Metagenomic Analyses of the Autotrophic Fe(II)-Oxidizing, Nitrate-Reducing Enrichment Culture KS

Shaomei He[1,2,3,*], Claudia Tominski[4], Andreas Kappler[4], Sebastian Behrens[5,6], and Eric E. Roden[1,3*]

SUPPLEMENTARY INFORMATION

Including:

Supplementary Texts (1-9)

Supplementary Reference

Supplementary Figure Legends

Supplementary Figures (1-8)

Supplementary Tables (1-3)

SUPPLEMENTARY TEXTS

Supplementary Text 1. More discussion on metagenome binning

Although the coverage of several *Bradyrhizobium* contigs deviated from the majority coverage of the *Bradyrhizobium* bin (**Figure S2a**), these contigs indeed belong to the *Bradyrhizobium* bin, supported by other binning criteria such as sequence homology and tetranucleotide frequency, and these contigs likely originated from plasmids, as suggested by the presence of plasmid associated genes on these contigs.

Notably, *Rhizobium* bin was the second most abundant bin in the KS-Mad metagenome, yet was represented by many short fragments (**Table 1**), and consisted of at least two coverage groups, with the major coverage around 600 X (**Figure S2a**). In addition to the presence of plasmids, some of these low-coverage contigs are likely from very closely related strains of *Rhizobium*, as suggested by the presence of several redundant essential single-copy genes in low-coverage contigs. Indeed, the co-existence of closely related strains may have made it difficult to assemble their reads into long contigs (despite being present at very high abundances) when using short-read assemblers, due to the presence of identical regions among these strains. For

simplicity, we grouped all these *Rhizobium* contigs into a composite bin, with the majority of sequences contributed by the most abundant *Rhizobium* strain in the culture.

Supplementary Text 2. Comparison of the Mad and Tueb versions of *Gallionellaceae* sp. genomes

The *Gallionellaceae* genomes in the two enrichment cultures are nearly identical except for some very minor differences at the ends of contigs, which were likely due to the usual assembly artifacts at the ends of contigs. However, there are two major differences between these two genomes that we think are real. First, in KS-Mad, there is a deletion of a 170-bp fragment in the middle of a gene, which is annotated as "adenylate and guanylate cyclase catalytic domain/AsmA-like C-terminal region" in KS-Tueb (IMG gene OID 2566081571). Therefore, this gene was broken into two shorter fragments (2566094026 and 2566094027) in KS-Mad. The base coverage along Gene 2566081571 was consistent when mapping KS-Tueb metagenome reads to Gene 2566081571, however, when mapping KS-Mad metagenome reads to Gene 2566081571, the coverage dropped to zero at the 170-bp deletion region, confirming that the 170-bp deletion in KS-Mad genome is real, and is not due to assembly artifacts. The other difference lies in a 46805-bp contig in KS-Tueb (IMG Scaffold ID 2565963370), which has a coverage about four times higher than the average coverage of the rest contigs in KS-Tueb *Gallionellaceae* sp. genome. However, in KS-Mad, this fragment is part of a larger contig (IMG Scaffold ID 2565963526, 206118-bp long), which has an even coverage along the contig, and is consistent with the coverage of the rest contigs in KS-Mad *Gallionellaceae* sp. genome. Mapping KS-Mad metagenome reads to KS-Mad Contig 2565963526 has a consistent coverage along the contig, and is also consistent with other contigs in the genome, indicating that this contig is from chromosomal DNA in KS-Mad *Gallionellaceae* sp. genome. However, mapping KS-Tueb reads to KS-Mad Contig 2565963526 clearly indicates that the beginning ~46 kbp region on this contig has a coverage about 4 times higher than the remaining ~160 kbp region on this contig, confirming that these two regions are indeed two separate fragments with different copy numbers in KS-Tueb *Gallionellaceae* sp. genome. Therefore, contig

2565963370 in KS-Tueb is likely from a plasmid, which can have multiple copies in a cell, rather than from the chromosome, which has only one copy in the genome.

Supplementary Text 3. Taxonomy of Culture KS *Gallionellaceae* sp. and comparison to ES-1 and ES-2

Based on the 16S rRNA gene sequence, the *Gallionellaceae* species in Culture KS is more similar to ES-1 (95-96% identical) than to ES-2 (94% identical). However, its essential single-copy genes are more similar to ES-2 than to ES-1 (**Figure S1**), and the phylogenic tree constructed using conserved phylogenetic marker genes also indicated that the *Gallionellaceae* sp. in KS is more affiliated with ES-2 than with ES-1 (**Figure S3**). When considering the entire genome, about 51% and 52% of KS *Gallionellaceae* sp. proteins match with ES-1 and ES-2 respectively with a <1e-5 E-value, a >60% sequence identity and a >50% alignable region on the query sequence as the cutoff. These numbers increased to 71% and 68% respectively when the sequence identity cutoff was decreased to 30%. Recently, the percentage of conserved proteins (POCP) between two organisms defined by >40% sequence identity and >50% alignable region on the query sequence was used to assess evolutionary and phenotypic distance, and a POCP of 50% was proposed to define the genus boundary (1). According to that, the POCP is 56% between ES-1 and ES-2, 60% between KS *Gallionellaceae* sp. and ES-1, and 59% between KS *Gallionellaceae* sp. and ES-2. As ES-1 and ES-2 represent different genera, their POCP value suggests that the 50% cutoff is probably too low to distinguish genera within *Gallionellaceae*, and this is probably due to the relative small genomes (~3Mbp) of members within this family. Thus, the POCP cannot resolve the genus to which the *Gallionellaceae* sp. in Culture KS belongs. Overall, using the criteria discussed above, we cannot unambiguously determine whether *Gallionellaceae* sp. in KS belongs to either *Sideroxydans* or *Gallionella*, or represents a novel genus. More genome sequences from this family may help to resolve its phylogeny in the future. In this manuscript, we tentatively refer to the primary Fe(II)-oxidizer in Culture KS as *Gallionellaceae* sp., and comparison of its genetic features and metabolisms with ES-1 and ES-2 is presented in more detail below.

Supplementary Text 4. Other porin-periplasmic cyt*c*/MCO gene clusters

We also performed a broader search for gene clusters potentially encoding "porin-periplasmic cyt*c*" or "porin-periplasmic MCO" complexes that were not previously recognized, and found interesting cases in *Gallionellaceae* sp., *Bradyrhizobium* sp., *Comamonadaceae* sp., and *Rhodanobacter* sp.

We found cyt*c* or the cytochrome cd_1-containing nitrite reductase (NirS) genes next to porin-coding genes that are annotated as outer membrane receptors (OMR) involved in assimilatory iron uptake in *Gallionellaceae* sp., *Comamonadaceae* sp. and *Bradyrhizobium* sp. (**Figure S5a**), and we also found the clustering of NirS and OMR genes in ES-1 genome. Iron uptake with OMRs is usually facilitated by the TonB-dependent transporter system, but interestingly, in these cases, other genes in the TonB-dependent transporter system (including TonB and ExbB/D) are not in the vicinity of the OMR gene, probably suggesting that these OMRs are involved in functions other than iron uptake. As the cyt*c*/*nirS* is not always in the same operon as the OMR gene, it is not clear whether their clustering in the genome is merely a coincidence, or suggesting an interaction between them. If they are involved in EET, as these cyt*c*/NirS have only one to three heme-binding sites, they probably cannot form a conductive wire similar to MtoA.

In addition to porin-cyt*c*, we also searched for porin-MCO, since MCO is another type of enzyme involved in metal redox reactions, such as Mn(II) oxidation (2), Fe(III) reduction (3) and Cu(I) oxidation which confers copper resistance (4). We found a periplasmic MCO in the same operon as a porin-like protein in *Comamonadaceae* sp. and *Rhodanobacter* sp. from the KS cultures (**Figure S5b**). This type of porin-MCO was previously identified in Fe(II)-oxidizing *Bradyrhizobium japonicum* strain 22 as its best candidate system for EET, since neither outer membrane/extracellular redox active enzyme nor any porin-cytochrome system is present in its genome (L. Shi and E. Roden, unpublished data). However, porin-MCO was not present in *Bradyrhizobium* sp. in the KS metagenomes. We also identified such a porin-MCO system in several iron redox cycling bacteria isolated from Hanford subsurface (5), including *Cupriavidus necator* A5-1, *Bradyrhizobium japonicum* strain is5 and strain in8p8. In addition, we also found this porin-MCO in nitrate-reducing Fe(II)-oxidizing *Thiobacillus denitrificans* strains, and in a *Ralstonia* sp. genome constructed from a pyrite-oxidizing enrichment culture (Percak-Dennett et al., in preparation). In all these cases, the MCO and the associated porin belong to the PcoA and

PcoB protein families, respectively (**Figure S5b**). PcoAB were initially identified in *Escherichia coli* as components of a plasmid-borne copper-resistance operon (*pco*) (6), and are homologous to CopAB in a copper-resistance operon (*cop*) in *Pseudomonas syringae* (7). Among proteins encoded by the *pco* or *cop* operon, PcoAB or CopAB are essential to the resistance (8), and are more conserved in bacterial genomes than other components in the operon (9).

It is believed that PcoA detoxifies copper by oxidizing siderophores (which subsequently chelate and sequester Cu(I)), or by directly oxidizing the extremely toxic Cu(I) to less toxic Cu(II) (4, 9). Notably, PcoA may have a broad substrate spectrum and was suggested to oxidize compounds other than Cu(I) and siderophores. For example, PcoA in *P. aeruginosa* was demonstrated to exhibit a ferroxidase activity and play a role in iron acquisition by oxidizing Fe(II) to Fe(III) for subsequent transport using Fe(III) transporters (10). Therefore, it is plausible that PcoA in these iron redox cycling bacteria can also oxidize Fe(II), supported by the observation that some of their PcoA proteins have the EXXE iron-binding motif.

The role of PcoB in copper resistance is less understood, but PcoB may interact with PcoA and perform a putative function in exporting Cu(II) across the outer membrane (4). In these iron redox cycling bacteria, PcoB usually have 10-14 transmembrane motifs (**Figure 5b**), probably forming a small porin. Interestingly, one of the most highly expressed genes in *T. denitrificans* when Fe(II) was used as the electron donor for nitrate reduction encodes a hypothetical protein (locus tag Tbd_1320), which is upstream of *pcoAB* (Tbd_1324, Tbd_1325), and the authors attributed this to stress response and metal efflux (11), yet the culture medium did not seem to impose copper stress on this bacterium. This might lend some support for a potential role of PcoAB in Fe(II) oxidation. Taken together, we hypothesize that, in addition to conferring copper resistance as its previously identified role, PcoAB might form a porin-MCO system, in which PcoA oxidizes Fe(II). However, it is not clear whether PcoA is imbedded into PcoB to oxidize Fe(II) at the cell surface, or Fe(II) is oxidized in the periplasm by PcoA and exported through the pore formed by PcoB.

Supplementary Text 5. Additional discussion on the lack of NO and N_2O reductase genes in *Gallionellaceae* sp. genome

An alternative way to detoxify NO can be through a novel nitric oxide dismutase (NOD). Previously, a NOD was postulated for "*Candidatus* Mehylomicralillis oxyferea" as suggested by the observation that ^{18}O-labeled O_2 was formed when ^{18}O-labeled nitrite was added to this nitrite-dependent methane oxidation bacterium (12), and a candidate gene for such a dismutase was further proposed (13). If such an enzyme exists in *Gallionellaceae* sp., not only can it provide a means for NO detoxification, but also generate oxygen to serve as an alternative electron acceptor for biological Fe(II) oxidation (**Figure S6**), which can generates more PMF than denitrification (14) and thus increase energy capturing in NDFO. However, we did not find a homolog of "*Cand.* M. oxyferea" NOD in *Gallionellaceae* sp., but we cannot completely exclude the possibility of such a novel enzyme in *Gallionellaceae* sp. based on genome information alone.

Supplementary Text 6. Replacing of *b*-type cytochrome with heme-binding porin in the Group I hydrogenase gene cluster

In many bacteria (probably including *Rhizobium* sp. in Culture KS), the Group 1 [NiFe]-hydrogenase gene in the *hya* (or *hyd*, *hup, hox*) operon is often upstream of *hyaC* (or *hydC*, *hupC*, *hoxZ*), which encodes a *b*-type cytochrome subunit, and is necessary for transferring electrons from H_2 to the quinone pool and also for anchoring the membrane-bound hydrogenase complex to the periplasmic side of the inner membrane (15-18) (**Figure S8b**). However, the gene encoding this *b*-type cytochrome subunit is missing in *Gallionellaceae* sp., and is replaced with a gene encoding a beta barrel outer membrane protein with six PKD domains and 27 transmembrane motifs, likely forming a porin (**Figure S8a**). To the best of our knowledge, the insertion of such an outer membrane protein into the hydrogenase operon had not been reported. We temperately refer to this outer membrane protein as "HyaX", since its function is unknown. Interestingly, HyaX also contains one heme-binding motif, a signature of *c*-type cytochrome, which is intriguing given the expected EET feature of *Gallionellaceae* sp. BLASTP search against the NCBI nr database indicates that HyaX is about 30% identical to a number of cell surface proteins, which are often involved in binding extracellular carbohydrate polymer substrates. However these proteins don't share any homology with HyaX in the first ~100 amino acid region at the N-terminal region of HyaX, and they lack the heme-binding site, suggesting

the novelty of HyaX. It is not clear whether the replacement of the inner membrane *b*-type cytochrome with this outer membrane heme-binding HyaX can change the cellular location where the hydrogenase is attached from inner to outer membrane, which, if true, may alter the electron flow from between periplasm and inner membrane to between periplasm and outer membrane (as indicated in **Figure S8**). If HyaX does function as an outer membrane electron carrier and anchor, *Gallionellaceae* sp. Group 1 [NiFe]-hydrogenase may be able to transfer electrons from periplasmic H_2 to an extracellular electron acceptor (**Figure S8a**). However, preliminary experiments showed that Culture KS was not able to use H_2 to reduce ferrihydrite (unpublished data). Therefore, the function of HyaX and its implication in hydrogen metabolism remains to be elucidated.

Supplementary Text 7. Sulfur and phosphorus metabolisms

All genomes harbor genes for sulfate transporters and assimilatory sulfate reduction, yet only ES-1 has the genetic potential for dissimilatory sulfate reduction. Previously, Emerson and coworkers reported that sulfur oxidation (*sox*) genes were present in ES-1 but absent in ES-2, and they demonstrated that ES-1 was able to grow on thiosulfate (19). Similar to ES-2, *Gallionellaceae* sp. lacks *sox* genes, and thus cannot use reduced sulfur compounds as electron donors. However, *sox* genes are present in *Bradyrhizobium* sp., *Comamonadaceae* sp. and *Rhizobium* sp.

All genomes have high-affinity phosphate transporters, and some also contain low-affinity transporters. Most genomes possess polyphosphate kinase (PPK) and polyphosphatase (PPX) in polyphosphate accumulation and degradation.

Supplementary Text 8. Strategies to avoid encrustation

Neutrophilic Fe(II)-oxidizers often form distinctive extracellular structures such as stalks and sheaths to avoid cell encrustation with iron minerals. Previous genomic analyses revealed that ES-1 and ES-2 have the genetic machinery to produce extracellular polysaccharide substances (EPS), which was postulated to be a way to prevent encrustation (19). Clusters of genes encoding for EPS biosynthesis were identified in the *Gallionellaceae* sp. genome, probably serving a

similar role. Indeed, neither stalk nor cell encrustation was observed for Culture KS organisms under the microscope; cells largely remained free of Fe(III) minerals, and minerals only loosely associated with cell surface, indicating that Fe(III) minerals precipitated at a certain distance from the cell surface (20). Several mechanisms were proposed, including acidic pH in the microenvironment in the cell vicinity that allows the Fe(III) remain in solution; organics complexing with Fe(III) to increase its solubility; and cell surface modification to be positively charged to repel Fe(III) ions from the cell surface (20). Interestingly, a very large protein with 3596 amino acids was predicted to be an extracellular protein, annotated as "RTX toxins and related Ca^{2+}-binding protein", with a total of 44 Ca^{2+}-binding repeats. Previously, such RTX toxins and related Ca^{2+}-binding proteins were speculated to be part of bacterial EPS in a wastewater treatment bioreactor, given the affinity of EPS for cations (21). If this is also the case in *Gallionellaceae* sp., it is plausible that the binding of Ca^{2+} onto the EPS would make the cell surface more positively charged and therefore repel the Fe(III) ions from the immediate proximity of the cell membrane.

Supplementary Text 9. Environmental sensing and motility

Due to the kinetic favorability of abiotic Fe(II) oxidation by oxygen under neutral pH, neutrophilic biotic Fe(II) oxidation is expected to occur at microoxic or anoxic conditions where FeOB can compete with the rate of abiotic oxidation. Therefore, sensing oxygen concentration and locating themselves at a favorable biogeochemical gradient is critical for FeOB. Similar to ES-1 and ES-2, *Gallionellaceae* sp. possess a complete gene set for flagellar assembly, as well as 10 hemerythrin-like proteins, which may be involved in oxygen sensing in the environment (19), probably also be involved in NO sensing, as hemerythrins also form stable NO complexes HrNO. Like ES-1 and ES-2, *Gallionellaceae* sp. has an overrepresentation of genes encoding chemotaxis and signal transduction in its genome, suggesting an active life style adapted to moving towards a favorable biogeochemical gradient. For example, *Gallionellaceae* sp. genome contains NarXL, which is upstream of nitrate transporter and nitrate reductase (Nar) gene cluster. There is no NarX-Tar hybrid for chemotaxis to nitrate and nitrite (22). However, *Gallionellaceae* sp. has at least seven copies of nitrate- and nitrate sensing domain (NIT, pfam08376). It was proposed that NIT-containing receptors regulate gene expression and cell motility in response to

nitrate and/or nitrite (23). In *Gallionellaceae* sp., two methyl-accepting chemotaxis proteins (MCPs) contain a NIT domain in the N-terminal region, and three NIT-containing genes are next to MCP coding genes. Therefore, it is likely that these NIT-associated MCPs are involved in chemotaxis in response to nitrate and/or nitrite. These genes may facilitate *Gallionellaceae* sp. to move towards its favorable redox gradient in the sediment core.

SUPPLEMENTARY REFERENCES

1. **Qin QL, Xie BB, Zhang XY, Chen XL, Zhou BC, Zhou J, Oren A, Zhang YZ.** 2014. A proposed genus boundary for the prokaryotes based on genomic insights. J Bacteriol **196:**2210-2215.
2. **Soldatova A, Butterfield C, Oyerinde O, Tebo B, Spiro T.** 2012. Multicopper oxidase involvement in both Mn(II) and Mn(III) oxidation during bacterial formation of MnO2. JBIC Journal of Biological Inorganic Chemistry **17:**1151-1158.
3. **Mehta T, Childers SE, Glaven R, Lovley DR, Mester T.** 2006. A putative multicopper protein secreted by an atypical type II secretion system involved in the reduction of insoluble electron acceptors in Geobacter sulfurreducens. Microbiology **152:**2257-2264.
4. **Chaturvedi KS, Henderson JP.** 2014. Pathogenic adaptations to host-derived antibacterial copper. Frontiers in Cellular and Infection Microbiology **4:**3.
5. **Benzine J, Shelobolina E, Xiong MY, Kennedy DW, McKinley JP, Lin X, Roden EE.** 2013. Fe-phyllosilicate redox cycling organisms from a redox transition zone in Hanford 300 Area sediments. Frontiers in Microbiology **4:**388.
6. **Lee BTO, Brown NL, Rogers S, Bergemann A, Camakaris J, Rouch DA.** 1990. Bacterial Response to Copper in the Environment: Copper Resistance in Escherichia coli as a Model System, p 625-632. *In* Broekaert JAC, Güçer Ş, Adams F (ed), Metal Speciation in the Environment, vol 23. Springer Berlin Heidelberg.
7. **Cha JS, Cooksey DA.** 1991. Copper resistance in Pseudomonas syringae mediated by periplasmic and outer membrane proteins. Proc Natl Acad Sci U S A **88:**8915-8919.
8. **Mellano MA, Cooksey DA.** 1988. Nucleotide sequence and organization of copper resistance genes from Pseudomonas syringae pv. tomato. J Bacteriol **170:**2879-2883.
9. **Rensing C, Grass G.** 2003. Escherichia coli mechanisms of copper homeostasis in a changing environment. FEMS Microbiol Rev **27:**197-213.
10. **Huston WM, Jennings MP, McEwan AG.** 2002. The multicopper oxidase of Pseudomonas aeruginosa is a ferroxidase with a central role in iron acquisition. Mol Microbiol **45:**1741-1750.
11. **Beller HR, Zhou P, Legler TC, Chakicherla A, Kane S, Letain TE, P AOD.** 2013. Genome-enabled studies of anaerobic, nitrate-dependent iron oxidation in the chemolithoautotrophic bacterium Thiobacillus denitrificans. Front Microbiol **4:**249.
12. **Ettwig KF, Butler MK, Le Paslier D, Pelletier E, Mangenot S, Kuypers MM, Schreiber F, Dutilh BE, Zedelius J, de Beer D, Gloerich J, Wessels HJ, van Alen T, Luesken F, Wu ML, van de Pas-Schoonen KT, Op den Camp HJ, Janssen-Megens EM, Francoijs KJ, Stunnenberg H, Weissenbach J, Jetten MS, Strous M.** 2010. Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. Nature **464:**543-548.
13. **Ettwig KF, Speth DR, Reimann J, Wu ML, Jetten MSM, Keltjens JT.** 2012. Bacterial oxygen production in the dark. Frontiers in Microbiology **3:**273.
14. **Chen J, Strous M.** 2013. Denitrification and aerobic respiration, hybrid electron transport chains and co-evolution. Biochim Biophys Acta **1827:**136-144.

15. **Bernhard M, Benelli B, Hochkoeppler A, Zannoni D, Friedrich B.** 1997. Functional and structural role of the cytochrome b subunit of the membrane-bound hydrogenase complex of Alcaligenes eutrophus H16. Eur J Biochem **248:**179-186.
16. **Gross R, Pisa R, Sanger M, Lancaster CR, Simon J.** 2004. Characterization of the menaquinone reduction site in the diheme cytochrome b membrane anchor of Wolinella succinogenes NiFe-hydrogenase. J Biol Chem **279:**274-281.
17. **Meek L, Arp DJ.** 2000. The hydrogenase cytochrome b heme ligands of Azotobacter vinelandii are required for full H(2) oxidation capability. J Bacteriol **182:**3429-3436.
18. **Palagyi-Meszaros LS, Maroti J, Latinovics D, Balogh T, Klement E, Medzihradszky KF, Rakhely G, Kovacs KL.** 2009. Electron-transfer subunits of the NiFe hydrogenases in Thiocapsa roseopersicina BBS. FEBS J **276:**164-174.
19. **Emerson D, Field EK, Chertkov O, Davenport KW, Goodwin L, Munk C, Nolan M, Woyke T.** 2013. Comparative genomics of freshwater Fe-oxidizing bacteria: implications for physiology, ecology, and systematics. Front Microbiol **4:**254.
20. **Schädler S, Burkhardt C, Hegler F, Straub K, Miot J, Benzerara K, Kappler A.** 2009. Formation of cell-iron-mineral aggregates by phototrophic and nitrate-reducing anaerobic Fe (II)-oxidizing bacteria. Geomicrobiology Journal **26:**93-103.
21. **Garcia Martin H, Ivanova N, Kunin V, Warnecke F, Barry KW, McHardy AC, Yeates C, He S, Salamov AA, Szeto E, Dalin E, Putnam NH, Shapiro HJ, Pangilinan JL, Rigoutsos I, Kyrpides NC, Blackall LL, McMahon KD, Hugenholtz P.** 2006. Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities. Nat Biotechnol **24:**1263-1269.
22. **Ward SM, Delgado A, Gunsalus RP, Manson MD.** 2002. A NarX-Tar chimera mediates repellent chemotaxis to nitrate and nitrite. Mol Microbiol **44:**709-719.
23. **Shu CJ, Ulrich LE, Zhulin IB.** 2003. The NIT domain: a predicted nitrate-responsive module in bacterial sensory receptors. Trends in Biochemical Sciences **28:**121-124.

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. Taxonomic assignment of essential single-copy genes in metagenomes KS-Mad (**a**) and KS-Tueb (**b**) by MEGAN. The number next to the taxon at each node indicates the number of genes assigned to that taxon based on the lowest common ancestor algorithm. For the leaf nodes, this number is also reflected by the size of the circle. The bar with a number next to each leaf node taxon indicate the average fold coverage of contigs which contain genes assigned to that leaf node. Based on the expected numbers of essential single-copy genes for complete genomes (105-106) and the fold coverage pattern of these leaf nodes, essential single-copy genes were assigned to bins indicated by different colors, and these bins were assigned to the level of Family or Genus, with the Family or Genus name labeled with bold blue fonts.

Figure S2. Binning results of KS-Mad metagenome, with contigs and sequence fragments from different populations indicated by different colors. Contig GC contents and fold coverage (indicative of organism abundance in the community) are shown in (**a**), with each point representing a contig, and the point size reflecting the contig length. (**b**) is the tiled display of an Emergent Self-organizing Map (ESOM) based on the tetranucleotide frequency calculated with a window size of 5 kbp for contigs longer than 3 kbp, with each dot representing a 5-kbp fragment (or a contig if its length is shorter than 5 kbp). An ESOM (as outlined by the white rectangular) is a borderless map, continuous from the top to the bottom and from left to right. The elevations on the map (shown in white and brown) indicate large differences in tetranucleotide frequency, and therefore suggest divisions between taxonomic groups.

Figure S3. Phylogenetic tree constructed with conserved phylogenetic marker genes by using the PhyloPhlAn analysis pipeline. Protein sequences from all these genomes were input to PhyloPhlAn for extracting and individually aligning the conserved phylogenetic marker proteins. The alignments were concatenated for phylogenetic tree construction using the FastTree algorithm with default settings.

Figure S4. Phylogenetic tree constructed with *Gallionellaceae* sp. and *Rhodanobacter* sp. Cyc2-like proteins (in red) and sequences homologous to them. The Cyc2 protein from *A. ferroxidans* (in blue) and a Cyc2-like protein in *Bradyrhizobium* sp. genome recovered from a pyrite-oxidizing enrichment culture (in green) were also included as references. IMG Gene Object IDs are indicated in the brackets.

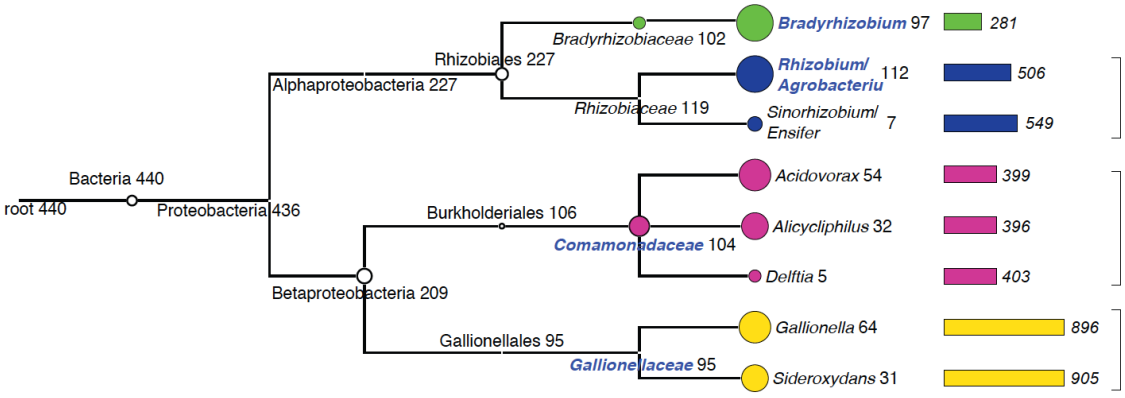
Figure S5. Other gene clusters encoding porin and periplasmic cyt*c* (**a**) or porin and periplasmic MCO (**b**). In (**a**), a periplasmic cyt*c*/*nirS* is clustered with a porin-coding gene annotated as outer membrane receptor (OMR). In (**b**), a periplasmic MCO (PcoA) is in the same operon with a porin (PcoB). In addition to bacteria in Culture KS (in blue), we also included a draft genome recovered from a pyrite-oxidizing enrichment culture and genomes of several iron redox cycling bacteria isolated from Hanford subsurface (in black) and publically available genomes (in grey). For each gene cluster, IMG Gene OID for the porin gene is listed in the parenthesis. Porin-coding genes are indicated with green arrows; cyt*c* are indicated with red arrows, and MCO are indicated with brown arrows. The number in the cyt*c*-coding genes indicates the number of heme-binding sites, and the number in the porin-coding genes indicates the number of transmembrane regions. Predicted cellular locations of porin, cyt*c* and MCO are indicated by different line types under the gene: thick solid lines for outer membrane, and dash lines for periplasmic, respectively.

Figure S6. A hypothesized electron transfer pathway proposed for NDFO by *Gallionellaceae* sp., if a nitric oxide dismutase (NOD) is present. Electron flows are indicated in yellow arrows, and chemical reduction of NO by Fe(II) is indicated in red arrows. Abbreviations: OM, outer membrane; IM, inner membrane; Nar, dissimilatory nitrate reductase complex; NarK, nitrate:nitrite antiporter; NirS/K: cytochrome cd_1- and copper-type nitrite reductase respectively; cbb_3, cbb_3-type cytochrome *c* oxidase.

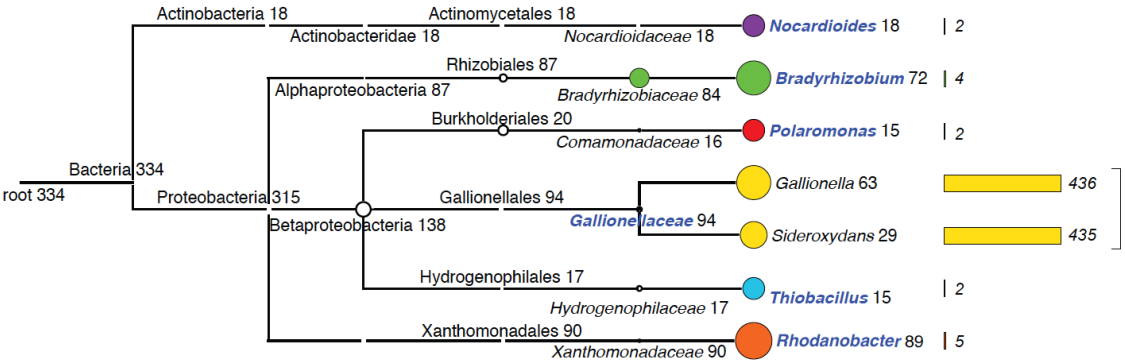
Figure S7. Phylogenetic tree of the large subunit of ribulose bisphosphate carboxylase (RuBisCO). RuBisCO sequences from the two metagenomes were colored (red: KS-Mad; blue: KS-Tueb), with their putative taxonomy assignments preceding their IMG Object IDs in the brackets.

Figure S8. Gene organization and predicted cellular location of the Group I uptake [NiFe]-hydrogenase encoded by the *hya* operon in *Gallionellaceae* sp. (**a**), as compared to the *hya* operon in *E. coli*, the *hyd* operon in *Wolinella succinogenes* and the Group I uptake [NiFe]-hydrogenase in the *Rhizobium* sp. draft genome in this study (**b**). Note that the hydrogenase operon in *Rhizobium* sp. was split into two contigs, which were manually joined in this figure, with a vertical dash line indicating the boundary of the two contigs. The novel porin-like protein (HyaX) in *Gallionellaceae* sp. has a heme-binding site (indicated by the vertical bar on the gene) (**a**), and it replaces *hyaC* in *E. coli, hydC* in *W. succinogenes* and *hyaC* in *Rhizobium* sp. (**b**). Electron flows are indicated with yellow arrows, and the generation and consumption of proton motive force are indicated with red and green arrows respectively. In (**a**), it is not clear whether electrons are passed to an extracellular electron acceptor (EEA), which is then transformed from an oxidized form, $EEA_{(ox)}$ to a reduced form, $EEA_{(red)}$, and therefore was marked with "???". Abbreviations: OM, outer membrane; PS, periplasmic space; IM, inner membrane; HoxS, soluble NAD-reducing hydrogenase encoded by the *hoxS* operon.

SUPPLEMENTARY FIGURES

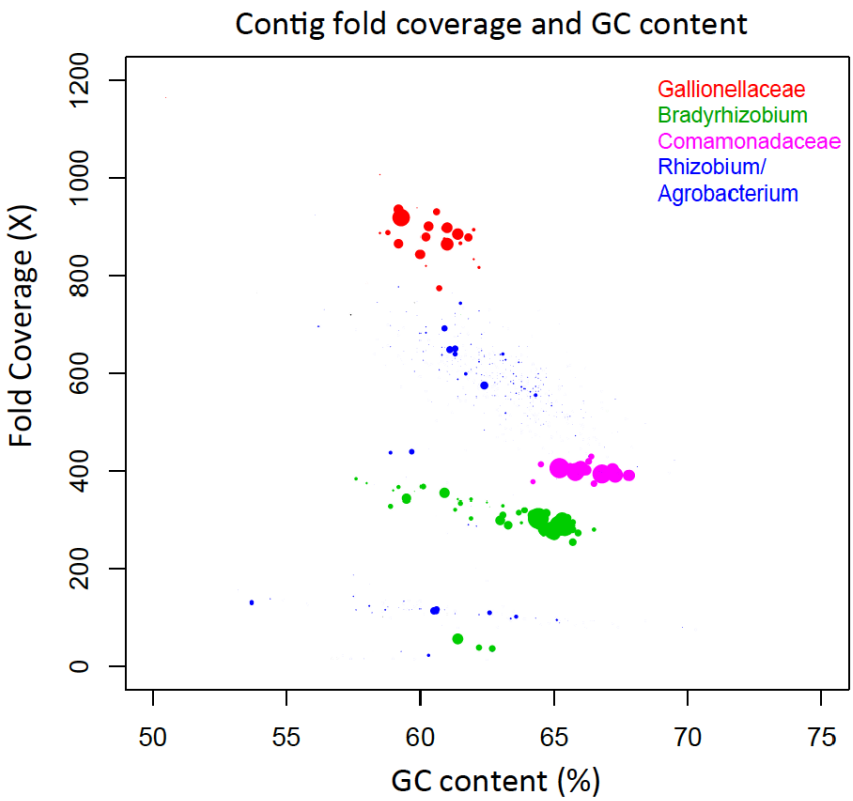


(**a**)

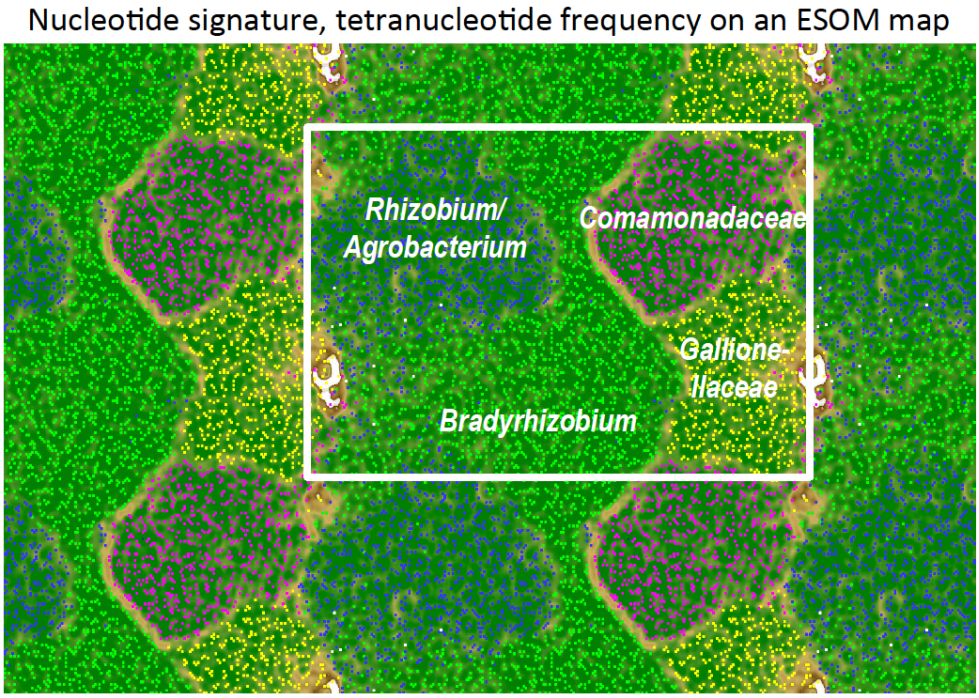


(**b**)

Figure S1.


Contig fold coverage and GC content
Gallionellaceae
Bradyrhizobium
Comamonadaceae
Rhizobium/
Agrobacterium
Fold Coverage (X)
0
200
400
600
800
1000
1200
50
55
60
65
70
75
GC content (%)

(a)


Nucleotide signature, tetranucleotide frequency on an ESOM map
Rhizobium/
Agrobacterium
Comamonadaceae
Gallione-
llaceae
Bradyrhizobium

(b)

Figure S2.

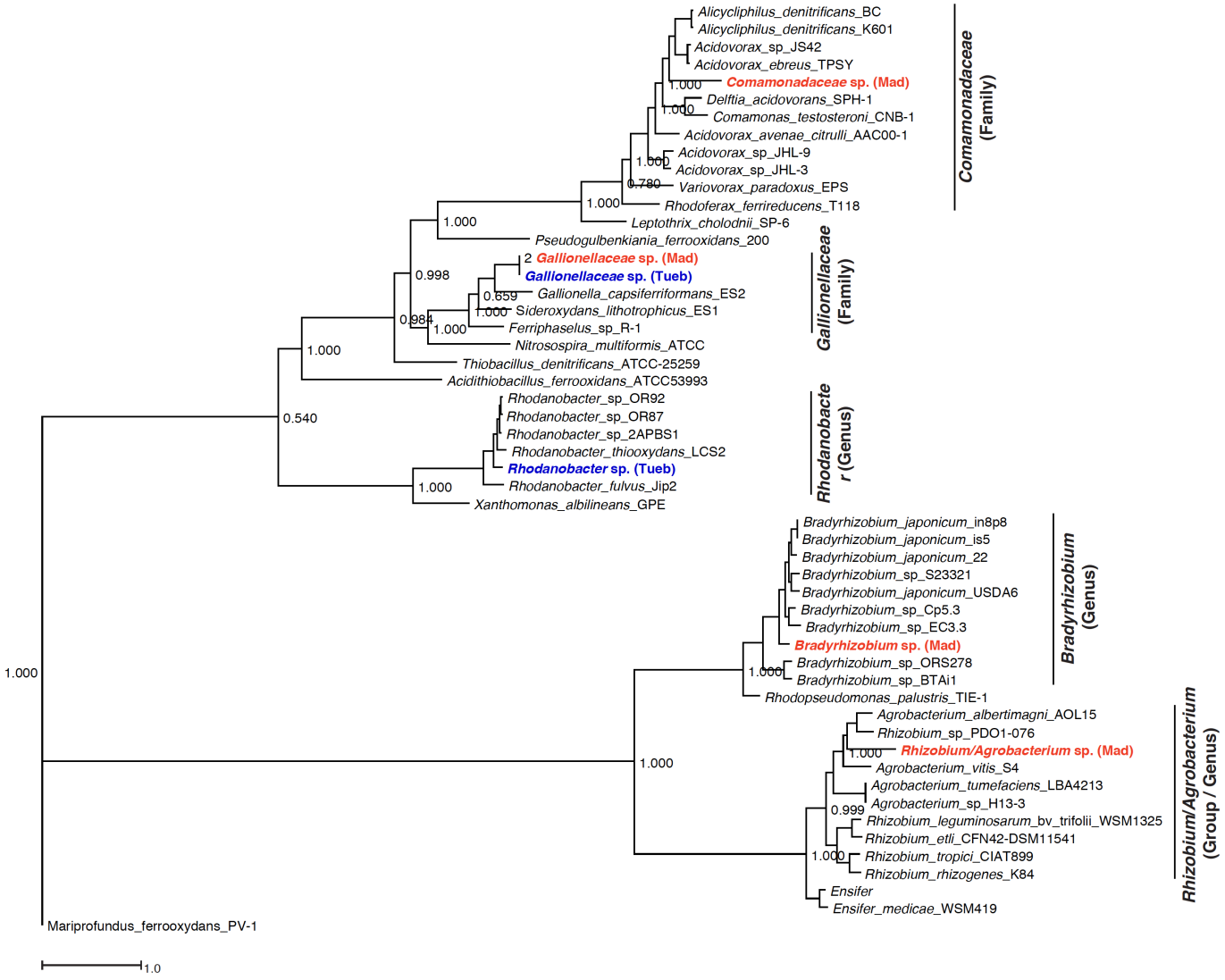

Alicycliphilus_denitrificans_BC
Alicycliphilus_denitrificans_K601
Acidovorax_sp_JS42
Acidovorax_ebreus_TPSY
1.000
Comamonadaceae sp. (Mad)
Delftia_acidovorans_SPH-1
1.000
Comamonas_testosteroni_CNB-1
Acidovorax_avenae_citrulli_AAC00-1
Acidovorax_sp_JHL-9
1.000
Acidovorax_sp_JHL-3
0.780
Variovorax_paradoxus_EPS
1.000
Rhodoferax_ferrireducens_T118
Comamonadaceae (Family)
1.000
Leptothrix_cholodnii_SP-6
Pseudogulbenkiania_ferrooxidans_200
2
Gallionellaceae sp. (Mad)
0.998
Gallionellaceae sp. (Tueb)
0.659
Gallionella_capsiferriformans_ES2
1.000
Sideroxydans_lithotrophicus_ES1
0.984
1.000
Ferriphaselus_sp_R-1
Nitrosospira_multiformis_ATCC
1.000
Thiobacillus_denitrificans_ATCC-25259
Acidithiobacillus_ferrooxidans_ATCC53993
Gallionellaceae (Family)
Rhodanobacter_sp_OR92
Rhodanobacter_sp_OR87
Rhodanobacter_sp_2APBS1
0.540
Rhodanobacter_thiooxydans_LCS2
Rhodanobacter sp. (Tueb)
1.000
Rhodanobacter_fulvus_Jip2
Xanthomonas_albilineans_GPE
Rhodanobacter (Genus)
Bradyrhizobium_japonicum_in8p8
Bradyrhizobium_japonicum_is5
Bradyrhizobium_japonicum_22
Bradyrhizobium_sp_S23321
Bradyrhizobium_japonicum_USDA6
Bradyrhizobium_sp_Cp5.3
Bradyrhizobium_sp_EC3.3
Bradyrhizobium sp. (Mad)
1.000
Bradyrhizobium_sp_ORS278
Bradyrhizobium_sp_BTAi1
Rhodopseudomonas_palustris_TIE-1
Bradyrhizobium (Genus)
1.000
Agrobacterium_albertimagni_AOL15
Rhizobium_sp_PDO1-076
1.000
Rhizobium/Agrobacterium sp. (Mad)
Agrobacterium_vitis_S4
Agrobacterium_tumefaciens_LBA4213
Agrobacterium_sp_H13-3
0.999
Rhizobium_leguminosarum_bv_trifolii_WSM1325
Rhizobium_etli_CFN42-DSM11541
1.000
Rhizobium_tropici_CIAT899
Rhizobium_rhizogenes_K84
Ensifer
Ensifer_medicae_WSM419
Rhizobium/Agrobacterium (Group / Genus)
1.000
Mariprofundus_ferrooxydans_PV-1
1.0

Figure S3.

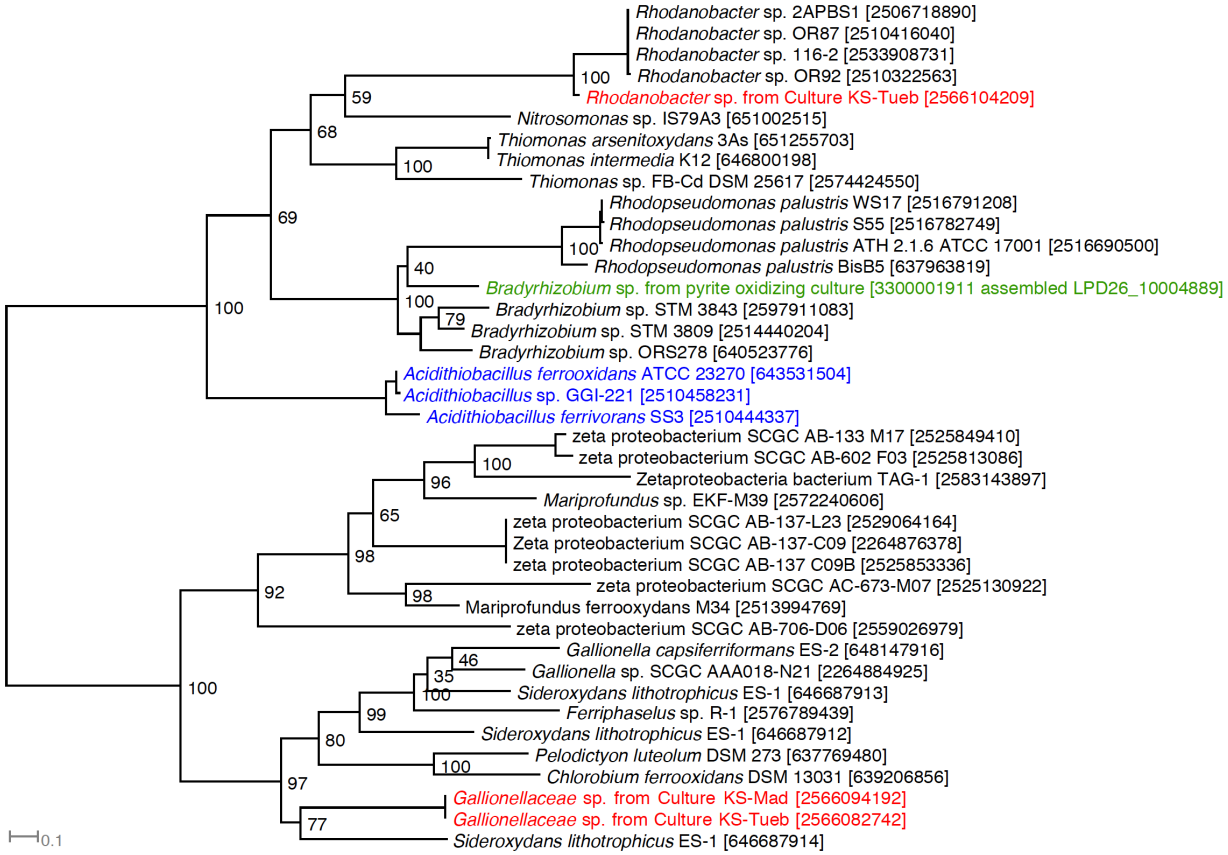

Rhodanobacter sp. 2APBS1 [2506718890]
Rhodanobacter sp. OR87 [2510416040]
Rhodanobacter sp. 116-2 [2533908731]
Rhodanobacter sp. OR92 [2510322563]
Rhodanobacter sp. from Culture KS-Tueb [2566104209]
Nitrosomonas sp. IS79A3 [651002515]
Thiomonas arsenitoxydans 3As [651255703]
Thiomonas intermedia K12 [646800198]
Thiomonas sp. FB-Cd DSM 25617 [2574424550]
Rhodopseudomonas palustris WS17 [2516791208]
Rhodopseudomonas palustris S55 [2516782749]
Rhodopseudomonas palustris ATH 2.1.6 ATCC 17001 [2516690500]
Rhodopseudomonas palustris BisB5 [637963819]
Bradyrhizobium sp. from pyrite oxidizing culture [3300001911 assembled LPD26_10004889]
Bradyrhizobium sp. STM 3843 [2597911083]
Bradyrhizobium sp. STM 3809 [2514440204]
Bradyrhizobium sp. ORS278 [640523776]
Acidithiobacillus ferrooxidans ATCC 23270 [643531504]
Acidithiobacillus sp. GGI-221 [2510458231]
Acidithiobacillus ferrivorans SS3 [2510444337]
zeta proteobacterium SCGC AB-133 M17 [2525849410]
zeta proteobacterium SCGC AB-602 F03 [2525813086]
Zetaproteobacteria bacterium TAG-1 [2583143897]
Mariprofundus sp. EKF-M39 [2572240606]
zeta proteobacterium SCGC AB-137-L23 [2529064164]
Zeta proteobacterium SCGC AB-137-C09 [2264876378]
zeta proteobacterium SCGC AB-137 C09B [2525853336]
zeta proteobacterium SCGC AC-673-M07 [2525130922]
Mariprofundus ferrooxydans M34 [2513994769]
zeta proteobacterium SCGC AB-706-D06 [2559026979]
Gallionella capsiferriformans ES-2 [648147916]
Gallionella sp. SCGC AAA018-N21 [2264884925]
Sideroxydans lithotrophicus ES-1 [646687913]
Ferriphaselus sp. R-1 [2576789439]
Sideroxydans lithotrophicus ES-1 [646687912]
Pelodictyon luteolum DSM 273 [637769480]
Chlorobium ferrooxidans DSM 13031 [639206856]
Gallionellaceae sp. from Culture KS-Mad [2566094192]
Gallionellaceae sp. from Culture KS-Tueb [2566082742]
Sideroxydans lithotrophicus ES-1 [646687914]
100
59
68
100
69
100
40
100
79
100
96
65
98
92
98
46
35
100
99
80
100
97
77
0.1

Figure S4.

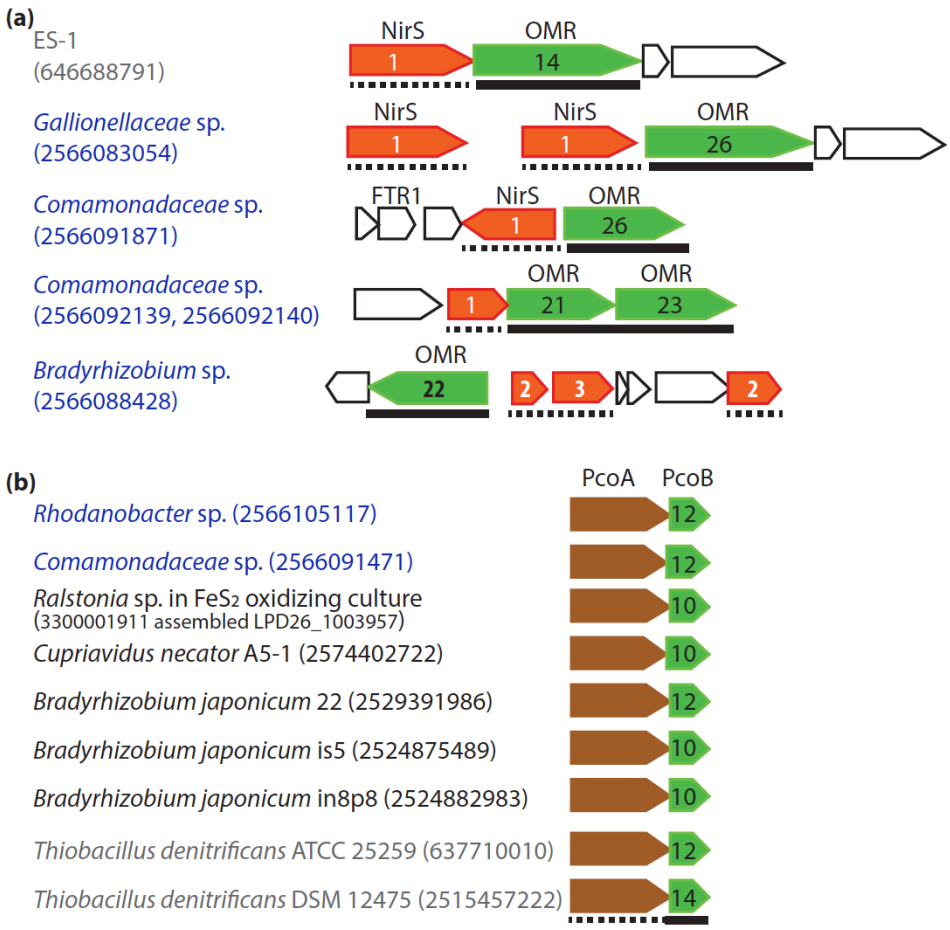

(a)
ES-1
(646688791)
NirS
1
OMR
14
Gallionellaceae sp.
(2566083054)
NirS
1
NirS
1
OMR
26
Comamonadaceae sp.
(2566091871)
FTR1
NirS
1
OMR
26
Comamonadaceae sp.
(2566092139, 2566092140)
OMR
1
21
OMR
23
Bradyrhizobium sp.
(2566088428)
OMR
22
2
3
2
(b)
PcoA
PcoB
Rhodanobacter sp. (2566105117)
12
Comamonadaceae sp. (2566091471)
12
Ralstonia sp. in FeS2 oxidizing culture
(3300001911 assembled LPD26_1003957)
10
Cupriavidus necator A5-1 (2574402722)
10
Bradyrhizobium japonicum 22 (2529391986)
12
Bradyrhizobium japonicum is5 (2524875489)
10
Bradyrhizobium japonicum in8p8 (2524882983)
10
Thiobacillus denitrificans ATCC 25259 (637710010)
12
Thiobacillus denitrificans DSM 12475 (2515457222)
14

Figure S5.

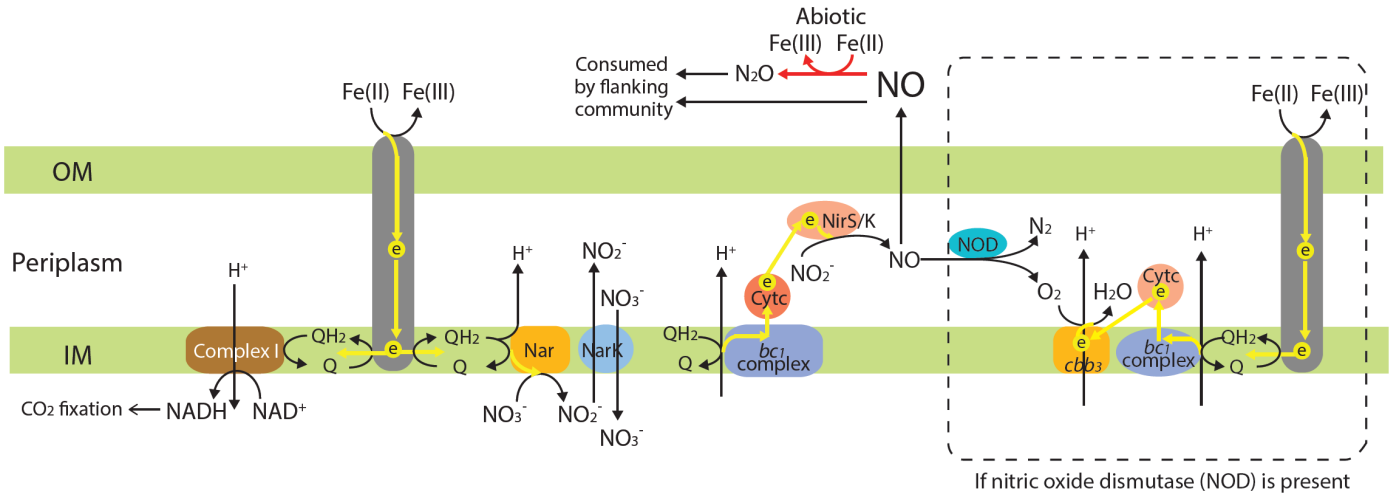

Abiotic
Fe(III) Fe(II)
Consumed by flanking community
N2O
NO
Fe(II) Fe(III)
Fe(II) Fe(III)
OM
Periplasm
IM
H+
e
NirS/K
NO2-
NO
NOD
N2
O2
H2O
Cytc
Complex I
QH2
Q
Nar
NarK
NO3-
NO2-
bc1 complex
cbb3
CO2 fixation
NADH
NAD+
If nitric oxide dismutase (NOD) is present

Figure S6.

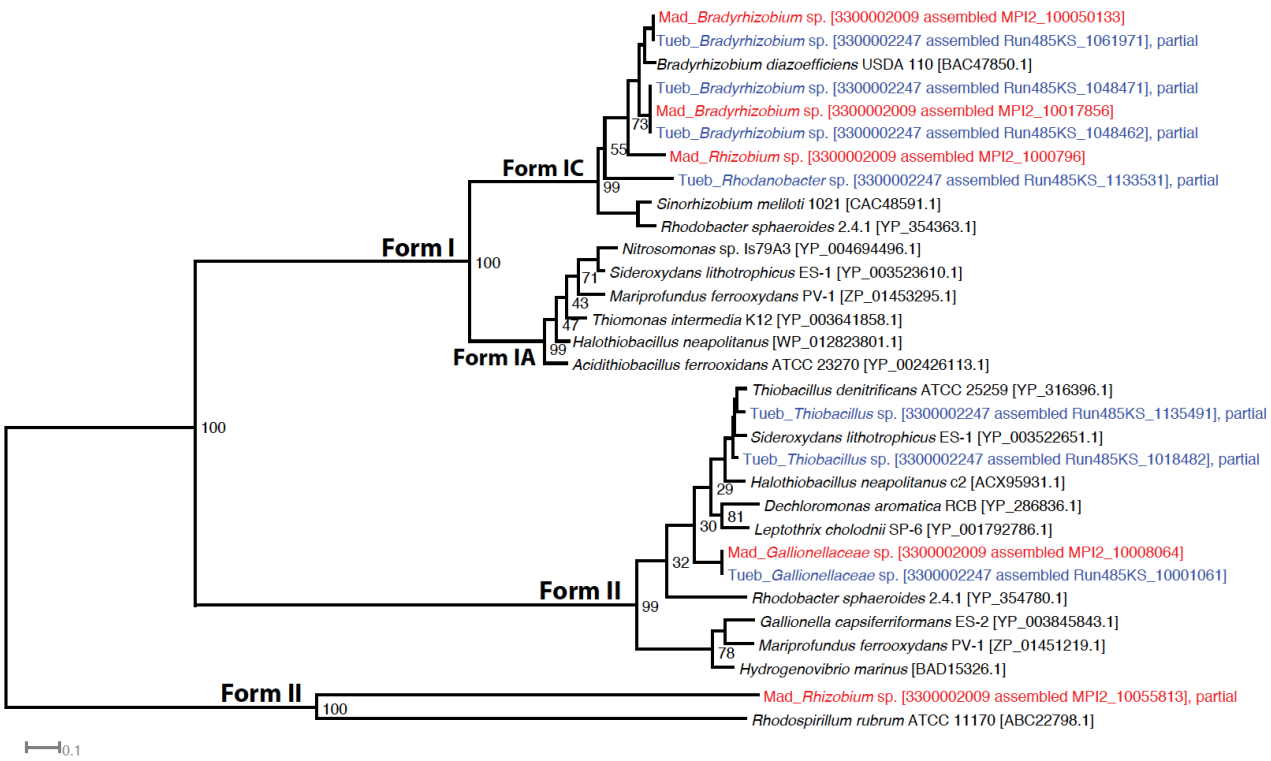

Mad_Bradyrhizobium sp. [3300002009 assembled MPI2_100050133]
Tueb_Bradyrhizobium sp. [3300002247 assembled Run485KS_1061971], partial
Bradyrhizobium diazoefficiens USDA 110 [BAC47850.1]
Tueb_Bradyrhizobium sp. [3300002247 assembled Run485KS_1048471], partial
Mad_Bradyrhizobium sp. [3300002009 assembled MPI2_10017856]
Tueb_Bradyrhizobium sp. [3300002247 assembled Run485KS_1048462], partial
Mad_Rhizobium sp. [3300002009 assembled MPI2_1000796]
Tueb_Rhodanobacter sp. [3300002247 assembled Run485KS_1133531], partial
Sinorhizobium meliloti 1021 [CAC48591.1]
Rhodobacter sphaeroides 2.4.1 [YP_354363.1]
Nitrosomonas sp. Is79A3 [YP_004694496.1]
Sideroxydans lithotrophicus ES-1 [YP_003523610.1]
Mariprofundus ferrooxydans PV-1 [ZP_01453295.1]
Thiomonas intermedia K12 [YP_003641858.1]
Halothiobacillus neapolitanus [WP_012823801.1]
Acidithiobacillus ferrooxidans ATCC 23270 [YP_002426113.1]
Thiobacillus denitrificans ATCC 25259 [YP_316396.1]
Tueb_Thiobacillus sp. [3300002247 assembled Run485KS_1135491], partial
Sideroxydans lithotrophicus ES-1 [YP_003522651.1]
Tueb_Thiobacillus sp. [3300002247 assembled Run485KS_1018482], partial
Halothiobacillus neapolitanus c2 [ACX95931.1]
Dechloromonas aromatica RCB [YP_286836.1]
Leptothrix cholodnii SP-6 [YP_001792786.1]
Mad_Gallionellaceae sp. [3300002009 assembled MPI2_10008064]
Tueb_Gallionellaceae sp. [3300002247 assembled Run485KS_10001061]
Rhodobacter sphaeroides 2.4.1 [YP_354780.1]
Gallionella capsiferriformans ES-2 [YP_003845843.1]
Mariprofundus ferrooxydans PV-1 [ZP_01451219.1]
Hydrogenovibrio marinus [BAD15326.1]
Mad_Rhizobium sp. [3300002009 assembled MPI2_10055813], partial
Rhodospirillum rubrum ATCC 11170 [ABC22798.1]
Form IC
Form I
Form IA
Form II
Form II
73
55
99
100
71
43
47
99
100
29
81
30
32
99
78
100
0.1

Figure S7.

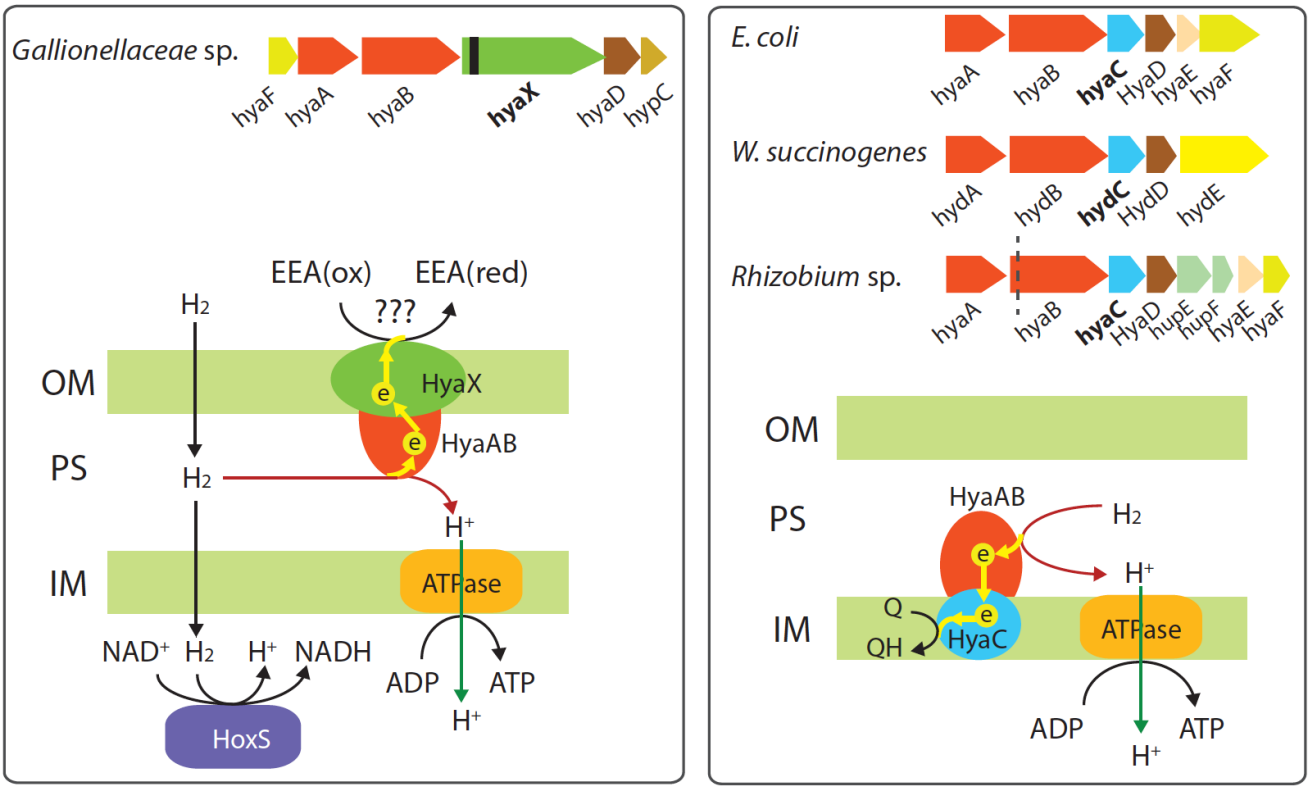

Gallionellaceae sp.
hyaF
hyaA
hyaB
hyaX
hyaD
hypC
EEA(ox)
EEA(red)
???
H2
OM
HyaX
e
HyaAB
PS
H2
H+
IM
ATPase
NAD+
H2
H+
NADH
ADP
ATP
H+
HoxS
(a)
E. coli
hyaA
hyaB
hyaC
HyaD
hyaE
hyaF
W. succinogenes
hydA
hydB
hydC
HydD
hydE
Rhizobium sp.
hyaA
hyaB
hyaC
HyaD
hupE
hupF
hyaE
hyaF
OM
PS
IM
HyaAB
H2
e
H+
Q
QH
e
HyaC
ATPase
ADP
ATP
H+
(b)

Figure S8.

SUPPLEMENTARY TABLES

Table S1. Metagenome sequencing and assembly statistics

Sequencing	**KS-Mad**	**KS-Tueb**
Platform	HiSeq 2000	MiSeq
Total No. of reads	111,887,076	5,915,054
Read length (bp)	100	250
Total raw sequence (Gb)	11.3	1.5
Assembly		
N75 (bp)	7,003	682
N50 (bp)	90,981	1,291
N25 (bp)	255,379	2,859
Minimum length (bp)	198	188
Maximum (bp)	606,530	435,491
Average (bp)	3,470	1,023
Count	6,238	14,698
Total length (bp)	21,648,289	15,029,786

Table S2. Classification and BLASTN search of 16S rRNA genes in the metagenomes

Locus Tag	Gene length (bp)[a]	Contig length (bp)[b]	Fold coverage (X)[c]	Classification by RDP classifier[d]	Top BLASTN hit to NCBI 16S rRNA database	Top BLASTN hit to clone libraries by Bloethe and Roden (2009)[e]
KS-Mad						
MPI2_1002824	1420	5046	2515	*Sideroxydans*	96% to ES-1 (94% to ES-2)	**100% to clone F30F68 (OTU1)**
MPI2_1004612	1366	5950	2981	*Rhizobium*	98% to *Rhizobium rosettiformans* W3	90% to *Parvibaculum* sp. MBNA2 (OTU3)
MPI2_1002014	1371	29769	580	*Bradyrhizobium*	99% to *Bradyrhizobium valentinum* LmjM3	89% to *Parvibaculum* sp. MBNA2 (OTU3)
MPI2_1010605	1415	5709	1777	*Comamonas*	97% to *Comamonas badia* IAM 14839	95% to *Comamonas* sp. MPI12 (OTU2)
MPI2_1060162	146	260	6	*Nocardioides*	100% to *Nocardioides* sp. JS614 JS614	93% to *Rhodanobacter* sp. MBNA3 (OTU4)
MPI2_1059442	89	271	4	*Alphaproteobacteria*	97% to *Nordella oligomobilis* N21	No hit
KS-Tueb						
Run485KS_1000224	1420	63602	650	*Sideroxydans*	96% to ES-1 (94% to ES-2)	**100% to clone F30F68 (OTU1)**
Run485KS_10002928	1081	25511	399	*Sideroxydans*	95% to ES-1 (94% to ES-2)	**100% to clone F29F67 (OTU1)**
Run485KS_1009693	1427	3882	6	*Rhodanobacter*	99% to *Rhodanobacter denitrificans* 2APBS1	**100% to Rhodanobacter sp. MBNA3 (OTU4)**
Run485KS_1083141	336	1644	5	*Bradyrhizobium*	100% to *Bradyrhizobium liaoningense* 2281	89% to *Parvibaculum* sp. MBNA2 (OTU3)
Run485KS_1118871	744	744	4	*Bradyrhizobium*	99% to *Bradyrhizobium valentinum* LmjM3	89% to *Parvibaculum* sp. MBNA2 (OTU3)
Run485KS_1098953	366	981	2	*Bradyrhizobiaceae*	97% to *Bradyrhizobium iriomotense* NBRC 102520	88% to *Rhodanobacter* sp. MBNA3 (OTU4)
Run485KS_1046253	372	2042	5	*Polaromonas*	99% to *Polaromonas* sp. JS666 JS666	90% to clone F32F70 (OTU1)
Run485KS_1124581	269	269	1	*Polaromonas*	100% to *Polaromonas* sp. JS666 JS666	95% to *Comamonas* sp. MPI12 (OTU2)
Run485KS_1118021	374	443	6	*Comamonadaceae*	99% to *Rhodoferax fermentans* FR2	90% to clone F33F71 (OTU1)
Run485KS_1095781	479	480	3	*Thiobacillus*	98% to *Thiobacillus thioparus* THI 111	88% to clone F33F71 (OTU1)
Run485KS_1123062	113	692	2	*Thiobacillus*	98% to *Thiobacillus thiophilus* D24TN	92% to clone F33F71 (OTU1)
Run485KS_1142341	468	468	2	*Nocardioides*	99% to *Nocardioides* sp. JS614 JS614	83% to *Parvibaculum* sp. MBNA2 (OTU3)
Run485KS_1142311	128	584	3	*Nocardioides*	98% to *Nocardioides* sp. JS614 JS614	No hit
Run485KS_1083691	250	250	1	*Betaproteobacteria*	96% to *Propionivibrio dicarboxylicus* CreMal1	**99% to clone F33F71 (OTU1)**

[a]The length of the recovered 16S rRNA genes/fragments
[b]The length of the 16S-rRNA gene containing contig
[c]Fold coverage of the 16S rRNA genes/fragments. This number is higher than the average contig fold coverage within the draft genome, due to multiple copies of *rrn* operon in the genome.
[d]Classification of 16S rRNA genes/fragments by the Ribosomal Database Project (RDP; http://rdp.cme.msu.edu/) classifier with a confidence level of 80%.
[e]Blothe M, Roden EE. 2009. Composition and activity of an autotrophic Fe(II)-oxidizing, nitrate-reducing enrichment culture. Appl Environ Microbiol 75:6937-6940.

Table S3. Counts of amino acid and carbohydrate transporters.

COG	COG Functions	ES-1	ES-2	*Gallionellaceae* (Tueb)	*Gallionellaceae* (Mad)	*Bradyrhizobium*	*Comamonadaceae*	*Rhizobium*	*Rhodanobacter*
Amino Acid transporters									
COG0765	ABC-type amino acid transport system, permease component	0	0	0	0	4	7	10	0
COG4597	ABC-type amino acid transport system, permease component	0	0	0	0	2	0	2	0
COG0410	ABC-type branched-chain amino acid transport system, ATPase component	2	1	0	0	33	18	7	0
COG0411	ABC-type branched-chain amino acid transport system, ATPase component	1	1	0	0	33	17	2	0
COG0683	ABC-type branched-chain amino acid transport system, periplasmic component	2	2	1	1	64	22	8	0
COG4177	ABC-type branched-chain amino acid transport system, permease component	1	1	0	0	27	15	6	0
COG3842	ABC-type Fe3+/spermidine/putrescine transport systems, ATPase components	1	0	1	1	2	1	11	0
COG1135	ABC-type methionine transport system, ATPase component	0	0	0	0	0	1	0	0
COG2011	ABC-type methionine transport system, permease component	0	0	0	0	0	1	0	1
COG4608	ABC-type oligopeptide transport system, ATPase component	0	0	0	0	3	1	1	0
COG4166	ABC-type oligopeptide transport system, periplasmic component	0	1	1	1	3	1	2	0
COG1126	ABC-type polar amino acid transport system, ATPase component	0	0	0	0	3	4	7	0
COG4175	ABC-type proline/glycine betaine transport system, ATPase component	0	0	0	0	0	1	5	0
COG1125	ABC-type proline/glycine betaine transport system, ATPase component	0	0	0	0	1	0	0	0
COG2113	ABC-type proline/glycine betaine transport system, periplasmic component	0	0	0	0	0	1	2	0
COG4176	ABC-type proline/glycine betaine transport system, permease component	0	0	0	0	0	1	2	0
COG1174	ABC-type proline/glycine betaine transport system, permease component	0	0	0	0	0	0	0	1
COG1176	ABC-type spermidine/putrescine transport system, permease component I	1	0	1	1	1	0	7	0
COG1177	ABC-type spermidine/putrescine transport system, permease component II	1	0	1	1	2	0	7	0
COG0747	ABC-type transport system, periplasmic component	0	0	0	0	6	2	9	0
COG0833	Amino acid permease	0	0	0	0	0	0	1	0
COG0531	Amino acid transporter	0	0	1	1	1	0	2	5
COG1279	Arginine exporter protein ArgO	0	0	0	0	0	0	1	0
COG0559	Branched-chain amino acid ABC-type transport system, permease component	2	1	0	0	35	18	9	0
COG3104	Dipeptide/tripeptide permease	0	0	0	0	0	0	0	4
COG1115	Na+/alanine symporter	0	0	0	0	0	0	1	0
COG0591	Na+/proline symporter	1	1	1	1	1	1	1	0
COG1296	Predicted branched-chain amino acid permease (azaleucine resistance)	0	0	0	0	1	2	1	0
COG2095	Small neutral amino acid transporter SnatA, MarC family	1	0	1	1	1	2	1	1
COG0687	Spermidine/putrescine-binding periplasmic protein	2	0	1	1	1	0	10	0
COG5006	Threonine/homoserine efflux transporter RhtA	3	1	2	2	0	5	0	1
COG1280	Threonine/homoserine/homoserine lactone efflux protein	0	0	0	0	5	3	10	0
COG0444	ABC-type dipeptide/oligopeptide/nickel transport system, ATPase component	0	1	1	1	4	1	1	0
COG0601	ABC-type dipeptide/oligopeptide/nickel transport system, permease component	1	1	1	1	6	1	12	0
COG1173	ABC-type dipeptide/oligopeptide/nickel transport system, permease component	1	1	1	1	6	1	11	0
COG0834	ABC-type amino acid transport/signal transduction system, periplasmic component	1	1	0	0	8	5	11	0
	Sum	**22**	**15**	**17**	**18**	**258**	**138**	**167**	**21**

Carbohydrate transporters									
COG1653	ABC-type glycerol-3-phosphate transport system, periplasmic component	0	0	0	0	8	3	17	0
COG0395	ABC-type glycerol-3-phosphate transport system, permease component	0	0	0	0	7	3	17	1
COG3839	ABC-type sugar transport system, ATPase component	0	0	0	0	7	5	21	0
COG1129	ABC-type sugar transport system, ATPase component	0	0	0	0	5	1	10	0
COG1879	ABC-type sugar transport system, periplasmic component, contains N-terminal xre family HTH domain	0	3	0	0	3	1	14	0
COG1175	ABC-type sugar transport system, permease component	0	0	0	0	7	3	18	0
COG4213	ABC-type xylose transport system, periplasmic component	0	0	0	0	0	0	2	0
COG4214	ABC-type xylose transport system, permease component	0	0	0	0	0	0	3	0
COG0471	Di- and tricarboxylate transporter	0	0	0	0	0	1	0	0
COG0738	Fucose permease	0	0	0	0	0	0	1	1
COG0580	Glycerol uptake facilitator and related aquaporins (Major Intrinsic Protein Family)	0	0	0	0	2	0	0	0
COG2211	Na+/melibiose symporter or related transporter	0	0	0	0	0	0	0	1
COG2893	Phosphotransferase system, mannose/fructose-specific component IIA	1	1	1	1	1	1	1	0
COG2814	Predicted arabinose efflux permease, MFS family	4	4	3	3	35	13	11	5
COG1172	Ribose/xylose/arabinose/galactoside ABC-type transport system, permease component	0	0	0	0	5	2	12	0
COG2271	Sugar phosphate permease	1	1	1	1	9	0	0	0
COG1593	TRAP-type C4-dicarboxylate transport system, large permease component	2	1	0	0	7	7	6	0
COG1638	TRAP-type C4-dicarboxylate transport system, periplasmic component	2	1	0	0	8	9	6	0
COG3090	TRAP-type C4-dicarboxylate transport system, small permease component	2	1	0	0	5	7	5	0
COG0697	Permease of the drug/metabolite transporter (DMT) superfamily	0	3	0	0	9	5	4	0
COG1682	ABC-type polysaccharide/polyol phosphate export permease	1	1	1	1	3	1	1	1
COG1134	ABC-type polysaccharide/polyol phosphate transport system, ATPase component	0	0	1	1	2	1	1	1
COG2610	H+/gluconate symporter or related permease	0	0	0	0	0	0	0	1
COG1762	Phosphotransferase system mannitol/fructose-specific IIA domain (Ntr-type)	1	1	1	1	1	1	0	0
	Sum	**14**	**17**	**8**	**8**	**124**	**64**	**150**	**11**