 doi:10.20944/preprints202205.0070.v1

Article

Taxonomic Re-Evaluation and Genomic Comparison of Novel Extracellular Electron Uptake-Capable *Rhodovulum visakhapatnamense* and *Rhodovulum sulfidophilum* isolates

Davenport, Emily J.[1], Bose, Arpita[1*]

[1] Washington University in St. Louis, demily@wustl.edu

[2]Washington University in St. Louis, abose@wustl.edu

[*] Corresponding author, abose@wustl.edu

Abstract: *Rhodovulum* spp. are anoxygenic photosynthetic purple bacteria with versatile metabolisms, including the ability to obtain electrons from minerals in their environment to drive photosynthesis, a relatively novel process called phototrophic extracellular electron uptake (pEEU). Recently, our group isolated 15 strains of *R. sulfidophilum* to observe this metabolism in marine phototrophs. Our group previously observed carbon dioxide fixation coupled to phototrophic iron oxidation (photoferrotrophy) and pEEU in AB26 and identified a novel di-heme *c*-type cytochrome EeuP important for pEEU but not photoferrotrophy. Taxonomic re-evaluation based on 16S and *pufM* phylogenetic analyses led us to re-classify two isolates, AB26 and AB19, as *Rhodovulum visakhapatnamense*. The AB26 genome consists of 4,380,746 base-pairs, including two plasmids, and encodes 4,296 predicted protein-coding genes. AB26 contains 22 histidine

 doi:10.20944/preprints202205.0070.v1

kinases, 20 response regulators, and dedicates ~16% of its genome to transport. Transcriptomic data under aerobic, photoheterotrophy, photoautotrophy, and pEEU reveals how gene expression varies between metabolisms. Lastly, we use transcriptomic data for a comparative genomic analysis of potential pEEU-relevant genes between all 15 isolates. With these data we identify potential pEEU capable phototrophs within these isolates, and likely molecular mechanisms of pEEU.

Keywords: phototrophic bacteria 1; phototrophic extracellular electron uptake 2; comparative genomics 3; transcriptomics 4; environmental microbiology 5

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

1. Introduction

Bacterial metabolisms are diverse, allowing for growth in nearly any environment. Some bacteria are capable of obtaining electrons from insoluble extracellular sources within their environments via a process called extracellular electron uptake (EEU)[1]. Various forms of this metabolism exist with the common theme being that microbes interact with solid-phase conductive substances as sources of electrons[2]. Through EEU, microbes drive biogeochemical cycles in all corners of the earth at undetermined rates[3]. This metabolism also has potential to improve bioremediation of contaminated sites and the biosynthesis of sustainable fuels[4-6].

EEU was first observed in species such as mineral-reducing, *Shewanella oneidensis* MR1 and the photoautotroph, *Rhodopseudomonas palustris* TIE-1[7-9]. One well-characterized electron uptake system is known; the PioABC operon studied in TIE-1[10]. Studies of pEEU in TIE-1 revealed the Calvin-Benson-Bassham cycle as an electron sink during pEEU[11, 12]. This link to carbon dioxide fixation connects pEEU to the carbon cycle. Furthermore, because TIE-1 can use various soluble and insoluble forms of iron, pEEU connects to the cycling of iron and other elements that are bound to iron minerals[2, 13]. Until recently pEEU has been studied solely in freshwater phototrophs, leaving a knowledge gap regarding the prevalence of pEEU in other ecosystems.

In an effort to address this, we isolated bacteria from a microbial mat in a marine estuary in Woods Hole, Massachusetts (USA)[14, 15]. Fifteen of the isolates were initially identified as *Rhodovulum sulfidophilum*, a purple non-sulfur bacterium capable of many diverse metabolisms including anoxygenic photosynthesis (non-oxygen-evolving). We

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

sequenced and assembled the genomes of 12 of the 15 isolates based on these data (3 isolates namely AB14, AB26, AB30 were assembled previously[14]). We have previously reported that all 15 strains were capable of heterotrophic, photoautotrophic, and photoheterotrophic growth as well as photoferrotrophy[11]. A representative strain, AB26, was selected for further growth studies and found to be capable of pEEU in bioelectrochemical systems (BES) using a novel uptake pathway consisting of a di-heme *c*-type cytochrome (EeuP), with carbon dioxide fixation as a cellular sink for acquired electrons[11]. EeuP is found in diverse bacterial lineages without other characterized pEEU uptake systems in their genomes, which may suggest its role as a novel uptake system[11].

Here, we furthered our understanding of this novel pEEU uptake system in marine *R. sulfidophilum* and *R. visakhapatnamense* via comparative genomic studies of the 15 isolates relative to the AB26 transcriptomic data. This dataset produces an improved understanding of pEEU and its contributions to geochemical processes in marine environments. Through this study we identified other potential genes important to pEEU in *Rhodovulum* species, and potentially utilize these species in sustainable synthesis of biofuels.

2. Materials and Methods

Genome sequencing

Genome analysis was performed as previously described[14]. Genomic DNA was isolated from mid-log phase cell culture in marine broth using the DNeasy Blood and

Tissue Kit (Qiagen, USA). Samples were prepped for Illumina 250-bp paired-end sequencing using Nextera sample prep kit (Illumina, Inc., San Diego, CA, USA), and sequenced on Illumina MiSeq platform with V2 chemistry. Reads were quality checked and adapter trimmed using Trimmomatic version 0.33 using default parameters for paired-end reads[16]. Reads were *de novo* assembled using CLC Genomics Workbench (CLC Bio-Qiagen, Aarhaus, Denmark), and scaffolds generated using MeDuSa[17] and *R. sulfidophilum* DSM 2351 as an alignment guide. Alignment of reads to *R. sulfidophilum* DSM 2351 was performed using Bowtie2 version 2.2.29[18] short read mapper, and annotated with National Center for Biotechnology Information Prokaryotic Genome Annotation Pipeline[19].

Phylogenetic Analysis, Taxonomic Re-evaluation, and Comparative Genomics

ANI between strains was performed using JSpeciesWS[20] and visualized with Morpheus (Morpheus, https://software.broadinstitute.org/morpheus). Phylogenetic trees were built using MEGA11[21]. For strain-level phylogenetic analysis, 16S rRNA sequences and the photosynthetic reaction center subunit M (pufM) protein sequences were used to build trees. The 16S tree used Kimura 2-paramter model and gamma distribution of 5. The pufM tree used Jones-Taylor-Thornton (JTT) model and gamma distribution of 5. For phylogenetic analysis of EeuP protein sequences among *R. sulfidophilum* strains, the JTT model was used with a gamma distribution of 5. All sequences were aligned using the ClustalW algorithm. BLASTp was performed using a local protein database (Supplemental Materials File 2).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

RNA isolation, RNA sequencing, and differential expression analysis

As previously described[11], Cell cultures were sampled anaerobically and immediately mixed with 1:1 with RNAlater® (Qiagen, Germantown, MD, USA). RNA was extracted from cells using the RNeasy Mini Kit (Qiagen, USA) and DNA removal performed using Turbo DNA-*free* Kit (Ambion, Austin, Texas, USA). RNA samples tested for purity using PCR. Illumina single-end 50-bp libraries were prepared and sequenced at Washington University's Genome Technology Access Center on an Illumina HiSeq3000 (Illumina Inc., Madison, WI, USA). Reads were mapped to the AB26 genome using TopHat2 version 2.1.1 and the gff3 annotation file as a guide for sequence alignment. Bowtie 2 version 2.3.3.1 was used to index the reference genome FASTA file. The number of reads mapping to each feature were counted by HTSeq version 0.9.1. Differentially expressed genes were determined in DESEQ2 version 1.16.1 using the HTSeq read counts. Significantly differentially expressed genes were identified using an adjusted p-value cutoff of 0.05.

3. Results and Discussion

3.1 Taxonomic re-evaluation of AB26 and AB19

Fifteen isolates were cultured from a marine microbial mat in Woods Hole, MA, USA (henceforth referred to as Woods Hole isolates) and identified as belonging to *R. sulfidophilum* through 16S sequencing[11, 14]. Prior to sequencing the Woods Hole isolates, few *R. sulfidophilum* genomes were publicly available. Whole genome

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

alignments of AB26 showed only 57.8% of the genome sequence aligned to the available reference genome *R. sulfidophilum* DSM2351[14]. Growth experiments of select isolates revealed differences in doubling times in the same growth conditions[11], and therefore we sought to further investigate the phylogenetic relatedness of the isolated strains. 16S and photosynthetic reaction center subunit M (*pufM*) phylogenetic trees were produced to determine the relatedness between isolates from Woods Hole (Figure 1).

16S phylogenetic analysis shows AB26 and AB19 reside in a separate branch from all other Woods Hole isolates, with a second clade consisting of AB33, AB16, and AB30. The remaining Woods Hole isolates generally nest within a group together. The pattern of the 16S phylogenetic tree is mirrored in the photosynthetic reaction center subunit M (*pufM*) tree in that AB26 and AB19 again reside in a separate branch from the remaining Woods Hole isolates. Further investigation using average nucleotide identities (ANI) revealed AB26 and AB19 ANI values are far below the cut off value for same-species relatedness (>~94%, Figure 2)[22]. Remaining Woods Hole isolates all show ANI values indicative of same-species relatedness. Two groups of strains, AB18 and AB33, and AB35 and AB15 show highest similarity to each other. One other isolate, AB28, showed lower ANI values to all other strains (94-95% ANI) but were still within same species-values (Figure 2).

To determine the proper species designation of AB26 and AB19, ANI comparisons between AB26 and AB28 were performed to include the closely related species *Rhodovulum viride* and *Rhodovulum visakhapatnamese* (Supplemental Figure 1)[23, 24]. AB28 did not show any increased similarities to other *Rhodovulum* genomes, suggesting the current taxonomy is correct, but may change as more *Rhodovulum* genomes are

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

assembled. ANI analysis shows AB26 and AB19 share ANI values of 97.95% with *R. visakhapatnamense*, placing AB26 and AB19 in a new species (Figure 3).

3.2 AB26 Genome Features and Expression Analysis

AB26 was chosen as a representative strain for this study due to its preferable growth characteristics[11]. The genome of AB26 consists of 4,380,746 bp of DNA and a GC content of 67.9%, assembled into 3 scaffolds. AB26 has 4,375 genes, with 4,296 protein-encoding genes (Table 1). The AB26 genome contains genes allowing for diverse metabolic capabilities that purple non-sulfur bacteria are known for. This is in line with previous observations that AB26 is capable of aerobic heterotrophic growth in the dark, photoheterotrophy using succinate and acetate for carbon and electrons, and photoautotrophy, using inorganic substances, including thiosulfate and hydrogen (H_2).

3.2.1 Phototrophy

R. visakhapatnamense is a member of the purple non-sulfur bacteria capable of anoxygenic photosynthesis[24]. The genes necessary for growth via photosynthesis occur in a ~50-kb region of the genome spanning 4 separate operons. The first operon spans a 13-kb region and encodes the reaction center, bacteriochlorophyll synthesis, carotenoid synthesis, and light harvesting proteins. Reaction center subunits M

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

(BV509_00330) and L (BV509_00335) share 70% identity with the reaction center genes in the model anoxygenic phototroph *Cereibacter sphaeroides[25]*. The second photosynthetic gene cluster (PGC) consists of tetrapyrrole biosynthesis genes *bchEJ*. The third PGC is homologous to the *puc* operon, which encodes the structural polypeptides of the light-harvesting-II peripheral antenna complex, and also contains a PucC family protein[26]. PucC encoded in model *puc* operons regulates response to oxygen tension and light intensity[25]. The fourth PGC spans 32-kb and contains further pigment synthesis genes, the photosynthetic reaction center subunit H, and various response regulators.

Transcriptomic data shows photosynthetic reaction center subunit M is significantly expressed during pEEU compared to aerobic growth (log_2fold change ~1; $P < 0.0001$), but not when grown with H_2 compared to aerobic growth (log_2fold change ~0.5; $P > 0.05$). The same is true for photosynthetic reaction center subunit L during pEEU compared to aerobic (log_2fold change ~1; $P < 0.05$), but not growth with H_2 compared to aerobic growth (log_2fold change ~0.5; $P > 0.05$).

3.2.2 Carbon Dioxide Fixation

The AB26 genome contains form I (BV509_05525) and form II (BV509_05520) of RuBisCO, the enzyme necessary for carbon dioxide fixation via the Calvin-Benson-Bassham (CBB) cycle[25]. The two forms reside in an operon together with the CbbQ RuBisCO activating protein (BV509_05515), and a LsyR family transcriptional regulator (BV509_15210), which shares similarity with the RuBisCO regulator CbbR. The genes

required for photorespiration via the Calvin-Benson-Bassham cycle are housed 2-Mb upstream from the RuBisCO enzyme gene cluster. As previously described[11], transcriptomic data reveals upregulation of form I RuBisCO (log_2fold change ~2, $P < 0.05$) but downregulation of form II RuBisCO (log_2fold change ~-0.35, $P > 0.1$) during pEEU, and upregulation of a formate dehydrogenase operon (log_2fold change ~4; $P < 0.0001$)[11]. The high upregulation of formate dehydrogenase may suggest an involvement of this enzyme in catalyzing CO_2 fixation[27].

3.2.3 Reducing Power

AB26 is metabolically diverse, capable of aerobic heterotrophy, anaerobic photoheterotrophy, or anaerobic photoautotrophy using various inorganic compounds for energy synthesis (Figure 4). AB26 contains a 16-kb operon coding for the synthesis and assembly of a nickel-dependent hydrogenase (BV509_06105-BV509_06195, Figure 4a). Expression of this nickel-dependent hydrogenase was highest during growth with thiosulfate compared to aerobic growth, with expression of genes within the operon ranging from ~1-3.5-fold ($P < 0.001$, Figure 4a). AB26 is also capable of growth using thiosulfate via *sox*-like genes organized into 3 clusters. The first cluster contains two SoxXY-like carrier proteins (BV509_04530- BV509_04535) and a single metallohydrolase SoxH (BV509_04575). The second cluster contains the canonical *soxABCXYZ* genes (BV509_09605-BV509_09655). *soxR* (BV509_15000) is 1-Mb downstream, and is a putative transcriptional regulator of the *soxABCXYZ* genes[28]. *soxR* is slightly upregulated during growth with thiosulfate compared to aerobic (log_2fold change ~1.3; $P < 0.0001$, Figure 4b), which corresponds with the increased expression of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

other genes in the *sox* operon. But *soxR* is not upregulated during growth with acetate as compared to aerobic growth (log_2fold change ~-0.3; $P < 0.0001$), despite increased expression patterns of the other *sox* genes (1-2-fold; $P < 0.0001$, Figure 4b). A carbon monoxide dehydrogenase (BV509_12715-BV509_12730, Figure 4a) and a formate dehydrogenase operon (BV509_15145-BV509_15165, Figure 4a) are also present in the AB26 genome.

3.2.4 Regulation and Signaling

To be successful in transient environments requires genomic resources to sense and respond. AB26 dedicates ~5% of its genome to signal transduction and regulation. Typical bacterial genomes dedicate 5-6% to regulation and sensing. Some purple non-sulfur bacteria such as *R. palustris* TIE-1 allot ~10% of their genome to environmental sensing, similar to other soil bacteria[29]. AB26 contains 22 histidine kinases, 4 of which contain an HAMP domain (domain present in Histidine kinases, Adenylyl cyclases, Methyl-accepting proteins and Phosphatases) involved with chemotaxis[30]. Thirteen PAS domain-containing proteins are likely involved in light and oxygen sensing[31], and 2 GAF domains are potentially associated with photoreceptors[32]. AB26 contains 19 RNA polymerase sigma factors and 121 transcriptional regulators belonging to 30 different families. During pEEU in AB26, transcripts encoding histidine kinases responding to envelope stress (BV509_04670, log_2fold change ~2, $P < .0001$, Figure 5a), nitrate assimilation (BV509_11520, log_2fold change ~3, $P < .0001$, Figure 5b), a sigma-70 factor involved in exocytoplasmic stress (BV509_00135, log_2fold change ~2, $P < .01$, Figure 5e),and an anti-sigma antagonist (BV509_19770 log_2fold change ~3, $P < .0001$, Figure

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

5d) were upregulated. This suggests that AB26 is potentially under stress during pEEU compared to growth under H_2. Also upregulated during pEEU is a flagellar switching response regulator (BV509_14575, (log_2fold change ~2, $P < .0001$, Figure 5c).

3.2.5 Biodegradation and Carbon Storage

The diverse metabolisms of purple non sulfur phototrophs allow for proliferation in industrial waste environments[33]. *Rhodopseudomonas palustris* and *Cereibacter sphaeroides* are two purple phototrophs known for their ability to metabolize diverse compounds including fatty acids, amino acids, aromatic compounds, and lignin. They are, therefore, important species in the bioremediation of wastewater and polluted environments[34-37]. Some purple phototrophs are also capable of synthesizing carbon storage polymers such as polyhydroxalkanoates (PHA), which they degrade in carbon-starved environments. The genome of AB26 encodes a variety of genes for the metabolism of aromatic compounds: a benzoate transporter (BV509_16875) and a 4-hydroxybenzoate octaprenyltransferase (BV509_17190) for the first step of benzoate degradation, a salicylate esterase (BV509_RS17105), a biphenyl-2,3-diol 1,2-dioxygenase III-related protein (BV509_11920) for the degradation of biphenyl, a type II 3-dehydroquinate dehydratase (BV509_13850) for the degradation of quinate, and for the degradation of gentisate, a fumarylacetoacetate hydrolase family protein (BV509_03940) and a maleylacetoacetate isomerase (BV509_16060). AB26 contains two operons for the degradation of phenylacetate (BV509_03065-BV509_03095, BV509_05275-BV509_05285)[38]. Like *R. palustris* TIE-1[37], AB26 encodes genes for the synthesis and catabolism of polyhydroxybutyrate (PHB). *phaA*, which encodes a beta-ketothiolase,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

and *phaB*, which encodes an acetoacetyl-CoA reductase, are organized separately from each other and the remaining PHB gene cluster. The remainder of the PHB cluster contains the PHA depolymerase gene *phaZ*, the class 1 poly(R)-hydroxyalkanoic acid synthase *phaC*, which encodes the phasin protein affecting PHB granule accumulation and utilization, and the PHA synthesis repressor gene *phaR* (BV509_06270-BV509_06285). Unlike TIE-1, AB26 only has single gene copies of each gene for the synthesis and utilization of PHB. As previously described[11], the PHB biosynthesis enzymes were downregulated during pEEU as compared to aerobic growth (log_2fold change ~0, $P > 0.05$), and the repressor *phaR* was slightly upregulated (log_2fold change ~1, $P < 0.01$).

3.2.6 Transporters

The AB26 genome encodes 561 transport system genes, dedicating ~16% of its genome to transport. Of these, 347 transport genes are involved in primary transport, with 45 different ABC transport systems and 23 ATPase genes. AB26 contains type III secretion system genes for flagella biosynthesis, and seven type II secretion system families (Fig. 6).

Sixty of the ABC transport genes were identified as various amino acid transport systems, with 22 genes dedicated to the transport of branched chain amino acids. For iron acquisition, AB26 contains 11 *tonB* genes and five iron transport systems. The lack of siderophore biosynthesis genes suggests that AB26 may transport siderophores produced by other bacteria into the cell. AB26 contains two ferrous iron

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

transport genes (BV509_02645-BV509_02650) for aerobic iron acquisition. Nutrient acquisition (Figure 6) for AB26 includes urea transport (BV509_09920-BV509_09945) and sulfate transport (BV509_13630-13645). Located on a plasmid are various metal uptake systems[14] including zinc (BV509_08850, BV509_19470, BV509_19480, BV509_20850), nickel (BV509_20695-BV509_20715, BV509_11055), and manganese (BV509_03550, BV509_20820-20825, Figure 6), which may help maintain metal ion homeostasis. To compete within their niche, AB26 contains four microcin transporters, 14 multidrug transporters, and 9 efflux transporters. The Sec pathway was upregulated compared to aerobic growth only during pEEU, supporting the requirement of protein export in the periplasmic space to facilitate extracellular electron transfer across the membranes (Figure 6)

AB26 contains 176 passive transport genes, with two multidrug and toxic compound extrusion (MATE) genes, 10 resistance-nodulation-cell division (RND) pumps, 23 major facilitator superfamily (MFS) genes, and 59 tripartite ATP-independent periplasmic (TRAP) transport genes. The TRAP family transporters are common in marine bacteria such as AB26 due to the natural Na^+ gradient in their environment[39].

3.2.7 Nitrogen Assimilation And Fixation

AB26 uses the glutamine synthetase and glutamate synthase for ammonia assimilation in the cell. AB26 genome contains two ammonia transporters, three glutamine synthetases, and a glutamine-synthetase adenyltransferase to regulate activity of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

glutamine synthetases. AB26 also encodes genes to convert urea to ammonia (BV509_09920-BV50-0995, Fig. 7). AB26 genome contains structural genes for a single nitrogenase (BV509_14955-BV509_14965), assembly and cofactor genes, and the nitrogen fixation-specific sigma factor RpoN to fix nitrogen during anaerobic growth (Figure 7). During pEEU, one glutamine synthetase is highly upregulated (BV509_16830, $\log_2$fold change ~2.5, $P < .0001$), a second glutamine synthetase is slightly upregulated (BV509_16820, $\log_2$fold change ~0.75, $P < .001$; Figure 7a), and both are downregulated during photoheterotrophic growth compared to aerobic growth. Conversely, during pEEU the glutamate synthase is downregulated as compared to aerobic growth (Figure 7a).

3.2.8 Extracellular Electron Uptake (EEU) Genes

To perform pEEU, gram-negative bacteria require uptake systems consisting of electron transfer proteins spanning the outer membrane, periplasmic space, and inner membrane. These systems may also include porin proteins to span membranes and facilitate the exchange of electrons from the cell exterior to the interior. While some characterized electron uptake systems exist[10, 40], these uptake systems or homologs of them are not present in all bacteria that are hypothesized to be able to perform pEEU. Our group recently identified a novel protein important to phototrophic extracellular electron uptake in AB26, which does not encode any known EEU pathways in its genome; the di-heme *c*-type cytochrome EeuP[11]. As previously reported, EeuP is a novel pEEU-important protein found in A26 but also in species spanning the Proteobacteria and Acidobacteria. So far, no other components of the novel EEU pathway have been identified in AB26 other than EeuP. To identify other potential proteins that might be important in pEEU, a

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

RNASeq guided comparative genomic approach with specific targets such as cytochrome c containing proteins would be useful. This is because these proteins are often associated with extracellular electron transfer[2].

3.3 RNA-Seq Led Genomic Comparison Between Rhodovulum spp.

Our group previously performed RNA-Seq in AB26 between photoautotrophy with H_2/CO_2 as electron donor/carbon source compared to pEEU via poised electrodes[11] in an effort to characterize the molecular mechanisms of pEEU in this novel isolate and identify the cellular sinks of electrons derived from extracellular sources. With these data, and the full genome sequences of the 14 other isolates comprising two *Rhodovulum* species, we can identify homologs of pEEU-important transcripts and infer ability to perform pEEU based on their presence or absence in each strain. Supplementary Table 1 shows the general genome features of all 15 Woods Hole isolates and publicly available genomes for both *R. visakhapatnamense*[24] and *R. sulfidophilum* DSM 2351[41] and DSM1374[42].

3.3.1 *C*-type cytochromes (cyt_c)

1. Diheme cyt_c EeuP

RNA-Seq identified a significantly upregulated transcript (log_2fold change ~2, P <0.0001) in AB26 encoding the hypothetical protein later identified to be the di-heme EeuP (BV509_10070). BLASTp analysis identified homologous sequences of EeuP in all isolates except AB21. BLASTp results for AB21 returned two partial cyt_c sequences and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

one partial hypothetical protein sequence. These partial sequences separately align with the intact EeuP homologs and may be an artifact of assembly. From this we can hypothesize that all the isolates are pEEU-capable. Maximum likelihood phylogenetic analysis of the EeuP protein sequence (Figure 8) reveals EeuP in the isolates reflects the same relationship shown in the 16S and *pufM* phylogenetic trees (Figure 1), in that AB19 and AB26 EeuP sequences group outside of the *R. sulfidophilum* EeuP sequences. As previously reported[11], EeuP homologous sequences were also identified in Alpha-, Beta-, and Gammaproteobacteria orders and in one *Acidobacter* species. Further analysis of the species containing EeuP homologs may reveal the relationship between EeuP conservation and metabolic/environmental adaptation.

2. Hypothetical protein BV509_18570

Another significantly upregulated transcript (log_2fold change ~6, P <0.0001) in AB26 encodes a hypothetical protein (BV509_18570). BV509_18570 is a 413-residue long protein containing a single heme-binding motif (CXXCH). InterProScan[43] predicts the molecular functions as electron transfer and heme binding. PRED-TAT[44] did not predict a secretory signal peptide (Sec), which would be expected for an electron transport protein involved in trans-membrane transport. Surprisingly, BLASTp analysis revealed no homologs of this open reading frame (ORF) within any of the *R. sulfidophilum* isolates. Homologs of BV509_18570 were found in *R. visakhapatnamense* AB19 (JMJ92_02255) and *R. visakhapatnamense* JA181 (EV657_1244). The homologous sequences are each 625 amino acid residues in length. Protein sequence alignments in MAUVE[45] show

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

JMJ92_02255 and EV657_1244 sequences align 212 residues before aligning to BV509_18570, and these additional residues contain a Sec signal peptide sequence.

3. Diheme cyt_c BV509_14915

This transcript (log_2fold change ~6, *P* <0.0001) encodes a cyt_c with two heme binding motifs. BV509_14915 is 347 residues long and contains a Sec signal peptide. BLASTp analysis shows all 15 marine *Rhodovulum* spp. isolates contain a homolog of this diheme cyt_c . Given the high fold change and the presence of homologs within the other isolates' genomes, this cytochrome could be a conserved component of the pEEU molecular pathway in the *Rhodovulum* spp. isolates.

4. cyt_c BV509_15885

The transcript encoding the gene BV509_15885 is a cyt_c. RNA-Seq data did not identify this cyt_c as significantly upregulated during pEEU (log_2fold change -0.48), but it was identified through heme staining (membrane fraction) followed by mass spectroscopy compared to hydrogen grown cultures[11]. This cyt_c contains 5 heme-binding motifs and a predicted Sec signal peptide. BLASTp analysis identified homologs of this cyt_c in AB19 and *R. visakhapatnamense* JA181. Homologs in the *R. sulfidophilum* isolates show 71-75% identity.

3.3.2 Transcriptional Regulators

1. Mar Family transcription regulator BV509_01095

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

BV509_01095 encodes a hypothetical protein identified as a member of the multi antibiotic resistance family transcriptional regulator (log_2fold change ~6, *P* <0.001). InterProScan predicts phosphorelay signal transduction function. BV509_01095 is present only in AB19 and AB26 via BLASTp analysis, but not *R. visakhapatnamense* JA181 or any of the *R. sulfidophilum* isolates.

2. Transcriptional regulator BV509_15825

This transcript is another transcriptional regulator upregulated during pEEU in AB26 (log_2fold change ~4, *P* <0.0001). This regulator is also predicted to be involved in signal transduction response regulation via InterProScan, with DNA-binding wing-helix domain and tetratricopeptide repeat domain. BLASTp reveals this transcription regulator is unique to AB19, AB26, and *R. visakhapatnamense* JA181 but not the *R. sulfidophilum* isolates.

4. Conclusions

pEEU is an ecologically significant metabolism in that microbes can interact with multiple geochemical cycles in their environment while growing under these conditions. pEEU is a relatively novel metabolism with few molecular uptake systems characterized[10, 40, 46]. Furthermore, less is known about the distribution of this metabolism in bacterial lineages, and the prevalence of this metabolism in natural environments. We sought to understand the environmental occurrence of pEEU through the isolation of the 15 Woods Hole isolates. We previously showed that one representative isolate, AB26, is capable of pEEU using a novel pEEU mechanism (EeuP)[11]. Using phylogenetic analyses, we reclassify AB26 as *Rhodovulum visakhapatnamense* along with AB19. Using these data,

 doi:10.20944/preprints202205.0070.v1

paired with the assembled genomes AB19 and 13 *Rhodovulum sulfidophilum* that we previously isolated, we sought to use RNA-Seq led comparative genomics to determine the potential of the remaining isolates to perform pEEU. We also determined whether these potentially pEEU-related genes are conserved between the *R. visakhapatnamense* and the *R. sulfidophilum* isolates.

From the RNA-Seq data, we identified that all 15 *Rhodovulum* isolates contain homologs of EeuP, suggesting this cyt_c is conserved within *Rhodovulum* for phototrophic extracellular electron uptake. A second highly upregulated diheme cyt_c conserved in all 15 isolates may also be integral to pEEU. We also identified proteins unique to *R. visakhapatnamense* that were highly upregulated during pEEU, including a diheme cyt_c and two transcription regulators. Despite *R. visakhapatnamense* and *R. sulfidophilum* being closely related, our findings reveal some species-specific adaptations may be occurring within their environment. Further genetic studies guided by these findings will determine the putative roles these regulators and proteins may play in pEEU.

Supplementary Materials: Supplementary Figure 1, Supplementary Table 1, database files for phylogenetic trees, RNASeq Heatmaps base files.

Author Contributions: EJD did all analysis and writing. AB contributed to writing.

Funding: This work was supported by the following grants to A.B.: The David and Lucile Packard Foundation Fellowship (201563111), the U.S. Department of Energy (grant number DESC0014613), and the U.S. Department of Defense, Army Research

Office (grant number W911NF-18-1-0037), Gordon and Betty Moore Foundation, National Science Foundation (Grant Number 2021822, Grant Number 2124088, and Grant Number 2117198), the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DEAC5207NA27344 (LLNL-JRNL-812309), and a DEPSCoR grant (FA9550-21-1-0211). A.B. was also funded by a Collaboration Initiation Grant, an Office of the Vice-Chancellor of Research Grant, and an International Center for Energy, Environment, and Sustainability Grant from Washington University in St. Louis. E.J.D. is supported by an Institutional Training Grant in Genomic Science from the NIH (T32 HG000045-18).

Data Availability Statement: Sequencing reads were deposited in the NCBI database under BioProject PRJNA546270, BioProject PRJNA692994 and BioProject PRJNA693004. Files for trees and heatmaps in supplement.

Acknowledgments: We thank Eric Conners for his assistance with manuscript editing.

Conflicts of Interest: The authors declare no conflict of interests.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

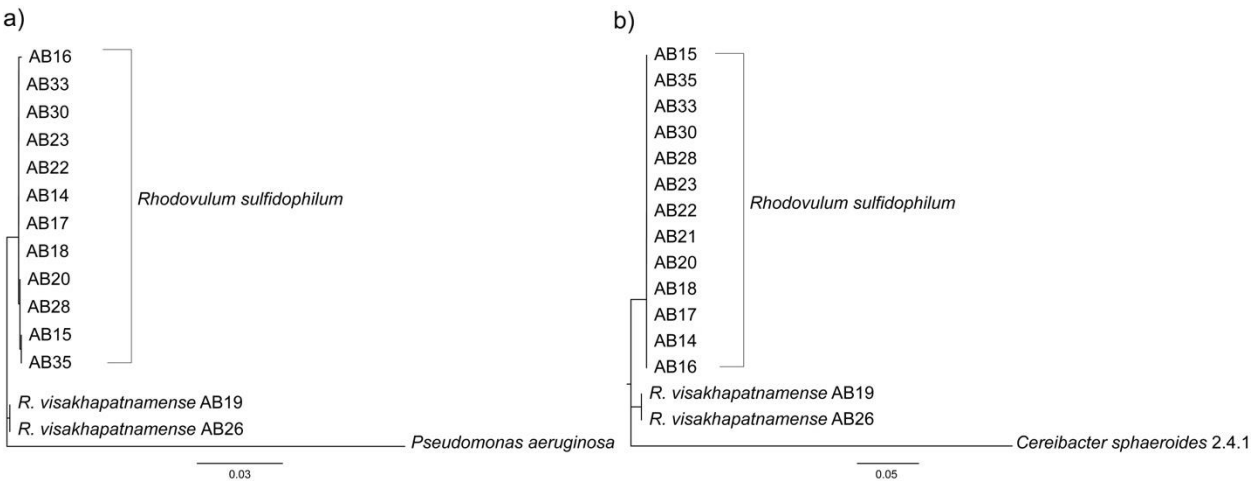


Figure 1. 16S rRNA (a) and *pufM* (b) phylogenetic trees between Woods Hole *R. sulfidophilum* isolates. (a) 16S phylogenetic tree constructed by maximum-likelihood method using the Kimura-2 parameter model[47], tested by bootstrapping (500 re-samplings) in MEGA11[21] with *Pseudomonas aeruginosa* as outgroup. (b) *pufM* tree constructed by maximum-likelihood using the Jones-Taylor-Thornton parameter model[48], tested by bootstrapping (500 re-samplings) in MEGA11[21], using *Cereibacter sphaeroides* 2.4.1 as outgroup.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

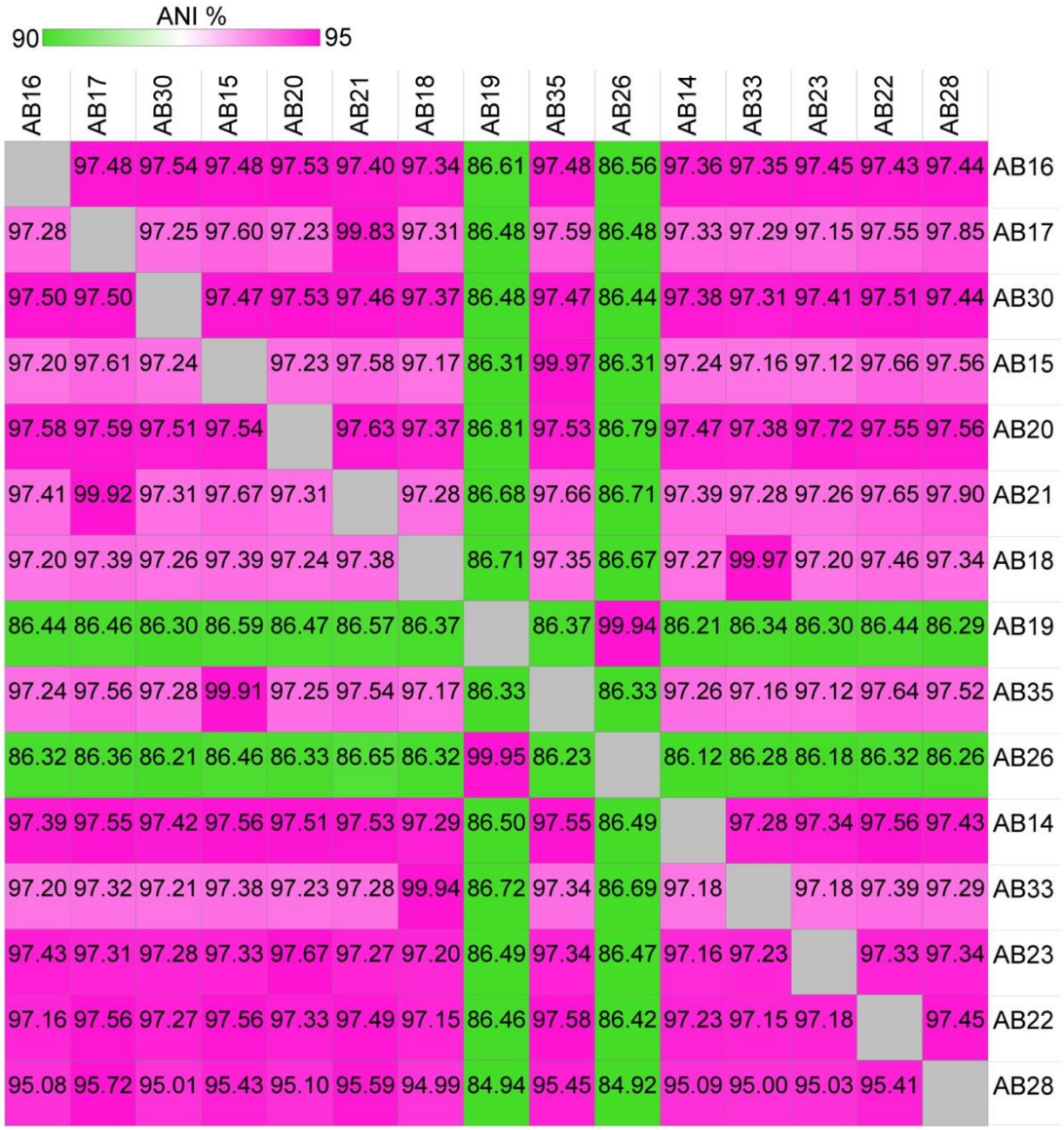


Figure 2. Heatmap showing average nucleotide identities (ANI) values from pairwise comparisons of the 15 Woods Hole isolates. ANI values were calculated using JSpeciesWSd[20] and visualized as a heatmap using Morpheus (Morpheus, https://software.broadinstitute.org/morpheus).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

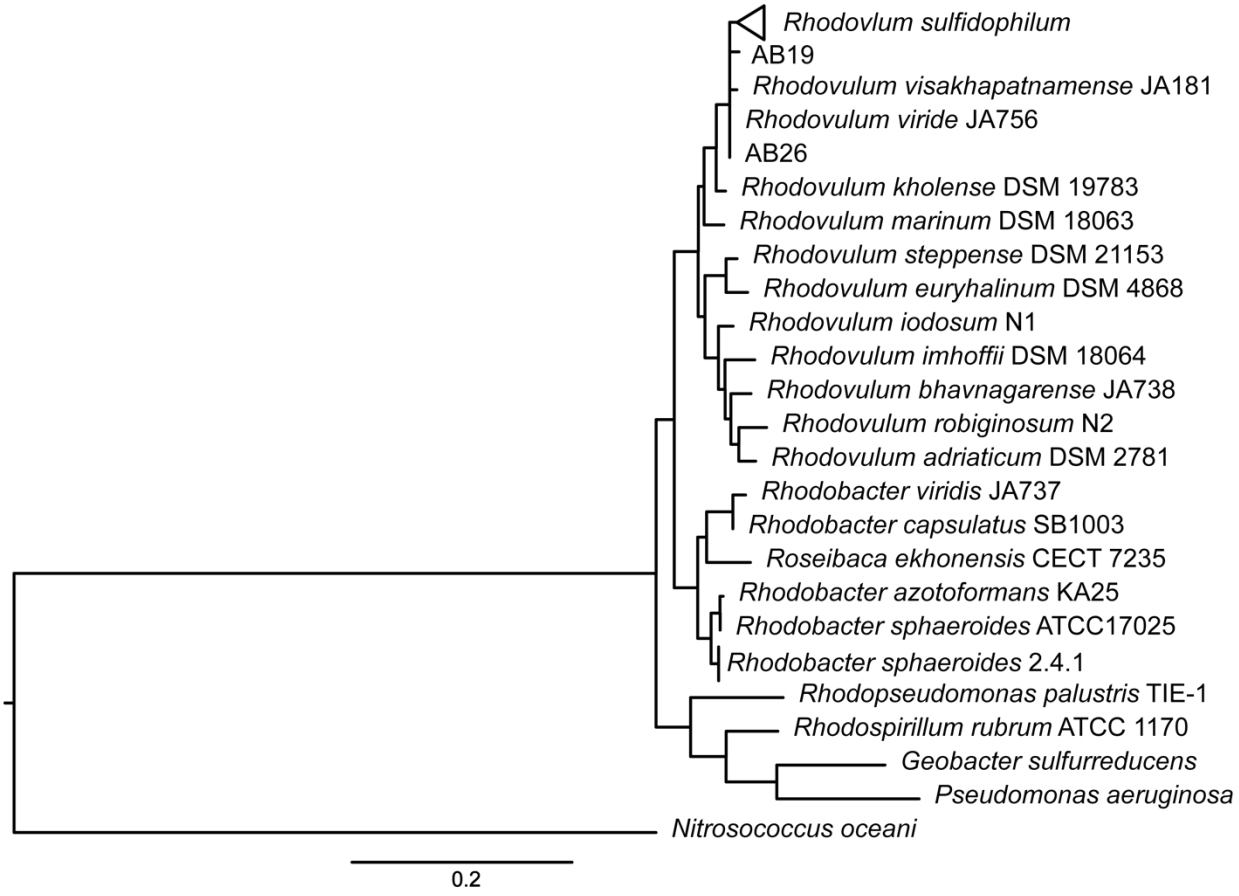


Figure 3. 16S phylogenetic tree showing the new taxonomic assignment of AB19 and AB26 as *Rhodovulum visakhapatnamense*. Tree constructed by maximum-likelihood method using the Kimura-2 parameter model[47], tested by bootstrapping (500 resamplings) in MEGA11[21].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

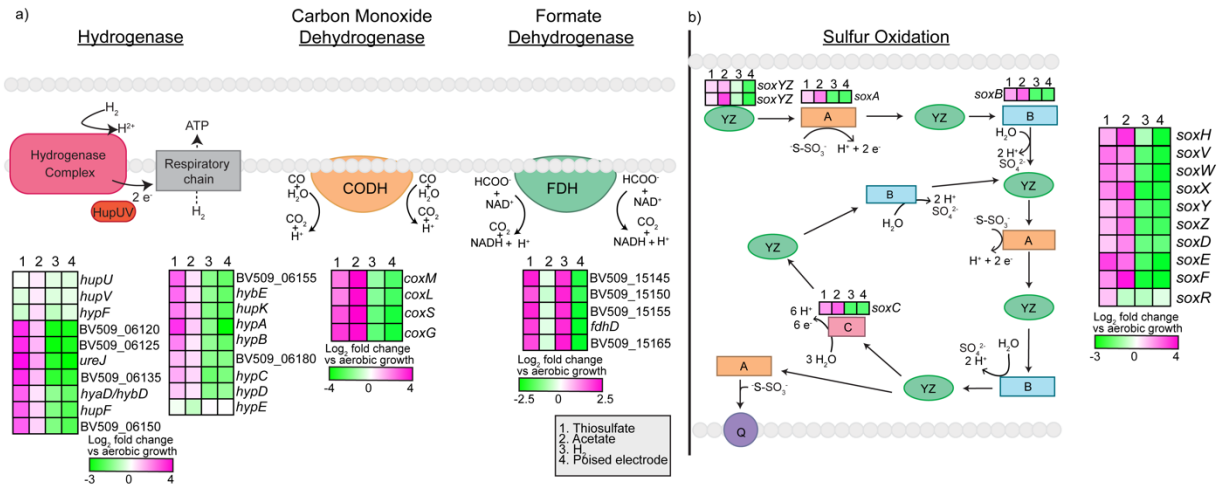
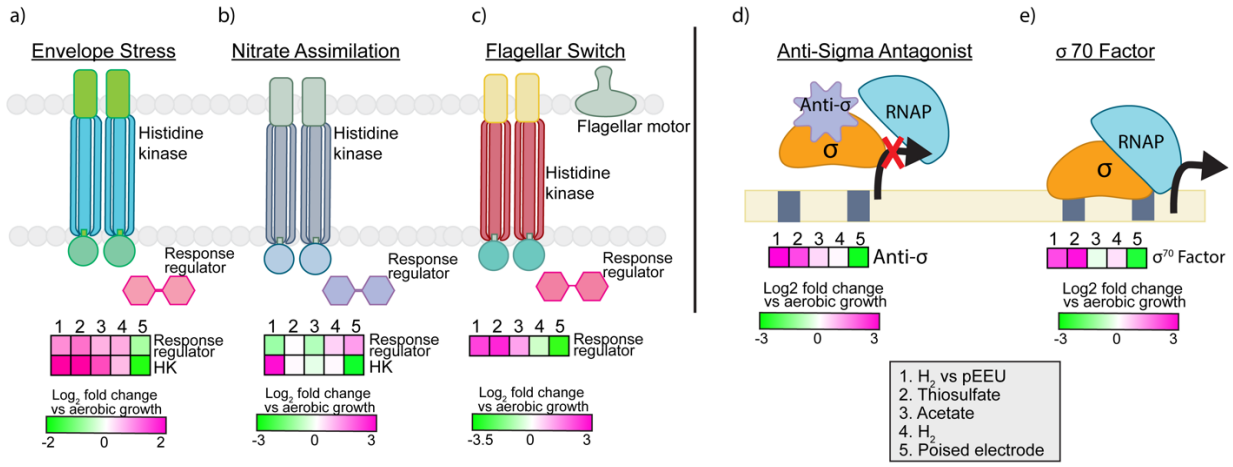


Figure 4. Genes involved in oxidation of inorganic compounds as a source of reducing power and their expression analysis via RNASeq. (a) nickel-dependent hydrogenase, carbon monoxide dehydrogenase, and formate dehydrogenase and their expression based on RNASeq. (b) sulfur oxidation pathway[49] and expression of the putative genes based on RNASeq. Heatmaps describe expression of four modes of metabolism (thiosulfate, acetate, H_2, and pEEU) compared to aerobic growth. (a) CODH (Carbon monoxide dehydrogenase), FDH (Formate dehydrogenase), CO (carbon monoxide), $HCOO^-$ (formate) (b) Q (quinone pool).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

479



480

481 **Figure 5.** Expression analysis of significantly upregulated environmental response
482 systems (a) and transcriptional regulators (b)under pEEU compared to photoautotrophy
483 with H_2 (growth condition 1). Heatmaps also describe expression of four modes of
484 metabolism (thiosulfate, acetate, H_2, and pEEU) compared to aerobic growth. (d) RNAP
485 (RNA polymerase).

486

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

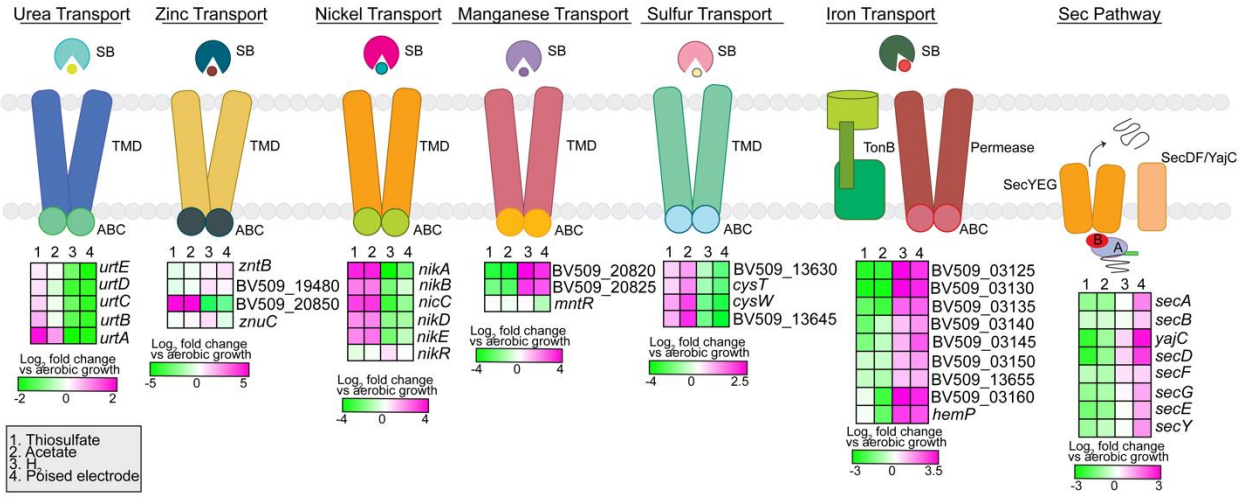
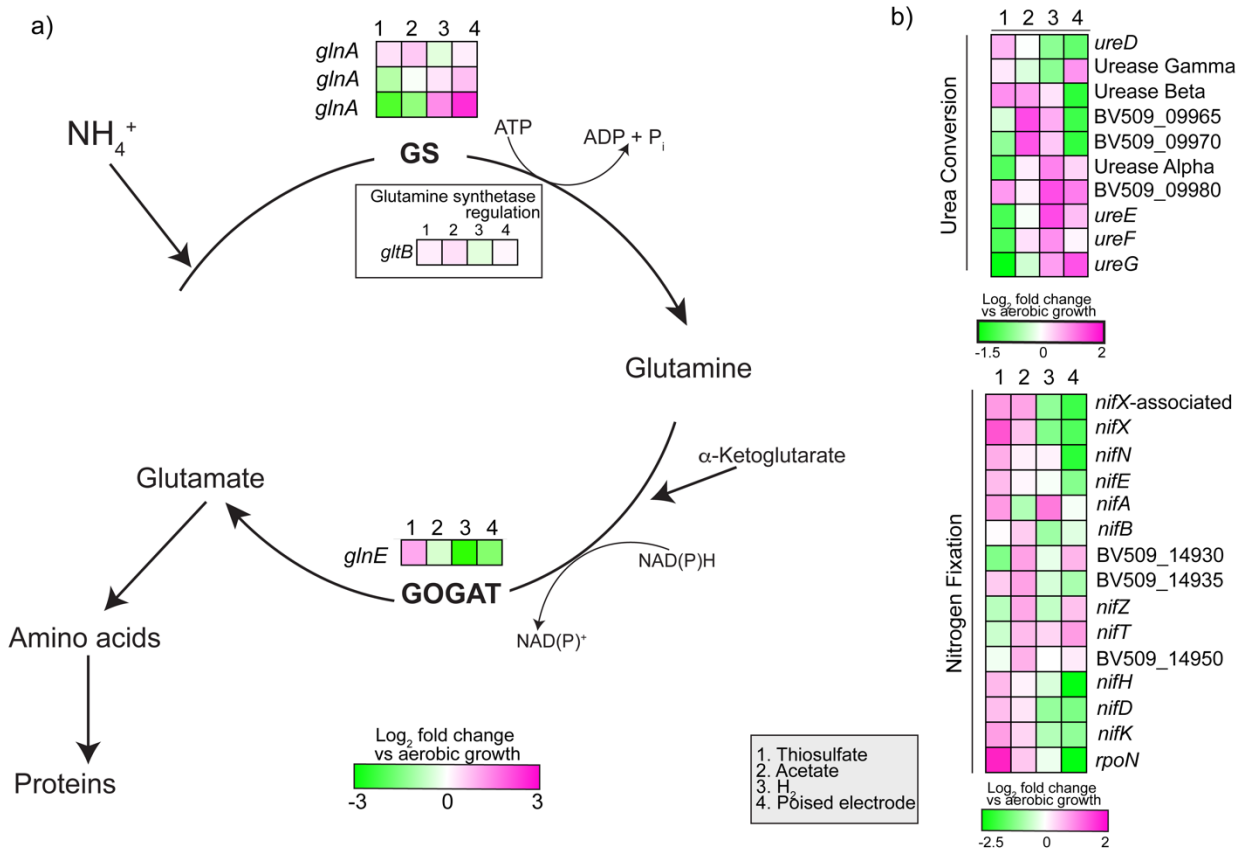


Figure 6. Expression analysis of transport systems in AB26. The Sec pathway genes are all upregulated during pEEU, which supports the need for protein export into the periplasmic space to facilitate extracellular electron transport. Heatmaps describe expression of four modes of metabolism (thiosulfate, acetate, H_2, and pEEU) compared to aerobic growth. SB (substrate binding domain), TMD (transmembrane domain), ABC (ATP cassette binding domain).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

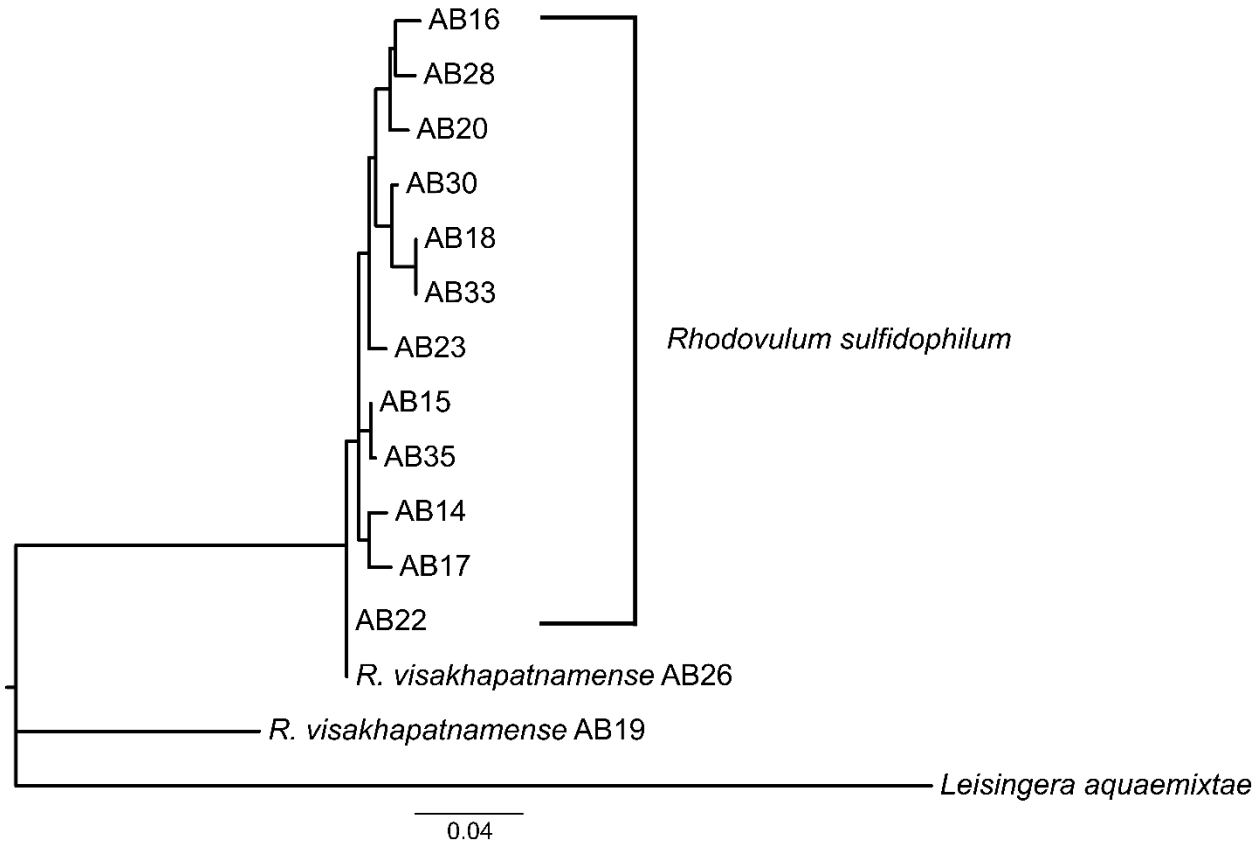


494

495 **Figure 7.** Expression analysis of genes involved in the glutamine synthetase/glutamate
496 synthase (GS/GOGAT) pathway for nitrate assimilation in AB26 (a). Gene expression in
497 urea conversion and nitrogen fixation (b). Heatmaps describe expression of four modes
498 of metabolism (thiosulfate, acetate, H_2, and pEEU) compared to aerobic growth. (a) GS
499 (glutamine synthetase), GOGAT (glutamate synthase) (b) BV509_09965 (hypothetical
500 protein), BV509_09970 (DUF1127 domain-containing protein), BV509_09980 (DUF1127
501 domain-containing protein), BV509_14930 (4Fe-4S dicluster domain-containing protein),
502 BV509_14935 (hypothetical protein), BV509_14950 (SIR2 family protein).

503

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1



504

505 **Figure 8.** Phylogenetic tree showing protein sequence similarity of the diheme EeuP in
506 Woods Hole isolates. Tree constructed by maximum-likelihood method using the Kimura-
507 2[47] parameter model, tested by bootstrapping (500 resamplings) in MEGA11[21].

508

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2022 doi:10.20944/preprints202205.0070.v1

509 **Table 1.** Genome features of AB26. Data for table derived from NCBI and JGI IMG.

510 AB26 is also capable of photosynthesis via pEEU, using electrons obtained from a poised
511 electrode (pEEU) or photoferrotrophy, obtaining electrons by oxidizing Fe(II)[11].

Genome size (bp)	4380746	
GC%	67.9	
Scaffolds	3	
Genes total number	4375	**% Total**
Protein coding genes	4296	98.2
rRNA genes	3	
tRNA genes	49	
Other RNA genes	7	
Protein coding genes with function prediction	3212	73.4
Without function prediction	1084	24.8
Biosynthetic Gene Clusters	5	
Genes in Biosynthetic Clusters	115	2.6
Protein coding genes coding signal peptides	370	8.5
Protein coding genes coding transmembrane proteins	919	21.0
COG clusters	1869	56.4
Pfam clusters	2235	65.0
TIGRfam clusters	1083	83.6

512

513

514

References:

1. Rosenbaum, M., Aulenta, F., Villano, M., and Angenent, L. T. Cathodes as electron donors for microbial metabolism: which extracellular electron transfer mechanisms are involved? *Bioresour Technol.* **2011**, 102, 324-333; DOI: 10.1016/j.biortech.2010.07.008
2. Gupta, D., Guzman, M. S., and Bose, A. Extracellular electron uptake by autotrophic microbes: physiological, ecological, and evolutionary implications. *J Ind Microbiol Biotechnol.* **2020,** 47, 863-876; DOI: 10.1007/s10295-020-02309-0
3. Falkowski, P. G., Fenchel, T., and Delong, E. F. The Microbial Engines That Drive Earth's Biogeochemical Cycles. *Science.* **2008,** 320, 1034-1039; DOI: 10.1126/science.1153213
4. Lovley, D. R., and Nevin, K. P. Electrobiocommodities: powering microbial production of fuels and commodity chemicals from carbon dioxide with electricity. *Curr Opin Biotechnol.* **2013,** 24, 385-390; DOI: 10.1016/j.copbio.2013.02.012
5. Rabaey, K., Rodriguez, J., Blackall, L. L., Keller, J., Gross, P., Batstone, D., Verstraete, W., and Nealson, K. H. Microbial ecology meets electrochemistry: electricity-driven and driving communities. *ISME J.* **2007**, 1, 9-18; DOI: 10.1038/ismej.2007.4
6. Rabaey, K., and Rozendal, R. A. Microbial electrosynthesis—revisiting the electrical route for microbial production. *Nat Rev Microbiol.* **2010**, 8, 706-716; DOI: 10.1038/nrmicro2422

7. Bose, A., Gardel, E. J., Vidoudez, C., Parra, E., and Girguis, P. R. Electron uptake by iron-oxidizing phototrophic bacteria. *Nat Commun.* **2014**, 5, 1-7; DOI: 10.1038/ncomms4391

8. Rowe, A. R., Rajeev, P., Jain, A., Pirbadian, S., Okamoto, A., Gralnick, J. A., El-Naggar, M. Y., and Nealson, K. H. Tracking electron uptake from a cathode into *Shewanella* cells: implications for energy acquisition from solid-substrate electron donors. *MBio.* **2018**, 9, e02203-02217; DOI: 10.1128/mBio.02203-17

9. Tefft, N. M., and TerAvest, M. A. Reversing an extracellular electron transfer pathway for electrode-driven acetoin reduction. *ACS Synth Biol.* **2019**, 8, 1590-1600; DOI: 10.1021/acssynbio.8b00498

10. Jiao, Y., and Newman, D. K. The pio operon is essential for phototrophic Fe (II) oxidation in *Rhodopseudomonas palustris* TIE-1. *J Bacteriol.* **2007**, 189, 1765-1773; DOI: 10.1128/JB.00776-06

11. Gupta, D., Guzman, M. S., Rengasamy, K., Stoica, A., Singh, R., Ranaivoarisoa, T. O., Davenport, E. J., Bai, W., McGinley, B., and Meacham, J. M. Photoferrotrophy and phototrophic extracellular electron uptake is common in the marine anoxygenic phototroph *Rhodovulum sulfidophilum*. *ISME J.* **2021**, 15, 3384-3398; DOI: 10.1038/s41396-021-01015-8

12. Guzman, M. S., Rengasamy, K., Binkley, M. M., Jones, C., Ranaivoarisoa, T. O., Singh, R., Fike, D. A., Meacham, J. M., and Bose, A. Phototrophic extracellular electron uptake is linked to carbon dioxide fixation in the bacterium *Rhodopseudomonas palustris*. *Nat Commun.* **2019**, 10, 1-13; DOI: 10.1038/s41467-019-09377-6

559 13. Shi, L., Dong, H., Reguera, G., Beyenal, H., Lu, A., Liu, J., Yu, H.-Q., and
560 Fredrickson, J. K. Extracellular electron transfer mechanisms between
561 microorganisms and minerals. *Nature Rev Microbiol.* **2016**, 14, 651-662; DOI:
562 10.1038/nrmicro.2016.93
563 14. Guzman, M. S., McGinley, B., Santiago-Merced, N., Gupta, D., and Bose, A. J. G.
564 A. Draft genome sequences of three closely related isolates of the purple nonsulfur
565 bacterium *Rhodovulum sulfidophilum*. *Genome Announc.* **2017**, 5, e00029-17;
566 DOI: 10.1128/genomeA.00029-17
567 15. Guzman, M. S., and Bose, A. Draft genome sequences of four Rhodobacter
568 sphaeroides strains isolated from a marine ecosystem. *Microbiol Resourc*
569 *Announc.* **2019,** 8, e01648-01618; DOI: 10.1128/MRA.01648-18
570 16. Bolger, A. M., Lohse, M., and Usadel, B. Trimmomatic: a flexible trimmer for
571 Illumina sequence data. *Bioinformatics.* **2014**, 30, 2114-2120; DOI:
572 10.1093/bioinformatics/btu170
573 17. Bosi, E., Donati, B., Galardini, M., Brunetti, S., Sagot, M.-F., Lió, P., Crescenzi, P.,
574 Fani, R., and Fondi, M. MeDuSa: a multi-draft based scaffolder. *Bioinformatics.*
575 **2015**, 31, 2443-2451; DOI: 10.1093/bioinformatics/btv171
576 18. Langmead, B., and Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat*
577 *Methods.* **2012**, 9, 357-359; DOI: 10.1038/nmeth.1923
578 19. Zhao, Y., Wu, J., Yang, J., Sun, S., Xiao, J., and Yu, J. PGAP: pan-genomes
579 analysis pipeline. *Bioinformatics.* **2011**, 28, 416-418. DOI:
580 10.1093/bioinformatics/btr655

20. Richter, M., Rosselló-Móra, R., Oliver Glöckner, F., and Peplies, J. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics.* **2015**, 32, 929-931. DOI: 10.1093/bioinformatics/btv681

21. Tamura, K., Stecher, G., and Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol.* 2021, 38, 3022-3027. DOI: 10.1093/molbev/msab120

22. Konstantinidis, K. T., and Tiedje, J. M. Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci U S A.* **2005**, 102, 2567-2572; DOI: 10.1073/pnas.0409727102

23. Srinivas, A., Kumar, B. V., Sree, B. D., Tushar, L., Sasikala, C., Ramana, C. V. J. *Rhodovulum salis* sp. nov. and *Rhodovulum viride* sp. nov., phototrophic Alphaproteobacteria isolated from marine habitats. *Int J Syst Evol Microbiol*. **2014**, 64, 957-962; DOI: 10.1099/ijs.0.058974-0

24. Srinivas, T., Anil Kumar, P., Sasikala, C., Ramana, C. V., and Imhoff, J. F. *Rhodovulum visakhapatnamense* sp. nov. *Int J Syst Evol Microbiol.* **2007***,* 57, 1762-1764; DOI: 10.1099/ijs.0.65076-0

25. Blankenship, R. E. (2021) *Molecular mechanisms of photosynthesis*, John Wiley & Sons

26. Nickens, D. G., and Bauer, C. E. Analysis of the puc operon promoter from *Rhodobacter capsulatus*. *J Bacteriol.* **1998**, 180, 4270-4277; DOI: 10.1128/JB.180.16.4270-4277.1998

603 27. Hartmann, T., and Leimkühler, S. The oxygen-tolerant and NAD^+-dependent
604 formate dehydrogenase from *Rhodobacter capsulatus* is able to catalyze the
605 reduction of CO_2 to formate. *FEBS J.* **2013**, 280, 6083-6096; DOI:
606 10.1111/febs.12528
607 28. Lahiri, C., Mandal, S., Ghosh, W., Dam, B., and Roy, P. A novel gene cluster
608 *soxSRT* is essential for the chemolithotrophic oxidation of thiosulfate and
609 tetrathionate by *Pseudaminobacter salicylatoxidans* KCT001. *Curr Microbiol.*
610 **2006**, 52, 267-273; DOI: 10.1007/s00284-005-0176-x
611 29. Larimer, F. W., Chain, P., Hauser, L., Lamerdin, J., Malfatti, S., Do, L., Land, M.
612 L., Pelletier, D. A., Beatty, J. T., and Lang, A. S. Complete genome sequence of
613 the metabolically versatile photosynthetic bacterium *Rhodopseudomonas*
614 *palustris*. *Nat Biotechnol.* **2004**, 22, 55-61; DOI: 10.1038/nbt923
615 30. Aravind, L., and Ponting, C. P. The cytoplasmic helical linker domain of receptor
616 histidine kinase and methyl-accepting proteins is common to many prokaryotic
617 signalling proteins. *FEMS Microbiol Lett*. **1999**, 176, 111-116; DOI:
618 10.1111/j.1574-6968.1999.tb13650.x
619 31. Kyndt, J. A., Meyer, T. E., and Cusanovich, M. A. Photoactive yellow protein,
620 bacteriophytochrome, and sensory rhodopsin in purple phototrophic bacteria.
621 *Photochem Photobiol Sci*. **2004**, 3, 519-530. DOI: 10.1039/b315731h
622 32. Martinez, S. E., Beavo, J. A., and Hol, W. G. GAF domains: two–billion–year–old
623 molecular switches that bind cyclic nucleotides. *Mol Interv.* **2002**, 2, 317-323; DOI:
624 10.1124/mi.2.5.317

33. Hiraishi, A., and Kitamura, H. (1984) Distribution of phototrophic purple nonsulfur bacteria in activated sludge systems and other aquatic environments. *Nippon Suisan Gakkaishi* 50, 1929-1937. DOI: 10.2331/suisan.50.1929

34. Wu, P., Zhang, Y., Chen, Z., Wang, Y., Zhu, F., Cao, B., Wu, Y., and Li, N. The organophosphorus pesticides in soil was degradated by *Rhodobacter sphaeroides* after wastewater treatment. *Biochem Eng J.* **2019**,141, 247-251; DOI: 10.1016/j.bej.2018.07.019

35. Hassan, M. A., Shirai, Y., Kusubayashi, N., Karim, M. I. A., Nakanishi, K., and Hasimoto, K. The production of polyhydroxyalkanoate from anaerobically treated palm oil mill effluent by *Rhodobacter sphaeroides*. *J Biosci Bioeng.* **1997**, 83, 485-488; DOI: 10.1016/S0922-338X(97)83007-3

36. Brandl, H., Gross, R. A., Lenz, R. W., Lloyd, R., and Fuller, R. C. The accumulation of poly (3-hydroxyalkanoates) in *Rhodobacter sphaeroides*. *Arch Microbiol.* **1991**, 155, 337-340; DOI: 10.1007/BF00243452

37. Ranaivoarisoa, T. O., Singh, R., Rengasamy, K., Guzman, M. S., and Bose, A. Towards sustainable bioplastic production using the photoautotrophic bacterium *Rhodopseudomonas palustris* TIE-1. *J Ind Microbiol Biotechnol.* **2019**, 46, 1401-1417; DOI: 10.1007/s10295-019-02165-7

38. Fuchs, G., Boll, M., and Heider, J. Microbial degradation of aromatic compounds—from one strategy to four. *Nat Rev Microbiol.* **2011**, 9, 803-816; DOI: 10.1038/nrmicro2652

39. Mulligan, C., Kelly, D. J., and Thomas, G. H. Tripartite ATP-independent periplasmic transporters: application of a relational database for genome-wide

analysis of transporter gene frequency and organization. *J Mol Microbiol Biotechnol.* **2007**, 12, 218-226; DOI: 10.1159/000099643

40. Shi, L., Rosso, K. M., Zachara, J. M., and Fredrickson, J. K. Mtr extracellular electron-transfer pathways in Fe (III)-reducing or Fe (II)-oxidizing bacteria: a genomic perspective. *Biochem Soc Trans.* **2012**, 40, 1261-1267; DOI: 10.1042/BST20120098

41. Nagao, N., Hirose, Y., Misawa, N., Ohtsubo, Y., Umekage, S., and Kikuchi, Y. Complete genome sequence of *Rhodovulum sulfidophilum* DSM 2351, an extracellular nucleic acid-producing bacterium. *Genome Announc.* **2015**, 3, e00388-00315; DOI: 10.1128/genomeA.00388-15

42. Masuda, S., Hori, K., Maruyama, F., Ren, S., Sugimoto, S., Yamamoto, N., Mori, H., Yamada, T., Sato, S., and Tabata, S. Whole-genome sequence of the purple photosynthetic bacterium *Rhodovulum sulfidophilum* strain W4. *Genome Announc.* **2013**, 1, e00577-00513; DOI: 10.1128/genomeA.00577-13

43. Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., and Nuka, G. InterProScan 5: genome-scale protein function classification. *Bioinformatics*. **2014**, 30, 1236-1240; DOI: 10.1093/bioinformatics/btu031

44. Bagos, P. G., Nikolaou, E. P., Liakopoulos, T. D., and Tsirigos, K. D. Combined prediction of Tat and Sec signal peptides with hidden Markov models. *Bioinformatics.* **2010**, 26, 2811-2817; DOI: 10.1093/bioinformatics/btq530

669 45. Darling, A. C., Mau, B., Blattner, F. R., and Perna, N. T. (2004) Mauve: multiple
670 alignment of conserved genomic sequence with rearrangements. *Genome Res.*
671 **2004,** 14, 1394-403; DOI: 10.1101/gr.2289704.
672 46. Gupta, D., Sutherland, M. C., Rengasamy, K., Meacham, J. M., Kranz, R. G., and
673 Bose, A. Photoferrotrophs produce a PioAB electron conduit for extracellular
674 electron uptake. *MBio*, **2019**, 10, e02668-02619; DOI: 10.1128/mBio.02668-19
675 47. Kimura, M. Model of effectively neutral mutations in which selective constraint is
676 incorporated. *Proc Natl Acad Sci U S A.* **1979**, 76, 3440-3444. DOI:
677 10.1073/pnas.76.7.3440
678 48. Jones, D. T., Taylor, W. R., and Thornton, J. M. The rapid generation of mutation
679 data matrices from protein sequences. *Bioinformatics.* **1992**, 8, 275-282; DOI:
680 10.1093/bioinformatics/8.3.275
681 49. Grabarczyk, D. B., and Berks, B. C. Intermediates in the Sox sulfur oxidation
682 pathway are bound to a sulfane conjugate of the carrier protein SoxYZ. *PLoS One.*
683 **2017**,12, e0173395; DOI: 10.1371/journal.pone.0173395
684
685
686
687
688