environmental microbiology

Environmental Microbiology (2021) **00**(00), 00–00 doi:10.1111/1462-2920.15823

Cultivation of a vampire: '*Candidatus* Absconditicoccus praedator'

Michail M. Yakimov,[1*] **Alexander Y. Merkel**,[2] **Vasil A. Gaisin**,[3] **Martin Pilhofer**,[3] **Enzo Messina**,[4] **John E. Hallsworth**,[5] **Alexandra A. Klyukina**,[2] **Ekaterina N. Tikhonova**[2] **and Vladimir M. Gorlenko**[2]

[1]*Institute of Polar Sciences, ISP-CNR, Messina, Italy.*

[2]*Winogradsky Institute of Microbiology, Research Centre of Biotechnology, Russian Academy of Sciences, Moscow, Russia.*

[3]*Institute of Molecular Biology and Biophysics, Eidgenössische Technische Hochschule Zürich, Zürich, Switzerland.*

[4]*Institute for Marine Biological Resources and Biotechnology, IRBIM-CNR, Messina, Italy.*

[5]*Institute for Global Food Security, School of Biological Sciences, Queen's University Belfast, Belfast, Northern Ireland, UK.*

Summary

***Halorhodospira halophila*, one of the most-xerophilic halophiles, inhabits biophysically stressful and energetically expensive, salt-saturated alkaline brines. Here, we report an additional stress factor that is biotic: a diminutive Candidate-Phyla-Radiation bacterium, that we named '*Ca*. Absconditicoccus praedator' M39-6, which predates *H. halophila* M39-5, an obligately photosynthetic, anaerobic purple-sulfur bacterium. We cultivated this association (isolated from the hypersaline alkaline Lake Hotontyn Nur, Mongolia) and characterized their biology. '*Ca*. Absconditicoccus praedator' is the first stably cultivated species from the candidate class-level lineage Gracilibacteria (order-level lineage Absconditabacterales). Its closed-and-curated genome lacks genes for the glycolytic, pentose phosphate- and Entner–Doudoroff pathways which would generate energy/reducing equivalents and produce central carbon currencies. Therefore, '*Ca*. Absconditicoccus praedator' is dependent on host-derived building blocks for nucleic acid-, protein-, and peptidoglycan synthesis. It shares traits with (the uncultured) '*Ca*. Vampirococcus lugosii', which is also of the Gracilibacteria lineage. These are obligate parasitic lifestyle, feeding on photosynthetic anoxygenic Gammaproteobacteria, and absorption of host cytoplasm. Commonalities in their genomic composition and structure suggest that the entire Absconditabacterales lineage consists of predatory species which act to cull the populations of their respective host bacteria. Cultivation of vampire : host associations can shed light on unresolved aspects of their metabolism and ecosystem dynamics at life-limiting extremes.**

Introduction

Salt-saturated brines are extreme habitats for living systems due to their ionic, high-osmolarity, low water-activity conditions. Superficial brine lakes can also be subject to high levels of ultraviolet radiation and extremes of temperature, and some are very alkaline or acidic (Hallsworth, 2019). The specialized cellular microbes capable of life in such brines are also subject to infection by viruses (Atanasova *et al*., 2016; Uritskiy *et al*., 2021). Here, we reveal an additional biotic stress imposed by a diminutive but vampire-like bacterium that colonizes, feeds off, and kills other bacteria within alkaline brines.

Taxonomically, these bacterial 'vampires' occur within the Candidate Phyla Radiation (CPR), a large and phylogenetically diverse radiation of the bacterial tree with poorly understood ecological functions and metabolism (Wrighton *et al*., 2012; Rinke *et al*., 2013; Brown *et al*., 2015; Castelle *et al*., 2015; Castelle and Banfield, 2018). Recent phylogenetic and taxonomic analyses suggested the reclassification of the CPR as a single phylum, Patescibacteria (Parks *et al*., 2018). Great progress in studying biology of this group has been achieved by analysis of aquifer microbial communities near the town of Rifle (Colorado, USA), where Patescibacteria have been found to make up a large fraction of bacterial community with a so-called 'ultrasmall' (≤500 nm in average) cell size (Wrighton *et al*., 2012; Kantor *et al*., 2013; Luef *et al*., 2015). Besides the

Received 1 September, 2021; revised 12 October, 2021; accepted 14 October, 2021. *For correspondence. E-mail mikhail.iakimov@cnr.it. Tel. +39(090)6015437; Fax +39(090)669007.

diminutive size, small predicted genomes (<1 Mbp on average) are characteristic of the Patescibacteria located there.

Subsequently, small Patescibacteria genomes have been assembled from other environments worldwide including surface freshwater and marine habitats, deep-subsurface sediments, and animal and human microbiomes (for further references, see Méheust *et al*., 2019 and Jaffe *et al*., 2020). It is currently assumed that Patescibacteria are phylogenetically very heterogeneous and constitute up to a half of the Earth's bacterial diversity (Hug *et al*., 2016; Castelle and Banfield, 2018). Metagenome-inferred analyses of the small Patescibacteria genomes highlight the dramatic deficits in their capacity for *de-novo* biosynthesis. This predicted minimal biosynthetic and metabolic potential is assumed to represent ancestral adaptations to obligate symbiotic or parasitic lifestyles, making them fully reliant on hosts with more-complete biosynthetic pathways (Castelle and Banfield, 2018; Dombrowski *et al*., 2019).

However, this assumption is subject to debate, mainly due to the impossibility of cultivating Patescibacteria organisms at the present time, but also due to ambiguity in the interpretation of data obtained using culture-independent molecular 'omic' platforms. Currently, we cannot be certain of the nature of the relationship between Patescibacteria bacteria and host bacteria: predatory, mutualistically beneficial, or one of neutral coexistence (in effect, no relationship) (Beam *et al*., 2020; Chaudhari *et al*., 2021). Genome-resolved metagenomic and metaproteomic platforms can provide valuable information about Patescibacteria bacteria. However, understanding of their phenotype, behaviour, and ecological interactions can only be gained through cultivation experiments, which are currently extremely limited. To date, the cultivable representatives of the Saccharibacteria lineage (former TM7 phylum) are the only example of Patescibacteria to have been studied as a stable culture. They are obligate epibionts with reduced genomes, living attached to the surface of *Actinomyces* bacteria and they kill their host at some stage of the growth (He *et al*., 2015; Cross *et al*., 2019; Bor *et al*., 2020; Batinovic *et al*., 2021). Experiments with these strains have provided evidence for a parasitic/predatory (rather than mutualistic) relationship between the Patescibacteria and hosts. However, these data cannot readily be extrapolated to other members of a bacterial lineage as large and diverse as the CPR.

During the preparation of the current manuscript, an interesting study was published that reported on the genome-inferred biology and the environmental observation of a novel Absconditabacteria-related organism thriving in anaerobic athalassohaline saline microbial mats of the Salada de Chiprana in Spain, although a stable culture was not obtained (Moreira *et al*., 2021). Cells of the new epibiont were attached to the surface of halophilic anoxygenic phototrophic gammaproteobacterium *Halochromatium* sp. and their activity resulted in complete lysis of host cells. The authors noted the morphological similarity of the new bacteria to the genus *Vampirococcus* observed several decades ago as small predatory epibiotic bacteria attaching to cells of purple bacteria of the genus *Chromatium*, which form a thick pink layer in the chemocline of stratified sulfidic freshwater lakes (Guerrero *et al*., 1986). On this basis, it was named '*Ca*. Vampirococcus lugosii' (Moreira *et al*., 2021). An amplification of the whole genome of 10 individual cells allowed the assembly of a near-complete genome, which was greatly reduced for genes involved in biosynthetic metabolism. Therefore, it can be assumed that '*Ca*. Vampirococcus lugosii' as the Saccharibacteria strains, receives all necessary synthetic blocks (nucleotides, amino acids, lipids, etc.) directly from the host through cell-to-cell contact. However, direct contact between '*Ca*. Vampirococcus lugosii' (later on '*Ca*. Vampirococcus') and the host-cell envelope has not been documented (Moreira *et al*., 2021).

The cell-to-cell contact between epibiotic Patescibacteria and their hosts is important for molecule exchanges between their cells (Castelle *et al*., 2018; Castelle and Banfield, 2018; Sieber *et al*., 2019). Due to the paucity of culture-based studies, however, current knowledge about the organization of the cell-to-cell contact is relatively poor (He *et al*., 2015; Cross *et al*., 2019; Bor *et al*., 2020; Batinovic *et al*., 2021). Therefore, both the ultrasmall cells from the Rifle aquifer and the '*Ca*. Vampirococcus' : host associations have been analysed only within their respective environmental samples. This limits the amount, stability, and reproducibility of materials for microscope studies.

In the current study, we took on these challenges to learn more about interaction between CPR and their hosts in an extreme (biophysically constrained) ecosystem: an alkaline brine lake. First, we isolated a novel epibiotic Patescibacterium, which we named '*Candidatus* Absconditicoccus praedator', by creating a stable binary culture with its host, *Halorhodospira halophila*, an obligately anaerobic photosynthetic purple sulfur bacterium. *Ca*. Absconditicoccus is the first stably cultivated CPR species from the candidate phylum Absconditabacteria (formerly SR1). The division was named for clones recovered from microbial streamers on sulfidic sediments from the Sulphur River in Parkers Cave, Kentucky (Angert *et al*., 1998). Second, we studied interactions between this epibiont and its host by producing and analysing the fully assembled and annotated genome, coupled with transmission electron microscopy (TEM) and cryo-electron tomography (cryo-ET).

Results and discussion

Physicochemical properties of haloalkaline Lake Hotontyn Nur

Lake Hotontyn Nur (48°06′10″N; 113°45′37″E), from which samples were taken to isolate anaerobic sulfur bacteria, is the most saline of the numerous lakes located in northeastern Mongolia. This region is a flat steppe with depressions, many of which are occupied by shallow salt lakes (Rasskazov and Abramov, 1987) and is characterized by a sharply continental, semi-arid climate; at about 10 000 km^2, it covers an area similar to that of Northern Ireland or Jamaica. Seasonal temperatures range from $-40°C$ to $+40°C$ and the annual precipitation is about 200 mm, more than 70% of which falls in summer. The Lake Hotontyn Nur brine has a total mineral content of 220–390 g L^{-1}, alkalinity of 5–405 mM HCO_3^-, CO_3^{2-} concentration of 425–850 mM, and pH of 9.5–10.3 (Sorokin *et al*., 2004). The main salt in the lake is sodium chloride so the water activity likely approximates to 0.755 (Lee *et al*., 2018) but, due to the presence of soda, the high pH is stable. The alkaline pH, high salinity, and presence of sulfates favour the development of haloalkalophilic microbial communities that participate in the sulfur cycle of this lake (Sorokin *et al*., 2004, 2021).

Bringing Absconditabacterales into culture

In September 1999, an enrichment culture of the halophilic obligately anaerobic purple sulfur bacterium *H. halophila* M39 was obtained from salt-crust samples using the dilution-to-extinction technique. The composition of the medium for the isolation and further cultivation of purple sulfur bacteria, as well as cultivation conditions are given in the Experimental procedures section. After phase-contrast microscopic analysis, tiny coccoid cells were occasionally observed attached to the spiral or ovoid cells of *H. halophila* (Fig. 1A). The culture was maintained in an active condition in the Laboratory of Ecology and Geochemical Activity of Microorganisms (Winogradsky Institute of Microbiology) until 2008.

These small cocci could not be identified at that time so the culture was stored in the laboratory (both at room temperature and at $-80°C$) for more than a decade. Both cultures were reactivated in 2019 to carry out the current study. Using Illumina shotgun sequencing, the presence of sequences associated with *H. halophila* accounted for about two-thirds of all sequences analysed (65.7%), while sequences associated with Absconditabacterales (Patescibacteria) were found in the activated culture accounted for about one third (29.5%). There was also an insignificant contamination of uncultured *Halarsenatibacteraceae* (2.7%), *Alkalilimnicola* (1.1%) and *Izemoplasmatales* (1.0%) (Supporting Information Fig. S1A). Epifluorescent microscopy revealed that tiny cocci were associated exclusively with autofluorescent and FM5-95 evidenced cells, indicating that only

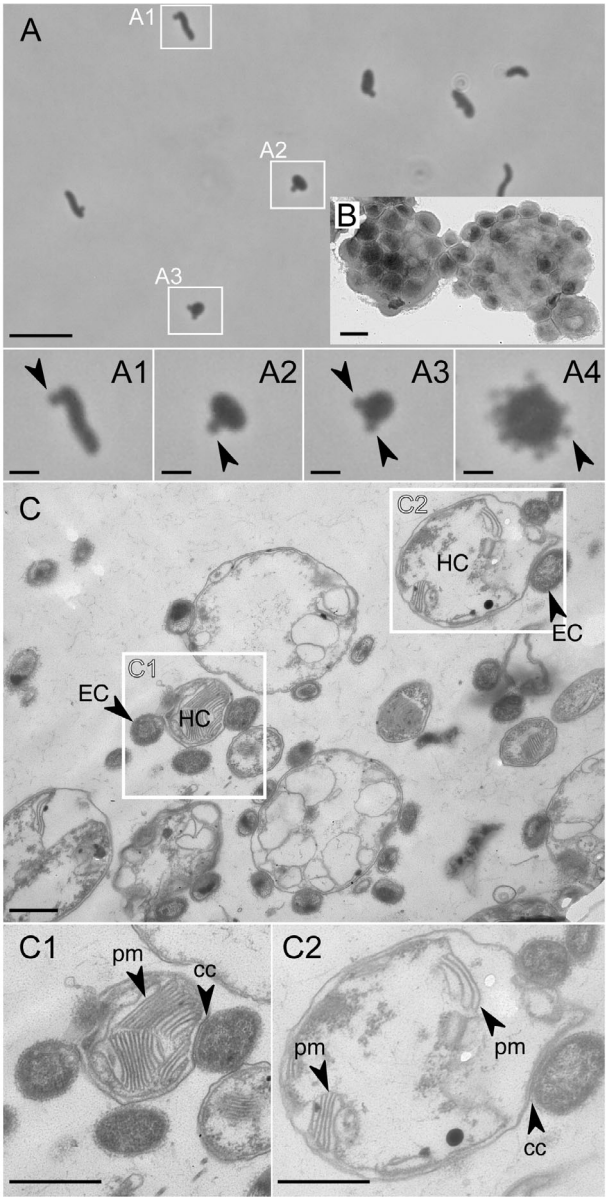

Fig. 1. Cells of '*Ca*. Absconditicoccus praedator' M39-6 colonizing the host cells.
A. Phase-contrast light micrograph of the epibiotic cells infecting host cells in the enrichment culture.
A1–A4. Examples of the infected host cells with diverse morphologies.
B. Transmission electron micrograph of the stained host cell with the multiple attaching epibiotic cells.
C. Transmission electron micrograph of the ultra-thin cross-section through the associated host and epibiotic cells.
C1. The infected host cell with intracellular contents.
C2. The inflated host cell with depredated intracellular content upon the infection. The arrowheads without abbreviations show epibiotic cells. Abbreviations: EC, epibiotic cell; HC, host cell; cc, cell-to-cell contact; and pm, photosynthetic membrane. Scale bars: A = 5 μm; A1–A4 = 1 μm; B, C and C1–C2 = 500 nm.

phototrophic purple sulfur *H. halophila* are likely their hosts (Supporting Information Fig. S2).

To get rid of minor contamination of non-phototrophs and to proceed with the analysis of this association, the grown culture was serially diluted (on a 10-fold basis) to 10^{-4} step, and the serial dilutions were spread anaerobically on a solid culture medium to obtain separated colonies (see Experimental procedures). Separate, coloured colonies that formed were further randomly selected and checked by specific PCR for the presence of Absconditabacterales. The PCR-positive colonies used for inoculation of the liquid medium and then spread again onto solid media, and the presence of the epibiont was again verified by '*Ca*. Abs. praedator'-specific PCR.

The PCR-positive colonies were used for the second iteration of the same procedures: plating, colony PCR, inoculating in the autotrophic medium, microscopy screening and the Illumina-generated 16S rRNA gene amplicon survey (Supporting Information Fig. S1B). The binary culture of epibiont : *H. halophila*, purified using this approach, was used for further studies. We propose this co-isolated Patescibacteria member as a novel genus and species, '*Ca*. Absconditicoccus praedator' M39-6 (later on '*Ca*. Abs. praedator'). An explanation/justification for these names is presented below.

General characteristics of the 'Ca*. Abs. praedator' and* Halorhodospira *genomes*

In order to assemble the genome of '*Ca*. Abs. praedator' M39-6, we carried out shotgun WGS sequencing of the enrichment culture. As a result, a circular chromosome of the '*Ca*. Abs. praedator' M39-6 was reconstructed. We found that the genome of '*Ca*. Abs. praedator' M39-6 is as small as many of the other environmental metagenome-assembled Patescibacteria genomes, and it consists of a single circular chromosome of 1 129 696 bp with a molGC content of 27.9% (Supporting Information Table S1a). The chromosome contains a single ribosomal RNA operon (23S-5S-16S rRNA genes) as well as 38 tRNA genes, one transfer-messenger RNA (tmRNA), and one RNase P class A ribozyme (non-coding RNA). The protein-coding genes were predicted using the genetic code number 25 (UGA translates as glycine) in agreement with previous studies of *Ca*. Gracilibacteria (Rinke *et al*., 2013; Hanke *et al*., 2014; Ivanova *et al*., 2014; Sieber *et al*., 2019), *Ca*. Absconditibacterales, and '*Ca*. Vampirococcus' (Campbell *et al*., 2013; Moreira *et al*., 2021).

Among the hypotheses that have been proposed to explain the emergence of alternative coding, one is a very low GC content (Ivanova *et al*., 2014), which is consistent with the mol.% G + C values for '*Ca*. Abs. praedator'. Of the 1054 protein-coding genes annotated (equivalent to 89.9% of genome coverage), only 58.4% were assigned to a *National Centre for Biotechnology Information* (NCBI) Clusters of Orthologous Group (COGs) category (Supporting Information Table S2a). Such a low percentage is typical for all currently known nano-sized prokaryotes. As the life cycle of Patescibacteria depends on other organisms, many of their unknown genes may be involved in interactions with cells of other organisms – a phenomenon that is poorly understood for uncultivated microbes.

The draft genome of the host bacterium *Halorhodospira* sp. M39-5 (Supporting Information Table S1b) consists of seven scaffolds with an estimated length of 2 767 651 bp, a molGC content of 68.4%, two 5S-23S-16S rRNA operons and 40 tRNA genes. Of the 1998 protein-coding genes annotated in M39-5, 1174 were assigned to COGs; i.e. 58.8% of the total (Supporting Information Table S2b). The type strain of the genus, *H. halophila* $SL1^T$ (CP000544), has similar a genome structure and similar arrangement of rRNA operons (Challacombe *et al*., 2013).

*'*Ca*. Abs. praedator' is a novel cultivable representative of the Gracilibacteria lineage*

We analysed the phylogeny of our epibiont using amino-acid sequences of 120 conservative proteins (Parks *et al*., 2020) as well as 16S rRNA gene sequence extracted from its genome. As a result of 16S rRNA gene-sequence analysis, we found that this diminutive bacterium belongs to a clade with uncultured Absconditabacteriales previously described in different saline and haloalkaline environments (Fig. 2), including '*Ca*. Vampirococcus' (Moreira *et al*., 2021). Pairwise comparison of 16S rRNA sequences from the *Halorhodospira* epibiont and '*Ca*. Vampirococcus' showed 87.6% identity, which is far below the cut-off value proposed for delineation of genus-level taxa (Yarza *et al*., 2008; Tindall *et al*., 2010).

Since members of Patescibacteria have been described mostly as environmental metagenome-assembled genomes (MAGs) and single-cell amplified genomes, we relied primarily on genome-based taxonomy and used the Genome Taxonomy Database (GTDB) (Chaumeil *et al*., 2020) during our phylogenetic analysis of 120 conservative proteins. In addition to genomes from GTDB (R05-RS95), we used the MAG of '*Ca*. Vampirococcus', which we assembled from genome-sequence data deposited by Moreira *et al*. (2021) (SRX9673531-SRX9673533). Our phylogenetic analysis showed that according to GTDB taxonomy together with previously described '*Ca*. Vampirococcus', the epibiont of *Halorhodospira* belongs to the order Absconditabacteriales within the class *Ca*.

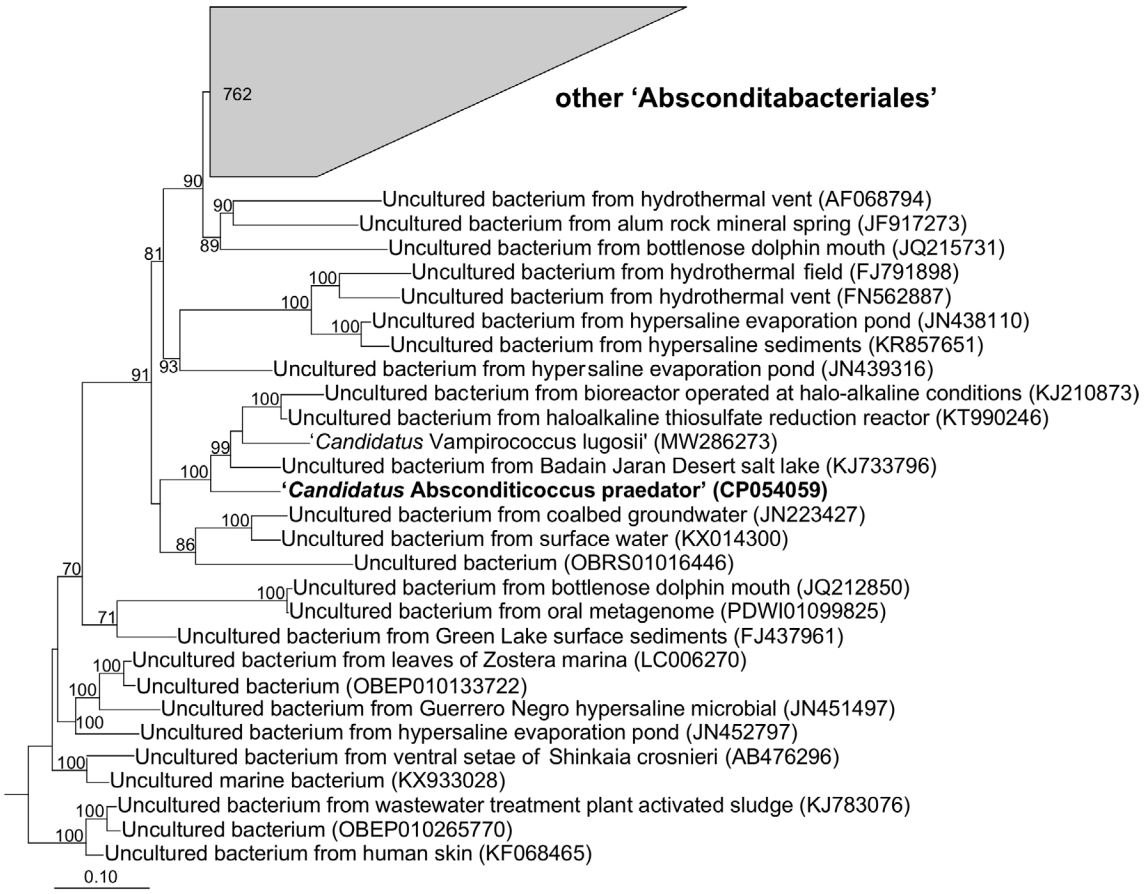

Fig. 2. Phylogenetic placement of '*Ca*. Absconditicoccus praedator' M39-6 based on 16S rRNA gene-sequence data. The trees were constructed using the IQ-TREE 2 program (Minh *et al*., 2020) with fast model selection via ModelFinder (Kalyaanamoorthy *et al*., 2017) and ultrafast bootstrap approximation (Minh *et al*., 2013) as well as an approximate likelihood-ratio test for branches (Anisimova and Gascuel, 2006). A bootstrap consensus tree is shown with values above 70% placed at the nodes. Bar, 0.10 changes per position. The names of the branches indicate the isolation source of the phylotype. The classification was based on the database SILVA r138.1 (Quast *et al*., 2013).

Gracilibacteria, which also accommodates the former BD1-5 and GN02 candidate divisions of the phylum Patescibacteria (Ley *et al*., 2006; Wrighton *et al*., 2012) (Fig. 3). These results and, especially, the deep branching on the genome-based phylogenetic tree motivated us to propose a novel genus name and species name for this new Absconditabacteriales bacterium : '*Candidatus* Absconditicoccus praedator' (see below).

Cell organization of 'Ca. Abs. praedator'

The availability of '*Ca*. Abs. praedator' in culture provided an opportunity to study its biology, including the direct observation of its cellular organization and interactions with the host. In electron micrographs, the host cells with photosynthetic intracellular membranes organized in stacks of lamellae were clearly distinguishable from those of contaminating (non-photosynthetic) bacteria in the enrichment culture (Supporting Information Fig. S3).

In light micrographs, '*Ca*. Abs. praedator' was clearly distinguished as relatively small coccoid cells (average diameter 0.3 μm, $n = 40$ cells) attaching to the host cells (Figs 1 and 4, Supporting Information Fig. S2). The numbers of '*Ca*. Abs. praedator' cells attaching to the host cell varied from 1 to 11 (Fig. 1A1–A4) but reached 20 cells per host cell on some TEM images (Fig. 1B). Interestingly, infected host cells can have an abnormal spherical morphology making them different from the typical spiral shape. Indeed, a weakly convoluted uninfected spirilla with a diameter of 0.8 μm and length of 3–5 μm becomes a spherical cell as a host, with a diameter of up to 3 μm (Fig. 1A1–A4). A similar change of host-cell morphology has also been reported for the association between Saccharibacterium TM7x and *Actinomyces odontolyticus* (Bor *et al*., 2018;2020) and is reminiscent of the cell gigantism of the hyperthermophilic and acidophilic archaeon *Sulfolobus islandicus* that is induced by viral infection (Liu *et al*., 2021).

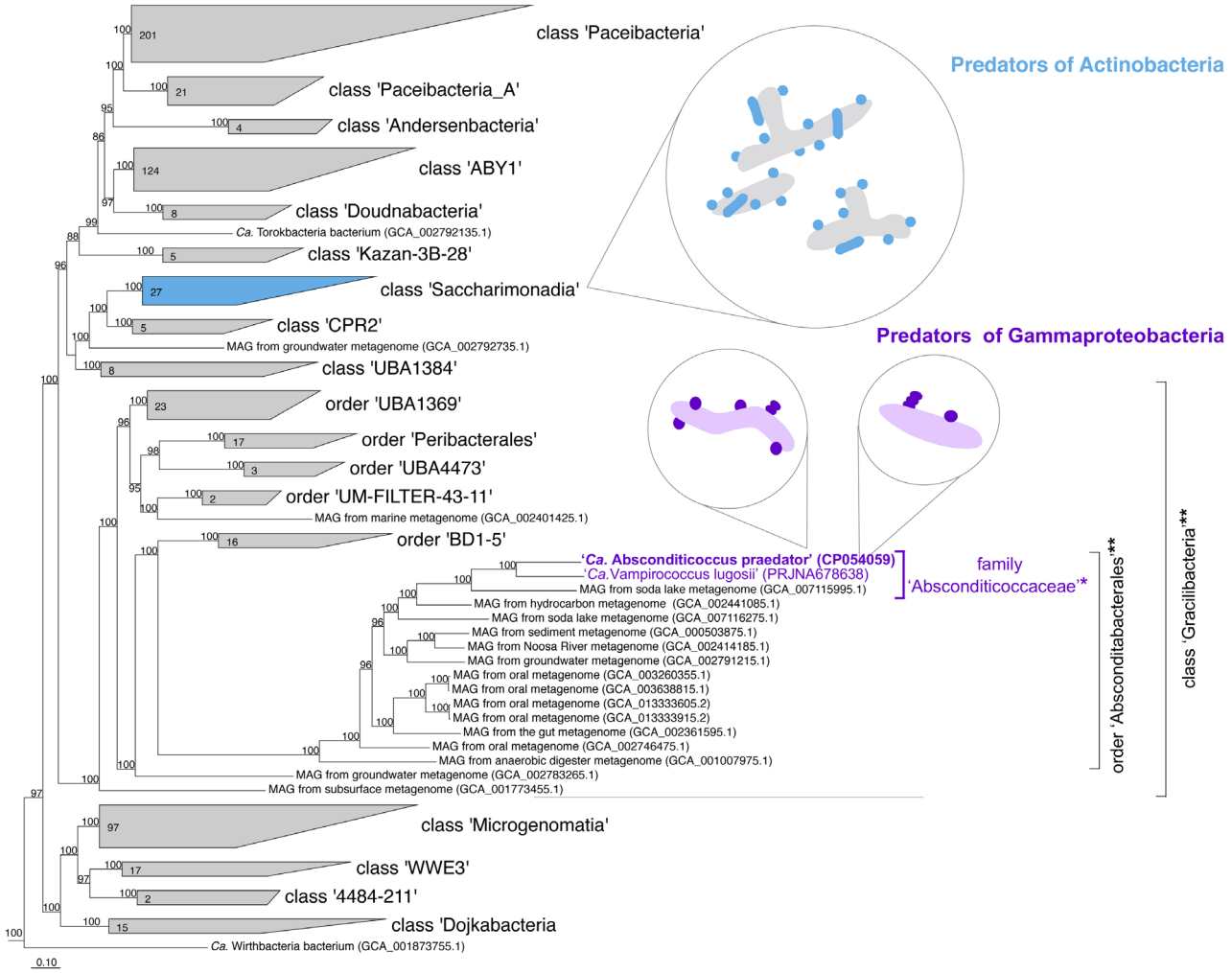

Fig. 3. Phylogenomic placement of '*Ca*. Absconditicoccus praedator' M39-6 and other representatives of the Absconditabacterales order among other groups of 'Patescibacteria'. This reconstruction is based on concatenated partial amino-acid sequences of 120 bacterial conservative proteins (Parks *et al*., 2018); taxonomic designations correspond with Genome Taxonomy DataBase Release 05-RS95. The trees were built using the IQ-TREE v. 2.0.3 (Minh *et al*., 2020). For rooting the tree sequences from 10 high-quality genomes of representatives of the *Thermotogota* phylum were taken. All assemblages of the order Absconditabacterales according to Genome Taxonomy DataBase Release 05-RS95 are shown on the tree. Bootstrap values above 85% are shown at the nodes. Bar, 0.10 changes per position. Asterisk means the name of family proposed through this study; double asterisk – GTDB defined taxa. [Color figure can be viewed at wileyonlinelibrary.com]

According to our cryo-ET data, the average diameter of all found '*Ca*. Abs. praedator' cells was 497 $\pm$ 75 nm ($n = 53$ cells), and the average diameter of attached cells was 485 $\pm$ 66 nm ($n = 48$ cells). We consider that these cryo-ET data to be the most accurate determinations for cell size.

Analysis of the cryo-ET data allowed us to describe the cell envelope of '*Ca*. Abs. praedator', which was composed of three layers that differed in their densities (Fig. 4A and E). The innermost layer likely corresponds to the cytoplasmic membrane, followed by a more diffuse (intermediate) layer, which may be a peptidoglycan cell wall. The outermost layer enveloping the cell appeared wavy in the two-dimensional slice through a tomogram (Fig. 4B). We segmented this outer layer (in three dimensions), using the isosurface function in Interferometric Modulator Display (IMOD) software and found it to comprise an arrangement of repeated protrusions (Fig. 4C). This organization of the structure was also observed in a planar slice that cut across the edge of the cell envelope (Fig. 4D). The morphology of this outer structure resembled the appearance of a bacterial S-layer in freeze-etched preparations (Sleytr *et al*., 1993). We also found pilus-like extracellular appendages protruding from the surface of the cell envelope (Fig. 4F). The overall organization of the cell envelope of '*Ca*. Abs. praedator' resembled those of the 'ultrasmall' CPR bacteria from aquifer microbial communities (Luef *et al*., 2015).

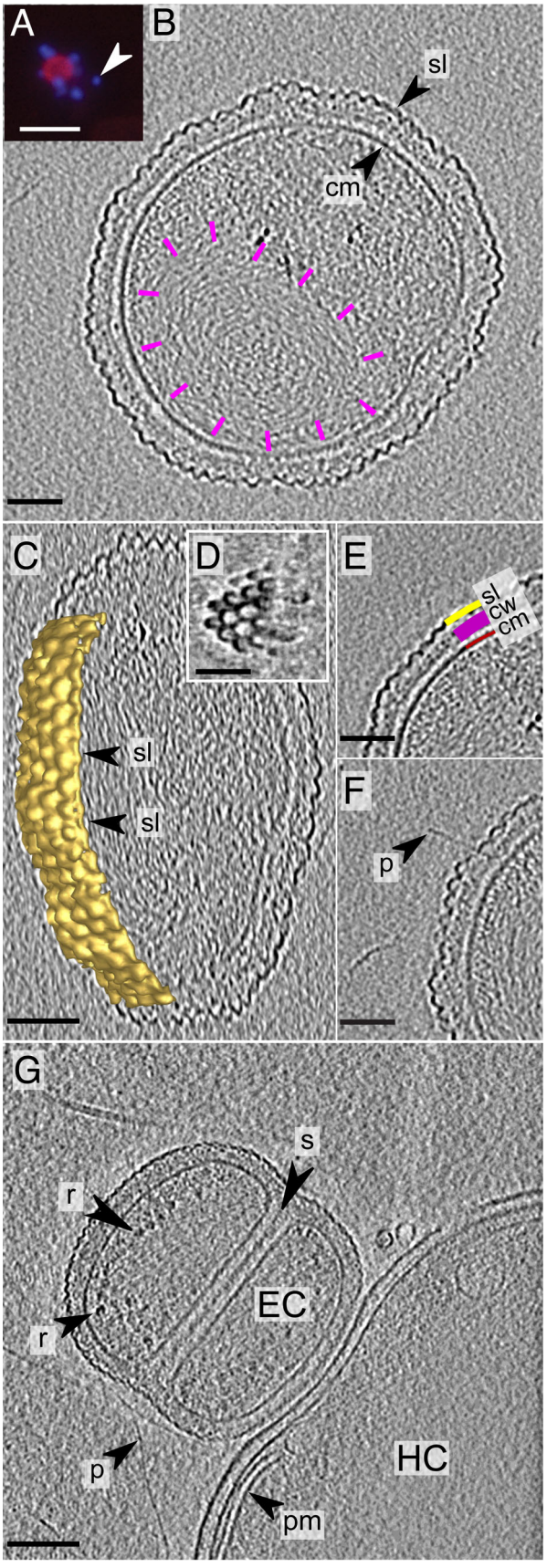

Fig. 4. Cell organization of '*Ca*. Absconditicoccus praedator' M39-6.
A. Fluorescence image of the epibiotic cells infecting host cells in the enrichment culture. The blue colour depicts fluorescence under excitation light 390 nm (stained DNA); the red colour depicts fluorescence under excitation light 510 nm (stained membrane + autofluorescence).
B. Slice through a filtered cryo-electron tomogram of an unattached cell. The dashed magenta lines contour a volume with differently organized densities.
C. Iso-surface segmentation of the S-layer-like structure in relation to the position in a tomogram.
D. Planar slice through the cryo-electron tomogram showing top view of the S-layer-like structure on the cell surface.
E. Slice through a filtered cryo-electron tomogram showing layers of the cell envelope.
F. Representative slice through a cryo-electron tomogram showing an extracellular appendage.
G. Slice through a cryo-electron tomogram showing an attached epibiotic cell. Abbreviations: EC, epibiontic cell; HC, host cell; cm, cytoplasmic membrane; cw, putative cell wall; om, outer membrane; pm, remnants of photosynthetic membrane; p, pilus-like extracellular appendage; sl, S-layer-like structure; s, cell division septum; and r, putative ribosome. Scale bars: A = 2.5 μm; B–G = 100 nm. [Color figure can be viewed at wileyonlinelibrary.com]

The cytoplasm of the unattached '*Ca*. Abs. praedator' cells exhibited a relatively homogeneous density in the tomograms, but with regions of fibrous material (Fig. 4B). In contrast, the cytoplasm of the epibiotic cells that were attached to the host-cell envelope appeared to contain ribosomes (Fig. 4G). The attached '*Ca*. Abs. praedator' cells formed diplococcus-like associations, divided by a septum-like ingrowth of the cell wall and the cytoplasmic membranes (Fig. 4G). This septum was parallel to the plane of cell–cell contact with the host ($n = 10$ diplococcus-like cellpairs).

'*Ca. Abs. praedator' colonizes host cells and affects their cytoplasm integrity*

Although some microbial predators consume their target cell via phagocytosis or by digesting them from inside (Evans *et al*., 2007), vampire CPR bacteria seem to gradually consume substances from the host-cell cytoplasm (He *et al*., 2015; Cross *et al*., 2019; Moreira *et al*., 2021). However, the mechanism via which this takes place is not fully understood. We used electron microscopy (and evidence from the genome biology, below) to analyse the '*Ca*. Abs. praedator' : *Halorhodospira* interaction and thereby gain some insight into its unconventional form of nourishment.

Electron micrographs of ultrathin cross-sections through the '*Ca*. Abs. praedator' : host-cell association showed that the host cells lost integrity of their cytoplasm, but remnants of their photosynthetic membranes were present (Fig. 1C). We assume that both the light micrographs and electron micrographs of the infected host cells show different stages of the host degradation

resulting from '*Ca*. Abs. praedator' infections; from freshly formed associations (Fig. 1C1) to the late-stage infections characterized by complete lysis of the host (Fig. 1C2). Based on these observations, we assume that '*Ca*. Abs. praedator' consumes the host cytoplasm in the same manner as does the '*Ca*. Vampirococcus' species (Guerrero *et al*., 1986; Moreira *et al*., 2021).

Genome of 'Ca*. Abs. praedator' encodes very limited catabolic capacity*

We performed annotation using integrated microbial genomes, Rapid Annotation using Subsystem Technology, KEGG Orthology Database, and manual curation of specific genes to accurately assess the completeness of central catabolic and anabolic pathways in the reconstructed genome of '*Ca*. Abs. praedator'. Biosynthetic pathways readily recognized in the '*Ca*. Abs. praedator' genome include those for nucleic acid synthesis, DNA repair and nucleotide interconversion, ribosome-based protein synthesis, synthesis and assembling of Fe-S proteins, cell division, production of cell envelope and cell wall components, Sec-type protein secretion, and Type IV pilus (T4P) formation (Fig. 5). In contrast, perhaps the most surprising feature of this parasitic epibiont is the complete lack of genes encoding core carbon metabolism (the Embden–Meyerhof–Parnas-, pentose phosphate-, Entner–Doudoroff-, and other pathways), which would generate energy/reducing equivalents and produce central carbon currencies.

This near-complete abandonment of central carbon metabolism was previously noted for some of the most minimalistic CPR genomes (Castelle *et al*., 2018; Sieber *et al*., 2019; Moreira *et al*., 2021). However, those organisms do possess at least the lower portion of the Embden–Meyerhof–Parnas pathway (glycolysis), which in the case of '*Ca*. Vampirococcus' begins with 3-phosphoglycerate, the only exploitable source for energy generation (Moreira *et al*., 2021). The conversion of this substrate to pyruvate, coupled with the formation of ATP, includes phosphoglycerate mutase, enolase and phosphoenolpyruvate kinase. We were unable to find the corresponding genes in the '*Ca*. Abs. praedator' genome. Only two enzymes that are presumably associated with reversible conversions of pyruvate were annotated: putative pyruvate phosphate dikinase and malic enzyme (Acpr_0159 and Acpr_0788). The absence of even partial substrate-level phosphorylation distinguishes the '*Ca*. Abs. praedator' genome from '*Ca*. Vampirococcus' and other Absconditabacteria.

The absence of core carbon metabolism raises the question of how general energy currencies (ATP, NADH, NADPH, and ferredoxin) are reduced and recycled in '*Ca*. Abs. praedator' cells. Like all sequenced Absconditabacteria, the '*Ca*. Abs. praedator' genome apparently lacks a standard electron transport chain. Addressing this issue, we have hypothesized and considered several reactions that could ultimately be involved in oxidation or reduction of these electron shuttles. In particular, ATP, which is required for all anabolic processes, is presumably formed via the F0F1-type ATP synthase complex (Acpr_0048-52, Acpr_0255, Acpr_0509 and Acpr_1048). Given the absence of an electron transport chain that could pump protons, PMF could be directly stolen from tightly attached *Halorhodospira* cells as has been suggested for some epibiotic CPR bacteria (Castelle *et al*., 2018; Castelle and Banfield, 2018), including Gracilibacterium BD1-5 (Sieber *et al*., 2019). Such close physical associations have also been reported for another CPR bacterial group, Saccharibacteria (TM7), which attaches to host (*Actinobacteria*)-cell surfaces (He *et al*., 2015), but were not observed for '*Ca*. Vampirococcus' (Moreira *et al*., 2021). In the case of the latter, the membranes of the ectoparasite and host cells were separated by a relatively large space of about 100 nm, which would allow the diffusion of protons thus preventing their efficient transfer between cells. However, accurate analysis of the binary CPR : host culture would be required to ascertain mechanistic details of their cell-to-cell exchanges.

Alternative pathways for the generation of PMF proposed in 'Ca*. Vampirococcus' and* 'Ca*. Abs. praedator'*

As an alternative for the generation of PMF in '*Ca*. Vampirococcus', Moreira *et al*. (2021) proposed the involvement of a protein with the atypical tripartite domain structure: the N-terminal region, the central part, and the C-terminal region were similar to flavocytochrome, and the rubredoxin-like non-heme iron-binding domain and NAD-binding motif of ferredoxin reductases, respectively. The study proposed conservation of this protein within Absconditabacteria, but we did not find it in the genome of '*Ca*. Abs. praedator'. Instead, another protein (Acpr_0793) was discovered that also has a tripartite structure, but is composed of completely different domains. This 73-kDa enzyme may be acting outside the cells, as indicated by the presence of the signal Sec/SPI peptide (cleavage site between positions 20 and 21) and the S-layer homology domain at the N-terminus. The central part of the protein resembles the blue copper proteins of the plastocyanin/azurin family. Along with various dehydrogenases and c-type cytochromes, these proteins are believed to act as mediators within multicomponent redox complexes (Chen *et al*., 1994). Surprisingly, the C-terminal region of Acpr_0793 containing the beta-propeller fold was similar to members of membrane-bound sugar dehydrogenases (SDH).

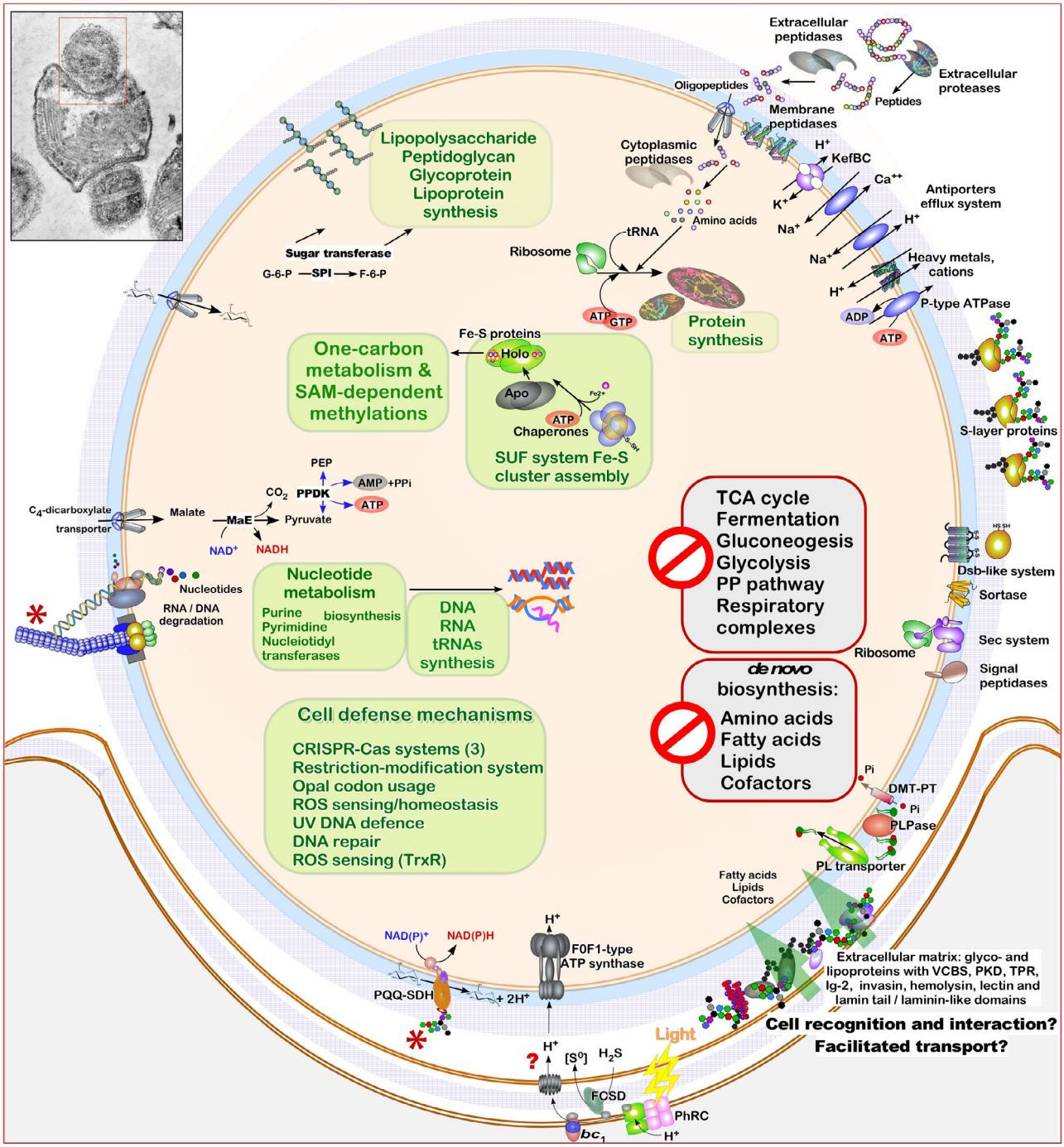

Fig. 5. Reconstruction of central metabolic and homeostatic functions of '*Ca*. Absconditicoccus praedator' M39-6. Bold green text indicates active and prominent functions; rectangles with a stop sign indicate the absence of metabolic and biosynthetic pathways; asterisks indicate the proposed models of electron transport chain and nutritional competence, depicted in details in Fig. 6; and the question mark indicates the stealing of proton motive force from attached host cells, as has been suggested for some CPR bacteria (Castelle *et al*., 2018; Castelle and Banfield, 2018; Sieber *et al*., 2019). Acquisition of all required biosynthetic precursors from the host cell is shown by the green arrows in the lower-right part of the figure. Details of genes and systems abbreviations are provided in the Supporting Information Table S3. Abbreviations: G-6-P, glucose-6-phosphate; F-6-P, fructose-6-phosphate; PEP, phosphoenolpyruvate; Pi, inorganic phosphate. [Color figure can be viewed at wileyonlinelibrary.com]

Members of this family possess pyrroloquinoline quinone (PQQ) as a prosthetic group and transfer electrons from aldose sugars, specially from glucose, mannose, xylose, and/or galactose, directly to the electron transport chain associated with the respiratory chain (Sode *et al*., 1995). It is noteworthy that the primary structure of

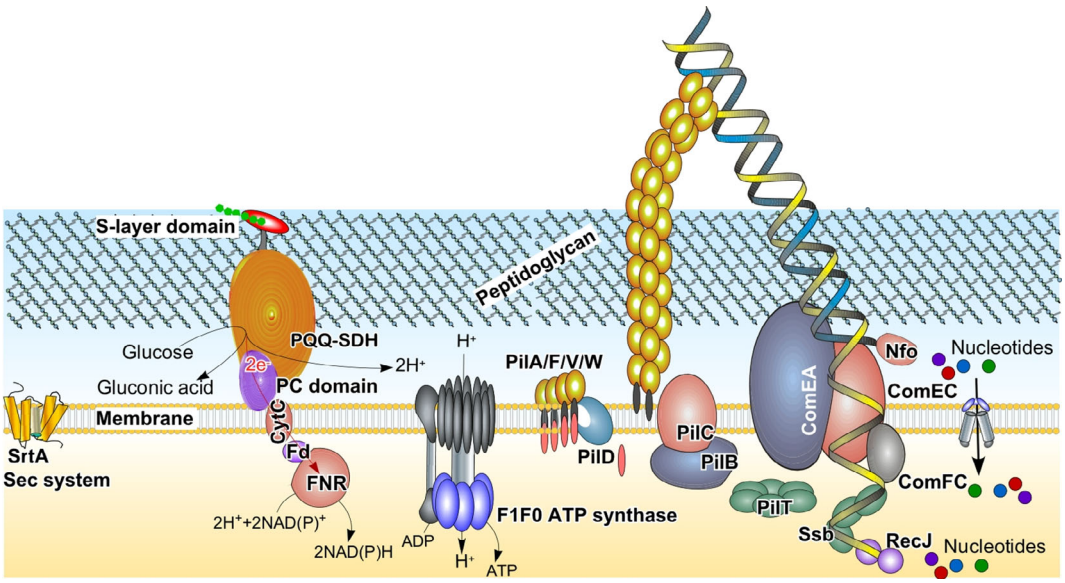

Fig. 6. Proposed models of electron transport chain and nutritional competence in *'Ca*. Absconditicoccus praedator' M39-6. Details of genes and systems abbreviations are provided in the text and in Supporting Information Table S3. [Color figure can be viewed at wileyonlinelibrary.com]

PQQ-SDH region of Acpr_0793 protein had significant BLASTp hits (*E*-value $<10^{-133}$, 57%–58% identity) against various natronoarchaeal glucose/sorbosone dehydrogenases. This similarity was observed only at the amino-acid sequence level and Acpr_0793 gene has a low molGC% content (34.5%) by comparison with the moderate molGC% of natronoarchaeal PQQ dehydrogenases (64.8%). Moreover, the transcribed region of Acpr_0614 gene contains 39 opal codons (UGA); all this suggests that if its acquisition by horizontal gene transfer (HGT) originated from natronoarchaea, then this event apparently occurred a long time ago.

Given that we also found both C5-cytochrome (Acpr_0948) and ferredoxin (Acpr_0251), we hypothesize that these heme- and iron–sulfur proteins may serve as an interface between plastocyanin-containing PQQ-SDH and cytoplasmic ferredoxin-$NADP^+$ oxidoreductase (FNR, Acpr_0614). Thus, in this inferred four-component PMF-generating system, the last step of electron transport ends with the reoxidation of reduced ferredoxin and the conversion of $NADP^+$ to NADPH catalysed by FNR (Fig. 6). The system seems to be based solely on externally derived aldose sugars, which are most likely supplied by the photosynthetic host. Alternatively, and/or under conditions of sugar deficiency, PMF could be generated by the Na^+/K^+ proton antiporters, or through cytoplasmic consumption of H^+ performed either by superoxide dismutase (Acpr_1058) or via reactions involved in breakdown of amino acids and other compounds. NADH may be regenerated by malic enzyme-catalysed conversion of malate to pyruvate. Like sugars, malate can be obtained from external sources and transported into the cytoplasm using a C_4-dicarboxylate/malic acid transporter (Acpr_0014). Other reactions, such as those involved in interconversion of tetrahydrofolate compounds and peptidoglycan synthesis, may also be involved in the recycling of energy currencies.

Genes for cell surface elaboration enrich the 'Ca. Abs. praedator' genome

Interestingly, the reduction of the '*Ca*. Abs. praedator' genome does not appear to have affected the genomic repertoire responsible for cell-surface elaboration. In fact, from the perspective of the cell envelope, the biosynthetic pathway for peptidoglycan is complete, although we predict the need for the UDP-*N*-acetylglucosamine precursor from host cells for use as a precursor. The genome contains 14 glycosyltransferases belonging to different CAZy families, many genes for the synthesis of (lipo)polysaccharides (e.g. Acpr_0618, 0624, 0752 and 1013) and for proteins with S-layer domains (Acpr_0078, 0102 and 0620). This prediction was in agreement with the observed S-layer-like surface of the 'Ca. Abs. praedator' cells (Fig. 4). Intriguingly, the '*Ca*. Abs. praedator' Acpr_0102 S-layer protein has an alpha-2-macroglobulin domain, which may function as an inhibitor of host cell protease, thereby disarming the host's antimicrobial defence system (Budd *et al*., 2004), as was proposed for '*Ca*. Vampirococcus' (Moreira *et al*., 2021).

'*Ca*. Abs. praedator' also possesses an impressive arsenal of proteins to deploy during hostile interactions with the host, most likely aimed at destroying the host's

cell wall and plasma membrane. These putative virulent factors include membrane-bound and secreted lytic murein transglycosylases (Acpr_0262, 086 and 0896), hemolysin (Acpr_0681), large protein with predicted invasin domain (Acpr_0660), membrane-bound phospholipid phosphatase (Acpr_1059), and phospholipase (Acpr_0481). It has been speculated that some of these enzymes are involved in mechanical deformation of the host cell. Acpr_1059 and Acpr_0481 can break down the lipids in the outer membrane of the host, which might increase permeability and release cytoplasm that is then drawn directly into the '*Ca*. Abs. praedator' at the epibiont cell : host–cell interface. This loss of cytoplasm might also cause the deformation of the cell via alterations in cell turgor and/or plasma membrane behaviour, causing the creation of edges. Geometrical deformation of the host cell can also be caused due to degradation of peptidoglycan, which, indeed, was not visible in infected host cells.

Other large membrane-bound and secreted proteins have annotations suggestive of their involvement in the formation of the very complex extracellular matrix; these may aid '*Ca*. Abs. praedator' attachment to the *Halorhodospira* surface. All these glyco- and lipoproteins contain either polycystic kidney disease-, laminin-like/lectin-, immunoglobulin- domains or tetratricopeptide repeat (Supporting Information, Table S4). Thus, based on genome analysis, we anticipated a cell wall-containing peptidoglycan with a periodic S-layer, pili, many large extracellular proteins and polymeric substances, which are likely to take part in surface interactions between predatory '*Ca*. Abs. praedator' and *Halorhodospira* host cells (Fig. 5).

*Degradation of the host DNA can be involved in '*Ca*. Abs. praedator' lifecycle*

Phosphorus is a key element for life, and in many ecosystems is the limiting inorganic nutrient. However, a significant amount of dissolved environmental phosphorus is often present in environmental DNA (eDNA) (Dell'Anno and Danovaro, 2005) and hypersaline environments are reported to have the highest concentrations of eDNA (Corinaldesi *et al*., 2008). Therefore, it is not surprising that microorganisms have developed mechanisms to thrive in such habitats, by using not only phosphate ions but also phosphorus sources such as eDNA (Hua *et al*., 2021).

A notable feature of '*Ca*. Abs. praedator' is its eventual nutritional competence – the ability to utilize environmental double-stranded DNA (edsDNA) molecules as a source of phosphorus. Based on models of natural competence in other bacteria, we have identified several genes that may be important for this process in '*Ca*. Abs. praedator', and suggested that single-stranded DNA is degraded following entry into the '*Ca*. Abs. praedator' cytoplasm (Fig. 6). In addition to a SrtA sortase (Acpr_0180), which can catalyse the assembly of pilins into pili, we have identified at least 10 genes, some organized into small operons and some in multicopy, which presumably encode for the assembly proteins of T4P. Note that pilus-like structures have been seen in cryo-ET images (Fig. 4F and G). It remains to be determined whether these extracellular appendages can be involved in either the motility of unattached of '*Ca*. Abs. praedator' cells or their attachment to the host cell.

We believe it is most likely that these proteins are involved in assembly of the competence (pseudo)pili; filamentous structures are similar to T4P, and the main function of which is to provide access for the membrane-bound translocation complex to edsDNA (Chen and Dubnau, 2004). A membrane-bound DNA receptor ComEA (Acpr_0617) transfers the bound DNA toward the membrane channel ComEC (Acpr_0619), which, in concert with ComFC (Acpr_0610), transports ssDNA into the cytoplasm. One strand enters the cytosol, while the remaining strand is likely degraded by one of the predicted nucleases (Acpr_0441, 0485, 0702, 0847) and the degradation products can either be released into the extracellular milieu or transported inside the cell using one of 27 predicted transporters. Inside the cytoplasm, ssDNA-binding protein Ssb (Acpr_1004) and DNA-protecting protein DprA (Acpr_0341) interact with the ssDNA to facilitate attachment of ssDNA-specific exonuclease RecJ (Acpr_0315), which ultimately degrades the ssDNA to nucleotides.

In addition to direct use for the synthesis of nucleic acids, these important building blocks can be sequentially converted into nucleoside monophosphates and other phosphates by nucleotide pyrophosphohydrolase (Acpr_0099), and soluble inorganic pyrophosphatase (Acpr_0997). Considering that conversion of one pyrophosphate molecule into two phosphate ions is strongly exergonic, this reaction may be associated with the synthesis and/or decomposition of lipids, and other unfavourable biochemical transformations (Carman and Han, 2006).

*Repertoire of '*Ca*. Abs. praedator' genome-defence mechanisms*

Nutritional competence, by definition, implies the active uptake and transport of exogenous DNA. Like other bacteria, the genome-reduced '*Ca*. Abs. praedator' had to develop defence mechanisms against plasmids, viruses, and other potential mobile agents that could threaten the integrity of its genome. Similar to what was found in '*Ca*. Vampirococcus' (Moreira *et al*., 2021), '*Ca*. Abs. praedator' possesses the AbiEii Type IV toxin–antitoxin

system (Acpr_1018-19) and two operons of the three-component Type I restriction–modification system (Acpr_0059-66 and Acpr_0907-12). Unlike the majority of CPR phyla that are impoverished in the CRISPR-Cas systems (Burstein *et al*., 2016, 2017), and, again similarly to '*Ca*. Vampirococcus' (Moreira *et al*., 2021), the genome of '*Ca*. Abs. praedator' encodes multiple CRISPR-Cas loci (Supporting Information Table S5 and Fig. S4). However, while '*Ca*. Vampirococcus' possess only two CRISPR-Cas systems, belonging to Type III-A and Type V-A, the genome of '*Ca*. Abs. praedator' possesses an additional one, resembling those of Type I-B.

The main function proposed for CRISPR-Cas systems relates to protection against genomic mobile agents, especially phages, so we checked the similarity between '*Ca*. Abs. praedator' spacers and virus sequences deposited in the NCBI database. Among a total of 81 spacers, only one, associated with a Type V-A CRISPR-Cas system, was found to be similar to the corresponding DNA fragment from the marine virus AFVG_117M50; 93.55% identity with 100% of coverage. The rest of the spacers were more or less similar to DNA fragments of other organisms, regardless of the domain of life (Supporting Information Table S5), suggesting an additional role of '*Ca*. Abs. praedator' CRISPR-Cas systems in the (intracellular) degradation of eDNA (Moreira *et al*., 2021). Notably, none of the '*Ca*. Abs. praedator' CRISPR-Cas spacers was similar to any sequences of the *Halorhodospira* genome. It is also noteworthy that '*Ca*. Abs. praedator' M39-6 and *H. halophila* M39-5 differ significantly in their GC genome contents (27.9 and 65.4 molGC% respectively).

Thus, recent HGT between these associated organisms can be detectable as an abnormality in the GC content(s) of particular region(s) and can further be analysed by parametric methods (Ravenhall *et al*., 2015). Eight genes with GC contents above 40% but lower 58% were found in the '*Ca*. Abs. praedator' genome (Supporting Information Table S6). All but one, Acpr_0894 (40.23 GC mol%), lack an opal codon (UGA) in their transcribed regions. The HGT-detection tool of the T-REX online portal (Boc *et al*., 2010, 2012) showed that only four out of the eight genes with atypically high GC contents can be considered HGT candidates. Among them there is an Acpr_0948 gene encoding a 130-aa long C5-cytochrome, mentioned above. At the N-terminus, this protein possesses Sec/SPI signal peptide with the cleavage site between positions 19 and 20. Interestingly, the amino-acid sequence of the trimmed part of Acpr_0948 resembles a 102-amino-acid-long c-type cytochrome of *Halorhodospira* sp. M39-5 and *H. halophila* $SL1^T$ (WP_011812935) (in each case, with 56% identity).

In the context of our hypothesis that the C5-cytochrome Acpr_0948 is involved in the functioning of the suggested respiratory chain then the use of the host cytochrome by '*Ca*. Abs. praedator' electron-transporting system is a realistic assumption. An additional candidate recognized by the HGT detection tool was the predicted extracellular phosphatase of the PAP2 family, Acpr_0497 (38.35 molGC%). Like other '*Ca*. Abs. praedator' HGT candidates, this gene lacks the opal codons (TGA) and its translated primary structure was found to be similar to that of the corresponding phosphatase Hhal_0806 of *H. halophila* $SL1^T$ (WP_011813605) (Supporting Information Table S6). Apart from these two, no other protein has been found in common, indicating very limited HGT between '*Ca*. Abs. praedator' and its host.

As mentioned previously (Campbell *et al*., 2013), in all members of the Gracilibacteria lineage the expression of any transcribed foreign genetic material possessing TGA could result in read-through of these stop codons, potentially affecting protein folding and enzymatic activity. Thus, we believe that, together with the reassignment of the TGA codon, variation in GC content exacerbates the genetic incompatibility between the vampire and its host. We also believe that imported host DNA is used by '*Ca*. Abs. praedator' only as a source of nucleotides, phosphates, and energy.

To further test this hypothesis, the '*Ca*. Abs. praedator' genome was checked for the presence of genomic islands using IslandViewer 4 (Bertelli *et al*., 2017). None of the five putative genomic islands predicted in '*Ca*. Abs. praedator' indicated any specific relationship with either the host *H. halophila* M39-5 or its phylogenetic neighbours. All of them contained genes for transposase or integrases and likely represented CPR-specific mobile elements (the mobilome) (Supporting Information Table S7). Subsequently, the '*Ca*. Abs. praedator' genome was the subject of virus/prophage searches using two different versions of the PHAge Search Tool (Zhou *et al*., 2011; Arndt *et al*., 2016). One genomic fragment was identified as a putative prophage fragment at position 179 110–187 562 bp (Supporting Information Table S7). Based on the absence of key proteins of the viral structure (capsid, head, plate, tail, coat, etc.); DNA-regulation genes encoding integrase, transposase and terminase; and functional genes (such as those coding for lysin and/or bacteriocin), a fragment was classified as an incomplete prophage with a slightly lower GC content (23.7 molGC%) than that of '*Ca*. Abs. praedator' genome.

Genome-inferred nutritional dependence of 'Ca*. Abs. praedator' on the host*

Extensive auxotrophy makes '*Ca*. Abs. praedator' completely dependent on the host for essential metabolic

precursors such as amino acids, vitamins, cofactors, and (phospho)lipids. This dependence is also evidenced by the presence of the 27 putative transporters and 110 membrane-associated proteins (which each has three or more transmembrane domains), many of which are hypothetical and likely play roles in the acquisition of essential metabolites (i.e. those required from the host). Of these proteins, 103 possess both SP(Sec/SPI) and LIPO(Sec/SPII) signal peptides and are believed to be transported across the membrane via the general Sec secretion system. Accordingly, six different genes have been recognized as components of the Sec protein-trafficking system: SecA (Acpr_0488), SecD (Acpr_0186), SecF (Acpr_0711), SecG (Acpr_1094), SecY (Acpr_0749), and membrane protein insertase YidC (Acpr_0256). Signal peptidase (SPase) I (Acpr_0680) and lipoprotein signal peptidase (SPase) II (Acpr_0462) are the principal intra-membrane peptidases responsible for the processing of most of the exported '*Ca*. Abs. praedator' proteins to amino acids used further in variety of biosynthetic processes.

Description of 'Candidatus *Absconditicoccus gen. nov'*

Absconditicoccus: (Abs.con.di.ti.coc'cus. L. masc. adj. *absconditus*, hidden; N.L. masc. n. *coccus* (from Gr. masc. n. *kokkos*), grain, seed; N.L. masc. n. *Absconditicoccus*, hidden coccus). Coccoid cell shape. Epibiotic lifestyle. Obligatory dependence on the host cells. Close cell–cell contact with host cells (<100 nm). Monoderm cell envelope with S-layer-like outer layer. Binary division through the formation of a septum. Phylogenetically belong to the order '*Ca*. Absconditabacteriales' within the class '*Ca*. Gracilibacteria' of the phylum Patescibacteria according to GTDB taxonomy. Identity of 16S rRNA sequences with '*Candidatus* Vampirococcus lugosii' is less than 90%. The type species of the genus is '*Candidatus* Absconditicoccus praedator'.

Description of 'Candidatus *Absconditicoccus praedator, sp. nov.' (prae.da'tor. L. masc. n. praedator, the pillager)*

Description of the species is identical to the genus description with following additional details. The host organism is *H. halophila*. Degradation of the host cells likely commences upon attachment. Average cell diameter is 497 nm; minimum cell diameter is around 323 nm and the DNA G + C content is 27.9%. The complete genome sequence has been deposited in GenBank under the accession number CP054059. Reliable morphological and ecological descriptions of the novel phylogenetically related representatives within the GTDB-inferred Absconditabacterales order, namely, '*Ca*. Absconditicoccus' praedator' and '*Ca*. Vampirococcus lugosii' allowed us to propose a family name encompassing these candidate genera and species: Absconditicoccaceae. This name refers to the morphology of both described representatives and incorporates them into the existing genome-based Patescibacteria taxonomy.

Description of 'Candidatus *Absconditicoccaceae' fam. nov.*

(Abs.con.di.ti.coc.ca.ce'ae. N.L. masc. n. *Absconditicoccus*, a Candidatus bacterial genus; *-aceae*, ending to denote a family; N.L. fem. pl. n. *Absconditicoccaceae*, the family of the genus *Absconditicoccus*). Epibiotic coccoid small Patescibacteria belonging to the order '*Ca*. Absconditabacterales' within the class '*Ca*. Gracilibacteria'. The known hosts are halophilic phototrophic obligate anaerobic purple Gammaproteobacteria.

Conclusion

'*Ca*. Absconditicoccus praedator' is the first long-term cultivable representative of the Gracilibacteria lineage of the Patescibacteria phylum. As with a few other cultivable Patescibacteria, this bacterium exhibits an ectoparasitic/predatory lifestyle rather than an ectosymbiotic one. If sufficient time is not taken to observe the progression of this interaction, it could be mistaken for symbiotic. However, the reduced genome of '*Ca*. Absconditicoccus praedator' alludes to dependency on and exploitation of the host; the gradual consumption of host cytoplasm constitutes vampire-like behaviour; and the resulting death of host cells mean that this CPR member is definitely a predator. While representing two separate genera, '*Ca*. Absconditicoccus' and '*Ca*. Vampirococcus' appear to have a similar biology; an obligate parasitic lifestyle, feeding on photosynthetic anoxygenic Gammaproteobacteria, and completely absorbing the host cytoplasm. Moreover, observed similarities in their genomic composition and structure suggest that the entire Absconditabacterales lineage consists of predatory species.

We were able to uncover and dissect the mechanics of (what turned out to be) the predator : host association, providing evidence that vampire-like CPR bacteria may play a potent role by controlling the size of host's population (which is in our case a primary producer) in high-density ecosystems, including some hypersaline habitats. Top-down control of bacterial populations via a 'kill-the-winner' mechanism is not unprecedented in microbial ecosystems. Indeed, a recent study reported the activities of another predatory bacterium, '*Ca*. Mycosynbacter amalyticus' (Batinovic *et al*., 2021). What is unusual about '*Ca*. Abs. praedator' is that it culls bacteria in an

extreme habitats and may represent a final turn of the screw for their host populations if they are already stressed/constrained by life-limiting conditions. Cultivation of predatory CPR bacteria can shed light on unresolved issues of CPR : host metabolism and other aspects of their predatory biology in the context of ecosystem dynamics at life-limiting extremes. However, much of the biology of the vampire : host interaction remains enigmatic.

Unresolved questions include: can this epibiont obtain nutrients other than those taken from the host; does *'Ca*. Abs. praedator' achieve osmotic adjustment using salt-in or salt-out strategies (possibly by uptaking glycine betaine from the host); what are the catabolic capabilities of *'Ca*. Abs. praedator' when assimilating host macromolecules; we know that *'Ca*. Abs. praedator' lives in alkaline brine, and survived long-term storage, but how robust is its stress biology in relation to factors such as ultraviolet radiation, dehydration–rehydration events, and hypo-osmotic shock; has *'Ca*. Abs. praedator' in some way(s) enhanced the fitness of the host during its evolutionary trajectory; how does *'Ca*. Abs. praedator' identify and anchor to the host cell (or does the host 'find' and attach to the '*Ca*. Abs. praedator'); and otherwise, what is the mechanism that determines epibiont : host specificity?

Experimental procedures

Bringing Absconditabacterales into enrichment culture

Enrichment culture of *H. halophila* M39-5 and *'Ca*. Absconditicoccus praedator' M39-6 was obtained from sediments of hypersaline alkaline Lake Hotontyn Nur (northeastern Mongolia) in September 1999 using medium following composition (g L^{-1}): NH_4Cl – 0.50, KH_2PO_4 – 0.50, $MgCl_2{\cdot}6H_2O$ – 0.50, $CaCl_2{\cdot}2H_2O$ – 0.05, $Na_2S_2O_3{\cdot}5H_2O$ – 1.00, $Na_2S{\cdot}9H_2O$ – 0.50, NaCl – 150, $NaHCO_3$–25.0, and Na_2CO_3–25.0, vitamin B12–20 μg L^{-1}, and trace metal solution – 1 ml (Bryantseva *et al*., 1999). The pH was adjusted to 9.0–9.5. Cultivation was carried out in glass serological bottles with rubber stoppers or sealed Hungate tubes and the culture was incubated under illumination (bulb lamp) at 30–37°C. The association of *H. halophila* M39-5 with the '*Ca*. Absconditicoccus praedator' M39-6 was enriched by serial dilutions and subsequential passages through the fresh media.

Bringing Absconditabacterales into binary culture

The enrichment culture was transferred to an autotrophic medium (modified DSM medium #253) following composition (g L^{-1}): KH_2PO_4 – 0.5, $CaCl_2{\cdot}2H_2O$ – 0.05, NH_4Cl – 0.8, $MgCl_2{\cdot}6H_2O$ – 0.30, NaCl – 150.00, $Na_2SO_4{\cdot}5H_2O$ – 1.00, Na_2CO_3 – 6.00, $NaHCO_3$ – 14.00 and $Na_2S{\cdot}9H_2O$ – 1.00, and trace element solution SL-8 – 1 ml, extended Basal Medium Eagle (BME) vitamin solution – 2 ml [Sigma-Aldrich #B6891-100ML, supplied additionally with biotin (0.4 μg ml^{-1}), thiamine HCl (10 μg ml^{-1}), and cyanocobalamin (10 μg ml^{-1})]. The pH was adjusted to 9.0–9.5. The medium was prepared according to the published recommendation for DSM medium #253 (https://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium253.pdf). The culture was incubated at 43°C with illumination by halogen lamp (45–48 μmol m^{-2} s^{-1}).

Serial dilutions (10×, 100×, 500×, 1000×, 10 000×) were prepared with the enrichment culture on the autotrophic medium. Then, 30 μl of each dilution was spread onto agar plates of modified DSM medium #253; while the trace element solution SL-8 and the extended BME vitamin solution were used instead of the original recommended solutions. The agar plates were incubated in Oxoid AnaeroJar 2.5-l jar with anaerobic atmosphere generation bag (Sigma-Aldrich #68061-10SACHETS-F) and anaerobe indicator test (Sigma-Aldrich #59886-1PAK-F) at 43°C with illumination by halogen lamp (45–48 μmol m^{-2} s^{-1}).

Purple colonies of *H. halophila* M39-5 were screened for presence of 'Ca. Absconditicoccus praedator' by colony PCR after a 5-day incubation. The following primers were used for the colony PCR: VampirF 5′-GCAA GTCGAGCGATCCCTAC-3′, VampirR 5′-GCGTTAGCT TACGGCACTCC-3′. PCR conditions were as follows: 95°C for 30 s, (95°C for 10 s; 52°C for 30 s; 72°C for 30 s) for a total of 30 cycles, followed by a final incubation at 72°C for 2 min. Infected and non-infected colonies were used for inoculation vials with the fresh autotrophic medium. The vials were incubated at 43°C with illumination by a halogen lamp (45–48 μmol m^{-2} s^{-1}) for 9 days and then screened for the presence '*Ca*. Absconditicoccus praedator' by fluorescence light microscopy. The infected culture was used for the second iteration of the same procedures: plating, colony PCR, inoculating in the autotrophic medium, and microscopy screening. Both enrichment and binary culture are freely available for use in academic research from the authors and will be provided upon request.

Light microscopy

Phase-contrast microscopy was performed using Olympus BX41 light microscope. Fluorescent microscopy was performed using Leica Thunder Imager 3D Cell Culture equipped with a Leica DFC9000 GTC CMOS camera. DNA of the cells was strained by NucBlue™ LiveReadyProbes™ Reagent (Life Technologies #R37605) according to the manufacturer's recommendations, prior to microscope analysis. To stain the

cytoplasmic membrane, FM 5-95 dye (Invitrogen #T23360) was added to cell suspension to a final concentration 5 μg ml^{-1} and incubated for 1 min before analysis. Cell diameter was measured manually by taking a random diameter using Fiji software (Schindelin *et al*., 2012).

Transmission electron microscopy of stained cells

The cells were fixed by mixing the cell suspension with 2% wt./vol. paraformaldehyde solution (on a 1:1 vol./vol. basis) following incubation at 4°C for 1 h. A 5-μl aliquot of the fixed cells were applied to a copper grid. The grid was then incubated for 10 min at room temperature and blotted by filter paper. Then a drop of aqueous solution of uranyl acetate (1% wt./vol.) was applied to the grid with following incubation for 30 s and blotting away. One drop of distilled water was applied to the stained grid and then the excess was blotted away. Finally, the grid was dried at room temperature.

For ultra-thin cross-sections, the cells were fixed in an OsO_4 solution (1% wt./vol. OsO4 + 0.7% wt./vol. ruthenium red solution in cacodylate buffer) (followed by incubation at 4°C for 90 min). After fixation, the cells were embedded into 2% wt./vol. agar-agar and successively incubated in 3% wt./vol. solution of uranyl acetate at 4°C: first, 4 h in the 30% vol./vol. ethanol-based solution, and second, 12 h in the 70% vol./vol. ethanol-based solution. The stained cells were dehydrated in 96% vol./vol. ethanol (two times for 15 min) and absolute acetone (three times for 10 min). The dehydrated cells were incubated for 1 h first in 1:1 (vol./vol.), and then for 1 h in 2:1 (vol./vol.) EPON-acetone mix (Epoxy Embedding Medium EPON® 812, Sigma-Aldrich). After this incubation, the cells were embedded into capsule with EPON resin. The resin was polymerized at 37°C for 24 h, and then at 60°C for 24 h. Ultra-thin cross-sections were made using an LKB-III microtome (LKB), and then contrasted by incubation in 3% vol./vol. aqueous solution of uranyl acetate and 4% vol./vol. aqueous solution of lead citrate for 30 min each. Images of the cells were obtained using a JEM 100CXII transmission electron microscope (JEOL) equipped with a Morada G2 digital camera using accelerating voltage of 80 kV and a working magnification of 5000–50 000×.

Plunge freezing and cryo-electron tomography

For cryo-electron tomography, the cell suspension was collected from Hungate tubes using a syringe with a 0.3 mm needle. The cell suspension was mixed with Protein A – 10 nm Gold Conjugate (Cytodiagnostics #AC-10-05-05) in a proportion of 1:10 vol./vol. before freezing. A 3.5-μl aliquot of the mixture was applied to a glow-discharged copper grid (200HQ R2/1, Quantifoil). PELCO easiGlow™ (Ted Pella) was used for glow discharging (15 mA for 15 s). The grid was blotted from the backside for 6 s (22°C, relative humidity 100%) and plunge-frozen automatically by immersion into liquid ethane-propane using Vitrobot Mark IV (Thermo Fisher Scientific). For the blotting, a Teflon™ pad was used on the front side of the grid where the sample was applied. The grid was stored in liquid nitrogen before image collection.

Images of the plunge-frozen cells were collected on a Titan Krios transmission electron microscope (Thermo Fisher Scientific) equipped with imaging filters and K3 electron detectors (Gatan) using an accelerating voltage of 300 kV. Automated tilt series were collected using SerialEM (Mastronarde, 2005) in the range of −60° to +60°C with 2°C increments. The defocus was −8 or −9 μm with a total dose of 154 e-/$Å^2$. IMOD program package was used to reconstruct tomograms and prepare segmentations (Mastronarde and Held, 2017). The images used to construct figures for the current article were retrieved from tomograms that have been processed with a deconvolution filter (Tegunov and Cramer, 2019) with the following parameters: angpix = 10.74, defocus = −8, snrfalloff = 1, deconvstrength = 0.5, highpassnyquist = 0.05, phaseflipped = true, phaseshift = 0. The cell diameter was measured manually by taking random diameter using 3dmod in the IMOD program package (Mastronarde and Held, 2017). Representative tomograms were deposited in EMDB under accession numbers EMD-13729 and EMD-13730.

DNA extraction, 16S rRNA gene amplicon sequencing and analysis

In order to verify axenic conditions of the binary culture, genomic DNA of the culture was extracted using QIAamp DNA Mini Kit (QIAGEN #51304). 16S rDNA next-generation sequencing library preparations and Illumina MiSeq sequencing were conducted at GENEWIZ (South Plainfield, NJ, USA). Sequencing library was prepared using a MetaVx™ 16S rDNA Library Preparation kit (GENEWIZ). The DNA was used to generate amplicons that cover V4 hypervariable regions of bacteria and archaea 16S rDNA. Indexed adapters were added to the ends of the 16S rDNA amplicons by limited cycle PCR. The samples were sequenced using a 2 × 250 paired-end configuration on the Illumina MiSeq instrument according to the manufacturer's instructions (Illumina, San Diego, CA, USA). Amplicon sequence variant table was produced using Dadaist2 (Ansorge *et al*., 2021). The sequencing data are deposited in NCBI BioProject PRJNA758801.

Metagenome library preparation, sequencing, and analysis

Shotgun metagenome library preparation and sequencing were done by BioSpark (Moscow, Russia) using KAPA HyperPlus Kits (Roche, Switzerland) according to the manufacturer's protocol and NovaSeq 6000 system (Illumina) using a 2 × 100 paired-end configuration. Raw reads were processed with Trimmomatic (Bolger *et al*., 2014) for adapter removal and quality filtering. Reads were then assembled using MetaSPAdes (Nurk *et al*., 2017). Raw reads have been submitted to GenBank Sequence Read Archive under BioProject number PRJNA758801. Reads were mapped back to the assembly using Bowtie2 software (Langmead and Salzberg, 2012) for coverage calculation. Binning of contigs into the MAGs was done by MaxBin 2 (Wu *et al*., 2016), MetaBAT 2 (Kang *et al*., 2019) and CONCOCT (Alneberg *et al*., 2014). Further aggregation and dereplication of these three sets of bins was done using the DAS Tool (Sieber *et al*., 2018). Bin completeness and contamination were evaluated using CheckM (Parks *et al*., 2015). Taxonomies were assigned to each bin using GTDBtk (Parks *et al*., 2018).

In order to improve the previous assembly, employing joined reads and to create a *de-novo* assembly graph file (gfa) to be viewed by Bandage program (Wick *et al*., 2015), the resulting 21 291 066 paired reads, and 1 318 648 joined, were assembled by SPAdes 3.14.1 software (Bankevich *et al*., 2012) with default parameters at first, then re-assembled by the Unicycler program (Wick *et al*., 2017). The produced MAGs were binned by MetaBAT 2 (Kang *et al*., 2019), while Geneious Prime 2020 software (Biomatters, New Zealand) was used to extract, map and join binned MAGs generated from the above-mentioned assembly procedure, verify consistency, and close the genome. Final genome coverage was about 510× for host *H. halophila* M39-5 and 252× for '*Ca*. Absconditicoccus praedator' M39-6. The genome of '*Ca*. Absconditicoccus praedator' M39-6 was deposited in the GenBank under accession number CP054059.

Phylogenetic analysis

A genome-based phylogenetic tree was constructed using the list of 120 bacterial marker genes taken from the GTDB (Parks *et al*., 2018). These marker genes were identified in studied bins using Prodigal v2.6.3 (Hyatt *et al*., 2010), aligned using MAFFT v7.427 (Nakamura *et al*., 2018), trimmed using trimAl 1.2rev59 (Capella-Gutiérrez *et al*., 2009) and then concatenated. A 16S rRNA-based phylogenetic tree was built using sequences from the SILVA database. Alignment and trimming were done as described above. The trees were built using the IQ-TREE 2 program (Minh *et al*., 2020) with fast model selection via ModelFinder (Kalyaanamoorthy *et al*., 2017) and ultrafast bootstrap approximation (Minh *et al*., 2013) as well as approximate likelihood-ratio test for branches (Anisimova and Gascuel, 2006).

Gene prediction and annotation

Prediction of Candidate division SR1 bacterium M39-6 protein-coding genes was performed using Glimmer 3.02 (Delcher *et al*., 2007) plugin inside Geneious Prime 2020 software, after arrangement of stop codons to match the Candidate Division SR1 and Gracilibacteria code (transl_table = 25). The resulting protein sequences were manually annotated via BLASTP and BLASTX searches against the NCBI NR database, integrated with information got from COGs Galperin *et al*., 2015), KEGG pathways (Kanehisa *et al*., 2017, http://www.genome.jp/tools/blast/), Pfam (El-Gebali *et al*., 2019), Conserved Domain Database (Lu *et al*., 2020), and TIGRFAMs (Haft *et al*., 2013).

An additional annotation check was performed by the eggNOG-mapper online tool (Huerta-Cepas *et al*., 2017 and Huerta-Cepas *et al*., 2019, http://eggnog-mapper.embl.de/). The tRNAscan-SE 2.0 online tool (Lowe and Chan, 2016) was used to predict tRNA genes, whereas RNAmmer 1.2 online tool (Lagesen *et al*., 2007) was used for rRNA genes. N-glycosylation sites were predicted using the NetNGlyc 1.0 Server (https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0). Prediction of transmembrane regions was carried out by the TMHMM 2.0 Server (Krogh *et al*., 2001) (https://services.healthtech.dtu.dk/service.php?TMHMM-2.0). Signal peptides prediction were performed by the SignalP 5.0 online tool (Almagro Armenteros *et al*., 2019) (https://services.healthtech.dtu.dk/service.php?SignalP-5.0). Prediction of genomic islands was performed using the IslandViewer4 online tool (Bertelli *et al*., 2017). Identification of phage regions inside genomes was performed by PHAST online search tool (Zhou *et al*., 2011). CRISPR-Cas loci were identified by CRISPRFinder online program (Grissa *et al*., 2007). Detection of orthologous genes from M39-6 and its host M39-5 was performed by Proteinortho 5 program (Lechner *et al*., 2011) with default parameters.

Acknowledgements

We are grateful to Aharon Oren (The Hebrew University of Jerusalem, Israel) for taxonomy- and species description-related discussions. V.M.G. was supported by research funds from the Russian Foundation for Basic Research (RFBR grant 19-04-00423). V.M.G., A.A.K. and A.Y.M. were supported by research funds from Russian Ministry of Science and Higher Education. M.P. was supported by the Swiss National Science Foundation (31003A_179255), the

European Research Commission (679209) and the NOMIS Foundation. The ETH facility ScopeM is acknowledged for instrument access. E.M. and M.M.Y. were supported by the European Commission's Horizon 2020 Program under INMARE Project (Contract 634486) and FUTURENZYMES Project (Contract 101000327).

Author Contributions

Experimental work and microscopy were carried out by A.A.K., E.N.T., V.M.G. and V.A.G. (who did the cryo-electron tomography and fluorescent microscopy, and preparation of pure cultures of host and Patescibacterium : host association). Bioinformatics analyses were carried out by A.Y.M., E.M. and M.M.Y. The data were interpreted and the manuscript was written by J.E.H., M.M.Y., V.A.G. and V.M.G. M.P. had an advisory role in the aspects of experimental work and input into the writing of the manuscript.

References

Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., *et al*. (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* **37**: 420–423.

Alneberg, J., Bjarnason, B.S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., *et al*. (2014) Binning metagenomic contigs by coverage and composition. *Nat Methods* **11**: 1144–1146.

Angert, E.R., Northup, D.E., Reysenbach, A.-L., Peek, A.S., Goebel, B.M., and Pace, N.R. (1998) Molecular phylogenetic analysis of a bacterial community in Sulphur River, Parker cave, Kentucky. *Am Mineral* **83**: 1583–1592.

Anisimova, M., and Gascuel, O. (2006) Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst Biol* **55**: 539–552.

Ansorge, R., Birolo, G., James, S.A., and Telatin, A. (2021) Dadaist2: a toolkit to automate and simplify statistical analysis and plotting of metabarcoding experiments. *Int J Mol Sci* **22**: 5309.

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., *et al*. (2016) PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* **44**: W16–W21.

Atanasova, N.S., Bamford, D.H., and Oksanen, H.M. (2016) Virus-host interplay in high salt environments. *Environ Microbiol Rep* **8**: 431–444.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., *et al*. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**: 455–477.

Batinovic, S., Rose, J., Ratcliffe, J., Seviour, R.J., and Petrovski, S. (2021) Cocultivation of an ultrasmall environmental parasitic bacterium with lytic ability against bacteria associated with wastewater foams. *Nat Microbiol* **6**: 703–711.

Beam, J.P., Becraft, E.D., Brown, J.M., Schulz, F., Jarett, J. K., Bezuidt, O., *et al*. (2020) Ancestral absence of electron transport chains in Patescibacteria and DPANN. *Front Microbiol* **11**: 1848.

Bertelli, C., Laird, M.R., Williams, K.P., Simon Fraser University Research Computing Group, Lau, B.Y., Hoad, G., *et al*. (2017) IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res* **45**: W30–W35.

Boc, A., Diallo, A.B., and Makarenkov, V. (2012) T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. *Nucleic Acids Res* **40**: W573–W579.

Boc, A., Philippe, H., and Makarenkov, V. (2010) Inferring and validating horizontal gene transfer events using bipartition dissimilarity. *Syst Biol* **59**: 195–211.

Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Bor, B., Collins, A.J., Murugkar, P.P., Balasubramanian, S., To, T.T., Hendrickson, E.L., *et al*. (2020) Insights obtained by culturing Saccharibacteria with their bacterial hosts. *J Dent Res* **99**: 685–694.

Bor, B., McLean, J.S., Foster, K.R., Cen, L., To, T.T., Serrato-Guillen, A., *et al*. (2018) Rapid evolution of decreased host susceptibility drives a stable relationship between ultrasmall parasite TM7x and its bacterial host. *Proc. Natl. Acad. Sci. U.S.A.* **115**: 12277–12282.

Brown, C.T., Hug, L.A., Thomas, B.C., Sharon, I., Castelle, C.J., Singh, A., *et al*. (2015) Unusual biology across a group comprising more than 15% of domain bacteria. *Nature* **523**: 208–211.

Bryantseva, I., Gorlenko, V.M., Kompantseva, E.I., Imhoff, J.F., Süling, J., and Mityushina, L. (1999) *Thiorhodospira sibirica* gen. nov., sp. nov., a new alkaliphilic purple sulfur bacterium from a Siberian soda lake. *Int J Syst Bacteriol* **49**: 697–703.

Budd, A., Blandin, S., Levashina, E.A., and Gibson, T.J. (2004) Bacterial alpha2-macroglobulins: colonization factors acquired by horizontal gene transfer from the metazoan genome? *Genome Biol* **5**: R38.

Burstein, D., Harrington, L.B., Strutt, S.C., Probst, A.J., Anantharaman, K., Thomas, B.C., *et al*. (2017) New CRISPR-Cas systems from uncultivated microbes. *Nature* **542**: 237–241.

Burstein, D., Sun, C.L., Brown, C.T., Sharon, I., Anantharaman, K., Probst, A.J., *et al*. (2016) Major bacterial lineages are essentially devoid of CRISPR-Cas viral defence systems. *Nat Commun* **7**: 10613.

Campbell, J.H., O'Donoghue, P., Campbell, A.G., Schwientek, P., Sczyrba, A., Woyke, T., *et al*. (2013) UGA is an additional glycine codon in uncultured SR1 bacteria from the human microbiota. *Proc Natl Acad Sci U S A* **110**: 5540–5545.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**: 1972–1973.

Carman, G.M., and Han, G.S. (2006) Roles of phosphatidate phosphatase enzymes in lipid metabolism. *Trends Biochem Sci* **31**: 694–699.

Castelle, C.J., and Banfield, J.F. (2018) Major new microbial groups expand diversity and alter our understanding of the tree of life. *Cell* **172**: 1181–1197.

Castelle, C.J., Brown, C.T., Anantharaman, K., Probst, A.J., Huang, R.H., and Banfield, J.F. (2018) Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. *Nat Rev Microbiol* **16**: 629–645.

Castelle, C.J., Wrighton, K.C., Thomas, B.C., Hug, L.A., Brown, C.T., Wilkins, M.J., *et al*. (2015) Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr Biol* **25**: 690–701.

Challacombe, J.F., Majid, S., Deole, R., Brettin, T.S., Bruce, D., Delano, S.F., *et al*. (2013) Complete genome sequence of *Halorhodospira halophila* SL1. *Stand Genomic Sci* **8**: 206–214.

Chaudhari, N.M., Overholt, W.A., Figueroa-Gonzalez, P.A., Taubert, M., Bornemann, T.L.V., Probst, A.J., *et al*. (2021) The economical lifestyle of CPR bacteria in groundwater allows little preference for environmental drivers. *bioRxiv*. https://doi.org/10.1101/2021.07.28.454184.

Chaumeil, P.A., Mussig, A.J., Hugenholtz, P., and Parks, D. H. (2019) GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* **36**: 1925–1927.

Chen, I., and Dubnau, D. (2004) DNA uptake during bacterial transformation. *Nat Rev Microbiol* **2**: 241–249.

Chen, L., Durley, R.C., Mathews, F.S., and Davidson, V.L. (1994) Structure of an electron transfer complex: methylamine dehydrogenase, amicyanin, and cytochrome c551i. *Science* **264**: 86–90.

Corinaldesi, C., Beolchini, F., and Dell'Anno, A. (2008) Damage and degradation rates of extracellular DNA in marine sediments: implications for the preservation of gene sequences. *Mol Ecol* **17**: 3939–3951.

Cross, K.L., Campbell, J.H., Balachandran, M., Campbell, A. G., Cooper, S.J., Griffen, A., *et al*. (2019) Targeted isolation and cultivation of uncultivated bacteria by reverse genomics. *Nat Biotechnol* **37**: 1314–1321.

Delcher, A.L., Bratke, K.A., Powers, E.C., and Salzberg, S.L. (2007) Identifying bacterial genes and endosymbiont DNA with glimmer. *Bioinformatics* **23**: 673–679.

Dell'Anno, A., and Danovaro, R. (2005) Extracellular DNA plays a key role in deep-sea ecosystem functioning. *Science* **309**: 2179.

Dombrowski, N., Lee, J.H., Williams, T.A., Offre, P., and Spang, A. (2019) Genomic diversity, lifestyles and evolutionary origins of DPANN archaea. *FEMS Microbiol Lett* **366**: fnz008.

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., *et al*. (2019) The Pfam protein families database in 2019. *Nucleic Acids Res* **47**: D427–D432.

Evans, K.J., Hobley, L., Lambert, C., and Sockett, R.E. (2007) *Bdellovibrio*: lone hunter "cousin" of the "pack hunting" myxobacteria. In *Myxobacteria: Multicellularity and Differentiation*, Whitworth, D.E. (ed). Washington, DC: ASM Press, pp. 351–362.

Galperin, M.Y., Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2015) Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res* **43**: D261–D269.

Grissa, I., Vergnaud, G., and Pourcel, C. (2007) The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* **8**: 172.

Guerrero, R., Pedros-Alio, C., Esteve, I., Mas, J., Chase, D., and Margulis, L. (1986) Predatory prokaryotes: predation and primary consumption evolved in bacteria. *Proc Natl Acad Sci U S A* **83**: 2138–2142.

Haft, D.H., Selengut, J.D., Richter, R.A., Harkins, D., Basu, M.K., and Beck, E. (2013) TIGRFAMs and genome properties in 2013. *Nucleic Acids Res* **41**: D387–D395.

Hallsworth, J.E. (2019) Microbial unknowns at the saline limits for life. *Nat Ecol Evol* **3**: 1503–1504.

Hanke, A., Hamann, E., Sharma, R., Geelhoed, J.S., Hargesheimer, T., Kraft, B., *et al*. (2014) Recoding of the stop codon UGA to glycine by a BD1-5/SN-2 bacterium and niche partitioning between alpha- and Gammaproteobacteria in a tidal sediment microbial community naturally selected in a laboratory chemostat. *Front Microbiol* **5**: 231.

He, X., McLean, J.S., Edlund, A., Yooseph, S., Hall, A.P., Liu, S.Y., *et al*. (2015) Cultivation of a human-associated TM7 phylotype reveals a reduced genome and epibiotic parasitic lifestyle. *Proc Natl Acad Sci U S A* **112**: 244–249.

Hua, Z., Ouellette, M., Makkay, A.M., Papke, R.T., and Zhaxybayeva, O. (2021) Nutrient supplementation experiments with saltern microbial communities implicate utilization of DNA as a source of phosphorus. *ISME J* **15**: 2853–2864. https://doi.org/10.1038/s41396-021-00960-8.

Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering, C., *et al*. (2017) Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol Biol Evol* **34**: 2115–2122.

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., *et al*. (2019) eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* **47**: D309–D314.

Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., *et al*. (2016) A new view of the tree of life. *Nat Microbiol* **1**: 16048.

Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F. W., and Hauser, L.J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**: 119.

Ivanova, N.N., Schwientek, P., Tripp, H.J., Rinke, C., Pati, A., Huntemann, M., *et al*. (2014) Stop codon reassignments in the wild. *Science* **344**: 909–913.

Jaffe, A.L., Castelle, C.J., Matheus Carnevali, P.B., Gribaldo, S., and Banfield, J.F. (2020) The rise of diversity in metabolic platforms across the candidate phyla radiation. *BMC Biol* **18**: 69.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T., von Haeseler, A., and Jermiin, L.S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**: 587–589.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**: D353–D361.

Kang, D.D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., *et al*. (2019) MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**: e7359.

Kantor, R.S., Wrighton, K.C., Handley, K.M., Sharon, I., Hug, L.A., Castelle, C.J., *et al*. (2013) Small genomes and sparse metabolisms of sediment-associated bacteria from four candidate phyla. *mBio* **4**: e00708–e00713.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**: 567–580.

Lagesen, K., Hallin, P., Rødland, E.A., Staerfeldt, H.H., Rognes, T., and Ussery, D.W. (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* **35**: 3100–3108.

Langmead, B., and Salzberg, S.L. (2012) Fast gapped-read alignment with bowtie 2. *Nat Methods* **9**: 357–359.

Lechner, M., Findeiss, S., Steiner, L., Marz, M., Stadler, P. F., and Prohaska, S.J. (2011) Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinformatics* **12**: 124.

Lee, C., McMullan, P.E., O'Kane, C.J., Stevenson, A., Santos, I.C., Roy, C., *et al*. (2018) NaCl-saturated brines are thermodynamically moderate, rather than extreme, microbial habitats. *FEMS Microbiol Rev* **42**: 672–693.

Ley, R.E., Harris, J.K., Wilcox, J., Spear, J.R., Miller, S.R., Bebout, B.M., *et al*. (2006) Unexpected diversity and complexity of the Guerrero Negro hypersaline microbial mat. *Appl Environ Microbiol* **72**: 3685–3695.

Liu, J., Cvirkaite-Krupovic, V., Baquero, D.P., Yang, Y., Zhang, Q., Yulong Shen, Y., *et al*. (2021) Virus-induced cell gigantism and asymmetric cell division in archaea. *Proc Natl Acad Sci U S A* **118**: e2022578118.

Lowe, T.M., and Chan, P.P. (2016) tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res* **44**: W54–W57.

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., *et al*. (2020) CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res* **48**: D265–D268.

Luef, B., Frischkorn, K.R., Wrighton, K.C., Holman, H.Y., Birarda, G., Thomas, B.C., *et al*. (2015) Diverse uncultivated ultra-small bacterial cells in groundwater. *Nat Commun* **6**: 6372.

Mastronarde, D.N. (2005) Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**: 6–51.

Mastronarde, D.N., and Held, S.R. (2017) Automated tilt series alignment and tomographic reconstruction in IMOD. *J Struct Biol* **197**: 102–113.

Méheust, R., Burstein, D., Castelle, C.J., and Banfield, J.F. (2019) The distinction of CPR bacteria from other bacteria based on protein family content. *Nat Commun* **10**: 4173.

Minh, B.Q., Nguyen, M.A., and von Haeseler, A. (2013) Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* **30**: 1188–1195.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., *et al*. (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* **37**: 1530–1534.

Moreira, D., Zivanovic, Y., López-Archilla, A.I., Iniesto, M., and López-García, P. (2021) Reductive evolution and unique predatory mode in the CPR bacterium *Vampirococcus lugosii*. *Nat Commun* **12**: 2454.

Nakamura, T., Yamada, K.D., Tomii, K., and Katoh, K. (2018) Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics* **34**: 2490–2492.

Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P.A. (2017) metaSPAdes: a new versatile metagenomic assembler. *Genome Res* **27**: 824–834.

Parks, D.H., Chuvochina, M., Chaumeil, P.A., Rinke, C., Mussig, A.J., and Hugenholtz, P. (2020) A complete domain-to-species taxonomy for bacteria and archaea. *Nat Biotechnol* **38**: 1079–1086.

Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.A., *et al*. (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* **36**: 996–1004.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043–1055.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., *et al*. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590–D596.

Rasskazov, A.A., and Abramov, A.B. (1987) The soda lakes of Mongolia. *Lithol Mineral Resour* **6**: 88–99 (in Russian).

Ravenhall, M., Skunca, N., Lassalle, F., and Dessimoz, C. (2015) Inferring horizontal gene transfer. *PLoS Comput Biol* **11**: e1004095.

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.F., *et al*. (2013) Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**: 431–437.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., *et al*. (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**: 676–682.

Sieber, C., Paul, B.G., Castelle, C.J., Hu, P., Tringe, S.G., Valentine, D.L., *et al*. (2019) Unusual metabolism and hypervariation in the genome of a Gracilibacterium (BD1-5) from an oil-degrading community. *mBio* **10**: e02128-19.

Sieber, C., Probst, A.J., Sharrar, A., Thomas, B.C., Hess, M., Tringe, S.G., *et al*. (2018) Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol* **3**: 836–843.

Sleytr, U.B., Messner, P., Pum, D., and Sára, M. (1993) Crystalline bacterial cell surface layers. *Mol Microbiol* **10**: 911–916.

Sode, K., Sugimoto, S., Watanabe, M., and Tsugawa, W. (1995) Effect of PQQ glucose dehydrogenase overexpression in *Escherichia coli* on sugar-dependent respiration. *J Biotechnol* **43**: 41–44.

Sorokin, D.Y., Gorlenko, V.M., Namsaraev, B.B., Namsaraev, Z.B., Lysenko, A.M., Eshinimaev, B.T., *et al*.

(2004) Prokaryotic communities of the north-eastern Mongolian soda lakes. *Hydrobiologia* **522**: 235–248.

Sorokin, D.Y., Roman, P., and Kolganova, T.V. (2021) Halo (natrono)archaea from hypersaline lakes can utilize sulfoxides other than DMSO as electron acceptors for anaerobic respiration. *Extremophiles* **25**: 173–180.

Tegunov, D., and Cramer, P. (2019) Real-time cryo-electron microscopy data preprocessing with warp. *Nat Methods* **16**: 1146–1152.

Tindall, B.J., Rosselló-Móra, R., Busse, H.J., Ludwig, W., and Kämpfer, P. (2010) Notes on the characterization of prokaryote strains for taxonomic purposes. *Int J Syst Evol Microbiol* **60**: 249–266.

Uritskiy, G., Tisza, M.J., Gelsinger, D.R., Munn, A., Taylor, J., and DiRuggiero, J. (2021) Cellular life from the three domains and viruses are transcriptionally active in a hypersaline desert community. *Environ Microbiol* **23**: 3401–3417.

Wick, R.R., Judd, L.M., Gorrie, C.L., and Holt, K.E. (2017) Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* **13**: e1005595.

Wick, R.R., Schultz, M.B., Zobel, J., and Holt, K.E. (2015) Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* **31**: 3350–3352.

Wrighton, K.C., Thomas, B.C., Sharon, I., Miller, C.S., Castelle, C.J., VerBerkmoes, N.C., *et al*. (2012) Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. *Science* **337**: 1661–1665.

Wu, Y.W., Simmons, B.A., and Singer, S.W. (2016) MaxBin 2.0: an automated binning algorithm to recover genomes from multiple meta-genomic datasets. *Bioinformatics* **32**: 605–607.

Yarza, P., Richter, M., Peplies, J., Euzeby, J., Amann, R., Schleifer, K.H., *et al*. (2008) The all-species living tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst Appl Microbiol* **31**: 241–250.

Zhou, Y., Liang, Y., Lynch, K.H., Dennis, J.J., and Wishart, D.S. (2011) PHAST: a fast phage search tool. *Nucleic Acids Res* **39**: W347–W352.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Supporting Information.

Table S4. Ca. Absconditicoccus praedator' M39-6, 20 longest proteins.

Table S5. NCBI blastp best matches of CRISPR1-Cas proteins in 'Ca. Absconditicoccus praedator' M39-6.

Table S6. Groups of similar proteins identified in 'Ca. Absconditicoccus praedator' M39-6 and Halorhodspira halophila M39-5 genomes with ProteinOrtho. Halorhodspira halophila SL1 was used as reference.

Table S7. Ca. Absconditicoccus praedator' M39-6 annotations with identification of genomic islands, phage regions, and prediction of N-glycosylation, transmembrane helix and signal peptide sites.