# Supplementary Materials and Methods

## A) Site description and sampling.

The FW106 well from which groundwater was obtained is located in Area 3 of the Oak Ridge Field Research Center (FRC) near the Y-12 National Security Complex in Oak Ridge, TN (http://www.esd.ornl.gov/orifrc/). This site lies in the path of a highly-contaminated groundwater plume originating from the original S-3 Waste Disposal Ponds. Because of the proximity of the well to the S-3 ponds, uranium contamination levels are among the highest reported in the world. Contaminants present in FW106 groundwater are listed in Table S1 (compared to pristine FRC groundwater from well FW301) and include high concentrations of nitric acid (pH ~3.7), radionucleotides (technetium-99 and uranium) and volatile organics.

Direct bacterial counts in the source groundwater were between $10^4$ and $10^5$ cells/ml. To obtain sufficient biomass for sequencing, the FW106 well was extensively purged (several well volumes of water removed), ~1700 liters of groundwater was pumped from the matrix surrounding the screened area (~10-4m depth) using peristaltic pumps and was passed through sintered metal (T. J. Phelps, unpublished) (589 L) or 0.2 µm Supor® (Pall Corporation) filters (1126 L) to collect the biomass. Microbial cells were counted using LIVE/DEAD® stain (Invitrogen) and fluorescent microscopy according to the manufacturer's recommendations. Cells were recovered from the filters by shaking and/or brief sonification and were pelleted by ultracentrifugation. The pH of the pellet was adjusted to 7.0 prior to DNA extraction (Zhou et al, 1996). Approximately 300 µg of DNA was obtained from the extracted groundwater. Recovered DNA was treated with RNAse (Zhou et al, 1996) and sent to the DOE Joint Genome Institute (JGI) for constructing SSU rDNA clone and three genome libraries for sequencing. The SSU rRNA gene sequences were initially processed within BioEdit (v. 7.0.5.3) (Hall, 2001) by aligning the sequences using ClustalW, trimming to a shared ~1,300 bp region, and generating a distance matrix using DNADist. DOTUR (Schloss & Handelsman, 2004) was then used to categorize the sequences into operational taxonomic units (OTUs), based on the distance matrix. 13 OTUs were defined at the 98% cutoff with the majority of OTUs mapping to *Rhodanobacter*-like γ-proteobacteria and the remainder to *Azoarcus*-like β-proteobacteria.

### B) Shotgun Sequencing, Assembly and Binning.

Metagenomic sequencing was conducted at JGI using random shotgun sequencing. ~53 Mb of high-quality Q20 read sequences were obtained from ~78 Mb raw sequence (three clone libraries: 20.04 Mb small insert (3 kb) pUC library, 23.13 Mb medium insert (8kb) pMCL, 9.27 Mb large insert (40kb) pCCiFos). The sequencing reads (66220) were assembled into 421 contigs w/ >20 reads (2770 contigs total, ~8.3 Mb assembled DNA) using Phrap as previously described (Tyson et al, 2004), and the contigs were further assembled by paired end analysis into 224 scaffolds ranging in size from 1.8 kb to 2.4 Mb. To account for polymorphisms expected to occur in community DNA, alignment discrepancies beyond those expected for random sequencing errors were allowed if they were consistent with end-pairing constraints. A second assembly of the FW106 metagenome was conducted using Lucy (vector and quality trimming) (Chou & Holmes, 2001) and the Paracel Genome Assembler (pga) (Paracel, Pasadena, CA). Two independent annotations were performed on the Phrap assembly using the JGI-ORNL single genome and JGI-Integrated Microbial Resource (IMG) annotation pipelines, and the pga assembly was annotated using using the IMG pipeline. The pga assembly and associated IMG annotation are available at the IMG/m database (Markowitz et al, 2007) and the FW106 read library has been deposited in GenBank (accession number pending).

Previous analyses and visual inspection of recovered biomass suggest a community dominated by a few closely-related γ-proteobacteria*l* populations. Consistent with this observation, metagenomic bins could not be defined by differences in GC content. Preliminary taxonomic bins were defined for the Phrap assembly using 16S rRNA gene sequences derived from the previously described OTU analysis and from 16S rRNA gene fragments identified directly from the metagenome. Contigs associated with these genes through scaffold assembly were in turn added to the appropriate bins. Once the preliminary bins were established, additional conserved anchor genes (23S rRNA, *recA*, *rpoB*, *gyrB*, *fusA* and *ileS*) were identified. The contigs containing these genes and contigs associated through scaffolding assigned to the appropriate bins. Finally, the remaining contigs were assigned by comparing the contig sequences to the GenBank nr database via BLAST where possible. The pga assembly was binned using PhyloPythia (McHardy et al, 2007). While the PhyloPythia binning still showed dominance by γ-proteobacteria followed by β-proteobacteria, the proportion of contigs sorting to β-proteobacterial and taxonomically unassigned bins was higher than that observed for the manual binning method. In both binning methods, the largest assembled contigs tended to bin to the γ-proteobacteria.

## C) Sequence Analysis and Metabolic Reconstruction.

Metabolic reconstruction of the community was performed automatically as part of the IMG annotation pipeline. Manual examination of the community metabolic network was used to fill in gaps in pathways (when possible) and to assign metabolic activities to specific bins. Abundance profiles for FW106 genes were determined by calculating odds ratios for genes assigned to general COG and KEGG functional categories and to specific geochemically-relevant COGs. Odds ratios were calculated (i.e. for COG categories) as follows:

$$Odds\_Ratio = \frac{A/B}{C/D}$$

Where:

A = #FW106 genes assigned to a specific COG category

B = #FW106 genes assigned to all COG categories

C = # genes from all sequenced bacteria assigned to a specific COG category

D = # genes from all sequenced bacteria assigned to all COG categories

*p*-values were calculated for each odds ratio using one- and two-tailed Fisher's Exact Test to determine significant deviations from equilibrium (odds ratio = 1) (Rosner, 2005) . Values were plotted as ln (odds ratio) to better visualize positive and negative trends in the data. Abundance profiles of FW106 versus other metagenomes using the IMG abundance profiling tools showed similar results, suggesting little bias in the bacterial isolate genome database due to overrepresentation of pathogenic species.

## D) Evolutionary Analyses.

Evolutionary analyses were conducted using *ad hoc* Perl scripts (Deng *et al.*, unpublished results) except where explicitly stated. A nucleotide change was classified as a single nucleotide polymorphism (SNP) in a manner similar to that described in Wu et al., 2006:

(i) sequence quality score >20 in both the contig and read sequences; (ii) >4-fold coverage in the affected assembly column; (iii) there were differences among the other reads at that position; and (iv) the change was flanked by 3 invariable nucleotides on each side(Wu et al, 2006). Genes potentially under positive selection were identified by pairwise comparison of each FW106 CDS to the best BLAST (Altschul et al, 1997) match from GenBank (blastn, default options, local alignment >70% gene sequence) with the Nei-Gojobori method (Nei & Gojobori, 1986) implemented in BioPerl (Stajich et al, 2002) and the maximum likelihood methods implemented in PAML (Yang, 1997). $\omega$ values >1 indicate putative positive selection.

Laterally transferred genes, particularly those which have occurred recently and/or from a phylogenetically distinct donor, often display sequences characteristics from those of the host genome, including G+C content, dinucleotide frequencies, codon usage bias, and etc. Various statistical methods including iterative discriminant and hidden Markov model methods have been developed to identify regions based on these deviations from the genomic background characteristics. Potential problems with this method occur when alien regions display sequence characteristics indistinguishable from the genomic background (e.g. lateral transfer within similar populations) or when sufficient time has passed such that amelioration has occurred, i.e. the alien region has assumed the characteristics of the genomic background. Alternatively, phylogenetic methods involving comparison of gene family trees to well-defined species trees can reveal deviations in the gene family caused by lateral gene transfer. This method can detect more ancient transfers or alien genes that are indistinguishable by statistical methods, but may not be able to distinguish between closely-related orthologs and genes transferred between similar populations. For recent LGT events occurring in the FW106 population over the 50 years since introduction of contamination, amelioration of transferred regions should be minimal and thus is not a factor in this analysis. A more likely scenario in a low-diversity environment such as FW106 groundwater is the transfer of genes between closely related populations. In this scenario, visual examination of the genome may be sufficient to identify potential LGT events between similar populations. In many cases, laterally transferred genes are associated with transposable elements or phage operon genes; though indistinguishable by statistical or phylogenetic means, the association of gene clusters with transposons and phage elements is a strong indication of lateral transfer. A final issue unique to metagenomic sequences is that the possibility of inaccurate assembly of the metagenomic reads can produce false instances of LGT. This problem is expected to be minimized in a low diversity community where assembly of long (kb-Mb range) contigs/scaffolds at high read depths is possible. Comparison of multiple assemblies of the FW106 metagenome as well as multiple PCR experiments using FW106 metagenomic DNA as template with primers designed from the assembled FW106 metagenome suggest the assembly does in fact accurately reflect the true genomic organization of the constituent species (see below).

Multiple complementary methods were employed in this study to detect putative genomic islands (GI) resulting from lateral transfer. First, an iterative discriminant analysis designed to detect deviations from background GC content, dinucleotide frequencies and codon usage was conducted for the major FW106 scaffolds (>100 Mb) (Tu & Ding, 2003). A second method, SIGI-HMM (as implemented in the Colombo program) employs a hidden Markov model (HMM)-based method for identifying regions of aberrant codon usage while minimizing false positives (Merkl, 2004; Waack et al, 2006). Third, selected GI's detected by these methods were verified by visual inspection of the GC content of these regions in Artemis (Berriman & Rutherford, 2003) (2.5 SD cutoff) and by synteny with phage genes, transposons and other known mobile elements. In addition, protein phylogenies of the putatively selected genes of interest (i.e. geochemical resistance genes) were constructed and compared to 16S rRNA species trees that included FW106 16S rRNA genes. 16S rRNA gene-based phylogenetic trees were constructed by adding FW106 sequences to the GreenGenes (DeSantis et al, 2006) 16S dataset using the parsimony insertion function of ARB (DeSantis et al, 2006). Deviations from the 16S rRNA gene tree suggest possible LGT events and putative donor taxa. Emphasis in the text is given for LGT events (e.g. acetone carboxylase) predicted by multiple analysis methods.

## E) Population Genetics Analysis

Degenerate PCR primers were constructed to probe the diversity of *narK* and *czcD* genes directly from FW106 metagenomic DNA. *Taq* DNA polymerase from Invitrogen was used for amplification and the PCR reaction mixture was: 10 µl of 10X PCR buffer minus Mg; 2 µl of 10 mM dNTP mixture; 3 µl of 50 mM $MgCl_2$; 5 µl of primer mix (10 µM each); 100 ng of template DNA from FW106; 0.5 µl of *Taq* DNA polymerase (5U/µl) and Nanopure water to adjust the final volume to 100 µl. PCR amplification was performed with a GeneAmp PCR system 9600 thermal cycler (Applied Biosystems) and subjected to a 5 min denaturation step at 95°C followed by 35 cycles at 94°C for 45 s, 55°C for 30 s and 72°C for 1.5 min. The reaction mixture was then held at 72°C for 15 min. Amplicons were gel purified by a Qiagen kit according to the manufacturer's manual and subsequently cloned into the easy sequencing TA-cloning kit. 96 colonies from each clone set were picked and sequenced by the laboratory of Dr. Bruce Roe (University of Oklahoma) as previously described (Elshahed et al, 2003).

Codon-based multisequence alignments of each sequence set were constructed using the ClustalW (Stajich et al, 2002) algorithm implemented in MEGA 3.1 (Kumar et al, 2004).

Single nucleotide indels resulting in frameshifts were assumed to be the result of sequencing error and were manually corrected within the alignments. Phylogenetic trees for each alignment were constructed in MEGA 3.1 (Kumar et al, 2004) using the neighbor-joining algorithm (Tamura-Nei, heterogeneous lineages, $\gamma$=2, 500 bootstrap replicates, complete gap deletion). Sequences in clusters representing putative populations were extracted, realigned, and analyzed using the population genetics algorithms implemented in DnaSP 4.0 (Rozas & Rozas, 1999). Statistics calculated for each dataset include Tajima's D, Fu and Li's D*/F*/D/F, Fu's $F_S$, Fay and Wu's H and standard population genetics parameters ($\pi$, $\theta_w$, S, $\omega$, ZZ, etc.). The parameters were then analyzed to identify deviations from neutrality consistent with the effects of selection or demographic effects as described in the literature (Charlesworth et al, 1995; Kim, 2006; Nei & Kumar, 2000; Tajima, 1989).

## Supplementary Tables

**Table S1. Geochemistry of FRC sites FW301 (uncontaminated background) and FW106. Contaminant concentrations obtained from http://www.esd.ornl.gov/orifrc/.**

| | FW301 | FW106 |
|---|---|---|
| pH | ~7 | 3.7 |
| $NO_3^-$ (mg/L) | 1.5 | 2331 |
| $SO_4^{2-}$ (mg/L) | 6.3 | 1997 |
| Uranium (mg/L) | >0.0001 | 51 |
| Technetium-99 (pCi/L) | - | 3700 |
| Aluminum (mg/L) | 0.37 | 192 |
| Cadmium (mg/L) | 0 | 0.16 |
| Cobalt (mg/L) | 0 | 0.47 |
| Copper (mg/L) | 0 | 0.63 |
| Nickel (mg/L) | 0.01 | 7.8 |
| Zinc (mg/L) | 0.02 | 0.92 |
| cis-1,2-Dichloroethene (μg/mL) | 5 | 1216 |
| 1,2-Dichloroethene (μg/mL) | 5 | 1153 |
| Tetrachloroethene (μg/mL) | 5 | 810 |
| 1-Butanol (μg/mL) | - | 475 |
| Acetone (μg/mL) | 10 | 823 |

| | | |
|---|---|---|
| Benzoic Acid (μg/mL) | - | 1400 |
| Sodium (mg/L) | 1.96 | 826 |
| Chloride (mg/L) | 1.125 | 465 |
| Magnesium (mg/L) | 2.58 | 45.7 |
| Dissolved Oxygen (mg/L) | - | 0.26 |

**Table S2. Metagenome statistics of FW106**

| | | % Total |
|---|---|---|
| *DNA # bases* | | |
| Total bases | 9554544 | 100.00% |
| DNA coding # bases | 8076611 | 84.53% |
| DNA G+C # bases | 6011119 | 63.20% |
| Scaffolds[a] | 5698 | - |
| *Genes* | | |
| Total # Genes | 12420 | 100.00% |
| RNA Genes | 85 | 0.68% |
| rRNA Genes | 7 | 0.06% |
| 5S rRNA | 1 | 0.01% |
| 16S rRNA | 3 | 0.02% |
| 23S rRNA | 3 | 0.02% |
| tRNA Genes | 78 | 0.63% |
| Protein coding genes | 12335 | 99.32% |
| Genes w/ function prediction | 8689 | 69.96% |
| Genes w/o function prediction | 3646 | 29.36% |
| Genes assigned to enzymes | 1692 | 13.62% |
| Genes connected to KEGG pathways | 1423 | 11.46% |
| Genes in COGs | 7961 | 64.10% |
| Genes in Pfam | 7410 | 59.66% |

**Table S3. Geochemical resistance genes identified in FW106 metagenome. A total of 444 genes (~4.7%) were identified.**

**Table S3.  Geochemical Resistance Genes of FW106**

| gene_oid | Gene Symb | Product Name | Enzymes | COGs | COG Categori | Pfams | TIGRfams | Geochemical Relevance |
|---|---|---|---|---|---|---|---|---|
| 2007099070 | *nhaP* | NhaP-type Na+/H+ and K+/H+ antiporters | | COG002 | P | pfam00999 | | Acid Response |
| 2007108365 | *nhaP* | NhaP-type Na+/H+ and K+/H+ antiporters | | COG002 | P | pfam00999 | | Acid Response |
| 2007098629 | | Malate/lactate dehydrogenases | | COG003 | C | pfam02866 | TIGR01759 | Acid Response |
| 2007101082 | | Malate/lactate dehydrogenases | EC:1.1.1.37 | COG003 | C | pfam00056 pfam02866 | TIGR01759 | Acid Response |
| 2007096987 | | Glutathione synthase/Ribosomal protein S6 modific ation enzyme (glutaminyl transferase) | | COG018 | J H | pfam02955 | | Acid Response |
| 2007099553 | | Glutathione synthase/Ribosomal protein S6 modific ation enzyme (glutaminyl transferase) | EC:6.3.2.3 | COG018 | H J | pfam08443 | TIGR00768 | Acid Response |
| 2007101796 | | Glutathione synthase/Ribosomal protein S6 modific ation enzyme (glutaminyl transferase) | EC:6.3.2.3 | COG018 | H J | pfam02951 pfam08443 | TIGR01380 | Acid Response |
| 2007102504 | | Glutathione synthase/Ribosomal protein S6 modific ation enzyme (glutaminyl transferase) | | COG018 | H J | pfam08443 | | Acid Response |
| 2007096882 | *ack* | Acetate kinase | | COG028 | C | pfam00871 | TIGR00016 | Acid Response |
| 2007098128 | *ack* | Acetate kinase | | COG028 | C | pfam00871 | | Acid Response |
| 2007102049 | *ack* | Acetate kinase | EC:2.7.2.1 | COG028 | C | pfam00871 | TIGR00016 | Acid Response |
| 2007100967 | | Formate hydrogenlyase subunit 4 | | COG065 | C | | | Acid Response |
| 2007102902 | | Formate hydrogenlyase subunit 4 | | COG065 | C | | | Acid Response |
| 2007108359 | | Formate hydrogenlyase subunit 4 | | COG065 | C | | | Acid Response |
| 2007097305 | | Formate hydrogenlyase subunit 3/Multisubunit Na+/ H+ antiporter, MnhD subunit | | COG065 | P C | pfam00361 | | Acid Response |
| 2007098041 | | Formate hydrogenlyase subunit 3/Multisubunit Na+/ H+ antiporter, MnhD subunit | | COG065 | C P | pfam00361 | | Acid Response |
| 2007098527 | | Formate hydrogenlyase subunit 3/Multisubunit Na+/ H+ antiporter, MnhD subunit | | COG065 | C P | pfam00361 | | Acid Response |
| 2007100966 | | Formate hydrogenlyase subunit 3/Multisubunit Na+/ H+ antiporter, MnhD subunit | | COG065 | C P | pfam00361 | | Acid Response |
| 2007100969 | | Formate hydrogenlyase subunit 3/Multisubunit Na+/ H+ antiporter, MnhD subunit | EC:1.6.99.5 | COG065 | C P | pfam00361 | | Acid Response |
| 2007103228 | | Formate hydrogenlyase subunit 3/Multisubunit Na+/ H+ antiporter, MnhD subunit | | COG065 | P C | pfam00361 | | Acid Response |
| 2007107941 | | Formate hydrogenlyase subunit 3/Multisubunit Na+/ H+ antiporter, MnhD subunit | | COG065 | C P | pfam00361 | | Acid Response |
| 2007100792 | | 2 Lactate dehydrogenase and related dehydrogenase s | | COG105 | C R H | | | Acid Response |
| 2007104300 | | 2 Lactate dehydrogenase and related dehydrogenase s | EC:1.1.1.26 | COG105 | C H R | pfam02826 | | Acid Response |
| 2007105579 | | 2 Lactate dehydrogenase and related dehydrogenase s | | COG105 | H C R | pfam02826 | | Acid Response |
| 2007107669 | | 2 Lactate dehydrogenase and related dehydrogenase s | | COG105 | C H R | pfam02826 | | Acid Response |
| 2007108064 | | 2 Lactate dehydrogenase and related dehydrogenase s | | COG105 | H C R | pfam00389 pfam02826 | | Acid Response |
| 2007108460 | | 2 Lactate dehydrogenase and related dehydrogenase s | | COG105 | R H C | pfam00389 | | Acid Response |
| 2007100375 | *nhaD* | Na+/H+ antiporter NhaD and related arsenite perme ases | | COG105 | P | | | Acid Response |
| 2007101079 | *nhaD* | Na+/H+ antiporter NhaD and related arsenite perme ases | | COG105 | P | pfam03600 | | Acid Response |
| 2007102033 | *nhaD* | Na+/H+ antiporter NhaD and related arsenite perme ases | | COG105 | P | pfam02040 | | Acid Response |
| 2007102034 | *nhaD* | Na+/H+ antiporter NhaD and related arsenite perme ases | | COG105 | P | pfam02040 | | Acid Response |
| 2007103167 | *nhaD* | Na+/H+ antiporter NhaD and related arsenite perme ases | | COG105 | P | | | Acid Response |
| 2007105139 | *nhaD* | Na+/H+ antiporter NhaD and related arsenite perme ases | | COG105 | P | | | Acid Response |
| 2007100875 | | Formate hydrogenlyase subunit 6/NADH | EC:1.6.5.3 EC:1.6 | COG114 | C | pfam00037 | TIGR01971 | Acid Response |
| 2007104572 | | Formate hydrogenlyase subunit 6/NADH | EC:1.6.5.3 | COG114 | C | pfam00037 | TIGR01971 | Acid Response |
| 2007104007 | | Na+/H+-dicarboxylate symporters | | COG130 | C | pfam00375 | | Acid Response |
| 2007104782 | | Na+/H+-dicarboxylate symporters | | COG130 | C | pfam00375 | | Acid Response |
| 2007107612 | | Na+/H+-dicarboxylate symporters | | COG130 | C | pfam00375 | | Acid Response |
| 2007101649 | | L-lactate dehydrogenase (FMN-dependent) and relat ed alpha-hydroxy acid dehydrogenases | EC:1.1.2.3 | COG130 | C | pfam01070 | | Acid Response |
| 2007101920 | | L-lactate dehydrogenase (FMN-dependent) and relat ed alpha-hydroxy acid dehydrogenases | EC:1.1.2.3 | COG130 | C | pfam01070 | | Acid Response |
| 2007102075 | | L-lactate dehydrogenase (FMN-dependent) and relat ed alpha-hydroxy acid dehydrogenases | | COG130 | C | pfam01070 | | Acid Response |
| 2007106087 | | L-lactate dehydrogenase (FMN-dependent) and relat ed alpha-hydroxy acid dehydrogenases | | COG130 | C | pfam01070 | | Acid Response |
| 2007096978 | | Uncharacterized protein required for formate dehy drogenase activity | EC:1.2.1.2 | COG152 | C | pfam02634 | TIGR00129 | Acid Response |
| 2007108517 | | Uncharacterized protein required for formate dehy drogenase activity | | COG152 | C | pfam02634 | TIGR00129 | Acid Response |
| 2007101921 | *lctP* | L-lactate permease | | COG162 | C | pfam02652 | TIGR00795 | Acid Response |
| 2007106011 | *lctP* | L-lactate permease | | COG162 | C | pfam02652 | | Acid Response |
| 2007107605 | | Malate/L-lactate dehydrogenases | | COG205 | C | pfam02615 | | Acid Response |
| 2007108864 | | Multisubunit Na+/H+ antiporter, MnhB subunit | | COG211 | P | | | Acid Response |
| 2007108773 | | Putative NAD-dependent formate dehydrogenase delt a subunit (EC 1.2.1.2) | EC:1.2.1.2 | | | | | Acid Response |
| 2007099280 | *eriC* | Chloride channel protein EriC | | COG003 | P | pfam00654 pfam00571 | | Metal Resistance |
| 2007099616 | *eriC* | Chloride channel protein EriC | | COG003 | P | pfam00654 pfam00571 | | Metal Resistance |
| 2007101620 | *eriC* | Chloride channel protein EriC | EC:3.6.1.- | COG003 | P | pfam00654 | | Metal Resistance |
| 2007105682 | *eriC* | Chloride channel protein EriC | | COG003 | P | pfam00654 | | Metal Resistance |
| 2007106402 | *eriC* | Chloride channel protein EriC | | COG003 | P | pfam00654 | | Metal Resistance |
| 2007108983 | *eriC* | Chloride channel protein EriC | | COG003 | P | pfam00654 | | Metal Resistance |
| 2007109052 | *eriC* | Chloride channel protein EriC | | COG003 | P | pfam00654 | | Metal Resistance |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 2007101444 | | Predicted Co/Zn/Cd cation transporters | | COG005 | P | pfam01545 | | Metal Resistance |
| 2007096699 | *mgtA*-like | Cation transport ATPase | | COG047 | P | pfam00689 | TIGR01494 | Metal Resistance |
| 2007096701 | *mgtA*-like | Cation transport ATPase | | COG047 | P | pfam00690 pfam00122 | | Metal Resistance |
| 2007097237 | *mgtA*-like | Cation transport ATPase | | COG047 | P | pfam00122 | TIGR01494 | Metal Resistance |
| 2007097500 | *mgtA*-like | Cation transport ATPase | | COG047 | P | | | Metal Resistance |
| 2007097501 | *mgtA*-like | Cation transport ATPase | | COG047 | P | pfam00702 | TIGR01494 | Metal Resistance |
| 2007099179 | *mgtA*-like | Cation transport ATPase | EC:3.6.3.2 | COG047 | P | pfam00690 pfam00122 pfam00702 pfa | TIGR01494 TIGR01524 | Metal Resistance |
| 2007102037 | *mgtA*-like | Cation transport ATPase | | COG047 | P | pfam00690 pfam00122 | TIGR01494 | Metal Resistance |
| 2007102038 | *mgtA*-like | Cation transport ATPase | | COG047 | P | pfam00702 | TIGR01494 | Metal Resistance |
| 2007102834 | *mgtA*-like | Cation transport ATPase | | COG047 | P | pfam00122 | TIGR01494 | Metal Resistance |
| 2007103238 | *mgtA*-like | Cation transport ATPase | | COG047 | P | pfam00690 pfam00122 | TIGR01494 | Metal Resistance |
| 2007103679 | *mgtA*-like | Cation transport ATPase | | COG047 | P | pfam00702 pfam00689 | TIGR01494 | Metal Resistance |
| 2007103887 | *mgtA*-like | Cation transport ATPase | | COG047 | P | | | Metal Resistance |
| 2007105642 | *mgtA*-like | Cation transport ATPase | | COG047 | P | pfam00702 | TIGR01494 | Metal Resistance |
| 2007105644 | *mgtA*-like | Cation transport ATPase | | COG047 | P | | | Metal Resistance |
| 2007105915 | *mgtA*-like | Cation transport ATPase | | COG047 | P | pfam00690 pfam00122 | TIGR01494 | Metal Resistance |
| 2007109078 | *mgtA*-like | Cation transport ATPase | | COG047 | P | pfam00122 | TIGR01494 | Metal Resistance |
| 2007100148 | *corA*-like | Mg2+ and Co2+ transporters | | COG059 | P | pfam01544 | TIGR00383 | Metal Resistance |
| 2007103766 | *corA*-like | Mg2+ and Co2+ transporters | | COG059 | P | pfam01544 | | Metal Resistance |
| 2007107039 | *corA*-like | Mg2+ and Co2+ transporters | | COG059 | P | pfam01544 | | Metal Resistance |
| 2007097336 | FTR1 | High-affinity Fe2+/Pb2+ permease | | COG067 | P | pfam03239 | TIGR00145 | Metal Resistance |
| 2007098170 | FTR1 | High-affinity Fe2+/Pb2+ permease | | COG067 | P | pfam03239 | TIGR00145 | Metal Resistance |
| 2007098560 | FTR1 | High-affinity Fe2+/Pb2+ permease | | COG067 | P | pfam03239 | TIGR00145 | Metal Resistance |
| 2007098764 | FTR1 | High-affinity Fe2+/Pb2+ permease | | COG067 | P | pfam03239 | TIGR00145 | Metal Resistance |
| 2007101558 | FTR1 | High-affinity Fe2+/Pb2+ permease | | COG067 | P | pfam03239 | TIGR00145 | Metal Resistance |
| 2007103639 | FTR1 | High-affinity Fe2+/Pb2+ permease | | COG067 | P | | TIGR00145 | Metal Resistance |
| 2007103874 | FTR1 | High-affinity Fe2+/Pb2+ permease | | COG067 | P | pfam00034 pfam03239 | TIGR00145 | Metal Resistance |
| 2007104865 | FTR1 | High-affinity Fe2+/Pb2+ permease | | COG067 | P | pfam03239 | TIGR00145 | Metal Resistance |
| 2007106955 | FTR1 | High-affinity Fe2+/Pb2+ permease | | COG067 | P | | TIGR00145 | Metal Resistance |
| 2007098087 | *fur*-like | Fe2+/Zn2+ uptake regulation proteins | | COG073 | P | pfam01475 | | Metal Resistance |
| 2007100804 | *fur*-like | Fe2+/Zn2+ uptake regulation proteins | | COG073 | P | pfam01475 | | Metal Resistance |
| 2007101153 | *fur*-like | Fe2+/Zn2+ uptake regulation proteins | | COG073 | P | pfam01475 | | Metal Resistance |
| 2007105436 | *fur*-like | Fe2+/Zn2+ uptake regulation proteins | | COG073 | P | | | Metal Resistance |
| 2007097464 | *merD* | HTH-type transcriptional regulator merD (Mercuric resistance protein merD) | | COG078 | K | pfam00376 | TIGR02054 | Metal Resistance |
| 2007097485 | *merD* | HTH-type transcriptional regulator merD (Mercuric resistance protein merD) | | COG078 | K | pfam00376 | TIGR02054 | Metal Resistance |
| 2007098596 | *merD* | HTH-type transcriptional regulator merD (Mercuric resistance protein merD) | | COG078 | K | pfam00376 | TIGR02054 | Metal Resistance |
| 2007098780 | *merD* | HTH-type transcriptional regulator merD (Mercuric resistance protein merD) | | COG078 | K | pfam00376 | TIGR02054 | Metal Resistance |
| 2007107725 | *merD* | HTH-type transcriptional regulator merD (Mercuric resistance protein merD) | | COG078 | K | pfam00376 | TIGR02054 | Metal Resistance |
| 2007099932 | ACR3 | Arsenite efflux pump ACR3 and related permeases | EC:1.20.4.1 | COG079 | P | pfam01758 | TIGR00832 | Metal Resistance |
| 2007104920 | ACR3 | Arsenite efflux pump ACR3 and related permeases | | COG079 | P | pfam01758 | | Metal Resistance |
| 2007107717 | ACR3 | Arsenite efflux pump ACR3 and related permeases | | COG079 | P | pfam01758 | TIGR00832 | Metal Resistance |
| 2007108451 | ACR3 | Arsenite efflux pump ACR3 and related permeases | | COG079 | P | | | Metal Resistance |
| 2007096850 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007096976 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007097975 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007098155 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007098249 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007099722 | *czcA*-like | Cation/multidrug efflux pump | EC:6.3.4.3 | COG084 | V | pfam00873 | TIGR00915 | Metal Resistance |
| 2007099950 | *czcA*-like | Cation/multidrug efflux pump | EC:6.3.4.3 | COG084 | V | pfam00873 | | Metal Resistance |
| 2007102124 | *czcA*-like | Cation/multidrug efflux pump | EC:6.3.4.3 | COG084 | V | pfam00873 | TIGR00915 | Metal Resistance |
| 2007102125 | *czcA*-like | Cation/multidrug efflux pump | EC:6.3.4.3 | COG084 | V | pfam00873 | TIGR00915 | Metal Resistance |
| 2007102469 | *czcA*-like | Cation/multidrug efflux pump | EC:6.3.4.3 | COG084 | V | pfam00873 | | Metal Resistance |
| 2007102491 | *czcA*-like | Cation/multidrug efflux pump | EC:6.3.4.3 | COG084 | V | pfam00873 | TIGR00915 | Metal Resistance |
| 2007103027 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007103093 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007103231 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007103385 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007103499 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007103830 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007104155 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007104699 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007104701 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007104745 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007105040 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007106029 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007106210 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | | | Metal Resistance |
| 2007106632 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007106726 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007106784 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007106801 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007106928 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007107099 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | | | Metal Resistance |
| 2007107210 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007107269 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007107658 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007108189 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007108354 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007108596 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007108770 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007108840 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007108965 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007108966 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007109097 | *czcA*-like | Cation/multidrug efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007100515 | *czcA*-like | Putative silver efflux pump | EC:6.3.4.3 | COG084 | V | pfam00873 | TIGR00914 | Metal Resistance |
| 2007107871 | *czcA*-like | Putative silver efflux pump | | COG084 | V | pfam00873 | | Metal Resistance |
| 2007096783 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | | Metal Resistance |
| 2007097019 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | TIGR01730 | Metal Resistance |
| 2007097021 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | TIGR01730 | Metal Resistance |
| 2007097341 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00364 pfam00529 | TIGR01730 | Metal Resistance |
| 2007098567 | *czcB*-like | Membrane-fusion protein | | COG084 | M | | TIGR01730 | Metal Resistance |
| 2007098577 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | TIGR01730 | Metal Resistance |
| 2007098983 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | | Metal Resistance |
| 2007099400 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00364 pfam00529 | TIGR01730 | Metal Resistance |
| 2007099723 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | TIGR01730 | Metal Resistance |
| 2007099897 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | TIGR01730 | Metal Resistance |
| 2007099949 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | TIGR01730 | Metal Resistance |
| 2007099983 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | TIGR01730 | Metal Resistance |
| 2007100516 | *czcB*-like | Membrane-fusion protein | | COG084 | M | | TIGR01730 | Metal Resistance |
| 2007101413 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | TIGR01730 | Metal Resistance |
| 2007102126 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | TIGR01730 | Metal Resistance |
| 2007102470 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | TIGR01730 | Metal Resistance |
| 2007103986 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | TIGR01730 | Metal Resistance |
| 2007104419 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | TIGR01730 | Metal Resistance |
| 2007105222 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | | Metal Resistance |
| 2007105236 | *czcB*-like | Membrane-fusion protein | | COG084 | M | | TIGR01730 | Metal Resistance |
| 2007105929 | *czcB*-like | Membrane-fusion protein | | COG084 | M | | | Metal Resistance |
| 2007106211 | *czcB*-like | Membrane-fusion protein | | COG084 | M | | | Metal Resistance |
| 2007106603 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | | Metal Resistance |
| 2007107365 | *czcB*-like | Membrane-fusion protein | | COG084 | M | pfam00529 | TIGR01730 | Metal Resistance |
| 2007108188 | *czcB*-like | Membrane-fusion protein | | COG084 | M | | | Metal Resistance |
| 2007098248 | *czcB*-like | Membrane-fusion protein | | COG084 | M | | | Metal Resistance |
| 2007101172 | *terC* | Membrane protein TerC, possibly involved in tellu rium resistance | | COG086 | P | pfam03741 | | Metal Resistance |
| 2007104010 | *terC* | Membrane protein TerC, possibly involved in tellu rium resistance | | COG086 | P | pfam03741 | | Metal Resistance |
| 2007097301 | *czcD*-like | Co/Zn/Cd efflux system component | | COG123 | P | pfam01545 | | Metal Resistance |
| 2007097351 | *czcD*-like | Co/Zn/Cd efflux system component | | COG123 | P | pfam01545 | | Metal Resistance |
| 2007098561 | *czcD*-like | Co/Zn/Cd efflux system component | | COG123 | P | pfam01545 | | Metal Resistance |
| 2007098563 | *czcD*-like | Co/Zn/Cd efflux system component | EC:3.6.1.- | COG123 | P | pfam01545 | TIGR01297 | Metal Resistance |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 2007097340 | *czcC* -like | Heavy metal RND efflux outer membrane protein, Cz cC family | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007098578 | *czcC* -like | Heavy metal RND efflux outer membrane protein, Cz cC family | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007102830 | *czcC* -like | Heavy metal RND efflux outer membrane protein, Cz cC family | | COG153 | U M | | | Metal Resistance |
| 2007100517 | *czcC* -like | Metal ion efflux outer membrane protein family pr otein, putative | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007097094 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | TIGR01845 | Metal Resistance |
| 2007097974 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | TIGR01844 | Metal Resistance |
| 2007098568 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007099295 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | TIGR01844 | Metal Resistance |
| 2007099399 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | TIGR01845 | Metal Resistance |
| 2007099898 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007099948 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007099984 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | TIGR01845 | Metal Resistance |
| 2007100535 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | TIGR01845 | Metal Resistance |
| 2007101414 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007101680 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | TIGR01845 | Metal Resistance |
| 2007101971 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | TIGR01845 | Metal Resistance |
| 2007102485 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | TIGR01845 | Metal Resistance |
| 2007103425 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007103828 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007103844 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007103850 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | TIGR01845 | Metal Resistance |
| 2007103851 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007103869 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007104307 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | TIGR01845 | Metal Resistance |
| 2007104395 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007104440 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007104517 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007104700 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007104743 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007105225 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007105234 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007105303 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007106080 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007106331 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007106867 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007106942 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007107837 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007108486 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007108569 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007096951 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007097286 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007101398 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007102790 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007103122 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007106010 | *czcC* -like | Outer membrane protein | | COG153 | U M | | | Metal Resistance |
| 2007106572 | *czcC* -like | Outer membrane protein | | COG153 | U M | pfam02321 | | Metal Resistance |
| 2007107342 | *czcC* -like | Outer membrane protein | | COG153 | M U | | | Metal Resistance |
| 2007108740 | *czcC* -like | Outer membrane protein | | COG153 | M U | pfam02321 | | Metal Resistance |
| 2007097242 | *chrA* | Chromate transport protein ChrA | | COG205 | P | pfam02417 | TIGR00937 | Metal Resistance |
| 2007097347 | *chrA* | Chromate transport protein ChrA | | COG205 | P | pfam02417 | TIGR00937 | Metal Resistance |
| 2007098752 | *chrA* | Chromate transport protein ChrA | | COG205 | P | pfam02417 | TIGR00937 | Metal Resistance |
| 2007107566 | *chrA* | Chromate transport protein ChrA | | COG205 | P | pfam02417 | | Metal Resistance |
| 2007104580 | *chrA* | Cation transport ATPase | | COG205 | P | | | Metal Resistance |
| 2007097190 | *cadA* -like | Cation transport ATPase | | COG221 | P | pfam00122 pfam00702 | TIGR01494 TIGR01525 | Metal Resistance |
| 2007097726 | *cadA* -like | Cation transport ATPase | | COG221 | P | | | Metal Resistance |
| 2007097893 | *cadA* -like | Cation transport ATPase | | COG221 | P | pfam04945 | | Metal Resistance |
| 2007098579 | *cadA* -like | Cation transport ATPase | | COG221 | P | pfam00122 pfam00702 | TIGR01494 TIGR01511 TI | Metal Resistance |
| 2007098580 | *cadA* -like | Cation transport ATPase | | COG221 | P | pfam04945 | | Metal Resistance |
| 2007098785 | *cadA* -like | Cation transport ATPase | EC:3.6.3.4 | COG221 | P | pfam00403 pfam00122 pfam00702 | TIGR01494 TIGR01511 TI | Metal Resistance |
| 2007099878 | *cadA* -like | Cation transport ATPase | EC:3.6.3.4 | COG221 | P | pfam00403 pfam00122 pfam00702 | TIGR01494 TIGR01511 TI | Metal Resistance |
| 2007099905 | *cadA* -like | Cation transport ATPase | EC:3.6.3.3 | COG221 | P | pfam00122 pfam00702 | TIGR01494 TIGR01512 TI | Metal Resistance |
| 2007099986 | *cadA* -like | Cation transport ATPase | EC:3.6.3.4 | COG221 | P | pfam04945 pfam00122 pfam00702 | TIGR01494 TIGR01511 TI | Metal Resistance |
| 2007101196 | *cadA* -like | Cation transport ATPase | EC:3.6.3.4 | COG221 | P | pfam00122 pfam00702 | TIGR01494 TIGR01525 | Metal Resistance |
| 2007101436 | *cadA* -like | Cation transport ATPase | EC:3.6.3.4 | COG221 | P | pfam00122 pfam00702 | TIGR01494 TIGR01511 TI | Metal Resistance |
| 2007104073 | *cadA* -like | Cation transport ATPase | EC:3.6.3.4 | COG221 | P | pfam00122 pfam00702 | TIGR01494 TIGR01511 TI | Metal Resistance |
| 2007104864 | *cadA* -like | Cation transport ATPase | | COG221 | P | pfam00702 | TIGR01494 | Metal Resistance |
| 2007105276 | *cadA* -like | Cation transport ATPase | | COG221 | P | pfam00122 pfam00702 | TIGR01494 | Metal Resistance |
| 2007105310 | *cadA* -like | Cation transport ATPase | | COG221 | P | pfam00702 | TIGR01494 | Metal Resistance |
| 2007106460 | *cadA* -like | Cation transport ATPase | | COG221 | P | pfam00403 pfam00122 | | Metal Resistance |
| 2007107066 | *cadA* -like | Cation transport ATPase | | COG221 | P | pfam00122 | TIGR01494 | Metal Resistance |
| 2007108265 | *cadA* -like | Cation transport ATPase | | COG221 | P | | | Metal Resistance |
| 2007108418 | *cadA* -like | Cation transport ATPase | | COG221 | P | pfam00403 | | Metal Resistance |
| 2007108811 | *cadA* -like | Cation transport ATPase | | COG221 | P | | | Metal Resistance |
| 2007104726 | *cadA* -like | Cation transport ATPase | | COG221 | P | | | Metal Resistance |
| 2007100282 | *mgtE* | Mg/Co/Ni transporter MgtE (contains CBS domain) | EC:3.6.1.- | COG223 | P | pfam03448 pfam00571 pfam01769 | TIGR00400 | Metal Resistance |
| 2007100677 | | Mercuric transport protein periplasmic component precursor (Periplasmic mercury ion binding [...] | | COG260 | P | pfam00403 | TIGR02052 | Metal Resistance |
| 2007097329 | | Putative mercuric transport protein periplasmic b inding protein | | COG260 | P | pfam00403 | TIGR02052 | Metal Resistance |
| 2007099946 | | Putative mercuric transport protein periplasmic b inding protein | | COG260 | P | pfam00403 | TIGR02052 | Metal Resistance |
| 2007107241 | | Membrane carboxypeptidase/penicillin-binding prot ein PbpC | | COG495 | M | pfam00912 | | Metal Resistance |
| 2007104919 | | Arsenate reductase | | | | | | Metal Resistance |
| 2007097827 | *eriC* | Chloride channel protein EriC | | | | pfam00654 | | Metal Resistance |
| 2007100253 | *eriC* | Chloride channel protein EriC | | | | | | Metal Resistance |
| 2007102610 | | Membrane carboxypeptidase/penicillin-binding prot ein PbpC | | | | pfam06832 | | Metal Resistance |
| 2007097328 | | Mercuric transport protein (Mercury ion transport protein) | | | | pfam02411 | | Metal Resistance |
| 2007098570 | | Mercuric transport protein (Mercury ion transport protein) | | | | pfam02411 | | Metal Resistance |
| 2007099947 | | Mercuric transport protein (Mercury ion transport protein) | | | | pfam02411 | | Metal Resistance |
| 2007097469 | *merT* | MerT mercuric transport protein | | | | pfam02411 | | Metal Resistance |
| 2007098601 | *merT* | MerT mercuric transport protein | | | | pfam02411 | | Metal Resistance |
| 2007098775 | *merT* | MerT mercuric transport protein | | | | pfam02411 | | Metal Resistance |
| 2007099968 | *merT* | MerT mercuric transport protein | | | | pfam02411 | | Metal Resistance |
| 2007100678 | *merT* | MerT mercuric transport protein | | | | pfam02411 | | Metal Resistance |
| 2007097345 | | Putative mercury transport protein | | | | pfam03203 | | Metal Resistance |
| 2007097467 | | Putative mercury transport protein | | | | pfam03203 | | Metal Resistance |
| 2007097766 | | Putative mercury transport protein | | | | pfam03203 | | Metal Resistance |
| 2007098599 | | Putative mercury transport protein | | | | pfam03203 | | Metal Resistance |
| 2007098777 | | Putative mercury transport protein | | | | pfam03203 | | Metal Resistance |
| 2007100676 | | Putative mercury transport protein | | | | pfam03203 | | Metal Resistance |
| 2007103368 | | Putative mercury transport protein | | | | pfam03203 | | Metal Resistance |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 2007098564 | *czcD* -like | Co/Zn/Cd efflux system component | | COG123 | P | pfam01545 | | Metal Resistance |
| 2007099891 | *czcD* -like | Co/Zn/Cd efflux system component | | COG123 | P | pfam01545 | | Metal Resistance |
| 2007099893 | *czcD* -like | Co/Zn/Cd efflux system component | EC:3.6.1.- | COG123 | P | pfam01545 | TIGR01297 | Metal Resistance |
| 2007100806 | *czcD* -like | Co/Zn/Cd efflux system component | EC:3.6.1.- | COG123 | P | pfam01545 | TIGR01297 | Metal Resistance |
| 2007103144 | *czcD* -like | Co/Zn/Cd efflux system component | | COG123 | P | pfam01545 | | Metal Resistance |
| 2007103871 | *czcD* -like | Co/Zn/Cd efflux system component | EC:3.6.1.- | COG123 | P | pfam01545 | TIGR01297 | Metal Resistance |
| 2007103873 | *czcD* -like | Co/Zn/Cd efflux system component | | COG123 | P | pfam01545 | | Metal Resistance |
| 2007104735 | *czcD* -like | Co/Zn/Cd efflux system component | | COG123 | P | pfam01545 | | Metal Resistance |
| 2007105970 | *czcD* -like | Co/Zn/Cd efflux system component | | COG123 | P | pfam01545 | TIGR01297 | Metal Resistance |
| 2007107730 | *czcD* -like | Co/Zn/Cd efflux system component | | COG123 | P | pfam01545 | TIGR01297 | Metal Resistance |
| 2007108260 | *czcD* -like | Co/Zn/Cd efflux system component | | COG123 | P | pfam01545 | TIGR01297 | Metal Resistance |
| 2007108530 | *czcD* -like | Co/Zn/Cd efflux system component | | COG123 | P | pfam01545 | | Metal Resistance |
| 2007104724 | *merA* | Mercuric (Ion) reductase (Putative mercuric reduc tase) (EC 1.16.1.1) | | COG124 | C | | | Metal Resistance |
| 2007101992 | | Tellurite resistance protein and related permease s | | COG127 | P | pfam03595 | | Metal Resistance |
| 2007097108 | | Arsenate reductase and related proteins, glutared oxin family | EC:1.20.4.1 | COG139 | P | pfam03960 | TIGR00014 | Metal Resistance |
| 2007101482 | | Arsenate reductase and related proteins, glutared oxin family | | COG139 | P | pfam03960 | TIGR01617 | Metal Resistance |
| 2007105850 | | Arsenate reductase and related proteins, glutared oxin family | | COG139 | P | pfam03960 | TIGR00014 | Metal Resistance |
| 2007097461 | | Glr3618 protein | | COG153 | U M | | | Metal Resistance |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 2007100487 | *amtB* | Ammonia permease | | COG000 | P | pfam00909 | TIGR00836 | Nitrogen |
| 2007102571 | *amtB* | Ammonia permease | | COG000 | P | pfam00909 | TIGR00836 | Nitrogen |
| 2007106084 | *amtB* | Ammonia permease | | COG000 | P | pfam00909 | | Nitrogen |
| 2007106590 | *amtB* | Ammonia permease | | COG000 | P | pfam00909 | | Nitrogen |
| 2007096815 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007097006 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007097257 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | | | Nitrogen |
| 2007097263 | *nosF*-like | ABC-type multidrug transport system, ATPase component | EC:3.6.3.25 | COG113 | V | pfam00005 | | Nitrogen |
| 2007097797 | *nosF*-like | ABC-type multidrug transport system, ATPase component | EC:3.6.3.25 | COG113 | V | pfam00005 | | Nitrogen |
| 2007099105 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007099253 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007100212 | *nosF*-like | ABC-type multidrug transport system, ATPase component | EC:3.6.3.25 | COG113 | V | pfam00005 | | Nitrogen |
| 2007100823 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007101276 | *nosF*-like | ABC-type multidrug transport system, ATPase component | EC:3.6.3.25 | COG113 | V | pfam00005 | | Nitrogen |
| 2007101287 | *nosF*-like | ABC-type multidrug transport system, ATPase component | EC:3.6.3.25 | COG113 | V | pfam00005 | | Nitrogen |
| 2007101456 | *nosF*-like | ABC-type multidrug transport system, ATPase component | EC:3.6.3.30 | COG113 | V | pfam00005 | | Nitrogen |
| 2007101737 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007101968 | *nosF*-like | ABC-type multidrug transport system, ATPase component | EC:3.6.3.25 | COG113 | V | pfam00005 | | Nitrogen |
| 2007102020 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007102234 | *nosF*-like | ABC-type multidrug transport system, ATPase component | EC:3.6.3.- | COG113 | V | pfam00005 | | Nitrogen |
| 2007102840 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | | | Nitrogen |
| 2007103345 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007103755 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007103975 | *nosF*-like | ABC-type multidrug transport system, ATPase component | EC:3.6.3.25 | COG113 | V | pfam00005 | | Nitrogen |
| 2007104514 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | | | Nitrogen |
| 2007105603 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | | | Nitrogen |
| 2007105650 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007106554 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007107067 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007108269 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007108660 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007109135 | *nosF*-like | ABC-type multidrug transport system, ATPase component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007104960 | *nosF*-like | ABC-type multidrug transport system, permease component | | COG113 | V | pfam00005 | | Nitrogen |
| 2007096875 | *narY* | Nitrate reductase beta subunit | | COG114 | C | | | Nitrogen |
| 2007100158 | *narY* | Nitrate reductase beta subunit | EC:1.7.99.4 | COG114 | C | | TIGR01660 | Nitrogen |
| 2007102007 | *narY* | Nitrate reductase beta subunit | EC:1.7.99.4 | COG114 | C | | TIGR01660 | Nitrogen |
| 2007102764 | *narY* | Nitrate reductase beta subunit | EC:1.7.99.4 | COG114 | C | | TIGR01660 | Nitrogen |
| 2007105726 | *narY* | Nitrate reductase beta subunit | | COG114 | C | | | Nitrogen |
| 2007107352 | *narY* | Nitrate reductase beta subunit | | COG114 | C | | | Nitrogen |
| 2007100695 | *nirB* | NAD(P)H-nitrite reductase | EC:1.7.1.4 | COG125 | C | pfam07992 pfam04324 pfam03460 pfa | TIGR02374 | Nitrogen |
| 2007100697 | *nirB* | NAD(P)H-nitrite reductase | EC:1.7.1.4 | COG125 | C | pfam07992 | | Nitrogen |
| 2007102820 | *nirB* | NAD(P)H-nitrite reductase | | COG125 | C | pfam01077 | | Nitrogen |
| 2007103334 | *nirB* | NAD(P)H-nitrite reductase | | COG125 | C | pfam07992 | | Nitrogen |
| 2007104197 | *nirB* | NAD(P)H-nitrite reductase | | COG125 | C | pfam04324 pfam03460 pfam01077 | TIGR02374 | Nitrogen |
| 2007099104 | *nosY*-like | ABC-type transport system involved in multi-copper enzyme maturation, permease component | | COG127 | R | | | Nitrogen |
| 2007101455 | *nosY*-like | ABC-type transport system involved in multi-copper enzyme maturation, permease component | | COG127 | R | pfam03379 | | Nitrogen |
| 2007102019 | *nosY*-like | ABC-type transport system involved in multi-copper enzyme maturation, permease component | | COG127 | R | | | Nitrogen |
| 2007099547 | *nirD* | Ferredoxin subunits of nitrite reductase and ring-hydroxylating dioxygenases | | COG214 | R P | pfam00355 | | Nitrogen |
| 2007100497 | *nirD* | Ferredoxin subunits of nitrite reductase and ring-hydroxylating dioxygenases | | COG214 | R P | pfam00355 | | Nitrogen |
| 2007100696 | *nirD* | Ferredoxin subunits of nitrite reductase and ring-hydroxylating dioxygenases | | COG214 | P R | pfam00355 | TIGR02378 | Nitrogen |
| 2007102819 | *nirD* | Ferredoxin subunits of nitrite reductase and ring-hydroxylating dioxygenases | EC:1.7.1.4 | COG214 | P R | pfam00355 | TIGR02378 | Nitrogen |
| 2007103674 | *nirD* | Ferredoxin subunits of nitrite reductase and ring-hydroxylating dioxygenases | | COG214 | R P | pfam00355 | | Nitrogen |
| 2007105628 | *nirD* | Ferredoxin subunits of nitrite reductase and ring-hydroxylating dioxygenases | | COG214 | P R | pfam00355 | | Nitrogen |
| 2007107554 | *nirD* | Ferredoxin subunits of nitrite reductase and ring-hydroxylating dioxygenases | | COG214 | P R | pfam00355 | | Nitrogen |
| 2007100159 | *narJ* | Nitrate reductase delta subunit | EC:1.7.99.4 | COG218 | C | pfam02613 | TIGR00684 | Nitrogen |
| 2007102006 | *narJ* | Nitrate reductase delta subunit | EC:1.7.99.4 | COG218 | C | pfam02613 | TIGR00684 | Nitrogen |
| 2007106780 | *narJ* | Nitrate reductase delta subunit | | COG218 | C | pfam02613 | TIGR00684 | Nitrogen |
| 2007107351 | *narJ* | Nitrate reductase delta subunit | | COG218 | C | pfam02613 | TIGR00684 | Nitrogen |
| 2007100160 | *narI* | Nitrate reductase gamma subunit | EC:1.7.99.4 | COG218 | C | pfam02665 | TIGR00351 | Nitrogen |
| 2007102005 | *narI* | Nitrate reductase gamma subunit | EC:1.7.99.4 | COG218 | C | pfam02665 | TIGR00351 | Nitrogen |
| 2007105260 | *narI* | Nitrate reductase gamma subunit | | COG218 | C | pfam02665 | | Nitrogen |
| 2007106822 | *narI* | Nitrate reductase gamma subunit | | COG218 | C | pfam02665 | | Nitrogen |
| 2007104849 | *narK* | 77 Permeases of the major facilitator superfamily | | COG222 | P | pfam07690 | | Nitrogen |
| 2007101996 | *narK* | Nitrate/nitrite transporter | | COG222 | P | | | Nitrogen |
| 2007102002 | *narK* | Nitrate/nitrite transporter | | COG222 | P | pfam07690 | | Nitrogen |
| 2007102003 | *narK* | Nitrate/nitrite transporter | | COG222 | P | pfam07690 | | Nitrogen |
| 2007102977 | *narK* | Nitrate/nitrite transporter | | COG222 | P | pfam07690 | | Nitrogen |
| 2007103436 | *narK* | Nitrate/nitrite transporter | | COG222 | P | | | Nitrogen |
| 2007103437 | *narK* | Nitrate/nitrite transporter | | COG222 | P | pfam07690 | | Nitrogen |
| 2007103451 | *narK* | Nitrate/nitrite transporter | | COG222 | P | | | Nitrogen |
| 2007107032 | *narK* | Nitrate/nitrite transporter | | COG222 | P | | | Nitrogen |
| 2007107033 | *narK* | Nitrate/nitrite transporter | | COG222 | P | | | Nitrogen |
| 2007097482 | *norB* | Nitric oxide reductase large subunit | | COG325 | P | | | Nitrogen |
| 2007099623 | *norB* | Nitric oxide reductase large subunit | EC:1.7.99.7 | COG325 | P | | | Nitrogen |
| 2007101981 | *norB* | Nitric oxide reductase large subunit | EC:1.7.99.7 | COG325 | P | pfam00115 | | Nitrogen |
| 2007103255 | *norB* | Nitric oxide reductase large subunit | | COG325 | P | | | Nitrogen |
| 2007103438 | *norB* | Nitric oxide reductase large subunit | | COG325 | P | | | Nitrogen |
| 2007103571 | *norB* | Nitric oxide reductase large subunit | | COG325 | P | | | Nitrogen |
| 2007104770 | *norB* | Nitric oxide reductase large subunit | | COG325 | P | | | Nitrogen |
| 2007104771 | *norB* | Nitric oxide reductase large subunit | | COG325 | P | | | Nitrogen |
| 2007105567 | *norB* | Nitric oxide reductase large subunit | | COG325 | P | | | Nitrogen |
| 2007107416 | *norB* | Nitric oxide reductase large subunit | | COG325 | P | | | Nitrogen |
| 2007107808 | *norB* | Nitric oxide reductase large subunit | | COG325 | P | | | Nitrogen |
| 2007103284 | *norB* | Nitric oxide reductase, subunit B (EC 1.7.99.7) | | COG325 | P | | | Nitrogen |
| 2007097049 | *nosD* | Nitrous oxidase accessory protein | | COG342 | P | | | Nitrogen |
| 2007102021 | *nosD* | Nitrous oxidase accessory protein | | COG342 | P | pfam05048 | | Nitrogen |
| 2007103546 | *nosD* | Nitrous oxidase accessory protein | | COG342 | P | pfam05048 | | Nitrogen |
| 2007105832 | *nosD* | Nitrous oxidase accessory protein | | COG342 | P | | | Nitrogen |
| 2007104141 | *narQ* | Signal transduction histidine kinase, nitrate/nitrite-specific | | COG385 | T | pfam00672 pfam07730 | | Nitrogen |
| 2007102023 | *nosR* | Regulator of nitric oxide reductase transcription | | COG390 | K | pfam04205 | | Nitrogen |
| 2007107567 | *nosR* | Regulator of nitric oxide reductase transcription | | COG390 | K | | | Nitrogen |
| 2007097051 | *nosZ* | Nitrous oxide reductase | | COG426 | C | | | Nitrogen |
| 2007097052 | *nosZ* | Nitrous oxide reductase | | COG426 | C | | TIGR01409 | Nitrogen |
| 2007102022 | *nosZ* | Nitrous oxide reductase | EC:1.7.99.6 | COG426 | C | pfam00116 | TIGR01409 | Nitrogen |
| 2007106622 | *nosZ* | Nitrous oxide reductase | | COG426 | C | | | Nitrogen |
| 2007106889 | *nosZ* | Nitrous oxide reductase | | COG426 | C | | TIGR01409 | Nitrogen |
| 2007102018 | *nosL* | Predicted lipoprotein involved in nitrous oxide reduction | | COG431 | C | pfam05573 | | Nitrogen |
| 2007100157 | *narG* | Nitrate reductase alpha subunit | EC:1.7.99.4 | COG501 | C | pfam00384 pfam01568 | TIGR01580 | Nitrogen |
| 2007102008 | *narG* | Nitrate reductase alpha subunit | EC:1.7.99.4 | COG501 | C | pfam00384 pfam01568 | TIGR01580 | Nitrogen |
| 2007102765 | *narG* | Nitrate reductase alpha subunit | | COG501 | C | pfam01568 | | Nitrogen |
| 2007103450 | *narG* | Nitrate reductase alpha subunit | | COG501 | C | pfam00384 | | Nitrogen |
| 2007105162 | *narG* | Nitrate reductase alpha subunit | | COG501 | C | pfam00384 | | Nitrogen |
| 2007105618 | *narG* | Nitrate reductase alpha subunit | | COG501 | C | pfam00384 | | Nitrogen |
| 2007106221 | *narG* | Nitrate reductase alpha subunit | | COG501 | C | pfam01568 | | Nitrogen |
| 2007107353 | *narG* | Nitrate reductase alpha subunit | | COG501 | C | pfam00384 | | Nitrogen |
| 2007104618 | | ATPase component ABC-type nitrate transport system | | | | | | Nitrogen |
| 2007098985 | | Nitric oxide reductase, cytochrome c-containing subunit (EC 1.7.99.7) | | | | pfam00034 | | Nitrogen |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 2007099753 | | N-methylhydantoinase A/acetone carboxylase, beta subunit | EC:3.5.2.14 | COG014 | E Q | pfam05378 pfam01968 | | Organic Solvents |
| 2007101635 | | N-methylhydantoinase A/acetone carboxylase, beta subunit | EC:3.5.2.9 | COG014 | Q E | pfam05378 pfam01968 | | Organic Solvents |
| 2007101637 | | N-methylhydantoinase A/acetone carboxylase, beta subunit | EC:3.5.2.9 | COG014 | Q E | pfam05378 pfam01968 | | Organic Solvents |
| 2007102328 | | N-methylhydantoinase A/acetone carboxylase, beta subunit | EC:3.5.2.9 | COG014 | Q E | pfam05378 pfam01968 | | Organic Solvents |
| 2007103760 | | N-methylhydantoinase A/acetone carboxylase, beta subunit | | COG014 | E Q | pfam01968 | | Organic Solvents |
| 2007104903 | | N-methylhydantoinase A/acetone carboxylase, beta subunit | | COG014 | Q E | pfam01968 | | Organic Solvents |
| 2007108143 | | N-methylhydantoinase A/acetone carboxylase, beta subunit | | COG014 | E Q | pfam05378 | | Organic Solvents |
| 2007099754 | | N-methylhydantoinase B/acetone carboxylase, alpha subunit | EC:3.5.2.14 | COG014 | Q E | pfam02538 | | Organic Solvents |
| 2007101634 | | N-methylhydantoinase B/acetone carboxylase, alpha subunit | | COG014 | Q E | pfam02538 | | Organic Solvents |
| 2007102327 | | N-methylhydantoinase B/acetone carboxylase, alpha subunit | EC:3.5.2.9 | COG014 | Q E | pfam02538 | | Organic Solvents |
| 2007104306 | | N-methylhydantoinase B/acetone carboxylase, alpha subunit | | COG014 | Q E | pfam02538 | | Organic Solvents |
| 2007105545 | | N-methylhydantoinase B/acetone carboxylase, alpha subunit | | COG014 | Q E | | | Organic Solvents |
| 2007106039 | | N-methylhydantoinase B/acetone carboxylase, alpha subunit | | COG014 | Q E | pfam02538 | | Organic Solvents |
| 2007107248 | | N-methylhydantoinase B/acetone carboxylase, alpha subunit | | COG014 | E Q | pfam02538 | | Organic Solvents |
| 2007108637 | | N-methylhydantoinase B/acetone carboxylase, alpha subunit | | COG014 | E Q | pfam02538 | | Organic Solvents |
| 2007104840 | | Predicted hydrolases or acyltransferases (alpha/b eta hydrolase superfamily) | EC:3.8.1.5 | COG059 | R | pfam00561 | | Organic Solvents |
| 2007105556 | | Predicted hydrolases or acyltransferases (alpha/b eta hydrolase superfamily) | | COG059 | R | pfam00561 | | Organic Solvents |
| 2007096736 | | Transcriptional activator of acetoin/glycerol met abolism | | COG328 | K Q | pfam00158 | | Organic Solvents |
| 2007097573 | | Transcriptional activator of acetoin/glycerol met abolism | | COG328 | K Q | pfam01590 | | Organic Solvents |
| 2007099750 | | Transcriptional activator of acetoin/glycerol met abolism | | COG328 | K Q | pfam00158 | | Organic Solvents |
| 2007101056 | | Transcriptional activator of acetoin/glycerol met abolism | | COG328 | Q K | pfam01590 pfam00158 | | Organic Solvents |
| 2007101653 | | Transcriptional activator of acetoin/glycerol met abolism | | COG328 | K Q | pfam00158 | | Organic Solvents |
| 2007101730 | | Transcriptional activator of acetoin/glycerol met abolism | | COG328 | K Q | pfam01590 pfam00158 | | Organic Solvents |
| 2007103224 | | Transcriptional activator of acetoin/glycerol met abolism | | COG328 | K Q | pfam00158 | | Organic Solvents |
| 2007103719 | | Transcriptional activator of acetoin/glycerol met abolism | | COG328 | K Q | pfam00158 | | Organic Solvents |
| 2007107606 | | Transcriptional activator of acetoin/glycerol met abolism | | COG328 | Q K | pfam01590 | | Organic Solvents |
| 2007106069 | | Transcriptional activator of acetoin/glycerol met abolism | | COG382 | K T | pfam02954 | | Organic Solvents |
| 2007106238 | | Transcriptional activator of acetoin/glycerol met abolism | | COG382 | K T | pfam02954 | | Organic Solvents |
| 2007099755 | | Acetone carboxylase, gamma subunit | | COG464 | Q | pfam08882 | | Organic Solvents |
| 2007101636 | | Acetone carboxylase, gamma subunit | | COG464 | Q | pfam08882 | | Organic Solvents |
| 2007099940 | *adc* | Acetoacetate decarboxylase | | COG468 | Q | pfam06314 | | Organic Solvents |
| 2007099954 | *adc* | Acetoacetate decarboxylase | EC:4.1.1.4 | COG468 | Q | pfam06314 | | Organic Solvents |
| 2007105147 | *adc* | Acetoacetate decarboxylase | EC:4.1.1.4 | COG468 | Q | pfam06314 | | Organic Solvents |
| 2007099852 | *cysK* | Cysteine synthase | EC:2.5.1.47 | COG003 | E | pfam00291 | | Sulfur |
| 2007100366 | *cysK* | Cysteine synthase | EC:2.5.1.47 | COG003 | E | pfam00291 | | Sulfur |
| 2007100819 | *cysK* | Cysteine synthase | EC:2.5.1.47 | COG003 | E | pfam00291 | TIGR01136 TIGR01139 | Sulfur |
| 2007101373 | *cysK* | Cysteine synthase | EC:4.2.1.22 | COG003 | E | pfam00291 pfam00571 | TIGR01137 | Sulfur |
| 2007102899 | *cysK* | Cysteine synthase | EC:2.5.1.47 | COG003 | E | pfam00291 | | Sulfur |
| 2007103902 | *cysK* | Cysteine synthase | | COG003 | E | pfam00291 | | Sulfur |
| 2007105701 | *cysK* | Cysteine synthase | | COG003 | E | pfam00291 | | Sulfur |
| 2007105941 | *cysK* | Cysteine synthase | | COG003 | E | pfam00291 | | Sulfur |
| 2007101021 | *cysI* | Sulfite reductase, beta subunit (hemoprotein) | EC:1.8.1.2 | COG015 | P | pfam03460 pfam01077 | TIGR02041 | Sulfur |
| 2007103494 | *cysI* | Sulfite reductase, beta subunit (hemoprotein) | EC:1.8.1.2 | COG015 | P | pfam01077 pfam03460 | TIGR02041 | Sulfur |
| 2007106274 | *cysI* | Sulfite reductase, beta subunit (hemoprotein) | | COG015 | P | pfam03460 pfam01077 | | Sulfur |
| 2007106478 | *cysI* | Sulfite reductase, beta subunit (hemoprotein) | | COG015 | P | | | Sulfur |
| 2007106962 | *cysI* | Sulfite reductase, beta subunit (hemoprotein) | | COG015 | P | pfam01077 | | Sulfur |
| 2007108605 | *cysI* | Sulfite reductase, beta subunit (hemoprotein) | | COG015 | P | | | Sulfur |
| 2007108654 | *cysI* | Sulfite reductase, beta subunit (hemoprotein) | | COG015 | P | pfam03460 pfam01077 | | Sulfur |
| 2007097155 | *cysH* | 3'-phosphoadenosine 5'-phosphosulfate sulfotransf erase (PAPS reductase)/FAD synthetase and related enzymes | | COG017 | H E | | | Sulfur |
| 2007100393 | *cysH* | 3'-phosphoadenosine 5'-phosphosulfate sulfotransf erase (PAPS reductase)/FAD synthetase and re | EC:2.7.7.4 | COG017 | E H | pfam01507 | TIGR02039 | Sulfur |
| 2007101020 | *cysH* | 3'-phosphoadenosine 5'-phosphosulfate sulfotransf erase (PAPS reductase)/FAD synthetase and re | EC:1.8.4.8 | COG017 | E H | pfam01507 | TIGR00434 TIGR02057 | Sulfur |
| 2007103945 | *cysH* | 3'-phosphoadenosine 5'-phosphosulfate sulfotransf erase (PAPS reductase)/FAD synthetase and related enzymes | | COG017 | E H | pfam01507 | | Sulfur |
| 2007103946 | *cysH* | 3'-phosphoadenosine 5'-phosphosulfate sulfotransf erase (PAPS reductase)/FAD synthetase and re | EC:1.8.4.8 | COG017 | H E | pfam01507 | TIGR00434 TIGR02055 | Sulfur |
| 2007108606 | *cysH* | 3'-phosphoadenosine 5'-phosphosulfate sulfotransf erase (PAPS reductase)/FAD synthetase and related enzymes | | COG017 | E H | pfam01507 | TIGR00434 | Sulfur |
| 2007099630 | *cysJ* | Sulfite reductase, alpha subunit (flavoprotein) | | COG036 | P | pfam00258 | | Sulfur |
| 2007101022 | *cysJ* | Sulfite reductase, alpha subunit (flavoprotein) | EC:1.8.1.2 | COG036 | P | pfam00258 pfam00667 pfam00175 | TIGR01931 | Sulfur |
| 2007102871 | *cysJ* | Sulfite reductase, alpha subunit (flavoprotein) | | COG036 | P | pfam00258 | | Sulfur |
| 2007105878 | *cysJ* | Sulfite reductase, alpha subunit (flavoprotein) | | COG036 | P | | | Sulfur |
| 2007107180 | *cysJ* | Sulfite reductase, alpha subunit (flavoprotein) | | COG036 | P | pfam00258 | | Sulfur |
| 2007109019 | *cysJ* | Sulfite reductase, alpha subunit (flavoprotein) | | COG036 | P | pfam00258 | | Sulfur |
| 2007103209 | *cysC* | Adenylylsulfate kinase and related kinases | EC:2.7.1.25 | COG052 | P | pfam01583 | TIGR00455 | Sulfur |
| 2007106282 | *cysT* | ABC-type sulfate transport system, permease compo nent | | COG055 | O | pfam00528 | | Sulfur |
| 2007108989 | *cysT* | ABC-type sulfate transport system, permease compo nent | | COG055 | O | | | Sulfur |
| 2007103177 | | Sulfate permease and related transporters (MFS su perfamily) | | COG065 | P | pfam00916 | | Sulfur |
| 2007106340 | | Sulfate permease and related transporters (MFS su perfamily) | | COG065 | P | pfam00916 | | Sulfur |
| 2007107283 | | Sulfate permease and related transporters (MFS su perfamily) | | COG065 | P | pfam00916 pfam01740 | TIGR00377 | Sulfur |
| 2007097245 | | ABC-type sulfate/molybdate transport systems, ATP ase component | | COG111 | P | pfam00005 | | Sulfur |
| 2007098116 | | ABC-type sulfate/molybdate transport systems, ATP ase component | | COG111 | P | | | Sulfur |
| 2007107115 | | ABC-type sulfate/molybdate transport systems, ATP ase component | | COG111 | P | pfam00005 | | Sulfur |
| 2007106379 | *cysJ* | Sulfite reductase, alpha subunit (flavoprotein) | | COG114 | C | pfam00037 | | Sulfur |
| 2007101783 | *cysQ* | 3'-Phosphoadenosine 5'-phosphosulfate (PAPS) 3'-p hosphatase | EC:3.1.3.7 | COG121 | P | pfam00459 | TIGR01331 | Sulfur |
| 2007106899 | *cysQ* | 3'-Phosphoadenosine 5'-phosphosulfate (PAPS) 3'-p hosphatase | EC:3.1.3.7 | COG121 | P | pfam00459 | TIGR01331 | Sulfur |
| 2007106283 | | ABC-type sulfate transport system, periplasmic co mponent | | COG161 | P | | | Sulfur |
| 2007105300 | | Sulfite oxidase and related enzymes | EC:1.7.1.1 | COG204 | R | pfam00174 | TIGR01409 | Sulfur |
| 2007106296 | | Sulfite oxidase and related enzymes | | COG204 | R | pfam00174 | | Sulfur |
| 2007108093 | | Sulfite oxidase and related enzymes | | COG204 | R | pfam00174 | | Sulfur |
| 2007097156 | *cysD* | GTPases - Sulfate adenylate transferase subunit 1 | | COG289 | P | pfam00009 | TIGR00231 | Sulfur |
| 2007097727 | *cysD* | GTPases - Sulfate adenylate transferase subunit 1 | | COG289 | P | pfam00009 pfam03144 | TIGR02034 | Sulfur |
| 2007100394 | *cysD* | GTPases - Sulfate adenylate transferase subunit 1 | EC:2.7.1.25 | COG289 | P | pfam00009 pfam03144 pfam01583 | TIGR00455 TIGR02034 | Sulfur |
| 2007105200 | *cysD* | GTPases - Sulfate adenylate transferase subunit 1 | | COG289 | P | pfam00009 | TIGR02034 | Sulfur |
| 2007108542 | | Thiosulfate reductase cytochrome B subunit (membr ane anchoring protein) | | COG411 | C | | | Sulfur |
| 2007097246 | | ABC-type sulfate transport system, permease compo nent | | COG420 | P | pfam00528 | TIGR00969 TIGR02140 | Sulfur |
| 2007106782 | *cysK* | Cysteine synthase | | | | | | Sulfur |

**Table S4. Putative alien (laterally transferred) genes identified by SIGI-HMM.**

**Table S5. Putative alien genes in FW106 as detected by SIGI-HMM.**

| Scaffold | Contig | IMG Gene | goid[a] | COG Category | Function |
|---|---|---|---|---|---|
| 1 | 2747 | 5559 | 2007101043 | - | phage-related conserved hypothetical protein |
| 1 | 2747 | 5560 | 2007101044 | - | Putative gene predicted by FgeneshB |
| 1 | 2747 | 5563 | 2007101045 | - | Putative gene predicted by FgeneshB |
| 1 | 2758 | 6361 | 2007101112 | - | Putative gene predicted by FgeneshB |
| 1 | 2758 | 6362 | 2007101113 | - | Putative gene predicted by FgeneshB |
| 1 | 2758 | 6363 | 2007101114 | - | hypothetical protein |
| 1 | 2758 | 6364 | 2007101115 | C | cytochrome c, class I |
| 1 | 2758 | 6365 | 2007101116 | C | cytochrome c oxidase |
| 1 | 2758 | 6366 | 2007101117 | - | Cytochrome C |
| 1 | 2758 | 6367 | 2007101118 | - | putative cytochrome c1 precursor protein |
| 1 | 2758 | 6377 | 2007101128 | - | Putative gene predicted by FgeneshB |
| 1 | 2758 | 6378 | 2007101129 | R | Secretion chaperone CsaA |
| 1 | 2758 | 6434 | 2007101188 | O | thioredoxin reductase |
| 1 | 2758 | 6462 | 2007101216 | K | helix-turn-helix motif |
| 1 | 2758 | 6463 | 2007101217 | - | Putative gene predicted by FgeneshB |
| 1 | 2758 | 6464 | 2007101218 | L | DNA adenine methylase |
| 1 | 2758 | 6465 | 2007101219 | - | Putative gene predicted by FgeneshB |
| 215 | 2732 | 5253 | 2007101333 | M | Rhamnan synthesis F |
| 215 | 2732 | 5254 | 2007101334 | - | Putative gene predicted by FgeneshB |
| 215 | 2732 | 5255 | 2007101335 | M | glycosyl transferase, family 2 |
| 215 | 2737 | 5325 | 2007101329 | L | hypothetical protein |
| 215 | 2737 | 5326 | 2007101328 | - | hypothetical protein |
| 215 | 2737 | 5327 | 2007101327 | M | glycosyl transferase, family 2 |
| 215 | 2737 | 5328 | 2007101326 | - | Putative gene predicted by FgeneshB |
| 215 | 2737 | 5329 | 2007101325 | G | Xylose isomerase-like TIM barrel |
| 215 | 2737 | 5330 | 2007101324 | G | hypothetical protein |
| 215 | 2737 | 5331 | 2007101323 | R | FAD dependent oxidoreductase |
| 215 | 2737 | 5332 | 2007101322 | M | glycosyl transferase, family 2 |
| 215 | 2737 | 5333 | 2007101321 | - | Putative gene predicted by FgeneshB |
| 215 | 2750 | 5743 | 2007099759 | - | Putative gene predicted by FgeneshB |
| 215 | 2750 | 5744 | 2007099760 | R | conserved hypothetical protein |
| 215 | 2750 | 5745 | 2007099761 | R | exporters of the RND superfamily |
| 215 | 2750 | 5747 | 2007099762 | R | exporters of the RND superfamily |
| 215 | 2750 | 5748 | 2007099763 | - | putative regulatory protein, LysR |
| 215 | 2750 | 5749 | 2007099764 | - | Putative gene predicted by FgeneshB |
| 215 | 2761 | 6903 | 2007101469 | - | Putative gene predicted by FgeneshB |
| 215 | 2761 | 6904 | 2007101468 | M | Cyclopropane-fatty-acyl-phospholipid synthase |
| 215 | 2761 | 6931 | 2007101441 | - | hypothetical protein |
| 215 | 2761 | 6932 | 2007101440 | - | Putative gene predicted by FgeneshB |
| 215 | 2761 | 6949 | 2007101422 | - | Putative gene predicted by FgeneshB |
| 215 | 2761 | 6951 | 2007101420 | - | hypothetical protein |
| 215 | 2761 | 6952 | 2007101419 | C | cytochrome c, class I |
| 215 | 2761 | 6953 | 2007101418 | - | hypothetical protein |
| 215 | 2761 | 6954 | 2007101417 | - | Putative WD-repeat containing protein |
| 215 | 2761 | 6955 | 2007101416 | - | Cytochrome c5 |
| 215 | 2761 | 6956 | 2007101415 | H | protoporphyrinogen oxidase |
| 215 | 2761 | 6961 | 2007101413 | - | membrane fusion protein |
| 215 | 2761 | 6962 | 2007101412 | - | putative silver efflux pump |
| 215 | 2761 | 6963 | 2007101411 | O | cytochrome c biogenesis protein, transmembrane region |
| 215 | 2761 | 6985 | 2007101386 | S | conserved hypothetical protein |
| 215 | 2761 | 6986 | 2007101385 | L | Resolvase-like |
| 215 | 2761 | 6987 | 2007101384 | - | Putative gene predicted by FgeneshB |
| 215 | 2761 | 6988 | 2007101383 | - | hypothetical protein |
| 215 | 2761 | 6989 | 2007101382 | L | ATP-dependent exoDNAse (exonuclease V) alpha subunit - helicase superfamily I member-like |
| 215 | 2761 | 6990 | 2007101381 | - | Putative gene predicted by FgeneshB |
| 215 | 2761 | 6991 | 2007101380 | - | Replication protein A |
| 215 | 2761 | 7030 | 2007101341 | - | hypothetical protein |
| 215 | 2761 | 7031 | 2007101340 | M | glycosyl transferase, group 1 |
| 215 | 2764 | 7446 | 2007102076 | - | Putative gene predicted by FgeneshB |
| 215 | 2764 | 7447 | 2007102077 | Q | DSBA oxidoreductase |
| 215 | 2764 | 7476 | 2007102106 | - | hypothetical protein |
| 215 | 2764 | 7477 | 2007102107 | - | hypothetical protein |
| 215 | 2764 | 7478 | 2007102108 | - | conserved hypothetical protein |
| 215 | 2764 | 7479 | 2007102109 | - | hypothetical protein |
| 215 | 2764 | 7480 | 2007102110 | - | Putative gene predicted by FgeneshB |
| 215 | 2764 | 7481 | 2007102111 | - | Putative gene predicted by FgeneshB |
| 215 | 2764 | 7482 | 2007102112 | - | Similarities with unknown protein |
| 215 | 2764 | 7483 | 2007102113 | - | Putative gene predicted by FgeneshB |
| 215 | 2764 | 7487 | 2007102117 | O | AAA ATPase, central region |
| 215 | 2764 | 7488 | 2007102118 | - | putative modification methylase |
| 215 | 2764 | 7489 | 2007102119 | R | conserved hypothetical protein |

| | | | | | |
|---|---|---|---|---|---|
| 215 | 2764 | 7490 | 2007102120 | - | Bll0881 protein |
| 215 | 2764 | 7491 | 2007102121 | - | hypothetical protein |
| 215 | 2764 | 7539 | 2007102169 | - | Putative gene predicted by FgeneshB |
| 215 | 2764 | 7540 | 2007102170 | - | hypothetical protein |
| 215 | 2764 | 7541 | 2007102171 | - | conserved hypothetical protein |
| 215 | 2764 | 7542 | 2007102172 | - | hypothetical protein |
| 215 | 2764 | 7543 | 2007102173 | - | Putative gene predicted by FgeneshB |
| 215 | 2764 | 7544 | 2007102174 | - | hypothetical protein |
| 215 | 2764 | 7545 | 2007102175 | - | hypothetical protein |
| 215 | 2764 | 7550 | 2007102180 | - | Putative gene predicted by FgeneshB |
| 215 | 2764 | 7551 | 2007102181 | - | hypothetical protein |
| 215 | 2769 | 8642 | 2007101726 | - | Putative gene predicted by FgeneshB |
| 215 | 2769 | 8794 | 2007101884 | - | Putative gene predicted by FgeneshB |
| 215 | 2769 | 8911 | 2007102001 | - | hypothetical protein |
| 219 | 2742 | 5420 | 2007101283 | - | peptidase M50 |
| 219 | 2742 | 5423 | 2007101280 | - | Putative gene predicted by FgeneshB |
| 219 | 2742 | 5424 | 2007101279 | - | hypothetical protein |
| 219 | 2742 | 5425 | 2007101278 | S | protein of unknown function DUF181 |
| 219 | 2742 | 5426 | 2007101277 | - | Nitroreductase |
| 219 | 2742 | 5427 | 2007101276 | V | ABC transporter related |
| 219 | 2742 | 5428 | 2007101275 | - | Putative gene predicted by FgeneshB |
| 219 | 2744 | 5456 | 2007098212 | - | Putative gene predicted by FgeneshB |
| 219 | 2744 | 5457 | 2007098213 | - | hypothetical protein |
| 219 | 2744 | 5458 | 2007098214 | - | Putative gene predicted by FgeneshB |
| 219 | 2744 | 5459 | 2007098215 | - | hypothetical protein |
| 219 | 2744 | 5461 | 2007101000 | - | Putative gene predicted by FgeneshB |
| 219 | 2744 | 5462 | 2007101001 | - | Putative gene predicted by FgeneshB |
| 219 | 2744 | 5463 | 2007101002 | - | hypothetical protein |
| 219 | 2744 | 5464 | 2007101003 | - | hypothetical protein |
| 219 | 2744 | 5465 | 2007101004 | L | phage integrase |
| 219 | 2744 | 5470 | 2007101009 | - | Putative gene predicted by FgeneshB |
| 219 | 2744 | 5471 | 2007101010 | - | Uncharacterized protein conserved in bacteria |
| 219 | 2744 | 5472 | 2007101011 | S | conserved hypothetical protein |
| 219 | 2744 | 5473 | 2007101012 | - | Putative gene predicted by FgeneshB |
| 219 | 2744 | 5495 | 2007101038 | - | hypothetical protein |
| 219 | 2744 | 5496 | 2007101039 | - | Putative gene predicted by FgeneshB |
| 219 | 2744 | 5497 | 2007101040 | - | hypothetical protein |
| 219 | 2744 | 5498 | 2007101041 | - | hypothetical protein |
| 219 | 2744 | 5499 | 2007101042 | - | hypothetical protein |
| 219 | 2744 | 5500 | 2007101043 | - | phage-related conserved hypothetical protein |
| 219 | 2749 | 5640 | 2007101269 | - | Predicted transcriptional regulators |
| 219 | 2749 | 5641 | 2007101268 | K | helix-turn-helix motif |
| 219 | 2749 | 5667 | 2007101242 | - | Putative gene predicted by FgeneshB |
| 219 | 2754 | 5970 | 2007099760 | - | Putative gene predicted by FgeneshB |
| 219 | 2754 | 5971 | 2007099759 | - | Putative gene predicted by FgeneshB |
| 219 | 2754 | 5974 | 2007099756 | - | Acetyl CoA acetyltransferase |
| 219 | 2754 | 5975 | 2007099755 | Q | acetone carboxylase gamma subunit |
| 219 | 2754 | 6031 | 2007104094 | R | modification methylase |
| 219 | 2754 | 6032 | 2007104095 | V | protein of unknown function DUF450 |
| 219 | 2755 | 6081 | 2007102338 | - | Putative gene predicted by FgeneshB |
| 219 | 2755 | 6082 | 2007102337 | E | hydantoin racemase |
| 219 | 2755 | 6083 | 2007102336 | R | hypothetical protein |
| 219 | 2755 | 6084 | 2007102335 | C | L-carnitine dehydratase/bile acid-inducible protein F |
| 219 | 2755 | 6085 | 2007102334 | E | pyruvate carboxyltransferase |
| 219 | 2755 | 6086 | 2007102333 | R | conserved hypothetical protein |
| 219 | 2755 | 6087 | 2007102332 | R | exporters of the RND superfamily |
| 219 | 2755 | 6088 | 2007102331 | - | putative regulatory protein, LysR |
| 219 | 2755 | 6093 | 2007102326 | Q | isochorismatase hydrolase |
| 219 | 2755 | 6094 | 2007102325 | R | Alcohol dehydrogenase, zinc-binding |
| 219 | 2755 | 6332 | 2007099766 | - | Putative gene predicted by FgeneshB |
| 219 | 2757 | 6234 | 2007099859 | R | SMC protein-like |
| 219 | 2757 | 6235 | 2007099858 | - | conserved hypothetical protein |
| 219 | 2757 | 6254 | 2007099844 | S | protein of unknown function DUF125, transmembrane |
| 219 | 2757 | 6333 | 2007099765 | - | hypothetical protein |
| 219 | 2757 | 6334 | 2007099764 | L | putative ISXo8 transposase |
| 219 | 2757 | 6335 | 2007099763 | - | conserved hypothetical protein |
| 219 | 2757 | 6336 | 2007099762 | - | putative regulatory protein, LysR |
| 219 | 2759 | 6472 | 2007101292 | - | Hypothetical protein |
| 219 | 2759 | 6473 | 2007101293 | - | Putative gene predicted by FgeneshB |
| 219 | 2759 | 6474 | 2007101294 | - | Putative gene predicted by FgeneshB |
| 219 | 2759 | 6501 | 2007100927 | - | hypothetical protein |
| 219 | 2759 | 6502 | 2007100926 | R | beta-lactamase-like |
| 219 | 2759 | 6503 | 2007100925 | R | Alcohol dehydrogenase, zinc-binding |
| 219 | 2759 | 6504 | 2007100924 | K | LysR, substrate-binding |

| | | | | | |
|---|---|---|---|---|---|
| 219 | 2759 | 6505 | 2007100923 | - | Putative gene predicted by FgeneshB |
| 219 | 2759 | 6506 | 2007100922 | S | Carboxymuconolactone decarboxylase |
| 219 | 2759 | 6653 | 2007100767 | - | Putative gene predicted by FgeneshB |
| 219 | 2759 | 6654 | 2007100766 | - | Putative gene predicted by FgeneshB |
| 219 | 2759 | 6655 | 2007100765 | - | Putative gene predicted by FgeneshB |
| 219 | 2759 | 6656 | 2007100764 | - | hypothetical protein |
| 219 | 2759 | 6657 | 2007100763 | - | Putative gene predicted by FgeneshB |
| 219 | 2759 | 6771 | 2007100194 | - | Putative gene predicted by FgeneshB |
| 219 | 2762 | 7035 | 2007097429 | - | hypothetical protein |
| 219 | 2762 | 7036 | 2007097428 | - | hypothetical protein |
| 219 | 2762 | 7037 | 2007097427 | D | hypothetical protein |
| 219 | 2762 | 7038 | 2007097426 | - | hypothetical protein |
| 219 | 2762 | 7077 | 2007097387 | - | Putative gene predicted by FgeneshB |
| 219 | 2762 | 7078 | 2007097386 | - | Putative gene predicted by FgeneshB |
| 219 | 2762 | 7079 | 2007097385 | - | Putative gene predicted by FgeneshB |
| 219 | 2762 | 7080 | 2007097384 | - | Putative gene predicted by FgeneshB |
| 219 | 2762 | 7084 | 2007097380 | - | Putative gene predicted by FgeneshB |
| 219 | 2762 | 7085 | 2007097379 | - | hypothetical protein |
| 219 | 2762 | 7086 | 2007097378 | - | hypothetical protein |
| 219 | 2762 | 7121 | 2007099224 | - | Putative gene predicted by FgeneshB |
| 219 | 2763 | 7209 | 2007100998 | - | Putative gene predicted by FgeneshB |
| 219 | 2763 | 7210 | 2007100999 | - | hypothetical protein |
| 219 | 2763 | 7220 | 2007100988 | - | Putative gene predicted by FgeneshB |
| 219 | 2763 | 7246 | 2007100963 | - | Putative gene predicted by FgeneshB |
| 219 | 2763 | 7253 | 2007100956 | - | Putative gene predicted by FgeneshB |
| 219 | 2763 | 7288 | 2007102206 | - | hypothetical protein |
| 219 | 2763 | 7289 | 2007102207 | - | hypothetical protein |
| 219 | 2765 | 7567 | 2007099871 | - | Putative gene predicted by FgeneshB |
| 219 | 2765 | 7596 | 2007099872 | - | Putative gene predicted by FgeneshB |
| 219 | 2765 | 7597 | 2007099873 | - | Putative gene predicted by FgeneshB |
| 219 | 2765 | 7598 | 2007099874 | S | conserved hypothetical protein |
| 219 | 2765 | 7599 | 2007099875 | - | Putative gene predicted by FgeneshB |
| 219 | 2765 | 7600 | 2007099876 | S | hypothetical protein |
| 219 | 2765 | 7604 | 2007099880 | - | Putative gene predicted by FgeneshB |
| 219 | 2765 | 7605 | 2007099881 | - | Transposase and inactivated derivatives |
| 219 | 2765 | 7606 | 2007099882 | - | ISPsy26, transposase orfB |
| 219 | 2765 | 7607 | 2007099883 | - | conserved hypothetical protein |
| 219 | 2765 | 7672 | 2007099976 | S | protein of unknown function DUF6, transmembrane |
| 219 | 2765 | 7673 | 2007099977 | Q | DSBA oxidoreductase |
| 219 | 2765 | 7674 | 2007099978 | S | Alkylhydroperoxidase AhpD core |
| 219 | 2765 | 7675 | 2007099979 | O | OsmC-like protein |
| 219 | 2765 | 7676 | 2007099980 | K | regulatory protein, TetR |
| 219 | 2765 | 7763 | 2007100067 | - | hypothetical protein |
| 219 | 2765 | 7764 | 2007100068 | - | Putative gene predicted by FgeneshB |
| 219 | 2765 | 7765 | 2007100069 | - | Putative gene predicted by FgeneshB |
| 219 | 2766 | 7822 | 2007100365 | - | Putative gene predicted by FgeneshB |
| 219 | 2766 | 7849 | 2007100392 | - | Putative gene predicted by FgeneshB |
| 219 | 2766 | 7854 | 2007100397 | E | Transcriptional regulator of met regulon-like |
| 219 | 2766 | 8053 | 2007097432 | - | Putative gene predicted by FgeneshB |
| 219 | 2769 | 8642 | 2007101726 | - | Putative gene predicted by FgeneshB |
| 219 | 2769 | 8794 | 2007101884 | - | Putative gene predicted by FgeneshB |
| 219 | 2769 | 8911 | 2007102001 | - | hypothetical protein |
| 219 | 2769 | 8926 | 2007099065 | - | Putative gene predicted by FgeneshB |
| 219 | 2770 | 8927 | 2007099066 | L | Resolvase-like |
| 219 | 2770 | 8928 | 2007099067 | - | Putative gene predicted by FgeneshB |
| 219 | 2770 | 8929 | 2007099068 | L | Recombinase |
| 219 | 2770 | 8968 | 2007099107 | K | MT-A70 |
| 219 | 2770 | 8969 | 2007099108 | - | Putative gene predicted by FgeneshB |
| 219 | 2770 | 8970 | 2007099109 | - | hypothetical protein |
| 219 | 2770 | 9264 | 2007099394 | - | Putative serine protease |
| 219 | 2770 | 9300 | 2007099358 | R | Patatin |
| 219 | 2770 | 9301 | 2007099357 | - | hypothetical protein |
| 219 | 2770 | 9302 | 2007099356 | - | Putative gene predicted by FgeneshB |
| 219 | 2770 | 9303 | 2007099355 | K | response regulator receiver |
| 219 | 2770 | 9319 | 2007099339 | - | hypothetical protein |
| 219 | 2770 | 9320 | 2007099338 | - | Putative gene predicted by FgeneshB |
| 219 | 2770 | 9321 | 2007099337 | - | hypothetical protein |
| 219 | 2770 | 9322 | 2007099336 | - | Putative gene predicted by FgeneshB |
| 220 | 2748 | 5621 | 2007102672 | - | Putative gene predicted by FgeneshB |
| 220 | 2748 | 5624 | 2007102675 | G | ketose-bisphosphate aldolases |
| 220 | 2748 | 5625 | 2007102676 | G | Myo-inositol catabolism IolB region |
| 220 | 2748 | 5626 | 2007102677 | G | PfkB |
| 220 | 2748 | 5627 | 2007102678 | E | thiamine pyrophosphate enzyme, central region |
| 220 | 2748 | 5630 | 2007102681 | - | Putative gene predicted by FgeneshB |

|  |  |  |  |  |  |
|---|---|---|---|---|---|
| 220 | 2748 | 5631 | 2007102682 | K | periplasmic binding protein/LacI transcriptional regulator |
| 220 | 2748 | 5633 | 2007102684 | R | oxidoreductase-like |
| 220 | 2748 | 5634 | 2007102685 | G | Xylose isomerase-like TIM barrel |
| 220 | 2748 | 5635 | 2007102686 | R | oxidoreductase-like |
| 220 | 2748 | 5636 | 2007102687 | - | transposase |
| 220 | 2752 | 5810 | 2007102742 | - | Putative gene predicted by FgeneshB |
| 220 | 2752 | 5811 | 2007102741 | - | hypothetical protein |
| 220 | 2752 | 5812 | 2007102740 | - | Putative gene predicted by FgeneshB |
| 220 | 2752 | 5813 | 2007102739 | - | hypothetical protein |
| 220 | 2752 | 5814 | 2007102738 | - | hypothetical protein |
| 220 | 2752 | 5815 | 2007102737 | - | hypothetical protein |
| 220 | 2752 | 5816 | 2007102736 | - | hypothetical protein |
| 220 | 2752 | 5817 | 2007102735 | - | hypothetical protein |
| 220 | 2752 | 5818 | 2007102734 | - | hypothetical protein |
| 220 | 2752 | 5819 | 2007102733 | - | Putative gene predicted by FgeneshB |
| 220 | 2752 | 5820 | 2007102732 | - | Putative gene predicted by FgeneshB |
| 220 | 2752 | 5821 | 2007102731 | - | Putative gene predicted by FgeneshB |
| 220 | 2752 | 5822 | 2007102730 | - | Putative gene predicted by FgeneshB |
| 220 | 2752 | 5823 | 2007102729 | - | Putative gene predicted by FgeneshB |
| 220 | 2752 | 5824 | 2007102728 | - | Putative gene predicted by FgeneshB |
| 220 | 2752 | 5825 | 2007102727 | - | hypothetical protein |
| 220 | 2752 | 5826 | 2007102726 | - | Putative gene predicted by FgeneshB |
| 220 | 2752 | 5827 | 2007102725 | - | hypothetical protein |
| 220 | 2752 | 5857 | 2007102695 | G | PfkB |
| 220 | 2752 | 5858 | 2007102694 | G | RbsD or FucU transport |
| 220 | 2752 | 5861 | 2007102693 | G | periplasmic binding protein/LacI transcriptional regulator |
| 220 | 2752 | 5862 | 2007102692 | F | deoxyribose-phosphate aldolase |
| 220 | 2752 | 5863 | 2007102691 | C | Betaine-aldehyde dehydrogenase |
| 220 | 2752 | 5864 | 2007102690 | C | aldehyde dehydrogenase |
| 220 | 2752 | 5865 | 2007102689 | - | Putative gene predicted by FgeneshB |
| 220 | 2752 | 5866 | 2007102688 | R | Oxidoreductase-like |
| 220 | 2768 | 8404 | 2007102592 | - | Putative gene predicted by FgeneshB |
| 220 | 2768 | 8405 | 2007102591 | - | conserved hypothetical protein |
| 220 | 2768 | 8406 | 2007102590 | - | Putative gene predicted by FgeneshB |
| 220 | 2768 | 8407 | 2007102589 | - | Putative gene predicted by FgeneshB |
| 220 | 2768 | 8408 | 2007102588 | - | hypothetical protein |
| 220 | 2768 | 8409 | 2007102587 | - | conserved hypothetical protein |
| 220 | 2768 | 8410 | 2007102586 | - | conserved hypothetical protein |
| 220 | 2768 | 8411 | 2007102585 | - | Putative gene predicted by FgeneshB |
| 220 | 2768 | 8414 | 2007102582 | V | restriction modification system DNA specificity domain |
| 220 | 2768 | 8445 | 2007102581 | - | Type I restriction modification system methyltransfe |
| 220 | 2768 | 8446 | 2007102580 | V | Restriction endonuclease S subunits-like |
| 220 | 2768 | 8501 | 2007102495 | - | Cold shock proteins |
| 220 | 2768 | 8502 | 2007102494 | L | exodeoxyribonuclease III (xth) |
| 220 | 2768 | 8503 | 2007102493 | R | YgfB and YecA |
| 220 | 2768 | 8551 | 2007102437 | - | Putative gene predicted by FgeneshB |
| 220 | 2768 | 8599 | 2007102387 | - | Putative gene predicted by FgeneshB |

**Table S5. Mobile element distribution in FW106 compared to *Xanthomonas* species. Numbers of transposons for *Xanthomonas* species were obtained from the IMG database based on COG assignment.**

| **Organism** | **Mobile Elements/Mb (COG)**[a] | **%LGT**[b] |
|---|---|---|
| FW106 Metagenome (All Contigs) | 10.4 | - |
| FW106 Metagenome (Major Scaffolds[3]) | 11.7 | 7.1 |
| *X. axonopodis* pv. *citri* str. 306 | 11.0 | 10.1 |
| *X. campestris* pv. *campestris* str. 8004 | 9.1 | 12.2 |
| *X. campestris* pv. *campestris* str. ATCC 33913 | 9.8 | 11.5 |
| *X. campestris* pv. *vesicatoria* 85-10 | 17.2 | 9.9 |
| *X. oryzae* KACC 10331 | 63.1 | 5.6 |
| *X. oryzae* MAFF 311018 | 79.8 | 5.2 |

[a] COG functional groups used for counting: 0675, 1662, 1943, 2801, 2826, 2842, 2963, 3039, 3293, 3316, 3328, 3335, 3385, 3415, 3436, 3464, 3547, 3666, 3676, 3677, 4584, 4644, 5421, 5433, 5558 and 5659

[b] %LGT is the percentages of genes classified as putative alien genes by Colombo (Waack et al, 2006)

[c] Scaffolds > 100 kb

**Table S6. Makeup of genomic islands identified on the major scaffold 224 by discriminant analysis (see also Figure S6).** Shaded entries indicate the putative acetone carboxylase operon.

| Contig | IMG GOID[a] | Length (bp) | Gene Function | PUTAL[b] | Donor Taxon[c] |
|---|---|---|---|---|---|
| *Peak 1* | | | | | |
| 2742 | 2005742489 | 792 | ABC-type multidrug transport system, permease component | N | - |
| 2742 | 2005742490 | 1125 | Two-component system sensor kinase | N | - |
| 2742 | 2005742491 | 603 | Two-component system LuxR-family response regulator | N | - |
| 2742 | 2005742492 | 1011 | Hypothetical peptidase | Y | - |
| 2742 | 2005742493 | 588 | Integrase | N | - |
| 2742 | 2005742494 | 324 | Insertion element IS401 | N | - |
| 2742 | 2005742495 | 168 | Peptidase M50 | Y | Chloroflexi |
| 2742 | 2005742496 | 909 | Putative gene predicted by FgeneshB | Y | - |
| 2742 | 2005742497 | 1422 | Hypothetical protein | Y | Actinobacteria |
| 2742 | 2005742498 | 705 | Protein of unknown function DUF181 | Y | Actinobacteria |
| 2742 | 2005742499 | 963 | Nitroreductase | Y | - |
| 2742 | 2005742500 | 738 | ABC transporter related | Y | - |
| *Peak 2* | | | | | |
| 2743 | 2005742527 | 891 | Signal peptidase I | N | - |
| 2743 | 2005742528 | 384 | Hypothetical protein | N | - |
| 2743 | 2005742529 | 660 | Ribonuclease III | N | - |

| | | | | | |
|---|---|---|---|---|---|
| 2743 | 2005742530 | 933 | GTPase | N | - |
| 2743 | 2005742531 | 729 | DNA repair protein RecO | N | - |
| 2743 | 2005742532 | 726 | OmpR-Family response regulator | N | - |
| *Peak 3* | | | | | |
| 2754 | 2005743057 | 261 | Putative gene predicted by FgeneshB | Y | Bacilli |
| 2754 | 2005743058 | 192 | Putative gene predicted by FgeneshB | Y | Actinobacteria |
| 2754 | 2005743059 | 1029 | NADP-dependent oxidoreductase | N | - |
| 2754 | 2005743060 | 1566 | Acyl-CoA synthetase | N | - |
| 2754 | 2005743061 | 204 | Acetyl-CoA acetyltransferase | Y | Actinobacteria |
| 2754 | 2005743062 | 507 | Acetone carboxylase γ-subunit | Y | Bacilli |
| 2754 | 2005743063 | 2325 | Acetone carboxylase α-subunit | N | - |
| 2754 | 2005743064 | 2181 | Acetone carboxylase β-subunit | N | - |
| 2754 | 2005743065 | 1065 | Hypothetical protein | N | - |
| 2754 | 2005743066 | 1305 | Hypothetical protein | N | - |
| 2754 | 2005743067 | 2022 | Transcriptional activator of acetoin/glycerol metabolism | N | - |
| 2754 | 2005743068 | 1071 | Transposase | N | - |

[a] IMG Gene Object Identifier

[b] Putative alien genes (outlined in bold) as determined by SIGI-HMM

[c] Putative donor taxon as determined by SIGI-HMM

## Supplementary Figures

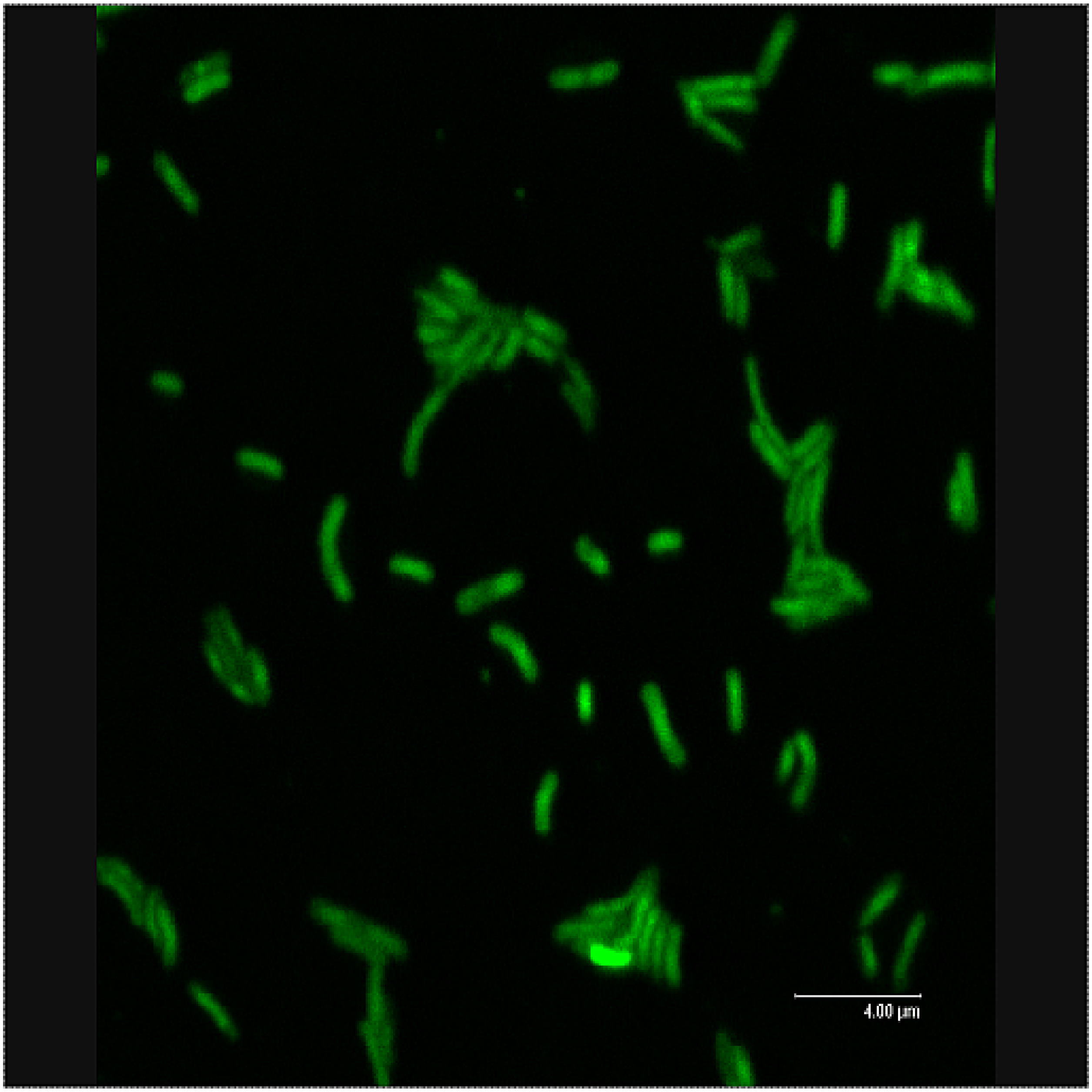


**Figure S1.** Confocal microscopy image of microbial cells obtained from FW106 groundwater.

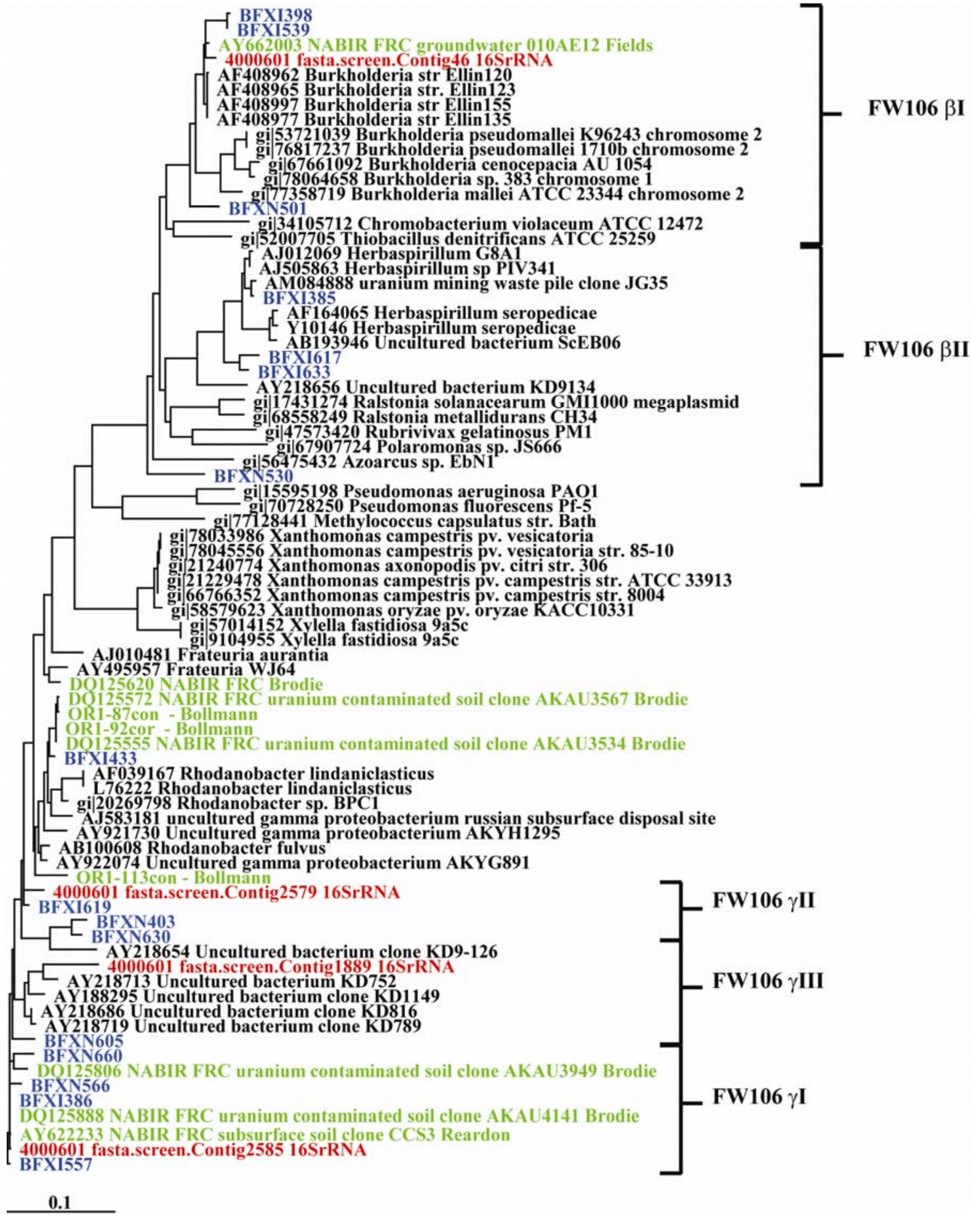
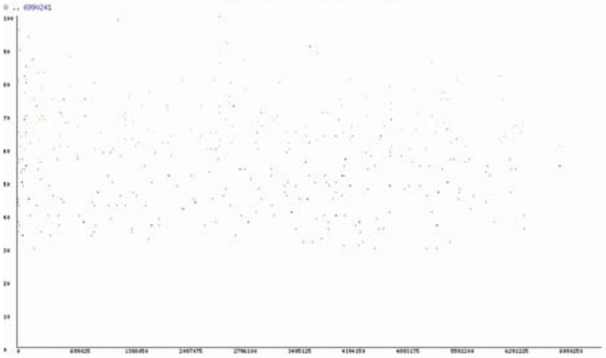
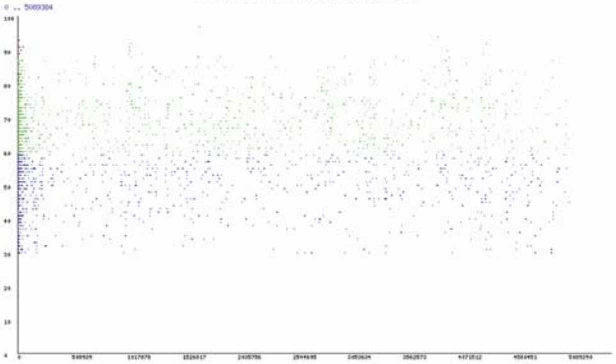

BFXI398
BFXI539
AY662003 NABIR FRC groundwater 010AE12 Fields
4000601 fasta.screen.Contig46 16SrRNA
AF408962 Burkholderia str Ellin120
AF408965 Burkholderia str. Ellin123
AF408997 Burkholderia str Ellin155
AF408977 Burkholderia str Ellin135
gi|53721039 Burkholderia pseudomallei K96243 chromosome 2
gi|76817237 Burkholderia pseudomallei 1710b chromosome 2
gi|67661092 Burkholderia cenocepacia AU 1054
gi|78064658 Burkholderia sp. 383 chromosome 1
gi|77358719 Burkholderia mallei ATCC 23344 chromosome 2
BFXN501
gi|34105712 Chromobacterium violaceum ATCC 12472
gi|52007705 Thiobacillus denitrificans ATCC 25259
AJ012069 Herbaspirillum G8A1
AJ505863 Herbaspirillum sp PIV341
AM084888 uranium mining waste pile clone JG35
BFXI385
AF164065 Herbaspirillum seropedicae
Y10146 Herbaspirillum seropedicae
AB193946 Uncultured bacterium ScEB06
BFXI617
BFXI633
AY218656 Uncultured bacterium KD9134
gi|17431274 Ralstonia solanacearum GMI1000 megaplasmid
gi|68558249 Ralstonia metallidurans CH34
gi|47573420 Rubrivivax gelatinosus PM1
gi|67907724 Polaromonas sp. JS666
gi|56475432 Azoarcus sp. EbN1
BFXN530
gi|15595198 Pseudomonas aeruginosa PAO1
gi|70728250 Pseudomonas fluorescens Pf-5
gi|77128441 Methylococcus capsulatus str. Bath
gi|78033986 Xanthomonas campestris pv. vesicatoria
gi|78045556 Xanthomonas campestris pv. vesicatoria str. 85-10
gi|21240774 Xanthomonas axonopodis pv. citri str. 306
gi|21229478 Xanthomonas campestris pv. campestris str. ATCC 33913
gi|66766352 Xanthomonas campestris pv. campestris str. 8004
gi|58579623 Xanthomonas oryzae pv. oryzae KACC10331
gi|57014152 Xylella fastidiosa 9a5c
gi|9104955 Xylella fastidiosa 9a5c
AJ010481 Frateuria aurantia
AY495957 Frateuria WJ64
DQ125620 NABIR FRC Brodie
DQ125572 NABIR FRC uranium contaminated soil clone AKAU3567 Brodie
OR1-87con - Bollmann
OR1-92cor - Bollmann
DQ125555 NABIR FRC uranium contaminated soil clone AKAU3534 Brodie
BFXI433
AF039167 Rhodanobacter lindaniclasticus
L76222 Rhodanobacter lindaniclasticus
gi|20269798 Rhodanobacter sp. BPC1
AJ583181 uncultured gamma proteobacterium russian subsurface disposal site
AY921730 Uncultured gamma proteobacterium AKYH1295
AB100608 Rhodanobacter fulvus
AY922074 Uncultured gamma proteobacterium AKYG891
OR1-113con - Bollmann
4000601 fasta.screen.Contig2579 16SrRNA
BFXI619
BFXN403
BFXN630
AY218654 Uncultured bacterium clone KD9-126
4000601 fasta.screen.Contig1889 16SrRNA
AY218713 Uncultured bacterium KD752
AY188295 Uncultured bacterium clone KD1149
AY218686 Uncultured bacterium clone KD816
AY218719 Uncultured bacterium clone KD789
BFXN605
BFXN660
DQ125806 NABIR FRC uranium contaminated soil clone AKAU3949 Brodie
BFXN566
BFXI386
DQ125888 NABIR FRC uranium contaminated soil clone AKAU4141 Brodie
AY622233 NABIR FRC subsurface soil clone CCS3 Reardon
4000601 fasta.screen.Contig2585 16SrRNA
BFXI557
FW106 βI
FW106 βII
FW106 γII
FW106 γIII
FW106 γI
0.1

**Figure S2.** 16S rRNA phylogeny of FW106 used to define metagenomic bins. Entries are colored as follows: red, 16S rRNA gene fragments identified from the assembled metagenome; blue, FW106 16S rRNA genes identified from the OTU analysis described in Materials and Methods; green, 16S rRNA genes independently isolated from contaminated FRC sites or cloned from FRC isolates. The initial tree was constructed from the GreenGenes 16S rRNA dataset using the neighbor-joining methods of ARB and metagenomic fragments were added to the tree using the ARB parsimony insertion method (Ludwig et al, 2004). Phylogeny supports previous results suggesting dominance of the community by γ- and β-proteobacterial populations. A single 16S rRNA fragment was identified corresponding to an *Afipia*-like α-proteobacterial species (not shown) but no other sequences could be assigned to this bin. No evidence was found in the sample for sulfate-reducing or iron-reducing δ-proteobacterial species.

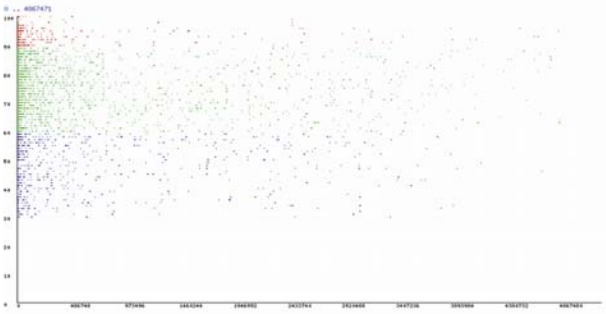
***Pseudomonadaceae***



***Xanthomonadaecae***



***Burkholderiaceae***



***Bradyrhizobiaceae***

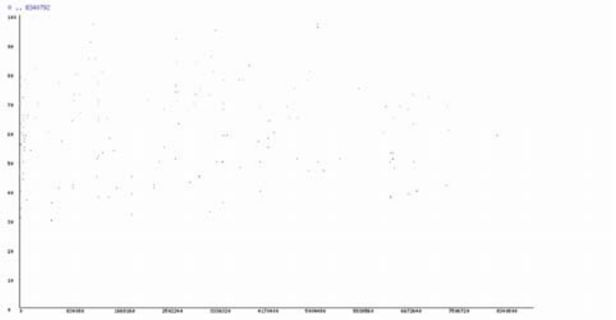


**Figure S3. Protein recruitment plots for FW106.** Colors are as follows: Red, >90% amino acid identity; Green, 60-90% aa identity; Blue, 30-60% aa identity.

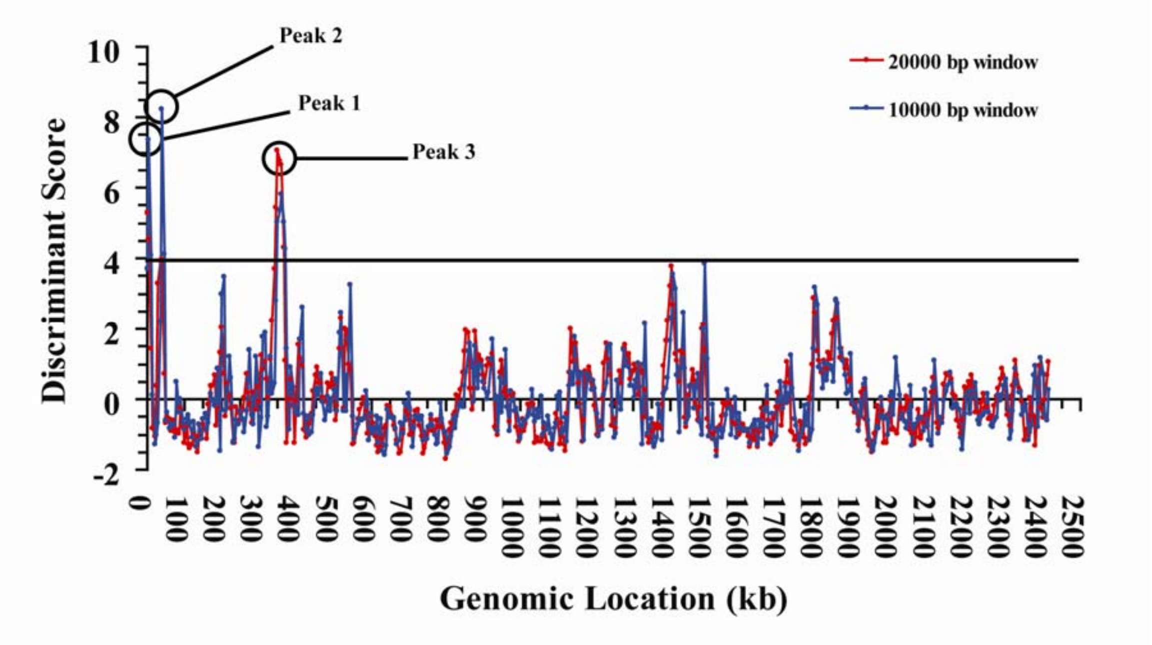


**Figure S4.** Identification of genomic islands (GI) in the major scaffold 219 (FW106 γ1). GI's were determined by iterative discriminant analysis (Tu & Ding, 2003). Analyses were conducted using sliding windows of length 10000 (blue) and 20000 (red) bp advanced in 2500 bp steps. A discriminant scores cutoff of 3.9 was used to determine significance (Tu & Ding, 2003). Peak 3 corresponds to acetone carboxylase Operon A (Fig. 4).

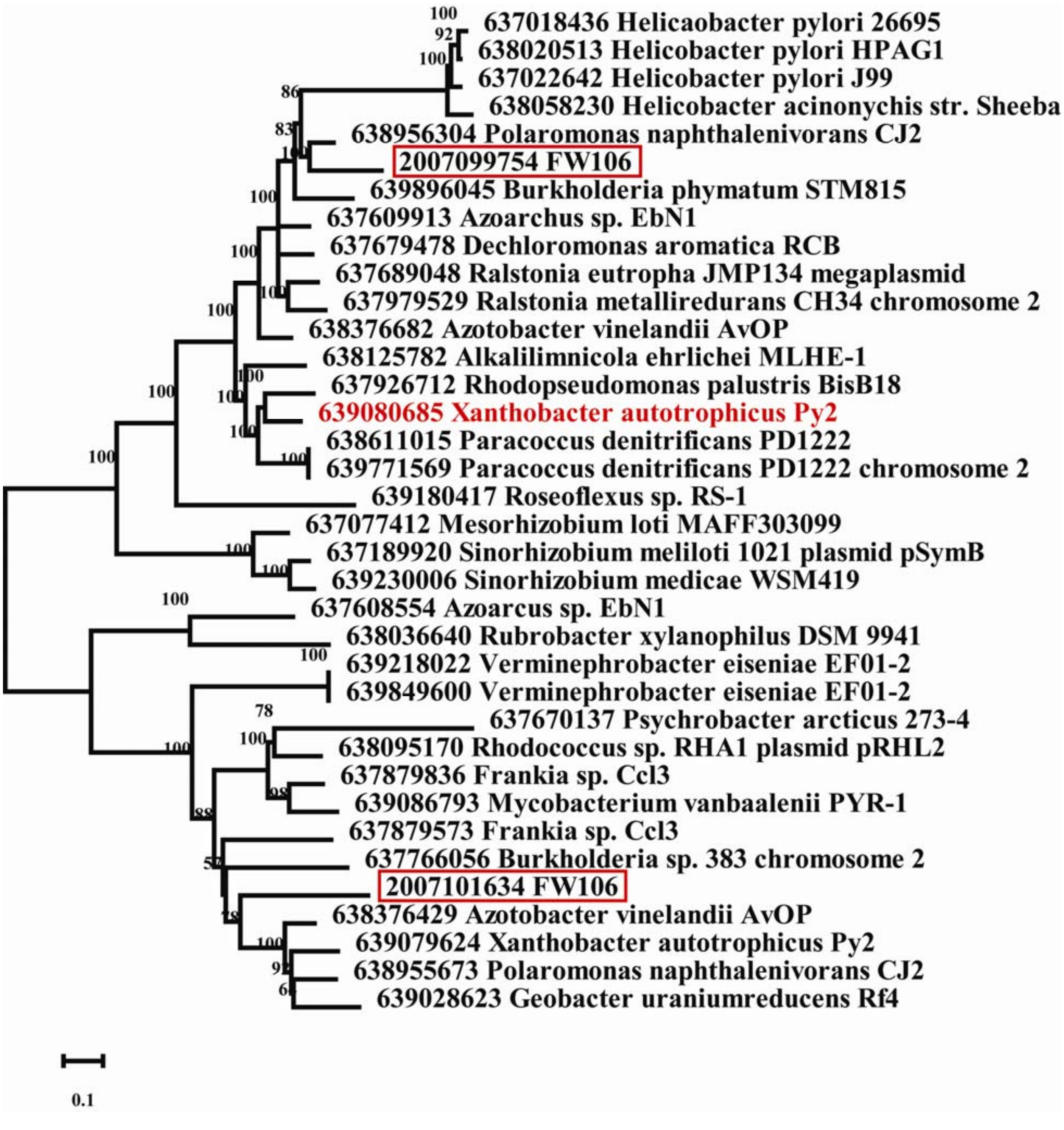

637018436 Helicaobacter pylori 26695
638020513 Helicobacter pylori HPAG1
637022642 Helicobacter pylori J99
638058230 Helicobacter acinonychis str. Sheeba
638956304 Polaromonas naphthalenivorans CJ2
2007099754 FW106
639896045 Burkholderia phymatum STM815
637609913 Azoarchus sp. EbN1
637679478 Dechloromonas aromatica RCB
637689048 Ralstonia eutropha JMP134 megaplasmid
637979529 Ralstonia metalliredurans CH34 chromosome 2
638376682 Azotobacter vinelandii AvOP
638125782 Alkalilimnicola ehrlichei MLHE-1
637926712 Rhodopseudomonas palustris BisB18
639080685 Xanthobacter autotrophicus Py2
638611015 Paracoccus denitrificans PD1222
639771569 Paracoccus denitrificans PD1222 chromosome 2
639180417 Roseoflexus sp. RS-1
637077412 Mesorhizobium loti MAFF303099
637189920 Sinorhizobium meliloti 1021 plasmid pSymB
639230006 Sinorhizobium medicae WSM419
637608554 Azoarcus sp. EbN1
638036640 Rubrobacter xylanophilus DSM 9941
639218022 Verminephrobacter eiseniae EF01-2
639849600 Verminephrobacter eiseniae EF01-2
637670137 Psychrobacter arcticus 273-4
638095170 Rhodococcus sp. RHA1 plasmid pRHL2
637879836 Frankia sp. Ccl3
639086793 Mycobacterium vanbaalenii PYR-1
637879573 Frankia sp. Ccl3
637766056 Burkholderia sp. 383 chromosome 2
2007101634 FW106
638376429 Azotobacter vinelandii AvOP
639079624 Xanthobacter autotrophicus Py2
638955673 Polaromonas naphthalenivorans CJ2
639028623 Geobacter uraniumreducens Rf4
100
92
100
86
83
100
100
100
100
100
100
100
100
100
100
100
100
100
100
100
100
78
100
98
88
52
78
100
92
64
0.1

**Figure S5.** Phylogeny of concatenated nucleotide sequences of the α-, β- and γ- subunits of acetone carboxylase. Entries outlined in red represent FW106 operons. The functionally verified *Xanthobacter autotrophicus* acetone carboxylase in indicated in red text and is orthologous to Operon A (Fig. 4A). Protein alignments were converted to codon-based nucleotide alignments prior to tree construction. The evolutionary history was inferred using the Neighbor-Joining method (Saitou & Nei, 1987). The bootstrap consensus tree inferred from 500 replicates (Felsenstein, 1985) is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) is shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al, 2004) and are in the units of the number of base substitutions per site. The rate variation among sites was modelled with a gamma distribution (shape parameter = 2). The differences in the composition bias among sequences were considered in evolutionary comparisons. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons (Pairwise deletion option). There were a total of 5400 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 (Tamura et al, 2007). Each branch is labelled with the IMG Gene Object Identifier of the α-subunit.

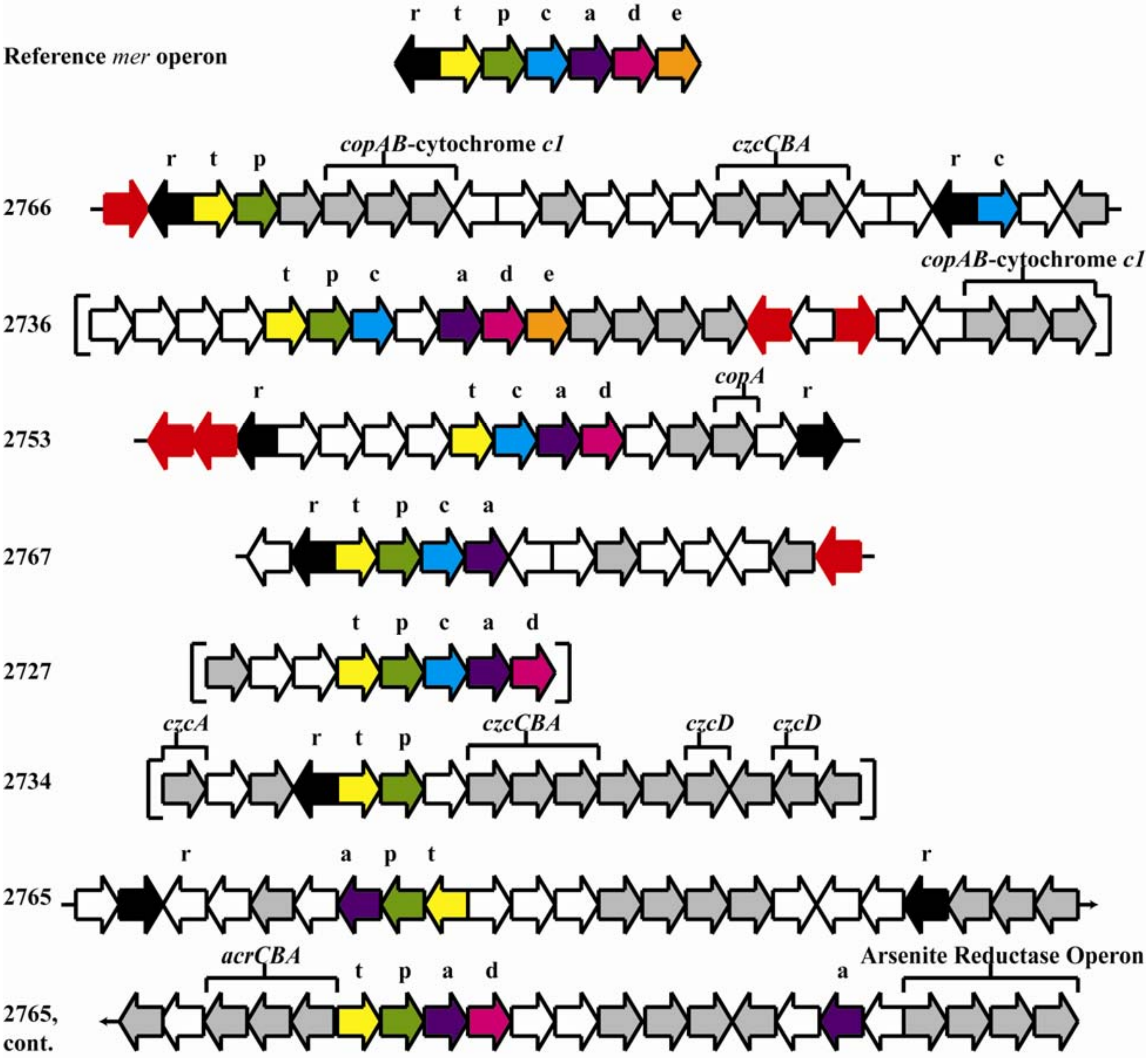

Reference *mer* operon
r t p c a d e
2766
r t p
*copAB*-cytochrome *c1*
*czcCBA*
r c
2736
t p c a d e
*copAB*-cytochrome *c1*
2753
r
t c a d
*copA*
r
2767
r t p c a
2727
t p c a d
2734
*czcA*
r t p
*czcCBA*
*czcD*
*czcD*
2765
r
a p t
r
2765, cont.
*acrCBA*
t p a d
a
Arsenite Reductase Operon

**Figure S6.** Presence and distribution of mercuric resistance genes of FW106. The typical arrangement of the *mer* operon is shown on the first line as a reference, with FW106 contigs aligned below (contig numbers are listed on the left). Contig 2765 is wrapped across multiple lines and contiguous regions of the contig are designated by small black arrows. Brackets at the ends of a contig represent breakpoints (i.e. the beginning or end of a contig) while black lines at the ends of the contigs indicates that the contig has been sequenced beyond that point. Genes are colored as follows: red, mobile elements; white, hypothetical proteins; grey, other functional protein-encoding genes. Orthologous FW106 *mer* genes are colored as in the reference operon.

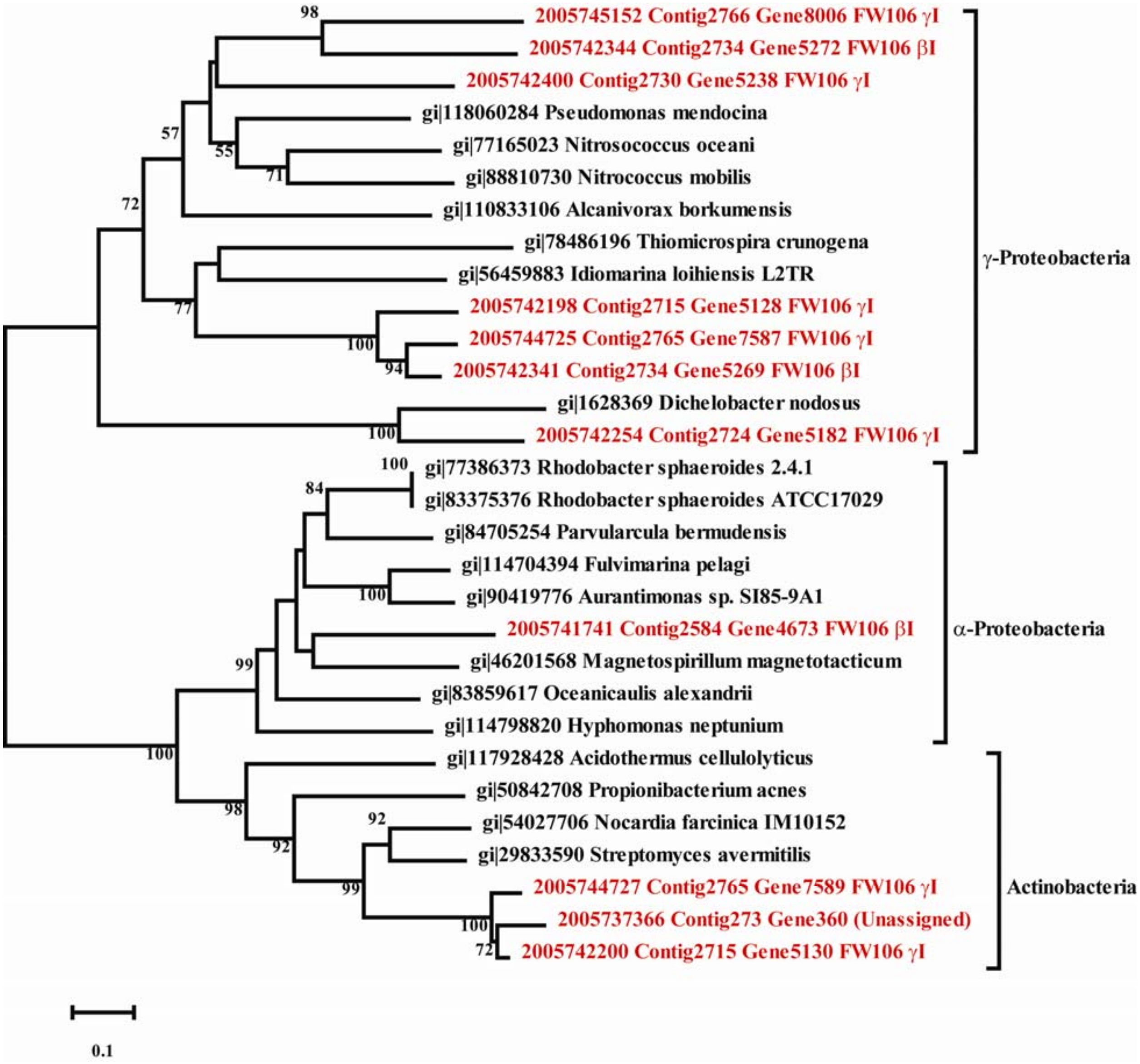

98
2005745152 Contig2766 Gene8006 FW106 γI
2005742344 Contig2734 Gene5272 FW106 βI
2005742400 Contig2730 Gene5238 FW106 γI
gi|118060284 Pseudomonas mendocina
57
gi|77165023 Nitrosococcus oceani
55
71
gi|88810730 Nitrococcus mobilis
72
gi|110833106 Alcanivorax borkumensis
gi|78486196 Thiomicrospira crunogena
gi|56459883 Idiomarina loihiensis L2TR
77
2005742198 Contig2715 Gene5128 FW106 γI
100
2005744725 Contig2765 Gene7587 FW106 γI
94
2005742341 Contig2734 Gene5269 FW106 βI
gi|1628369 Dichelobacter nodosus
100
2005742254 Contig2724 Gene5182 FW106 γI
γ-Proteobacteria
100
gi|77386373 Rhodobacter sphaeroides 2.4.1
84
gi|83375376 Rhodobacter sphaeroides ATCC17029
gi|84705254 Parvularcula bermudensis
gi|114704394 Fulvimarina pelagi
100
gi|90419776 Aurantimonas sp. SI85-9A1
2005741741 Contig2584 Gene4673 FW106 βI
α-Proteobacteria
99
gi|46201568 Magnetospirillum magnetotacticum
gi|83859617 Oceanicaulis alexandrii
gi|114798820 Hyphomonas neptunium
100
gi|117928428 Acidothermus cellulolyticus
gi|50842708 Propionibacterium acnes
98
92
gi|54027706 Nocardia farcinica IM10152
92
gi|29833590 Streptomyces avermitilis
99
2005744727 Contig2765 Gene7589 FW106 γI
Actinobacteria
100
2005737366 Contig273 Gene360 (Unassigned)
72
2005742200 Contig2715 Gene5130 FW106 γI
0.1

**Figure S7.** Phylogeny of CzcD-like metal efflux transporter protein sequences from FW106. Reference sequences are labelled with Genbank gi numbers. Additionally, FW106 sequences are colored red and are labelled with the IMG Gene Object Identifier number. The evolutionary history was inferred using the Neighbor-Joining method (Saitou & Nei, 1987). The bootstrap consensus tree inferred from 500 replicates (Felsenstein, 1985) is taken to represent the evolutionary history of the taxa analyzed (Felsenstein, 1985). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (Felsenstein, 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl & Pauling, 1965) and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total of 144 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 (Tamura et al, 2007).

## Supplementary References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl Acids Res* **25:** 3389-3402

Berriman M, Rutherford K (2003) Viewing and annotating sequence data with Artemis. *Brief Bioinform* **4:** 124-132

Charlesworth D, Charlesworth B, Morgan MT (1995) The Pattern of Neutral Molecular Variation Under the Background Selection Model. *Genetics* **141:** 1619-1632

Chou H-H, Holmes MH (2001) DNA Sequence Quality Trimming and Vector Removal. *Bioinformatics* **17:** 1093-1104

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL (2006) Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl Environ Microbiol* **72:** 5069-5072

Elshahed MS, Senko JM, Najar FZ, Kenton SM, Roe BA, Dewers TA, Spear JR, Krumholz LR (2003) Bacterial Diversity and Sulfur Cycling in a Mesophilic Sulfide-Rich Spring. *Appl Environ Microbiol* **69:** 5609-5621

Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39:** 783-791

Hall T (2001) BioEdit version 5.0.6. *Department of Microbiology, North Carolina State University*

Kim Y (2006) Allele Frequency Distribution Under Recurrent Selective Sweeps. *Genetics* **172:** 1967-1978

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment. *Briefings in Bioinformatics* **5:** 150-163

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, Buchner A, Lai T, Steppi S, Jobb G, Forster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, Konig A, Liss T et al (2004) ARB: a software environment for sequence data. *Nucl Acids Res* **32:** 1363-1371

Markowitz VM, Ivanova NN, Szeto E, Palaniappan K, Chu K, Dalevi D, Chen IMA, Grechkin Y, Dubchak I, Anderson I, Lykidis A, Mavromatis K, Hugenholtz P, Kyrpides NC (2007) IMG/M: a data management and analysis system for metagenomes. *Nucl Acids Res***:** gkm869

McHardy AC, Martin HG, Tsirigos A, Hugenholtz P, Rigoutsos I (2007) Accurate phylogenetic classification of variable-length DNA fragments. *Nat Meth* **4:** 63-72

Merkl R (2004) SIGI: score-based identification of genomic islands. *BMC Bioinformatics* **5:** 22

Nei M, Gojobori T (1986) Simple Methods for Estimating the Numbers of Synonymous and Nonsynonymous Nucleotide Substitutions. *Mol Biol Evol* **3:** 418-426

Nei M, Kumar S (2000) *Molecular Evolution and Phylogenetics*, New York, NY: Oxford University Press.

Rosner B (2005) *Fundamentals of Biostatistics*, 6 edn.: Duxbury Press.

Rozas J, Rozas R (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15:** 174-175

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4:** 406-425

Schloss PD, Handelsman J (2004) Introducing DOTUR, a Computer Program for Defining Operational Taxonomic Units and Estimating Species Richness. *Applied and Environmental Microbiology* **71:** 1501-1506

Stajich JE, Block D, Boulez K, Brenner SE, Chervitz SA, Dagdigian C, Fuellen G, Gilbert JGR, Korf I, Lapp H, Lehvaslaiho H, Matsalla C, Mungall CJ, Osborne BI, Pocock MR, Schattner P, Senger M, Stein LD, Stupka E, Wilkinson MD et al (2002) The Bioperl Toolkit: Perl Modules for the Life Sciences. *Genome Res* **12:** 1611-1618

Tajima F (1989) Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics* **123:** 585-595

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Mol Biol Evol***:** msm092

Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences* **101:** 11030-11035

Tu Q, Ding D (2003) Detecting pathogenicity islands and anomalous gene clusters by iterative discriminant analysis. *FEMS Microbiology Letters* **221:** 269-275

Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM, Solovyev VV, Rubin EM, Rokhsar DS, Banfield JF (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428:** 37-43

Waack S, Keller O, Asper R, Brodag T, Damm C, Fricke W, Surovcik K, Meinicke P, Merkl R (2006) Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. *BMC Bioinformatics* **7:** 142

Wu D, Daugherty SC, Van Aken SE, Pai GH, Watkins KL, Khouri H, Tallon LJ, Zaborsky JM, Dunbar HE, Tran PL (2006) Metabolic Complementarity and Genomics of the Dual Bacterial Symbiosis of Sharpshooters. *PLoS Biol* **4:** e188

Yang Z (1997) PAML: A Program Package for Phylogenetic Analysis by Maximum Likelihood. *Computer Applications in BioSciences* **13:** 555-556

Zhou J, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. *Applied and Environmental Microbiology* **62:** 316-322

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins, pp 97-166.