylobacterium* but with a fatty acid profile highly similar to that of *P. glaciei* 20VBR1^T.

423
424
425
426
427
428
429
430
431

Table 2. Comparison of whole-cell fatty acid profiles (% of the total) of strain HK31-G^T with type strains of closely related species.

Taxa : 1. HK31-G^T, 2. *P. glaciei* 20VBR1^T (DSM 111428^T) (Thomas et al., 2022), 3. *P. aquaticum* W2-3-4^T (KACC 18306^T) (Jo et al., 2016), 4. *P. haematophilum* LMG 11050^T (Abraham et al., 2008), 5. *P. conjunctum* FWC 21^T (LMG 24262^T) (Abraham et al., 2008).

Values are percentages of the fatty acids that were assigned to fatty acids in the peak-naming table of the MIS database (MIDI, Microbial ID, Newark, DE 19711 U.S.A.). The nomenclature is as follows: the first number indicates the number of carbon atoms in the molecule; 'OH' and 'cyclo' indicate hydroxy or cyclic fatty acids; the second number following the colon indicates the number of double bonds present. The position of the double bond is indicated by the carbon atom position starting from the methyl (*ω*) end of the molecule. *c*, cis isomer. Major fatty acids (>10% of the total fatty acids) are indicated in bold.

Fatty acids[a] (%)	1	2	3	4	5
Saturated					
$C_{10:0}$	–	–	2.1	–	–
$C_{11:0}$	0.18	–	0.69	TR	0.8
$C_{12:0}$	–	–	0.3	2.3	–
$C_{13:0}$	–	–	–	–	–
$C_{14:0}$	0.44	0.48	0.2	TR	0.5
$C_{15:0}$	TR	6.77	–	–	–
$C_{16:0}$	**11.62**	**10.2**	**30.3**	**17.4**	**20.5**
$C_{17:0}$	4.87	4.0	**10.67**	7.6	6.9
$C_{18:0}$	0.76	–	3.7	TR	0.5
$C_{20:0}$	0.68	–	–	1.3	–
Unsaturated					
$C_{12:1}\omega7c$	TR	–	–	–	–
$C_{16:1}\omega11c$	0.91	0.5	1.3	1.6	4.0
$C_{16:1}\omega7c$	**19.30**	–	–	–	–
$C_{17:1}\omega6c$	6.86	4.20	1.27	2.0	1.7
$C_{17:1}\omega8c$	1.76	1.56	1.09	0.9	0.8
$C_{18:1}\omega7c$	**44.54**	–	–	–	–
Methyl ester					
$C_{18:1}\omega5c$ *11-methyl*	–	–	–	–	2.0
$C_{18:1}\omega7c$ *11-methyl*	2.54	6.55	**21.7**	**28.0**	**10.8**
Branched-chain fatty acid					
Iso-$C_{15:0}$	0.32	–	–	–	–
Iso-$C_{17:0}$	0.96	0.5	0.75	TR	0.5
Anteiso- $C_{15:0}$	0.27	–	–	–	–
Hydroxyl fatty acids					
$C_{12:0}$ 3-OH	0.68	0.79	0.85	0.8	1.0
$C_{12:1}$ 3-OH	2.64	2.0	1.7	1.9	2.1
Cyclo fatty acids					
$C_{17:0}$ cyclo *ω8c*	–	–	0.57	2.7	–
Summed features					
1 ($C_{13:0}$ 3-OH and/or iso-$C_{15:1}$ H)	0.26	–	–	–	–
3 ($C_{16:1}\omega6c$ and/or $C_{16:1}$ *ω7c*)	**19.30**	**19.18**	1.88	5.9	8.9
8 ($C_{18:1}\omega7c$ and/or $C_{18:1}$ *ω6c*)	**44.54**	**40.25**	**16.86**	**25.4**	**38.8**
9 ($C_{17:1}$iso *ω9c* and/or $C_{16:0}$-methyl)	0.41	–	–	–	–

–, Not detected

TR, trace amount

As all strains were not grown under exactly the same conditions or in the same media, the fatty acid percentages must be compared with caution from one strain to another.

Iron oxidation metabolism

After three successive cultures in a medium targeting growth by iron oxidation, cell density and Fe(II) concentrations were determined immediately after inoculation in all cultures and after 5 and 6 days of incubation under microaerophilic and anaerobic (with nitrate as the terminal electron acceptor) conditions, respectively. For each condition, a cell growth greater than 0.5 Log2(Number of cells/mL) was obtained for the new strain HK31-G^T. Indeed, under the microaerophilic conditions, an increase of 0.53 Log2(Number of cells/mL) was evidenced while an even greater increase of 1.03 Log2(Number of cell/mL) was obtained under the anaerobic conditions (Fig. 2). As abiotic Fe(II) oxidation occurs under the microaerophilic conditions, it was not possible to evidence a significant biotic Fe(II) oxidation by cells of strain HK31-G^T in comparison with control by using the ferrozine bioassay (*P > 0.05;* data not shown). Conversely, under anaerobic condition, Fe(II) concentrations were significantly different between T0 (21.45 µmol.L^{-1} ± 1.57 µmol.L^{-1}) and T6 days (17.22 µmol.L-1 ± 1.39 µmol.L^{-1}) (*P < 0.001*) when cells of strain HK31-G^T were present in the medium whereas no significant difference was observed for negative controls between T0 and T6 days (*P > 0.05*). In addition, potential cell-mineral interactions necessary to the extracellular oxidation of Fe, were also observed by TEM (Fig. S5). Overall, cells of strain HK31-G^T are capable of oxidizing Fe as an energy source under both microaerophilic and anaerobic conditions. To our knowledge, *Phenylobacterium* species are not known to have a secondary metabolism based on iron oxidation. This shift in metabolism is consistent with the environment in which the strain was isolated. Indeed, following an injection of CO_2, environmental conditions become favorable to the development of chemolithoautotrophic microbial communities (Mu et al., 2015; O'Mullan et al., 2015) and Fe(II)-oxidizing bacteria (Trias et al., 2017).

Phylogenetic affiliation

Comparative analysis of 16S rRNA gene sequences of HK31-G^T (1372 bp) and closely related type strains with validated published names, indicated that its closest relatives were *P. aquaticum* W2-3-4^T (97.52%), *P. haematophilum* LMG 11050^T (97.37%) and *P. parvum* HYN0004^T (96.94%). In addition, the 16S rDNA gene sequence of strain *P. glaciei* 20VBR1^T, which does not have a validated published name, was 99.9% similar. The phylogenetic trees based on ML, MP and NJ algorithms and on 16S rDNA sequences of strains belonging to the genera *Phenylobacterium*, *Brevundimonas*, *Caulobacter*, *Aquidulcibacte*r, and *Terricaulis*, showed that strain HK31-G^T clusters within the genus *Phenylobacterium*, and forms a clade (MP, ML and NJ algorithms) with *P. glaciei* 20VBR1^T, *P. aquaticum* W2-3-4^T, *P. haematophilum* LMG 11050^T and *P. conjunctum* FWC 21^T (Fig. S6).

The marker nucleotides of the 16S rRNA gene of the order *Caulobacterales* and in particular of the genus *Phenylobacterium*, described in Abraham et al. (2008) were used to determine the nucleotide signature of strain HK31-G^T in relation to the type strains of the closest species, *P. glaciei* 20VBR1^T, *P. aquaticum* W2-3-4^T, *P. haematophilum* LMG 11050^T, *P. conjunctum* FWC 21^T, *P. parvum* HYN0004^T and *P. koreense* Slu-01^T (Table S1). We confirmed that the new strain and all compared species were missing nucleotides at positions 71–88, 183–190, 206–211 and 452–476 (*Escherichia coli* str. K12 subtr. MG1655 numbering), as expected for the order *Caulobacterales*. Strains HK31-G^T and *P. glaciei* 20VBR1^T have an A at position 1265 and a T at position 1270, which is different from the closest type strains (1265-T and 1270-A) (Table S1). It has been previously shown that *P. glaciei* 20VBR1^T has the same nucleotide signature as *P. falsum* AC-49^T and *P. panacis* DCY 109^T (Thomas *et al.,* 2022). These two nucleotides (1265-T and 1270-A) have been used so far to differentiate the genera

Phenylobacterium and *Caulobacter* from the genera *Brevundimonas* and *Asticcacaulis,* where they are absent according to Abraham et al. (2008).

Based on MaGe, the total genome size of strain HK31-G^T was 4.46 Mbp for 100% completion and 1.41% redundancy (seven markers were duplicated) and consisted of 104 contigs. The N50 and L50 were respectively 89,925 bp and 19 contigs. The G+C content of the genomic DNA of strain HK31-G^T was 67.95%, which is consistent with the closely related strains of the genus *Phenylobacterium* (G+C contents ranging from 66.73% to 68.87%) (Table 3). Annotations with MaGe resulted in 4,667 CDS, six rRNA operons and 48 tRNA genes (corresponding to the 20 essential amino acids) (Table 3; Fig. S7). Most of the CDS (82.65%) were assigned to at least one COG category (Table S2). Among major processes, COGs categories related to metabolism were dominant (27.06% of the CDS) and included (>4% of the CDS) (i) inorganic ion transport and metabolism (4.90%), (ii) amino acid transport and metabolism (4.58%), and (iii) energy production and conversion (4.51%). Then, 17.05% of the CDS were involved in cellular processes and signaling represented mainly by signal transduction mechanisms (4.36%) and cell wall, membrane and envelope biogenesis (4.19%). Finally, 13.41% of the COGs were dedicated to information storage and processing with mechanisms such as transcription (5.07%), replication, recombination and repair (4.45%) or signal transduction (4.36%).

Phylogenomic tree and Overall Genome Relatedness Indices (OGRI)

Calculation of ANI scores between genome of strain HK31-G^T and all the sequences of MAGs affiliated to *Phenylobacterium* sp. available on NCBI database allowed to identify six MAGs (accession numbers: GCA_030693625, GCA_030645635, GCA_030704785, GCA_030696765, GCA_030683775 and GCA_030652015) with a score above the

530 species delineation threshold and affiliated to species *Phenylobacterium* sp030693625
531 (GTDB) (Table S3). OrthoANIu values confirmed these results for the six MAGs
532 identified previously, with values ranging from 97.11% to 98.12% (Richter & Rossello-
533 Mora 2009). These six MAGs were therefore used to build the phylogenomic tree.
534 Similarly to the phylogenetic analysis based on the 16S rRNA gene, the phylogenomic
535 tree identified a cluster containing genomes of strain HK31-G^T, *P. glaciei* 20VBR1^T, *P.*
536 *aquaticum* W2-3-4^T and *P. haematophilum* DSM 21793^T and the six MAGs affiliated to
537 *Phenylobacterium* sp030693625. The genome of strain HK31-G^T being the most closely
538 related to that of the reference strain *P. glaciei* 20VBR1^T (Fig. 3 and Fig. S8.) and the
539 MAG with accession number GCA_030645635 (Fig. 3). Digital DNA-DNA hybridization
540 (dDDH) scores between the genome of strain HK31-G^T and the most closely related
541 genome of reference species ranged from 20.1 to 35.0% and were well below the
542 threshold for distinguishing two different species (Table 3) (Wayne et al., 1987;
543 Stackebrandt et al., 2006). In the same way, OrthoANIu values were all below the
544 generally accepted threshold of 95-96% for species delineation with values ranging from
545 75.14 to 88.97% (Richter & Rossello-Mora 2009). In addition, ANI values, calculated
546 with FastANI between the most closely related genome of reference species and the
547 genome of HK31-G^T ranged from 74.9 to 89.1% and were also well below the threshold
548 for species delineation while FastANI values calculated between the genome of HK31-
549 G^T and its six closely related MAGs ranged from 96.94 to 97.90 %, confirming that they
550 represent the same species (Table S4 and Fig. 4). Despite the lack of sufficient
551 resolution of 16S rRNA coding gene sequence comparisons to delineate the strain
552 HK31-G^T as a new species, OGRI (dDDH and ANI values) were powerful enough to
553 lead to the conclusion that strain HK31-G^T represents a new species of the genus
554 *Phenylobacterium*. Therefore, phylogenomic distances as well as OGRI values

555 confirmed that strain HK31-G^T and *P. glaciei* 20VBR1^T represent two distinct species of
556 the genus *Phenylobacterium*.
557
558 **Table 3. Genome statistics and Overall Genome Relatedness Indices (OGRI)**
559 **between the genome of strain HK31-G^T and the reference genomes of closely**
560 **related species of the genus *Phenylobacterium*.** Species: 1, HK31-G^T (data from this
561 study); 2, *P. glaciei* 20VBR1^T; 3, *P. aquaticum* W2-3-4^T; 4, *P. haematophilum* DSM
562 21793^T; 5, *P. immobile* strain E^T.
563

	1	2	3	4	5
Number of contigs	104	5	90	21	6
Size (Mbp)	4.46	4.24	5.25	4.43	3.33
G+C content (%)	67.95	67.86	68.87	67.90	66.73
Number of CDS	4667	4257	5552	4502	3316
rRNA	6	3	3	3	3
tRNA	48	43	46	47	45
16S rDNA similarity (%)	100	99.9	97.52	97.37	96.06
dDDH (%)	100	35.0	22.1	21.1	20.1
OrthoANIu (%)	100	88.97	78.87	77.64	75.14
FastANI (%)	100	89.1	82.2	80.9	79.4

564
565 *Central metabolism and energy production pathways*
566 Based on genomic predictions, strain HK31-G^T encodes complete pathways for the
567 biosynthesis of the 20 essential amino acids (alanine, arginine, asparagine, aspartate,
568 cysteine, glutamate, glutamine, glycine, histidine, isoleucine, leucine, lysine,
569 methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine and valine).
570 Complete pathways for organo-heterotrophic growth, namely the glycolysis, Entner-
571 Doudoroff, TCA cycle pathways and the pentose phosphate (oxidative and non
572 oxidative branches) pathways have been identified, as well as the degradation
573 pathways for several amino acids such as alanine, asparagine, glutamate, glutamine,
574 L-cysteine, L-serine and taurine. The genome also encodes complete pathways for
575 glycerol and acetoacetate degradation and for C1 compounds, including methanol and
576 formaldehyde oxidation, and for CO_2 fixation into oxaloacetate by anapleurotic reaction
577 (Hobmeier *et al.*, 2020). This last reaction could be the one used for chemoautotrophic
578 growth under Fe(II) oxidation conditions. Genomic predictions allowed identifying genes
579 involved in polyhydroxyalkanoates (PHAs) synthesis. Indeed, *phaB* (locus_tag (PGAP):

OCL97_08250), *phaC* (locus_tag (PGAP): OCL97_01260) and *phaR* (locus_tag (PGAP): OCL97_08240), encoding respectively for an acetoacetyl-CoA reductase, the poly(3-hydroxyalkanoate) polymerase and the PHA synthesis repressor were identified. In addition, genes encoding polyhydroxyalkanoic acid system family protein (locus_tag (PGAP): OCL97_12980) and a PHA depolymerase (locus_tag (PGAP): OCL97_14885) were also evidenced. The presence of these genes suggests that it is the class IV PHA synthase operon, widespread in bacteria belonging to the genus *Bacillus* (Tsuge et al., 2015), that is at work in strain HK31-G^T. These results, combined with the observations of intracellular granules by TEM, reinforce the hypothesis that the new strain produces PHAs storage granules to adapt to harsh environments where the organic matter is scarce. With regard to energy metabolism, genes encoding cytochrome *c* oxidase complex for aerobic respiration (*ctaC* (OCL97_16890), *ctaD* (OCL97_16895), *ctaG* (OCL97_16910), and *ctaE* (OCL97_16915)) are coded in the genome, confirming that the enzyme is at work in strain HK31-G^T, which has been demonstrated experimentally. Enzymes needed to resist oxidative stress have also been predicted in the genome of strain HK31-G^T (but not demonstrated experimentally). Indeed, catalase (OCL97_19865), catalase peroxidase (OCL97_00270), superoxide dismutase (OCL97_11735), and glutathione peroxidase *gpo* (OCL97_19215) were identified in its genome.

Interestingly, the genome of strain HK31-G^T encodes for the *cbb3* cytochrome (OCL97_16190 for the subunit I; OCL97_16185 for the subunit II; OCL97_16175 and OCL97_16180 for the subunit III and OCL97_16155 for the *cbb3*-type cytochrome oxidase assembly protein CcoS), which has a high affinity for dioxygen, so allows growth in microaerophilic environments. In addition, subunits I and II of cytochrome *bd* ubiquinol oxidase were identified (OCL97_15410 for the subunit I; OCL97_15415 for the subunit II) confirming the ability of the new strain to grow under microaerophilic

conditions (Jünemann et al., 1997) as experimentally shown. Finally, the *aa3* cytochrome oxidase was detected; it could be involved in the electron transfer mechanism from the extracellular Fe(II) to the intracellular O_2, as proposed by Peng et al. (2022). Predictions by MicroCyc on MaGe revealed the presence of a complete dissimilatory nitrate reduction pathway in the genome of strain HK31-G^T. Indeed, genes encoding for the A subunit of nitrate reductase (subunits alpha (narG (OLC97_05710)), beta (narH (OLC97_05715)), and gamma (narI (OLC97_05720)) subunits), for the chaperone protein (narJ (OLC97_05725)), and the nitrate transporter (*narK* (OLC97_05700)), as well as for the NADH: ubiquinone oxidoreductase subunits (*nuoB*, *nuoH*, *nuoI*, *nuoJ*, *nuoK*, *nuoM* and *nuoN*) were identified (for locus tag see Table S5). The missing NADH:ubiquinone oxidoreductase, and transporter subunits (A, E, F, G and L) were finally detected from the synteny map around genes annotated as *nuo*. The comparison to the Uniprot database allows identifying the *nuoA*, *nuoE*, *nuoF*, *nuoG* and *nuoL* genes confirming their putative function. The presence of this complete nitrate reduction pathway (Fig. 5) indicates the ability of strain HK31-G^T to respire nitrate, which is consistent with the anaerobic respiration demonstrated experimentally. This is in line with the strain's ability to change ecological niches and adapt easily.

Concerning iron metabolism, *i.e.* iron acquisition, storage and oxidation/reduction, several genes were identified by FeGenie and annotations were completed by PGAP (NCBI), Prokka and MaGe predictions. A total of 72 Fe-related genes were identified. Among them, 38 genes were identified by FeGenie (Table 4) and only 5 of them were found by all approaches (Table S6). A majority of them (56) were involved in iron assimilation, and included eight genes involved in iron(II)/(III) transport (including the Fe(II)-transporters *FeoA* and *FeoB* and the Fe(II)-permease *EfeU*), two genes involved in transport by hemes (namely *HmuU* (identified as *FpvE*, involved in siderophore

transport, by FeGenie) and *HmuV*), and 42 genes involved in transport by siderophores (with most of them encoding for *Ton-B* dependent receptor). Then, 15 genes were involved in iron regulation, one in iron storage, and two genes might be involved in dissimilatory iron reduction. The first iron reductase was only predicted with MaGe and was 69.3% similar to the gene NX02_p1185 present in the genome of *Sphingomonas sanxanigenens* NX02^T (Uniprot database). The second iron reductase was predicted with PGAP, and identified as a ferric reductase-like transmembrane domain-containing protein. This enzyme could be involved in Fe(III) respiration or in iron assimilation for metalloproteins. Various *c*-type cytochromes have been identified but none of them corresponds to the *c*-type cytochrome encoded by *cyc1* and *cyc2* genes involved in iron oxidation. However, the *cyc2* gene coding an outer-membrane cytochrome protein was identified, by FeGenie, in the genome of *P. glaciei* 20VBR1^T, the closest relative of strain HK31-G^T, and in the genome of *P. hankyongense* HKS-05^T. The co-transcribed *cyc1* and *cyc2* genes, encoding cytochrome *c*4 (or C_{552}) (Appia-Ayme et al., 1998), and known to be involved in Fe(II) oxidation (McAllister et al., 2020) were not identified in the HK31-G^T genome, by none of the annotation tools used, suggesting, either that an alternative Fe(II) oxidation pathway exists, or that the genome is not sufficiently complete to identify all the genes it encodes.

Table 4. List of genes involved in iron transport, storage and redox cycling identified in the genome of strain HK31-G^T and in the genomes of closely related species of *Phenylobacterium* using the FeGenie tool. Species: 1, HK31-G^T; 1a, GCA_030645635; 1b, GCA_030652015; 1c, GCA_030683775; 1d, GCA_030693625; 1e, GCA_030696765; 1f, GCA_030704785; 2, *P. glaciei* 20VBR1^T; 3, *P. aquaticum* W2-3-4^T; 4, *P. haematophilum* DSM 21793^T; 5, *P. immobile* strain E^T.

		1	1a	1b	1c	1d	1e	1f	2	3	4	5
Iron acquisition	**Iron transport**	4	0	4	4	4	4	4	0	4	2	2
	Heme transport	0	0	0	0	0	0	0	0	0	1	0
	Heme oxygenase	0	0	0	0	0	0	0	0	0	0	0
	Siderophore synthesis	0	0	0	0	0	0	0	0	0	0	3
	Siderophore transport	9	7	8	12	8	10	12	0	18	17	11
	Siderophore transport potential	9	6	6	6	6	6	10	0	6	6	8
Energetic metabolism	**Iron oxidation**	0	0	0	0	0	0	0	1	0	0	0
	Possible iron oxidation and possible iron reduction	0	0	0	0	0	0	0	0	0	0	0
	Probable iron reduction	0	0	0	0	0	0	0	0	0	0	0
	Magnetosome formation	0	0	0	0	0	0	0	0	0	0	0
Other	**Iron gene regulation**	15	11	17	9	16	12	13	18	20	16	10
	Iron storage	1	2	2	2	2	2	2	1	1	1	2
	Magnetosome formation	0	0	0	0	0	0	0	0	0	0	0

Conclusion

A bacterial strain representative of the MAG (GCA_030645635) belonging to the species *Phenylobacterium* sp030693625, according to GTDB database, has been isolated. The new strain HK31-G^T is a mesophilic, neutrophilic and chemoheterotrophic bacteria able to grow by dioxygen and nitrate respiration using a wide variety of organic substrates (Table 5). We also provided evidence that it is able to grow under chemolithoautotrophic conditions, by Fe(II) oxidation, using nitrate and low dioxygen concentrations as terminal electron acceptors, and CO_2 as the sole carbon source. To our knowledge, this secondary metabolism based on Fe(II) oxidation has not previously been demonstrated in this genus. Genomic predictions indicate the presence of genes

encoding for *c*-type and *cbb3*-type cytochrome oxidase and cytochrome *bd* ubiquinol oxidase as well as genes encoding for the complete nitrate reduction pathway, which is congruent with the strain's ability to develop under oxic, microoxic, and anoxic conditions in the presence of nitrate. Despite that Fe(II) oxidation metabolism has been proven experimentally, key genes involved in Fe(II) oxidation pathway could not be identified. Growth observed under different nutrients and physico-chemical conditions, suggest that HK31-G^T can adapt to different ecological niches. In this respect, strain HK31-G^T is thought to produce polyhydroxyalkanoate storage granules (PHA), compounds that confer an ecological advantage since PHAs can help to cope with different stress conditions. Phylogenomic relatedness indices, as well as phenotypic, metabolic and chemotaxonomic differences between strain HK31-G^T and the most closely related strain *P. glaciei* 20VBR1^T confirm that strain HK31-G^T is a novel species of the genus *Phenylobacterium* for which we suggest the name *Phenylobacterium ferrooxidans* HK31-G^T.

701 **Table 5. Description of *Phenylobacterium ferrooxidans* sp. nov.**

Parameters	*Phenylobacterium ferrooxidans* sp. nov.
Guiding Code for Nomenclature	ICNP
Nature of the type material	strain and genome sequence
Genus name	*Phenylobacterium*
Species name	*Phenylobacterium ferrooxidans*
Specific epithet	*ferrooxidans*
Species status	sp. nov.
Species etymology	fer.ro.ox'i.dans sp. nov. L neut. n. *ferrum,* iron; N.L. v. *oxido*, oxidize; N.L. part. adj. *ferroxidans,* iron-oxidizing
Designation of the Type Strain	HK31-G^T
Strain Collection Numbers	DSM 116432^T = UBOCC-M-3429^T = LMG 33376^T
Type Genome, MAG or SAG accession Nr. [INSDC databases]	JAOTJD000000000
Genome status	Complete
Genome size	4.46 Mbp
GC mol%	67.95%
16S rRNA gene accession nr.	OR652334
Description of the new taxon and diagnostic traits	*Phenylobacterium ferrooxidans* members are motile, Gram-negative bacilli (0.4 – 0.7 x 1.0 – 3.6 µm in size). Colonies appear smooth, circular, light creamish, convex and 0.5 – 1 mm in diameter. The type strain HK31-G^T is mesophilic and grows at circumneutral pH. Growth occurs at 10-30 °C (optimum, 25-30 °C), at pH 2-12 (optimum, 6) and with 0.5% NaCl concentration (optimum, 0% NaCl). They are oxidase-positive, catalase-negative, aerobic and chemoorganotrophic when cultivated on R2A but microaerophilic and anaerobic and chemolithoautotrophic when cultivated on Fe(II) oxidation liquid medium. Nitrate is reduced to nitrite. Capable of oxidizing iron under microaerophilic and anaerobic conditions and using CO_2 as carbon source on Fe(II) oxidation rich medium. The following carbon sources are used: L-alanine, L-proline, phenylalanine, β-hydroxybutyrate and malonate. The following carbon sources are used: L-alanine, L-proline, phenylalanine, β-hydroxybutyrate and malonate. On the contrary, unable to use: xylose, glycerol, aspartate, acetate, lactate, succinate, glutamate, propionate, D-glucose, L-arabinose, D-mannose, N-acetyl-glucosamine, D-maltose, citrate, gluconate, caprate, adipate, malate and phenylacetate, D-mannitol, inositol, D-sorbitol, L-rhamnose, D-saccharose, D-melibiose, amygdalin and L-arabinose. Cells produce β-glucosidase (aesculin hydrolysis), α-galactosidase, β-galactosidase (PNPG), N-acetyl-β-glucosaminidase, α-mannosidase, tryptophan deaminase, gelatinase, alkaline phosphatase, esterase, esterase lipase, leucine arylamidase, acid phosphatase and naphthol-AS-BI-phosphohydrolase but unable to produce lysine decarboxylase, ornithine decarboxylase, urease, arginine dihydrolase, valine arylamidase, β-Galactosidase, arginine dihydrolase, lipase, α-glucosidase, β-glucosidase, crystine arylamidase, chymotrypsine, β-glucuronidase, and α-fucosidase.
Country of origin	Iceland

Region of origin	Carbfix-1 site of an adjacent geothermal power plant at Hellisheidi
Date of isolation	08/06/2021
Source of isolation	Deep subsurface geothermal aquifer at an adjacent geothermal power plant at Hellisheidi
Sampling date	18/07/2019
Latitude	64° 02’ 14’’ N
Longitude	21° 24’ 03’’ W
Number of strains in study	1
Information related to the Nagoya Protocol	Not applicable

Fundings

This research was funded by the S4CE collaborative project (H2020-EU.3.3.2. - Low-cost, low-carbon energy supply -764810), the French National Research Agency for the project IRON2MI under the reference ANR-22-CE01-0013-01 to S.M., by the Sino-French IRP 1211 Microbsea to K.A. and by the UMR6197.

CRediT authorship contribution statement

Eva Pouder: Conceptualization, Formal analysis, Methodology, Investigation, Software, Visualization, Writing – original draft, Writing – review and editing. **Erwann Vince**: Investigation, Writing – review and editing. **Karen Jacquot**: Investigation, Writing – review and editing. **Maimouna batoma Traoré**: Investigation, Writing – review and editing. **Ashley Grosche**: Investigation, Methodology, Writing – review and editing. **Maria Ludwig**: Investigation, Writing – review and editing. **Mohamed Jebbar**: Resources, Funding acquisition, Writing – review and editing. **Lois Maignien**: Resources, Writing – review and editing. **Karine Alain**: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Visualization, Supervision, Validation, Writing – original draft, Writing – review and editing. **Sophie Mieszkin**: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology,

Visualization, Project administration, Supervision, Validation, Writing – original draft, Writing – review and editing.

Acknowledgements

We gratefully acknowledge financial support provided by the S4CE collaborative project, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement number 764810. AG was supported by a postdoctoral fellowship from the S4CE project. We are grateful to Claire Geslin (UBO, BEEP) and Philippe Eliès (UBO, Plateforme d'imagerie et de mesures en microscopie (PIMM)) for their technical assistance in TEM imaging. We acknowledge Nadège Bienvenu (UBO, BEEP) for her help concerning the deposit of the strain at the UBOCC and Olivier Rouxel for his guidelines concerning the ferrozine assay protocol. We are grateful to Lukáš V. F. Novák for the use of HMM profiles in the search for certain genes. We thank Ranime Safieddine and Léna Aillot (UBO, BEEP) for their technical help concerning the culture of the bacterial strains. The LABGeM (CEA/Genoscope and CNRS UMR8030), the France Génomique and French Bioinformatics Institute national infrastructures (funded as part of Investissement 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.

Declaration of competing interest

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

Appendix A. Supplementary data

References

Abraham, W. R., Macedo, A. J., Lunsdorf, H., Fischer, R., Pawelczyk, S., Smit, J., & Vancanneyt, M. (2008). Phylogeny by a polyphasic approach of the order *Caulobacterales*, proposal of *Caulobacter mirabilis* sp. nov., *Phenylobacterium haematophilum* sp. nov. and *Phenylobacterium conjunctum* sp. nov., and emendation of the genus Phenylobacterium. *IJSEM*, *58*(8), 1939-1949. https://doi.org/10.1099/ijs.0.65567-0

Alain, K., Marteinsson, V. T., Miroshnichenko, M. L., Bonch-Osmolovskaya, E. A., Prieur, D., & Birrien, J. L. (2002). *Marinitoga piezophila* sp. nov., a rod-shaped, thermo-piezophilic bacterium isolated under high hydrostatic pressure from a deep-sea hydrothermal vent. https://doi.org/10.1099/00207713-52-4-1331

Alfredsson, H. A., Hardarson, B. S., Franzson, H., & Gíslason, S. R. (2008). CO_2 sequestration in basaltic rock at the Hellisheidi site in SW Iceland: stratigraphy and chemical composition of the rocks at the injection site. Mineralogical Magazine, 72(1), 1-5. https://doi.org/10.1180/minmag.2008.072.1.1

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.

Appia-Ayme, C., Bengrine, A., Cavazza, C., Giudici-Orticoni, M. T., Bruschi, M., Chippaux, M., & Bonnefoy, V. (1998). Characterization and expression of the co-transcribed *cyc1* and *cyc2* genes encoding the cytochrome *c4* (*c552*) and a high-molecular-mass cytochrome *c* from *Thiobacillus ferrooxidans* ATCC 33020. *FEMS microbiology letters*, 167(2), 171-177. https://doi.org/10.1111/j.1574-6968.1998.tb13224.x

769 Aradóttir, E. S., Sigurdardóttir, H., Sigfússon, B., & Gunnlaugsson, E. (2011). CarbFix:
770 a CCS pilot project imitating and accelerating natural CO_2 sequestration. Greenhouse
771 Gases: Science and Technology, 1(2), 105-118. https://doi.org/10.1002/ghg.18

772 Aradóttir, E. S., Gunnarsson, I., Sigfússon, B., Gunnarsson, G., Júliusson, B. M.,
773 Gunnlaugsson, E., ... & Sonnenthal, E. (2015). Toward cleaner geothermal energy
774 utilization: capturing and sequestering CO_2 and H_2S emissions from geothermal power
775 plants. *Transport in Porous Media*, *108*(1), 61-84. https://doi.org/10.1007/s11242-014-
776 0316-5

777 Charbonnier, F.; Forterre, P.; Erauso, G.; Prieur, D. Purification of plasmids from
778 thermophilic and hyperthermophilic archaebacteria. In Archaea: A Laboratory Manual;
779 Robb, F.T., Place, A.R., DasSarma, S., Schreier, H.J., Fleischmann, E.M., Eds.; Cold
780 Spring Harbor Laboratory Press: Woodbury, NY, USA, 1995; pp. 87–90.

781 Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., ... & Gascuel,
782 O. (2008). Phylogeny. fr: robust phylogenetic analysis for the non-specialist. *Nucleic*
783 *acids research*, *36*(suppl_2), W465-W469. https://doi.org/10.1093/nar/gkn180

784 Eddy, S. (1992). HMMER user's guide. *Department of Genetics, Washington University*
785 *School of Medicine*, *2*(1), 13.

786 Emerson, D., & Floyd, M. M. (2005). Enrichment and isolation of iron-oxidizing bacteria
787 at neutral pH. *Methods in enzymology*, *397*, 112-123. https://doi.org/10.1016/S0076-
788 6879(05)97006-7

789 Eren, A. M., Kiefl, E., Shaiber, A., Veseli, I., Miller, S. E., Schechter, M. S., ... & Willis,
790 A. D. (2021). Community-led, integrated, reproducible multi-omics with anvi'o. *Nature*
791 *microbiology*, *6*(1), 3-6. https://doi.org/10.1038/s41564-020-00834-3

792 Gangiredla, J., Rand, H., Benisatto, D., Payne, J., Strittmatter, C., Sanders, J., ... &
793 Strain, E. (2021). GalaxyTrakr: a distributed analysis tool for public health whole
794 genome sequence data accessible to non-bioinformaticians. *BMC genomics*, 22, 1-11.
795 https://doi.org/10.1186/s12864-021-07405-8

796 Garber, A. I., Nealson, K. H., Okamoto, A., McAllister, S. M., Chan, C. S., Barco, R. A.,
797 & Merino, N. (2020). FeGenie: a comprehensive tool for the identification of iron genes
798 and iron gene neighborhoods in genome and metagenome assemblies. Frontiers in
799 Microbiology, 11, 499513. https://doi.org/10.3389/fmicb.2020.00037

800 Gouy, M., Guindon, S., & Gascuel, O. (2010). SeaView version 4: a multiplatform
801 graphical user interface for sequence alignment and phylogenetic tree building.
802 *Molecular biology and evolution*, *27*(2), 221-224.
803 https://doi.org/10.1093/molbev/msp259

804 Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O.
805 (2010). New algorithms and methods to estimate maximum-likelihood phylogenies:
806 assessing the performance of PhyML 3.0. *Systematic biology*, *59*(3), 307-321.
807 https://doi.org/10.1093/sysbio/syq010

808 Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment
809 tool for genome assemblies. Bioinformatics, 29(8), 1072–1075.
810 https://doi.org/10.1093/bioinformatics/btt086

Hernández-Plaza, A., Szklarczyk, D., Botas, J., Cantalapiedra, C. P., Giner-Lamia, J., Mende, D. R., ... & Huerta-Cepas, J. (2023). eggNOG 6.0: enabling comparative genomics across 12 535 organisms. *Nucleic Acids Research*, *51*(D1), D389-D394. https://doi.org/10.1093/nar/gkac1022

Hobmeier, K., Goëss, M. C., Sehr, C., Schwaminger, S., Berensmeier, S., Kremling, A., ... & Marin-Sanguino, A. (2020). Anaplerotic pathways in Halomonas elongata: the role of the sodium gradient. Frontiers in Microbiology, 11, 561800. https://doi.org/10.3389/fmicb.2020.561800

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., & Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature communications*, *9*(1), 5114. https://doi.org/10.1038/s41467-018-07641-9

Jo, J. H., Choi, G. M., Lee, S. Y., & Im, W. T. (2016). *Phenylobacterium aquaticum* sp. nov., isolated from the reservoir of a water purifier. *IJSEM*, *66*(9), 3519-3523. https://doi.org/10.1099/ijsem.0.001223

Jünemann, S. (1997). Cytochrome bd terminal oxidase. Biochimica et Biophysica Acta (BBA)-Bioenergetics, 1321(2), 107-127. https://doi.org/10.1016/S0005-2728(97)00046-7

Khan, I. U., Habib, N., Xiao, M., Huang, X., Khan, N. U., Im, W. T., ... & Li, W. J. (2018). *Phenylobacterium terrae* sp. nov., isolated from a soil sample of Khyber-Pakhtun-Khwa,

832 Pakistan. *Antonie van Leeuwenhoek*, *111*, 1767-1775. https://doi.org/10.1007/s10482-
833 018-1064-2

834 Kimura M. The Neutral Theory of Molecular Evolution. Cambridge: Cambridge
835 University Press; 1983.

836 Kuykendall, L. D., Roy, M. A., O'Neill, J. J., Devine, T. E. (1988). Fatty acids, antibiotic
837 resistance, and deoxyribonucleic acid homology groups of *Bradyrhizobium japonicum*.
838 *Int. J. Syst. Bact.* 1988, *38*, 358–361. https://doi.org/10.1099/00207713-38-4-358

839 Lingens, F., Blecher, R., Blecher, H., Blobel, F., Eberspächer, J., Fröhner, C., ... & Layh,
840 G. (1985). *Phenylobacterium immobile* gen. nov., sp. nov., a Gram-negative bacterium
841 that degrades the herbicide chloridazon. *IJSEM*, *35*(1), 26-39.
842 https://doi.org/10.1099/00207713-35-1-26

843 Matter, J. M., Broecker, W. S., Gislason, S. R., Gunnlaugsson, E., Oelkers, E. H., Stute,
844 M., ... & Wolff-Boenisch, D. (2011). The CarbFix Pilot Project–storing carbon dioxide in
845 basalt. *Energy Procedia*, *4*, 5579-5585. https://doi.org/10.1016/j.egypro.2011.02.546

846 McAllister, S. M., Polson, S. W., Butterfield, D. A., Glazer, B. T., Sylvan, J. B., & Chan,
847 C. S. (2020). Validating the Cyc2 neutrophilic iron oxidation pathway using meta-omics
848 of Zetaproteobacteria iron mats at marine hydrothermal vents. Msystems, 5(1), 10-
849 1128. https://doi.org/10.1128/msystems.00553-19

850 Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., & Göker, M. (2013). Genome sequence-
851 based species delimitation with confidence intervals and improved distance functions.
852 *BMC bioinformatics*, *14*, 1-14. https://doi.org/10.1186/1471-2105-14-60

853 Mieszkin, S., Pouder, E., Uroz, S., Simon-Colin, C., & Alain, K. (2021). *Acidisoma silvae*
854 sp. nov. and *Acidisoma cellulosilytica* sp. nov., two acidophilic bacteria isolated from
855 decaying wood, hydrolyzing cellulose and producing poly-3-hydroxybutyrate. 2053.
856 https://doi.org/10.3390/microorganisms9102053

857 Mu, A., & Moreau, J. W. (2015). The geomicrobiology of CO2 geosequestration: a
858 focused review on prokaryotic community responses to field-scale CO2 injection.
859 *Frontiers in microbiology*, *6*, 263. https://doi.org/10.3389/fmicb.2015.00263

860 O'Mullan, G., Dueker, M. E., Clauson, K., Yang, Q., Umemoto, K., Zakharova, N., ... &
861 Goldberg, D. (2015). Microbial stimulation and succession following a test well injection
862 simulating CO_2 leakage into a shallow Newark Basin aquifer. *PLoS One*, *10*(1),
863 e0117812. https://doi.org/10.1371/journal.pone.0121932

864 Oh, Y. S., & Roh, D. H. (2012). *Phenylobacterium muchangponense* sp. nov., isolated
865 from beach soil, and emended description of the genus Phenylobacterium. *International*
866 *journal of systematic and evolutionary microbiology*, *62*(Pt_4), 977-983.
867 https://doi.org/10.1099/ijs.0.028902-0

868 Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015).
869 CheckM: assessing the quality of microbial genomes recovered from isolates, single
870 cells, and metagenomes. *Genome research*, *25*(7), 1043-1055.
871 http://www.genome.org/cgi/doi/10.1101/gr.186072.114

872 Peng, Z., Liu, Z., Jiang, Y., Dong, Y., & Shi, L. (2022). In vivo interactions between *Cyc2*
873 and *Rus* as well as *Rus* and *Cyc1* of *Acidithiobacillus ferrooxidans* during extracellular
874 oxidization of ferrous iron. *International Biodeterioration & Biodegradation*, *173*, 105453.
875 https://doi.org/10.1016/j.ibiod.2022.105453

876 Pitcher, R. S., & Watmough, N. J. The bacterial cytochrome cbb 3 oxidases. *Biochim.*
877 *Biophys. Acta* 1655, 388–399 (2004). https://doi.org/10.1016/j.bbabio.2003.09.017

878 Price, M. N., Dehal, P. S., & Arkin, A. P. (2009). FastTree: computing large minimum
879 evolution trees with profiles instead of a distance matrix. *Molecular biology and*
880 *evolution*, *26*(7), 1641-1650. https://academic.oup.com/mbe/article/26/7/1641/1128976

881 Reasoner, D. J., & Geldreich, E. (1985). A new medium for the enumeration and
882 subculture of bacteria from potable water. *Applied and environmental microbiology*,
883 *49*(1), 1-7. https://doi.org/10.1128/aem.49.1.1-7.1985

884 Richter & Rossello-Mora (2009). Shifting the genomic gold standard for the prokaryotic
885 species definition. PNAS 106:19126-19131. https://doi.org/10.1073/pnas.0906412106

886 Rouxel, O., Toner, B., Germain, Y., & Glazer, B. (2018). Geochemical and iron isotopic
887 insights into hydrothermal iron oxyhydroxide deposit formation at Loihi Seamount.
888 *Geochimica et Cosmochimica Acta*, *220*, 449-482.
889 https://doi.org/10.1016/j.gca.2017.09.050

890 Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing
891 phylogenetic trees. Mol Biol Evol 1987;4:406–425.
892 https://doi.org/10.1093/oxfordjournals.molbev.a040454

893 Seemann T, Booth T. BARNAP: BAsic Rapid Ribosomal RNA Predictor [Internet].
894 Berlin: GitHub, 2013.

895 Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. Bioinformatics,
896 30(14), 2068-2069. https://doi.org/10.1093/bioinformatics/btu153

Snæbjörnsdóttir, S.Ó., Sigfússon, B., Marieni, C. *et al.* Carbon dioxide storage through mineral carbonation. *Nat Rev Earth Environ* 1, 90–102 (2020). https://doi.org/10.1038/s43017-019-0011-8

Stackebrandt, E. (2006). Taxonomic parameters revisited: tarnished gold standards. *Microbial Today*, *33*, 152. https://api.semanticscholar.org/CorpusID:155392058

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., ... & Ostell, J. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic acids research*, *44*(14), 6614-6624 https://doi.org/10.1093/nar/gkw569

Thomas, F. A., Sinha, R. K., Hatha, A. M., & Krishnan, K. P. (2022). *Phenylobacterium glaciei* sp. nov., isolated from Vestrebroggerbreen, a valley glacier in Svalbard, Arctic. *IJSEM*, *72*(5), 005375. https://doi.org/10.1099/ijsem.0.005375

Tindall, B. J. (1990a). A comparative study of the lipid composition of *Halobacterium saccharovorum* from various sources. *Syst. Appl. Microbiol.* 1990, *13*, 128–130. https://doi.org/10.1016/S0723-2020(11)80158-X

Tindall, B.J. (1990b). Lipid composition of *Halobacterium lacus profundi*. *FEMS Microbiol. Lett.* 1990, *66*, 199–202. https://doi.org/10.1111/j.1574-6968.1990.tb03996.x

Trias, R., Ménez, B., le Campion, P., Zivanovic, Y., Lecourt, L., Lecoeuvre, A., ... & Gérard, E. (2017). High reactivity of deep biota under anthropogenic CO_2 injection into basalt. *Nature communications*, *8*(1), 1063. https://doi.org/10.1038/s41467-017-01288-8

Tsuge, T., Hyakutake, M., & Mizuno, K. (2015). Class IV polyhydroxyalkanoate (PHA) synthases and PHA-producing Bacillus. *Applied microbiology and biotechnology*, *99*, 6231-6240. https://doi.org/10.1007/s00253-015-6777-9

Vallenet, D., Calteau, A., Dubois, M., Amours, P., Bazin, A., Beuvin, M., ... & Médigue, C. (2020). MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis. *Nucleic Acids Research*, *48*(D1), D579-D589. https://doi.org/10.1093/nar/gkz926

Viollier, E., Inglett, P. W., Hunter, K., Roychoudhury, A. N., & Van Cappellen, P. (2000). The ferrozine method revisited: Fe (II)/Fe (III) determination in natural waters. *Applied geochemistry*, *15*(6), 785-790. https://doi.org/10.1016/S0883-2927(99)00097-9

Wayne, L. G. (1987). Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int. J. Syst. Bacteriol*, *37*, 463464. https://doi.org/10.1099/00207713-37-4-463

Wee, S. K., & Yap, E. P. H. (2021). GALAXY Workflow for Bacterial Next-Generation Sequencing De Novo Assembly and Annotation. *Current Protocols*, *1*(9), e242. https://doi.org/10.1002/cpz1.242

Yoon, S. H., Ha, S. M., Kwon, S., Lim, J., Kim, Y., Seo, H., & Chun, J. (2017). Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *IJSEM*, *67*(5), 1613. https://doi.org/10.1099/ijsem.0.001755

List of figures:

Fig. 1. Transmission electron microscopy microphotographs of cells of strain HK31-G^T grown on R2A at 25 °C for 3 days. (a) Some cells harbor a cellular appendage, so-called prostheca or stalk (indicated by black arrow). (b) Intracellular granules. (c) A cell with its polar flagellum. (d) Enlargement of image (c).

Fig. 2. Cell concentrations of strain HK31-G^T and Fe(II) concentrations at T0 and T6 days in Fe(II)-oxidizing medium (cultures were performed in triplicate, negative controls in duplicate). Cellular concentrations are represented by full black bars. Fe(II) concentrations in the Fe(II)-oxidizing medium inoculated with cells of HK31-G^T are represented by striped bars while Fe(II) concentrations in negative controls (without inoculation of cells) are represented by dotted bars. A Mann-Whitman test was applied to show any significant difference of Fe(II) concentrations between T0 and T6 (*P<0.001*).

Fig. 3. **Phylogenomic tree showing the phylogenomic position of strain HK31-G^T with respect to *Phenylobacterium*, *Caulobacter* and *Brevundimonas* species having sequenced genomes, and the six closely related MAGs belonging to the species *Phenyloabcterium* sp030693625 (GTDB).** The genome sequence of *Sphingomonas paucimobilis* ZJSH1 was used as an outgroup. The tree was built using FastTree implemented in anvi'o and using the maximum-likelihood method. Bootstrap values below 0.5 are not shown. Bar, 0.07 substitutions per position.

Fig. 4 Heatmap showing FastANI scores between genomes and MAGs. FastANI values below 77.5% are not shown.

Fig. 5. Overview of the dissimilatory nitrate reduction pathway in strain HK31-G^T operating under anaerobic conditions. Black arrows represent the proton translocation and red arrows represent the electron fluxes in protein complexes. Here, NarI is used to oxidize the ubiquinol (UQH_2) into ubiquinone (UQ^+), then two electrons are transferred to NarG through NarH. The *bis*-molybdopterin guanine dinucleotide (MGD) allows the final nitrate reduction.

Supplementary data

Phenylobacterium ferrooxidans sp. nov., isolated from a sub-surface geothermal aquifer in Iceland

Eva Pouder, Erwann Vince, Karen Jacquot, Maimouna batoma Traoré, Ashley Grosche, Maria Ludwig, Mohamed Jebbar, Lois Maignien, Karine Alain and Sophie Mieszkin*

List of figures:

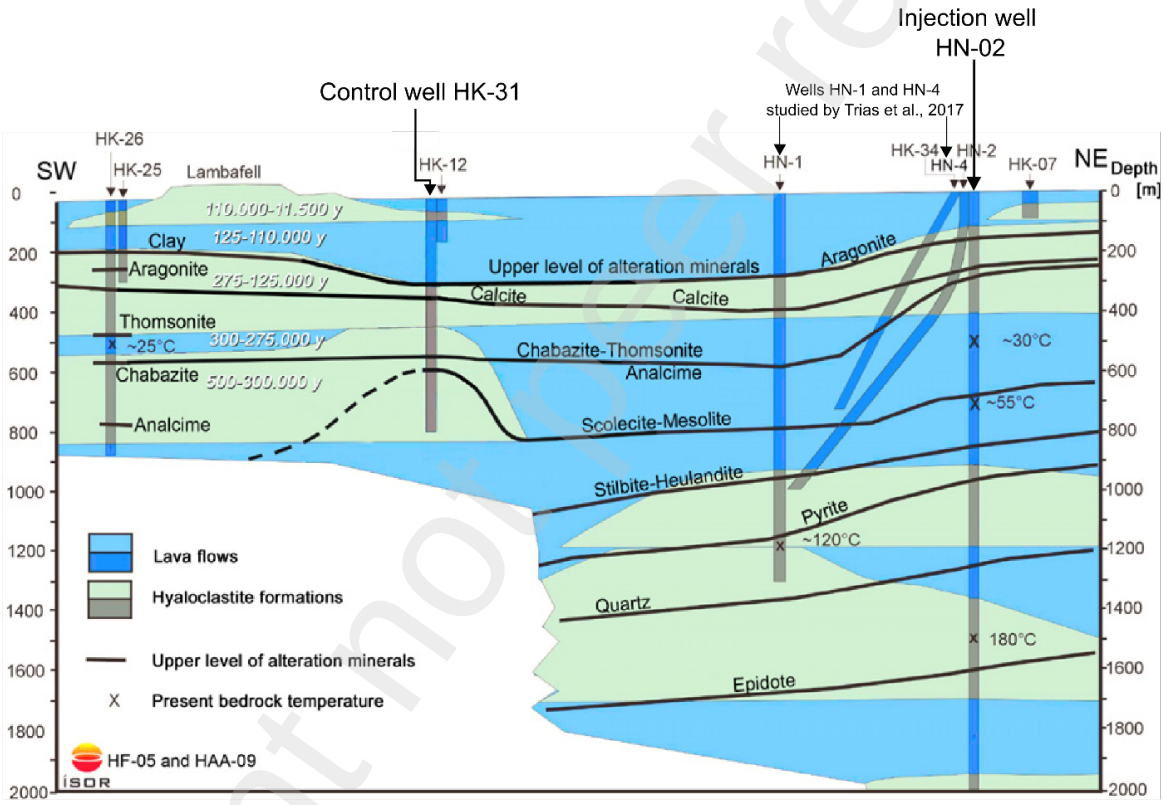

Fig. S1: North-South geological section at the CarFix-1 site showing the position of the control well HK-31 targeting in this study and injection wells studied by Trias et al., (2017) (Modified from Matter et al., 2011).

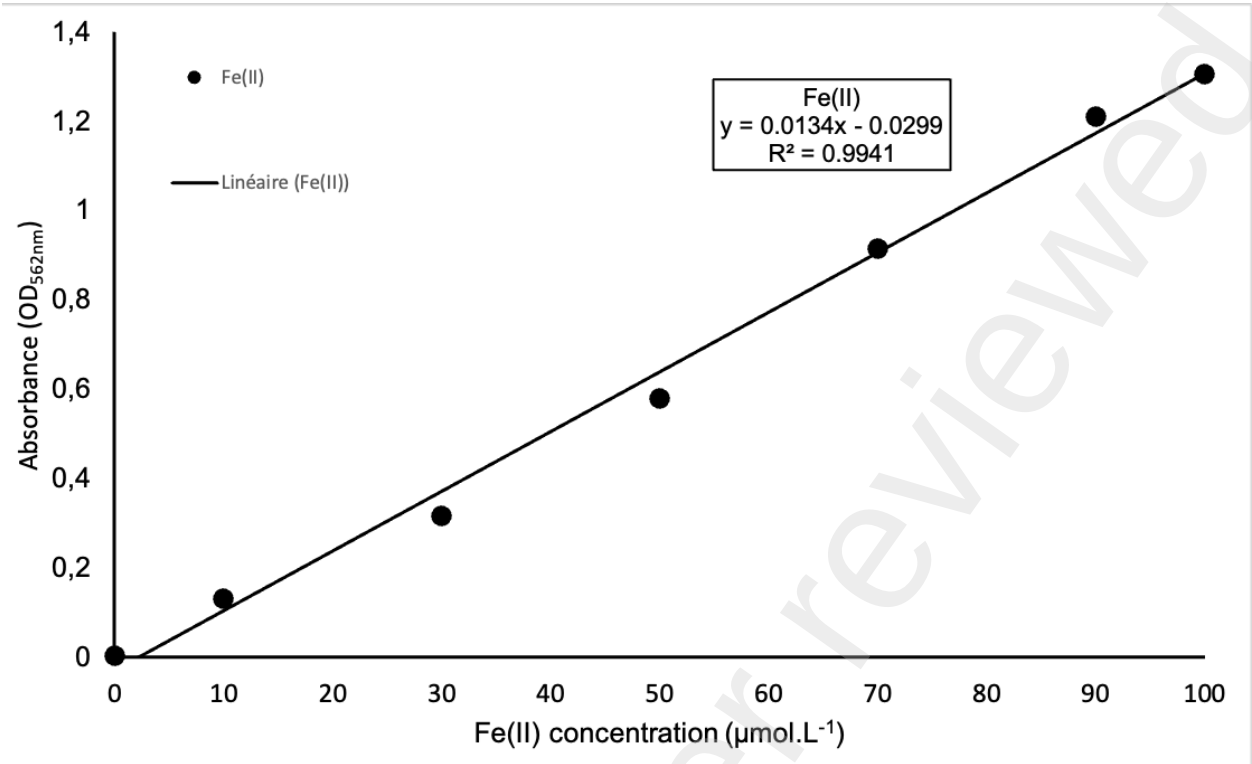

Fig. S2. Standard curves used for the determination of Fe(II) concentration in cultures.

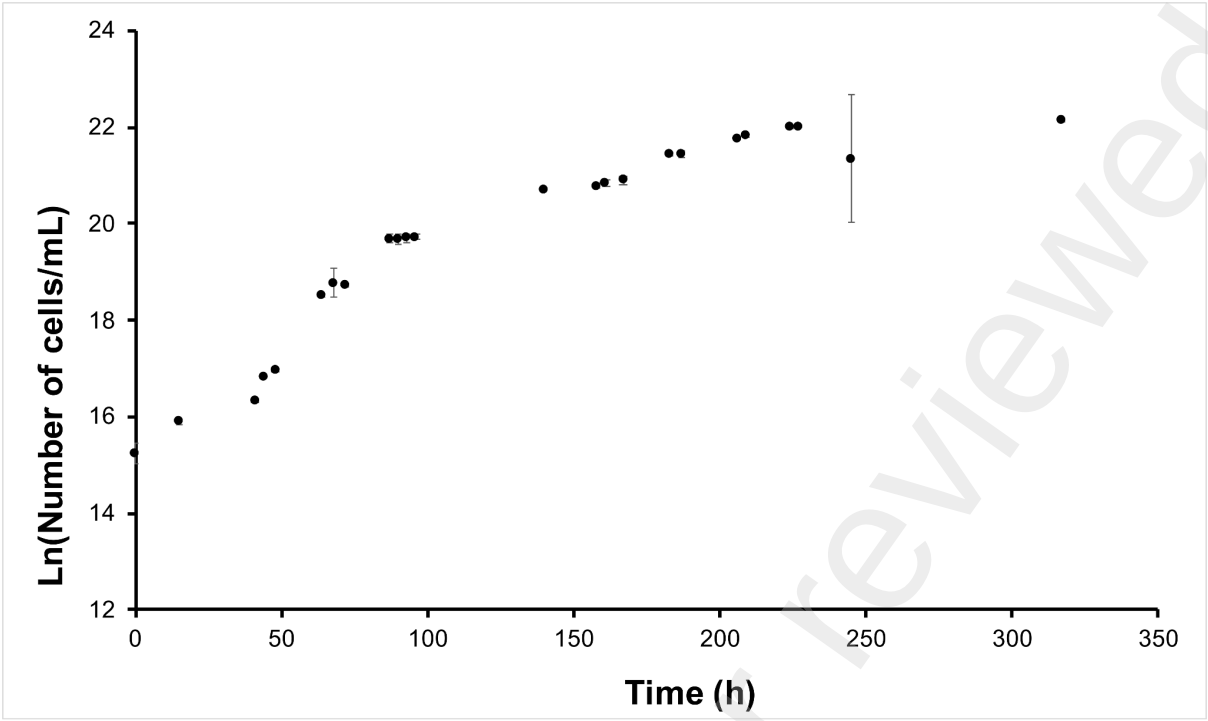

Fig. S3. Growth kinetics of strain HK31-G^T under optimal growth conditions. Each value is the mean of three independent replicates. Standard deviation for each mean value is indicated on the graph but is too small to be visible at most points.

PG = Phosphatidylglycerol

GL = Glycolipid

GPL = Glycophospholipid

PL = Phospholipid

L = Lipid

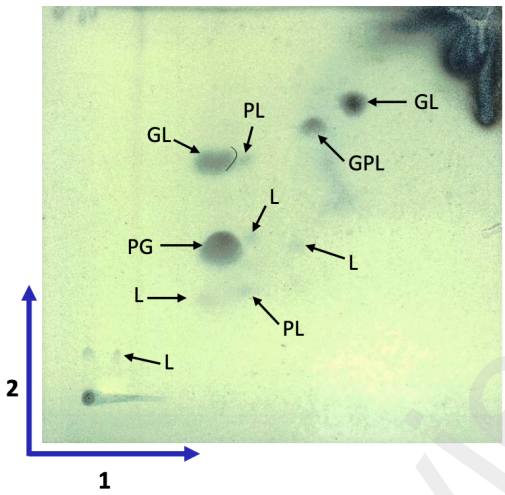

Fig. S4. Two-dimensional chromatogram of polar lipids of strain HK31-G^T.

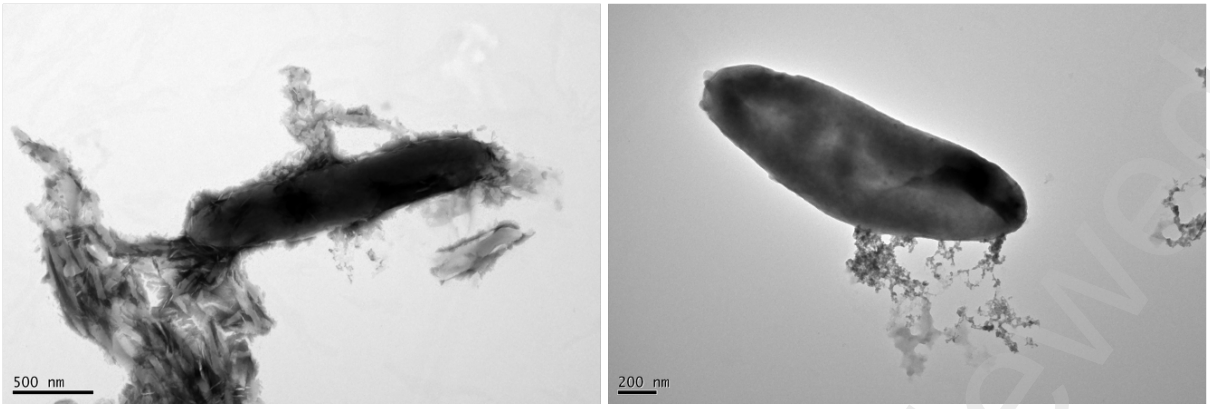

Figure S5. Transmission electron microscopy microphotographs of cells of strain HK31-G^T grown on iron-oxidizing medium at 25 °C under microaerophilic conditions for 5 days, showing potential cell-mineral interactions.

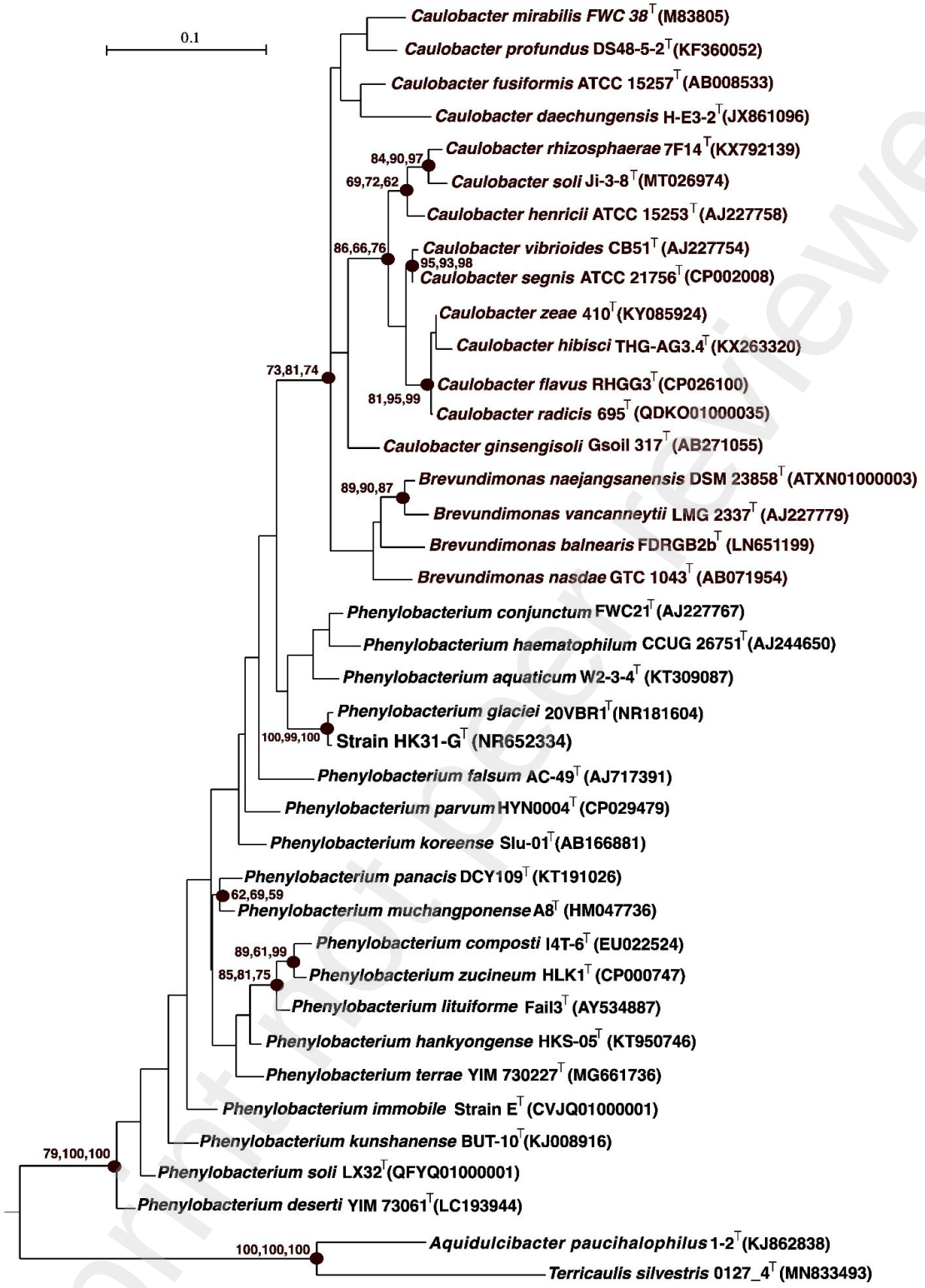

Fig. S6. Maximum-likelihood (ML) phylogenetic tree reconstructed from a comparative analysis of 16S rRNA gene sequences showing the relationships of strain HK31-G^T with the type strains of related species. Filled circles indicate that the corresponding nodes were also recovered using the maximum-parsimony (MP) and neighbor-joining (NJ) algorithms. Numbers at branch nodes indicate bootstrap values (%) as calculated by ML/MP/NJ algorithms. Only values greater than 50% are shown. *Sphingomonas paucimobilis* DSM 1098^T was used as an outgroup. Bar, 0.01 substitutions per nucleotide position.

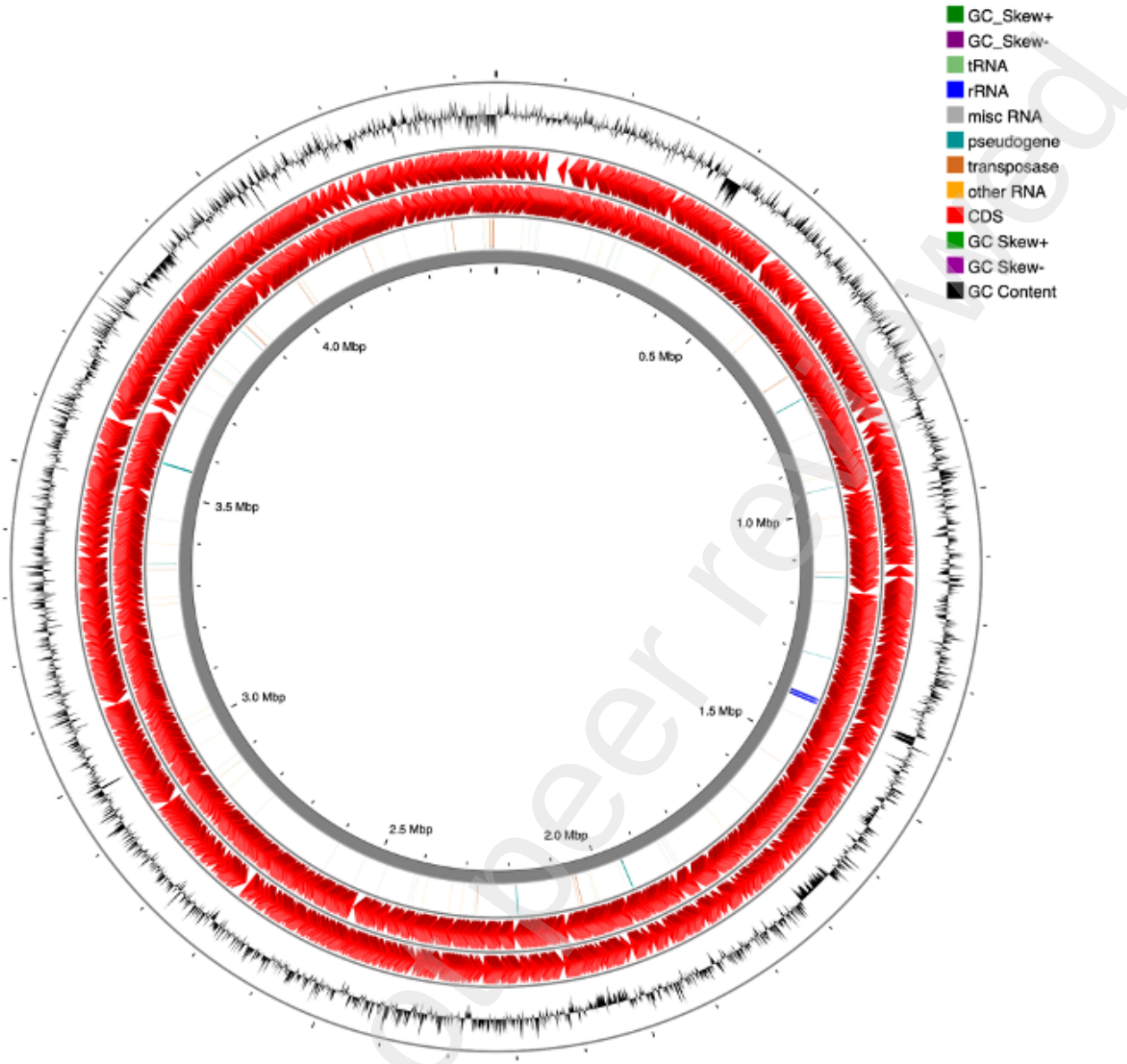

Fig. S7. Circular mapping of the genome of strain HK31-G^T (*Phenylobacterium ferrooxidans*) obtained from the Circular Genome Viewer of the MaGe platform.

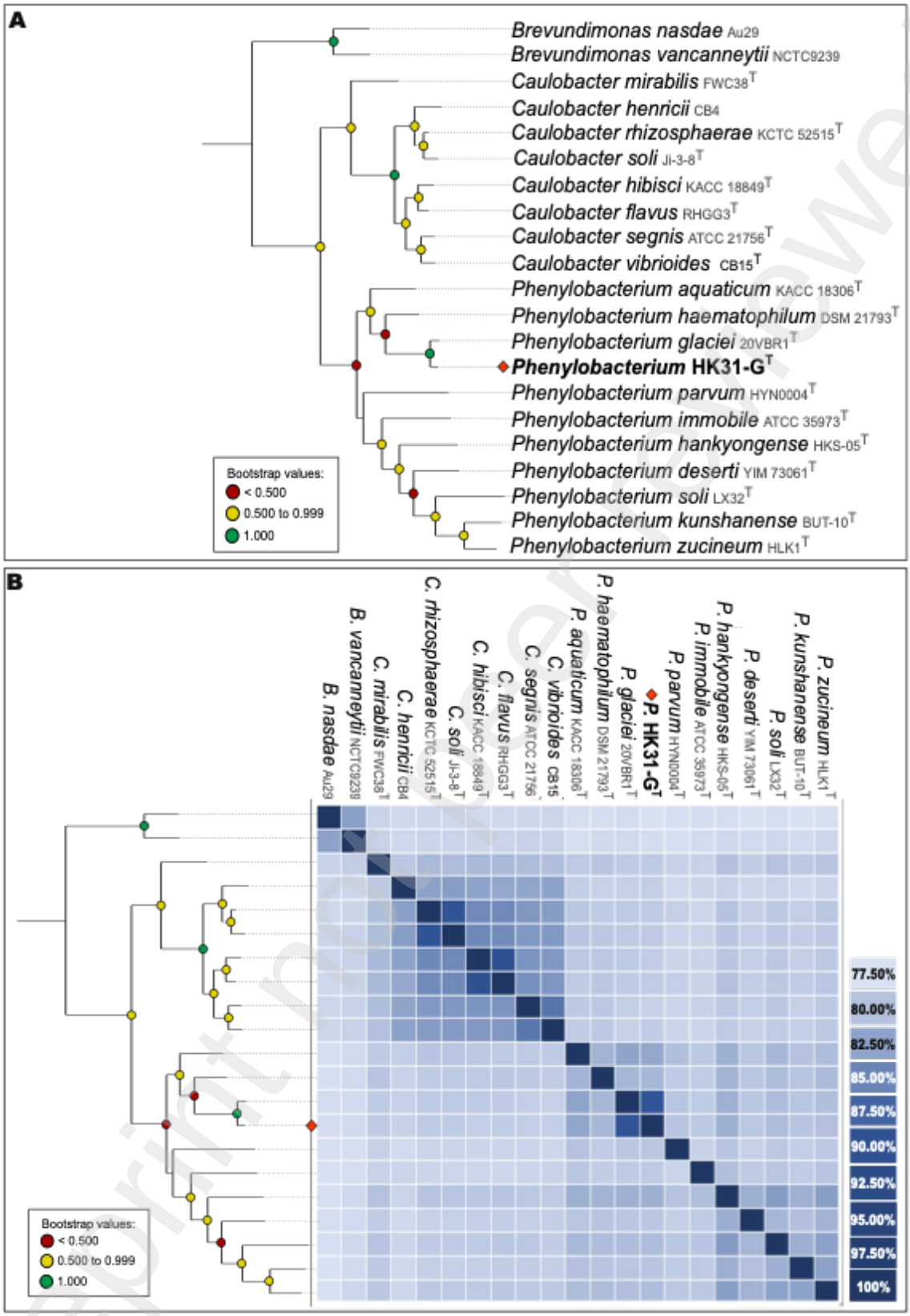

Fig. S8. Phylogenomic tree built with anvi'o 7.1 from a comparative analysis of genomes showing the relationships of strain HK31-G^T with the reference strains of closely related species and (b) heatmap showing FastANI scores between genomes. Bootstrap scores are represented by color circles (below 0.5 in red, between 0.5 and 0.999 in yellow and equal to 1 in green). The genome sequence of *Sphingomonas paucimobilis* ZJSH1 was used as an outgroup. In the heatmap, all FastANI values were greater than 77.5%.

List of tables:

Table S1. Marker nucleotides in the 16S rDNA genes of strain HK31-G^T and closest relatives. All species are lacking nucleotides 73–89, 199–212, 453–477 and 829–832 (*Escherichia coli* str. K12 subtr. MG1655 numbering). Species: 1, HK31-G^T (data from this study); 2, *P. glaciei* 20VBR1^T; 3, *P. aquaticum* W2-3-4^T; 4, *P. haematophilum* CCUG 26751^T; 5, *P. conjunctum* FWC 21^T; 6, *P. parvum* HYN0004^T; 7, *P. koreense* Slu-01^T.

***E. coli* no.**	***Phenylobacterium* genus** (Abraham *et al.,* 2008)	**1**	**2**	**3**	**4**	**5**	**6**	**7**
122	G	G	G	G	G	G	G	G
178	T (*P. falsum*: C)	T	T	T	T	T	T	T
359	A	A	A	A	A	A	A	A
610	G	G	G	G	G	G	G	G
639	G	G	G	G	G	G	G	G
823	G	G	G	G	G	G	G	G
877	C	C	C	C	C	C	C	C
1145	C	C	C	C	C	C	C	C
1265	T (*P. falsum*: A)	**A**	**A**	T	T	T	T	T
1270	A (*P. falsum*: T)	**T**	**T**	A	A	A	A	A

Table S2. Classification of the coding DNA sequences (CDS) of strain HK31-G^T in clusters of orthologous groups (COG) categories.

Process	Class ID	Description	CDS (nb)	CDS (%)
CELLULAR PROCESSES AND SIGNALING	D	Cell cycle control, cell division, chromosome partitioning	29	0.6203 %
	M	Cell wall/membrane/envelope biogenesis	196	4.1925 %
	N	Cell motility	51	1.0909 %
	O	Post-translational modification, protein turnover, chaperones	161	3.4439 %
	T	Signal transduction mechanisms	204	4.3636 %
	U	Intracellular trafficking, secretion, and vesicular transport	89	1.9037 %
	V	Defense mechanisms	65	1.3904 %
	Z	Cytoskeleton	1	0.0214 %
INFORMATION STORAGE AND PROCESSING	B	Chromatin structure and dynamics	1	0.0214 %
	J	Translation, ribosomal structure and biogenesis	181	3.8717 %
	K	Transcription	237	5.0695 %
	L	Replication, recombination and repair	208	4.4492 %
METABOLISM	C	Energy production and conversion	211	4.5134 %
	E	Amino acid transport and metabolism	214	4.5775 %
	F	Nucleotide transport and metabolism	70	1.4973 %
	G	Carbohydrate transport and metabolism	127	2.7166 %
	H	Coenzyme transport and metabolism	106	2.2674 %
	I	Lipid transport and metabolism	179	3.8289 %
	P	Inorganic ion transport and metabolism	229	4.8984 %
	Q	Secondary metabolites biosynthesis, transport and catabolism	129	2.7594 %
POORLY CHARACTERIZED	S	Function unknown	1228	26.2674 %

Table S3. Summary of Average Nucleotide Identity Scores, calculated between the genome of HK31-GT and the 262 fasta sequences affiliated to *Phenylobacterium* sp. available on NCBI database. - indicates that the score is < 77.5%. In bold, the accession number corresponding to MAGs having an ANI score above the threshold for species delineation

Accession number	FastANI value	Accession number	FastANI value	Accession number	FastANI value	Accession number	FastANI value
GCA_000017265.1	80.4135	GCA_004799515.1	80.3851	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097
GCA_001375595.1	79.3637	GCA_004799545.1	80.3999	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097
GCA_001425305.1	81.1073	GCA_013822795.1	80.9567	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097
GCA_001425915.1	81.1474	GCA_013911965.1	80.8698	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097
GCA_001428705.1	81.1373	GCA_014196295.1	80.9255	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097
GCA_001429025.1	82.0695	GCA_014359675.1	80.901	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097
GCA_001557235.1	80.4082	GCA_016124325.1	79.6619	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097
GCA_001557375.1	80.2531	GCA_016463085.1	79.772	GCA_018001015.1	89.5097	GCA_020402185.1	79.3394
GCA_001724605.1	79.4629	GCA_016772415.2	89.1632	GCA_018001015.1	89.5097	GCA_020402205.1	79.4272
GCA_001724985.1	81.73	GCA_016793225.1	79.1602	GCA_018001015.1	89.5097	GCA_020402225.1	79.3147
GCA_001824475.1	80.1946	GCA_016793285.1	79.6804	GCA_018001015.1	89.5097	GCA_020402245.1	79.4278
GCA_001825585.1	80.7296	GCA_017989235.1	88.2551	GCA_018001015.1	89.5097	GCA_020402265.1	79.3512
GCA_002221445.1	77.6254	GCA_017991675.1	81.2341	GCA_018001015.1	89.5097	GCA_020402285.1	79.4699
GCA_002693985.1	80.7369	GCA_017999155.1	88.7545	GCA_018001015.1	89.5097	GCA_020402305.1	79.4562
GCA_003136395.1	79.842	GCA_017999535.1	81.3524	GCA_018001015.1	89.5097	GCA_020402325.1	79.4838
GCA_003150835.1	79.6819	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097	GCA_020402345.1	79.439
GCA_003243355.1	79.4665	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097	GCA_020402365.1	79.2848
GCA_003254475.1	80.4205	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097	GCA_020402385.1	79.6131
GCA_003254505.1	81.3997	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097	GCA_020402405.1	79.5239
GCA_003254525.1	79.7363	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097	GCA_020402425.1	79.5426
GCA_003254705.1	79.5696	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097	GCA_020402435.1	79.4484
GCA_004297125.1	80.201	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097	GCA_020402445.1	79.4695
GCA_004299445.1	79.7903	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097	GCA_020402485.1	79.4446
GCA_004799395.1	79.8199	GCA_018001015.1	89.5097	GCA_018001015.1	89.5097	GCA_020402505.1	79.4965

Accession number	FastANI value	Accession number	FastANI value	Accession number	FastANI value	Accession number	FastANI value
GCA_020402525.1	79.6275	GCA_025350885.1	79.1024	GCA_030651185.1	90.0682	GCA_028291005.1	80.2739
GCA_020402545.1	79.4142	GCA_025351295.1	80.1573	GCA_030651755.1	80.3873	GCA_028291025.1	80.0331
GCA_020402565.1	79.4474	GCA_025925325.1	78.7597	GCA_030652015.1	97.0203	GCA_028291065.1	79.099
GCA_020402585.1	79.4551	GCA_025934275.1	79.0816	GCA_030654475.1	95.9416	GCA_028291215.1	80.2271
GCA_020402595.1	79.0445	GCA_026126275.1	79.14	GCA_030654915.1	80.9729	GCA_028291285.1	79.4742
GCA_020402625.1	79.4399	GCA_026411515.1	81.3651	GCA_030679875.1	80.5322	GCA_028291325.1	80.4289
GCA_020402645.1	79.4158	GCA_027286305.1	81.3681	GCA_030680185.1	80.6624	GCA_028291375.1	79.9727
GCA_020402665.1	79.4384	GCA_028290885.1	80.1266	GCA_030680595.1	79.7207	GCA_028698405.1	80.2998
GCA_020402685.1	79.4321	GCA_028290925.1	80.2246	GCA_030681155.1	80.6874	GCA_029256285.1	80.7689
GCA_020402695.1	79.411	GCA_028290945.1	80.2467	GCA_030682345.1	80.427	GCA_029977805.1	80.0858
GCA_020402725.1	79.4485	GCA_028290955.1	80.1925	GCA_030683235.1	80.7165	GCA_030148385.1	80.9119
GCA_020402745.1	79.4854	GCA_028290985.1	79.0543	GCA_030683775.1	97.1291	GCA_030645195.1	88.5519
GCA_020402755.1	79.3879	GCA_028291005.1	80.2739	GCA_024699945.1	79.769	**GCA_030645635.1**	**97.9033**
GCA_020402775.1	79.3862	GCA_028291025.1	80.0331	GCA_024699985.1	80.4964	GCA_030646615.1	80.9366
GCA_020402805.1	79.3289	GCA_028291065.1	79.099	GCA_025350885.1	79.1024	GCA_030651185.1	90.0682
GCA_020402825.1	79.3581	GCA_028291215.1	80.2271	GCA_025351295.1	80.1573	GCA_030651755.1	80.3873
GCA_020402835.1	79.4858	GCA_028291285.1	79.4742	GCA_025925325.1	78.7597	**GCA_030652015.1**	**97.0203**
GCA_021298495.1	79.5416	GCA_028291325.1	80.4289	GCA_025934275.1	79.0816	**GCA_030654475.1**	**95.9416**
GCA_021300115.1	79.2871	GCA_028291375.1	79.9727	GCA_026126275.1	79.14	GCA_030654915.1	80.9729
GCA_022402485.1	79.4452	GCA_028698405.1	80.2998	GCA_026411515.1	81.3651	GCA_030679875.1	80.5322
GCA_022695515.1	82.2104	GCA_029256285.1	80.7689	GCA_027286305.1	81.3681	GCA_030680185.1	80.6624
GCA_023260155.1	78.6622	GCA_029977805.1	80.0858	GCA_028290885.1	80.1266	GCA_030680595.1	79.7207
GCA_024298925.1	82.6141	GCA_030148385.1	80.9119	GCA_028290925.1	80.2246	GCA_030681155.1	80.6874
GCA_024508875.1	79.4173	GCA_030645195.1	88.5519	GCA_028290945.1	80.2467	GCA_030682345.1	80.427
GCA_024699945.1	79.769	GCA_030645635.1	97.9033	GCA_028290955.1	80.1925	GCA_030683235.1	80.7165
GCA_024699985.1	80.4964	GCA_030646615.1	80.9366	GCA_028290985.1	79.0543	**GCA_030683775.1**	**97.1291**

Accession number	FastANI value	Accession number	FastANI value	Accession number	FastANI value	Accession number	FastANI value
GCA_030693405.1	80.592	GCA_035324635.1	-	GCA_036383735.1	79.1561	GCA_040508585.1	79.1157
GCA_030693625.1	**97.0431**	GCA_035327745.1	-	GCA_036402295.1	79.6	GCA_040545335.1	80.9583
GCA_030693805.1	79.4087	GCA_035387405.1	81.4738	GCA_036403175.1	79.5834	GCA_902826855.1	79.3123
GCA_030694205.1	-	GCA_035423395.1	81.4056	GCA_036495085.1	78.6425	GCA_943327825.2	79.9248
GCA_030694925.1	79.8083	**GCA_035428425.1**	**96.3998**	GCA_036496685.1	79.4649	GCA_943328465.2	79.8183
GCA_030696555.1	80.1498	GCA_035431805.1	81.5166	GCA_036501465.1	79.9221	GCA_943328735.2	-
GCA_030696615.1	80.1592	GCA_035475775.1	79.4958	GCA_036502675.1	80.5061	GCA_943332615.1	79.828
GCA_030696765.1	**96.9398**	GCA_035539035.1	80.7029	GCA_036561045.1	79.4425	GCA_945859865.1	78.9728
GCA_030697305.1	79.5747	GCA_035539755.1	80.4084	GCA_036563285.1	80.1194	GCA_945874825.1	78.8177
GCA_030697925.1	82.1973	GCA_035569475.1	80.4032	GCA_036563385.1	79.9281	GCA_945883235.1	79.0263
GCA_030698145.1	82.3559	GCA_035627295.1	79.4153	GCA_036567285.1	81.1065	GCA_945907165.1	78.826
GCA_030704785.1	**97.1471**	GCA_035656755.1	79.9507	GCA_036567325.1	80.224	GCA_945952175.1	81.6429
GCA_031181185.1	79.7934	GCA_035656795.1	79.7229	GCA_036676115.1	82.5374	GCA_945952355.1	80.8263
GCA_031360265.1	79.6093	GCA_035694305.1	79.3646	GCA_036677185.1	82.4977	GCA_946222325.1	-
GCA_031412415.1	81.5827	GCA_035944335.1	80.5083	GCA_036723115.1	80.315	GCA_947371235.1	82.1591
GCA_031425575.1	93.2717	GCA_035997815.1	-	GCA_036788895.1	79.3609	GCA_947372525.1	82.0106
GCA_031429955.1	82.7383	GCA_036261715.1	-	GCA_036821595.1	80.6052	GCA_947375025.1	-
GCA_031984885.1	81.872	GCA_036262155.1	79.7713	GCA_036963605.1	80.637	GCA_947377555.1	82.0873
GCA_034004485.1	80.2883	GCA_036263695.1	79.567	GCA_037138395.1	78.8476	GCA_947377895.1	77.5927
GCA_034366425.1	86.4913	GCA_036264965.1	79.9383	GCA_037189685.1	79.55	GCA_947378305.1	77.5481
GCA_034376405.1	79.9212	GCA_036268175.1	80.0114	GCA_037968635.1	88.5732	GCA_947378715.1	82.097
GCA_034377185.1	79.0334	GCA_036268735.1	79.6197	GCA_038920875.1	79.9097	GCA_947379345.1	82.2768
GCA_034665885.1	79.8832	GCA_036270195.1	79.7641	GCA_038920955.1	78.8877	GCA_947379565.1	82.0526
GCA_035274295.1	79.9922	GCA_036273595.1	79.2811	GCA_039930425.1	79.4131	GCA_947380745.1	77.6672
GCA_035276905.1	80.008	GCA_036279815.1	82.3452	GCA_040391295.1	81.2432		
GCA_035309685.1	80.078	GCA_036383455.1	79.7484	GCA_040508045.1	80.1999		

Table S4. MAGs statistics and Overall Genome Relatedness Indices (OGRI) between the genome of strain HK31-G^T and the six MAGs affiliated to the species *Phenylobacterium* sp030693625. Species: 1, HK31-G^T (data from this study); 1a, GCA_030645635; 1b, GCA_030652015; 1c, GCA_030683775; 1d, GCA_030693625; 1e, GCA_030696765; 1f, GCA_030704785 (data from GTDB).

	1	1a	1b	1c	1d	1e	1f
Number of contigs	104	454	66	220	97	323	73
Size (Mbp)	4.46	3.75	4.29	4.01	4.12	4.05	4.35
Completeness	100	88.25	99.91	99.99	99.96	100	100
Contamination	1.41	3.69	1.64	3.29	1.37	4.81	8.67
G+C content (%)	67.95	68.20	68.06	68.06	68.13	68.18	68.17
Number of CDS	4667	4107	4271	4125	4149	4267	4395
tRNA	48	18	19	19	17	19	19
dDDH (%)	100	83.10	75.30	76.00	76,10	75,70	74,10
OrthoANIu (%)	100	98.12	97.20	97.28	97.22	97.11	97.15
FastANI (%)	100	97.90	97.02	97.13	97.04	96.94	97.14

Table S5. Table of locus-tags of genes of known function in strain HK31-G^T involved in nitrogen metabolism

PGAP: Product - gene	Locus_tag
Acetoacetyl-CoA reductase - *phaB*	OCL97_08250
Poly(3-hydroxyalkanoate) polymerase - *phaC*	OCL97_01260
Polyhydroxyalkanoate synthesis repressor - *phaR*	OCL97_08240
polyhydroxyalkanoic acid system family protein	OCL97_12980
Cytochrome *c* oxidase subunit 2 - *ctaC*	OCL97_16890
Cytochrome *c* oxidase subunit 1 - *ctaD*	OCL97_16895
Cytochrome *c* oxidase assembly protein - *CtaG*	OCL97_16910
Cytochrome *c* oxidase subunit 3 -*ctaE*	OCL97_16915
Nitrate reductase subunit alpha - *narG*	OLC97_05710
Nitrate reductase subunit beta - *narH*	OLC97_05715
Nitrate reductase subunit gamma - *narI*	OLC97_05720
Nitrate reductase molybdenum cofactor assembly chaperone - *narJ*	OLC97_05725
Family nitrate transporter - *narK*	OLC97_05700
NADH-quinone oxidoreductase subunit A - *nuoA*	OLC97_00445
NADH-quinone oxidoreductase subunit B - *nuoB*	OLC97_00450
NADH-quinone oxidoreductase subunit C - *nuoC*	OLC97_00455
NADH-quinone oxidoreductase subunit D - *nuoD*	OLC97_00460
NADH-quinone oxidoreductase subunit E - *nuoE*	OLC97_00470
NADH-quinone oxidoreductase subunit F - *nuoF*	OLC97_00480
NADH-quinone oxidoreductase subunit G - *nuoG*	OLC97_00485
NADH-quinone oxidoreductase subunit H - *nuoH*	OLC97_00490
NADH-quinone oxidoreductase subunit I - *nuoI*	OLC97_00500
NADH-quinone oxidoreductase subunit J - *nuiJ*	OLC97_00505
NADH-quinone oxidoreductase subunit K - *nuoK*	OLC97_00510
NADH-quinone oxidoreductase subunit L - *nuoL*	OLC97_00515
NADH-quinone oxidoreductase subunit M - *nuoM*	OLC97_00520
NADH-quinone oxidoreductase subunit N - *nuoN*	OLC97_00525

Table S6. List of genes involved in iron metabolism predicted by PGAP and FeGenie tools. The "-" indicates that the genes have not been detected thanks to the prediction tool.

Gene category	PGAP tag	PGAP gene product	FeGenie gene product	Prokka gene product
Iron aquisition - Iron(II)/(III) transport	OCL97_13350	dipeptide ABC transporter ATP-binding protein	FbpC-family-ATPase	Glutathione import ATP-binding protein GsiA
	OCL97_13360	ABC transporter ATP-binding protein	YfeB-family-membrane-proteins	Putative ABC transporter ATP-binding protein YxlF
	OCL97_15185	ferrous iron transporter B	FeoB-family-iron-transporter	Fe^{2+} transporter FeoB
	OCL97_15190	ferrous iron transport protein A	FeoA-family-iron-transporter	Hypothetical protein
	OCL97_15185	ferrous iron transporter B	FeoB-family-iron-transporter iron_aquisition-iron_transport	Fe^{2+} transporter FeoB
	OCL97_15190	ferrous iron transport protein A	FeoA-family-iron-transporter iron_aquisition-iron_transport	Hypothetical protein
	OCL97_07540	FTR1 family protein	-	Ferrous iron permease EfeU
	-	-	-	Ferrous iron permease EfeU
Iron aquisition - Siderophore transport	OCL97_09815	iron ABC transporter permease	FpvE-family-permease	Hemin transport system permease protein HmuU
	OCL97_09820	ABC transporter substrate-binding protein	HatD-family-substrate-binding-protein	Hypothetical protein
	OCL97_03805	efflux RND transporter periplasmic adaptor subunit	PvdR-family-siderophore-export	Multidrug resistance protein MdtA
	OCL97_03810	ABC transporter ATP-binding protein	PvdT-family-siderophore-export	Putative ABC transporter ATP-binding protein YknY
	OCL97_03815	ABC transporter permease	PvdT-family-siderophore-export	Macrolide export ATP-binding/permease protein MacB
	OCL97_05930	TonB-dependent receptor	PirA-family-siderophore-receptor	Vitamin B_{12} transporter BtuB
	-	-	LbtU-family-siderophore receptor	Vitamin B_{12} transporter BtuB
	OCL97_18855	TonB-dependent receptor	PirA-family-siderophore-receptor	Vitamin B_{12} - transporter BtuB
	OCL97_09385	TonB-dependent receptor	PirA-family-siderophore-receptor	Putative TonB-dependent receptor
	OCL97_09820	ABC transporter substrate-binding protein	HatD-family-substrate-binding-protein	Hypothetical protein
	OCL97_18855	TonB-dependent receptor	PirA-family-siderophore-receptor	Vitamin B_{12} - transporter BtuB
	OCL97_11255	protein TolQ	ExbB-family	Biopolymer transport protein ExbB
	OCL97_11260	ExbD/TolR family protein	ExbD-family	Tol-Pal system protein TolR
	-	-	ExbD-family	Biopolymer transport protein ExbD
	OCL97_11495	MotA/TolQ/ExbB proton channel family protein	ExbB-family	Tol-Pal system protein TolQ
	OCL97_11500	energy transducer TonB	TonB-family	Hypothetical protein
	-	-	TonB-family	Hypothetical protein
	OCL97_09390	MotA/TolQ/ExbB proton channel family protein	ExbB-family	Biopolymer transport protein ExbB
	OCL97_09395	biopolymer transporter ExbD	ExbD-family	Biopolymer transport protein ExbD

	OCL97_09400	energy transducer TonB	TonB-family	Hypothetical protein
	OCL97_00360	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	OCL97_02245	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	OCL97_02390	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	OCL97_02520	TonB-dependent receptor	-	Hypothetical protein
	OCL97_02730	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	OCL97_03605	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	-	-	-	Vitamin B_{12} transporter BtuB
	OCL97_05495	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	-	-	-	Vitamin B_{12} transporter BtuB
	OCL97_05800	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	OCL97_05825	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	OCL97_05935	TonB-dependent receptor	-	-
	OCL97_09005	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	OCL97_09385	TonB-dependent receptor	PirA-family-siderophore-receptor	Putative TonB-dependent receptor
	OCL97_09495	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	OCL97_10675	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	OCL97_16215	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	OCL97_17745	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	OCL97_19150	TonB-dependent receptor	-	Ferric aerobactin receptor
	OCL97_19355	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	OCL97_19465	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
	OCL97_19945	TonB-dependent receptor	-	Vitamin B_{12} transporter BtuB
Iron gene regulation	-	-	Sid_PvdS_regulator_Paeruginosa_PA2426_180620	ECF RNA polymerase sigma factor SigR
	OCL97_15730	RNA polymerase sigma factor	Sid_PvdS_regulator_Paeruginosa_PA2426_180620	ECF RNA polymerase sigma factor SigE
	OCL97_04690	PadR family transcriptional regulator	Sid_YqjI_regulator_for_YqjH_P64588_Escherichia_coli_180606	Hypothetical protein
	OCL97_02745	helix-turn-helix domain-containing protein	Sid_PchR_pyochelin_regulator_Pseudomonas_aeruginosa_PA4227_180623	HTH-type transcriptional regulator CdhR
	OCL97_01890	transcriptional repressor	PF01475-Iron_dependent_repressor-fur_family	Zinc uptake regulation protein
	OCL97_02175	FecR family protein	PF04773_FecR	Hypothetical protein
	OCL97_05900	transcriptional regulator FtrA	Sid_PchR_pyochelin_regulator_Pseudomonas_aeruginosa_PA4227_180623	HTH-type transcriptional regulator CdhR
	OCL97_05940	FecR domain-containing protein	PF04773_FecR	Protein FecR
	OCL97_05945	sigma-70 family RNA polymerase sigma factor	Sid_PvdS_regulator_Paeruginosa_PA2426_180620	Hypothetical protein

	OCL97_05290	helix-turn-helix domain-containing protein	Sid_PchR_pyochelin_regulator_Pseudomonas_aeruginosa_PA4227_180623	Hypothetical protein
	OCL97_12495	FecR domain-containing protein	PF04773_FecR	Protein FecR
	OCL97_12500	RNA polymerase sigma factor	Sid_PvdS_regulator_Paeruginosa_PA2426_180620	Putative ECF RNA polymerase sigma factor SigI
	OCL97_12725	transcriptional repressor	PF01475-Iron_dependent_repressor-fur_family	Ferric uptake regulation protein
	OCL97_20500	FecR domain-containing protein	PF04773_FecR	Hypothetical protein
	OCL97_21580	AraC family transcriptional regulator	Sid_PchR_pyochelin_regulator_Pseudomonas_aeruginosa_PA4227_180623	HTH-type transcriptional activator RhaS
Iron Storage	OCL97_15400	bacterioferritin	PF00210-Ferritin_like_domain	Bacterioferritin
Iron transport - Heme transport	OCL97_09810	ABC transporter ATP-binding protein	-	Hemin import ATP-binding protein HmuV
Unclassified	OCL97_00775	VIT family protein	-	Hypothetical protein
	OCL97_10360	FTR1 family protein	-	-
	-	-	-	Ferric aerobactin receptor
	OCL97_13230	cation diffusion facilitator family transporter	-	Ferrous-iron efflux pump FieF
	OCL97_10245	ferric reductase-like transmembrane domain-containing protein	-	Protein-methionine-sulfoxide reductase heme-binding subunit MsrQ

References:

Abraham, W. R., Macedo, A. J., Lunsdorf, H., Fischer, R., Pawelczyk, S., Smit, J., & Vancanneyt, M. (2008). Phylogeny by a polyphasic approach of the order *Caulobacterales*, proposal of *Caulobacter mirabilis* sp. nov., *Phenylobacterium haematophilum* sp. nov. and *Phenylobacterium conjunctum* sp. nov., and emendation of the genus Phenylobacterium. *IJSEM*, *58*(8), 1939-1949.

Matter, J. M., Broecker, W. S., Gislason, S. R., Gunnlaugsson, E., Oelkers, E. H., Stute, M., ... & Wolff-Boenisch, D. (2011). The CarbFix Pilot Project–storing carbon dioxide in basalt. *Energy Procedia*, *4*, 5579-5585.