Table S1. Similarity matrix of *Desulfobulbus* genomes based upon average nucleotide identities.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Dsb1 | Dsb2 | Dsb3 | Dsb4 | Dsb5 |
| Dsb1 | --- |  |  |  |  |
| Dsb2 | 99.11 | --- |  |  |  |
| Dsb3 | 98.00 | 97.79 | --- |  |  |
| Dsb4 | 97.02 | 96.62 | 96.89 | --- |  |
| Dsb5 | 95.86 | 96.03 | 97.60 | 97.72 | --- |

Table S2.Similarity matrix of *Desulfobulbus* genomes based upon tetranucleotide frequencies.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Dsb1 | Dsb2 | Dsb3 | Dsb4 | Dsb5 |
| Dsb1 | --- |  |  |  |  |
| Dsb2 | 0.990 | --- |  |  |  |
| Dsb3 | 0.989 | 0.992 | --- |  |  |
| Dsb4 | 0.983 | 0.989 | 0.987 | --- |  |
| Dsb5 | 0.978 | 0.984 | 0.982 | 0.993 | --- |

Table S3. Dsb1-5 contigs that contain large numbers of horizontally acquired polysaccharide metabolism genes. Information for the top BLASTP hit for each predicted gene is shown.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dsb1-5 Contig \_Gene Number** | **Information for top BLASTP hit** | | | |
|  | **E-value** | **Accession** | **Description** | **Organism** |
| Dsb1-5\_Contig22\_1 | 1.24621E-09 | ZP\_10104881 | glycosyl transferase family 2 | *Thiothrix nivea* DSM 5205 |
| Dsb1-5\_Contig22\_2 | 2.94302E-66 | ZP\_08821249 | glycosyl transferase family 2 | *Thiorhodococcus drewsii* AZ1 |
| Dsb1-5\_Contig22\_3 | 0.00747914 | EGH05187 | 4-oxalocrotonate tautomerase | *Pseudomonas syringae pv. aesculi* str. 0893\_23 |
| Dsb1-5\_Contig22\_4 | 7.16306E-44 | ZP\_07333891 | conserved hypothetical protein | *Desulfovibrio fructosovorans* JJ |
| Dsb1-5\_Contig22\_5 | 3.55497E-08 | ZP\_06298373 | hypothetical protein pah\_c004o242 | *Parachlamydia acanthamoebae* str. Hall's coccus |
| Dsb1-5\_Contig22\_6 | 6.82515E-20 | YP\_001230579 | Ig family protein | *Geobacter uraniireducens* Rf4 |
| Dsb1-5\_Contig22\_7 | 1.04521E-53 | YP\_003809026 | acetyltransferase | *Desulfarculus baarsii* DSM 2075 |
| Dsb1-5\_Contig22\_8 | 9.87963E-24 | YP\_003206524 | membrane protein | *Candidatus Methylomirabilis oxyfera* |
| Dsb1-5\_Contig22\_9 | 1.31344E-71 | YP\_003809023 | hypothetical protein Deba\_3076 | *Desulfarculus baarsii* DSM 2075 |
| Dsb1-5\_Contig22\_10 | 5.189E-140 | YP\_003806580 | family 2 glycosyl transferase | *Desulfarculus baarsii* DSM 2075 |
| Dsb1-5\_Contig22\_11 | 1.70997E-69 | YP\_645837 | glycosyl transferase family protein | *Rubrobacter xylanophilus* DSM 9941 |
|  |  |  |  |  |
| Dsb1-5\_Contig5\_1 | 4.105E-97 | YP\_002354782 | acetyl transferase protein | *Thauera* sp. MZ1T |
| Dsb1-5\_Contig5\_2 | 1.40426E-95 | ZP\_10662483 | glycosyl transferase (involved in LPS synthesis) | *Pseudomonas* sp. GM48 |
| Dsb1-5\_Contig5\_3 | 2.8352E-158 | YP\_728196 | glycosyl transferase | *Ralstonia eutropha* H16 |
| Dsb1-5\_Contig5\_4 | 1.8521E-158 | ZP\_08931531 | UDP-N-acetylglucosamine 2-epimerase | *Thioalkalivibrio thiocyanoxidans* ARh 4 |
| Dsb1-5\_Contig5\_5 | 1.2898E-113 | YP\_728198 | hypothetical protein H16\_B0027 | *Ralstonia eutropha* H16 |
| Dsb1-5\_Contig5\_6 | 1.7469E-139 | ZP\_09507281 | glycosyltransferase | *Marinobacterium stanieri* S30 |
| Dsb1-5\_Contig5\_7 | 2.2309E-168 | YP\_003506150 | asparagine synthase | *Meiothermus ruber* DSM 1279 |
| Dsb1-5\_Contig5\_8 | 4.24594E-44 | ZP\_04414198 | glycosyltransferase | *Vibrio cholerae* bv. *albensis* VL426 |
| Dsb1-5\_Contig5\_9 | 2.15494E-08 | ZP\_08271641 | O-antigen polymerase precursor | gamma proteobacterium IMCC3088 |
| Dsb1-5\_Contig5\_10 | 0 | YP\_983341 | LPS biosynthesis protein WbpG | *Polaromonas naphthalenivorans* CJ2 |
| Dsb1-5\_Contig5\_11 | 1.53953E-82 | YP\_983344 | polysaccharide biosynthesis protein | *Polaromonas naphthalenivorans* CJ2 |
| Dsb1-5\_Contig5\_12 | 0 | YP\_001339692 | DegT/DnrJ/EryC1/StrS aminotransferase | *Marinomonas* sp. MWYL1 |
| Dsb1-5\_Contig5\_13 | 2.5574E-107 | ZP\_09911538 | hexapaptide repeat-containing transferase | SAR324 cluster bacterium JCVI-SC AAA005 |
| Dsb1-5\_Contig5\_14 | 0 | ZP\_08931518 | oxidoreductase domain protein | *Thioalkalivibrio thiocyanoxidans* ARh 4 |
| Dsb1-5\_Contig5\_15 | 0 | YP\_847464 | UDP-glucose/GDP-mannose dehydrogenase | *Syntrophobacter fumaroxidans* MPOB |
| Dsb1-5\_Contig5\_16 | 0 | YP\_004194024 | polysaccharide biosynthesis protein CapD | *Desulfobulbus propionicus* DSM 2032 |
| Dsb1-5\_Contig5\_17 | 6.95693E-14 | CBL39994 | hypothetical protein CK3\_01200 | butyrate-producing bacterium SS3/4 |
| Dsb1-5\_Contig5\_18 | 0.000524226 | ZP\_08576528 | Gp54 protein | *Lactobacillus farciminis* KCTC 3681 |

Table S4. COGs unique to either human-associated or environmental *Desulfovibrio* genomes.

|  |  |
| --- | --- |
| **COGS unique to host-associated *Desulfovibrio*** | |
| COG0122 | 3-methyladenine DNA glycosylase/8-oxoguanine DNA glycosylase |
| COG1464 | ABC-type metal ion transport system, periplasmic component/surface antigen |
| COG2011 | ABC-type metal ion transport system, permease component |
| COG2759 | Formyltetrahydrofolate synthetase |
| COG3616 | Predicted amino acid aldolase or racemase |
| COG0311 | Predicted glutamine amidotransferase involved in pyridoxine biosynthesis |
| COG1058 | Predicted nucleotide-utilizing enzyme related to molybdopterin-biosynthesis enzyme MoeA |
| COG5496 | Predicted thioesterase |
| COG0214 | Pyridoxine biosynthesis enzyme |
| COG2326 | Uncharacterized conserved protein |
| COG4729 | Uncharacterized conserved protein |
| COG2315 | Uncharacterized protein conserved in bacteria |
| COG3735 | Uncharacterized protein conserved in bacteria |
|  |  |
| **COGS unique to environmental *Desulfovibrio*** | |
| COG1042 | Acyl-CoA synthetase (NDP forming) |
| COG1254 | Acylphosphatases |
| COG2177 | Cell division protein |
| COG0648 | Endonuclease IV |
| COG0731 | Fe-S oxidoreductases |
| COG0174 | Glutamine synthetase |
| COG1443 | Isopentenyldiphosphate isomerase |
| COG0346 | Lactoylglutathione lyase and related lyases |
| COG1762 | Phosphotransferase system mannitol/fructose-specific IIA domain (Ntr-type) |
| COG1075 | Predicted acetyltransferases and hydrolases with the alpha/beta hydrolase fold |
| COG2884 | Predicted ATPase involved in cell division |
| COG0702 | Predicted nucleoside-diphosphate-sugar epimerases |
| COG3635 | Predicted phosphoglycerate mutase, AP superfamily |
| COG5001 | Predicted signal transduction protein with membrane domain, an EAL and a GGDEF domain |
| COG1995 | Pyridoxal phosphate biosynthesis protein |
| COG1663 | Tetraacyldisaccharide-1-P 4'-kinase |
| COG0176 | Transaldolase |
| COG1690 | Uncharacterized conserved protein |
| COG1912 | Uncharacterized conserved protein |
| COG0011 | Uncharacterized conserved protein |
| COG0398 | Uncharacterized conserved protein |
| COG1900 | Uncharacterized conserved protein |
| COG2122 | Uncharacterized conserved protein |
| COG2509 | Uncharacterized FAD-dependent dehydrogenases |
| COG2231 | Uncharacterized protein related to Endonuclease III |