

SUPPLEMENTARY INFORMATION

Phylogenetic and environmental diversity of DsrAB-type dissimilatory (bi)sulfite reductases

Albert Leopold Müller, Kasper Urup Kjeldsen, Thomas Rattei, Michael Pester, and Alexander Loy

SUPPLEMENTARY MATERIALS AND METHODS

Recovery of *dsrAB* sequences from public databases using tblastx analysis

A bit score threshold for retrieving *dsrAB* sequences from public databases was defined by blasting entries of a *dsrAB* in-house 'seed' database (n=998) against each other with archaeal type *dsrAB* sequences serving as the outgroup for bacterial type *dsrAB* sequences and vice versa. The highest bit score of the outgroup entries + 10% (to make the search more conservative) was then used as a bit score threshold for each *dsrAB* seed entry. Subsequently, each *dsrAB* seed entry with its own threshold was blasted against NCBI's non-redundant and environmental databases (Benson *et al.*, 2014), the IMG/M database (Markowitz *et al.*, 2012) and the Camera database (Sun *et al.*, 2011). Metagenomic databases were filtered to exclude short single reads (<400 bp) in order to enhance search speed. Since we aimed at recovering continuous nucleic acid sequences of *dsrAB* with the intergenic region in addition to partial *dsrA* and *dsrB* sequences, we did not use the FunGene database (Fish *et al.*, 2013) that only collects individual *dsrA* and *dsrB* sequences from GenBank based on a Hidden Markov Model.

Ecological classification of *dsrAB* sequences

To analyze general habitat patterns, *dsrAB* sequences were assigned to broad environmental categories based on the qualitative description submitted with the sequence and/or the information in the publication. Categories based on the environmental origin of the sample (marine, estuarine, freshwater, soil, and industrial) were complemented by categories denoting special microbial lifestyles (thermophilic, alkali-/halophilic, and symbiotic). Sequences were not assigned to multiple categories, in case of sequences fitting in two or more categories, lifestyle categories were given precedence (i.e. a sequence from a marine thermophile would be classified as thermophilic and not as marine).

Tree calculations

Maximum likelihood (RAxML (Stamatakis, 2006), Dayhoff amino acid substitution model) and maximum parsimony (PHYLIP PROTPARS) trees were calculated in ARB. Neighbor joining trees (PHYLIP NEIGHBOR) were calculated in the PHYLIP software package (Felsenstein, 1989) based on the JTT matrix model (Jones *et al.*, 1992). The three trees were combined into a consensus tree by using the extended majority rule in PHYLIP CONSENSE. Branch lengths of the consensus tree were inferred by using the JTT matrix model (PHYLIP PROML). Bootstrap analysis was performed with all three methods with 100 or 1,000 re-samplings, depending on the size of the tree. In order to save computation time, sequences were clustered at 97% amino acid identity using the average neighbor algorithm in mother (Schloss *et al.*, 2009) and only a representative sequence of each cluster was used for calculation of the larger trees. The other sequences were subsequently added to the consensus tree using the ARB Parsimony interactive tool using the same alignment filter as for tree calculation. Trees were visualized using iTOL (Letunic & Bork, 2007).

For comparison of tree topologies, DsrAB and 16S rRNA trees were based on corresponding data sets, that is, the same number of organisms, to avoid sampling artifacts. Consensus trees were constructed using the strict consensus rule setting to allow for more polytomies and thus conservative, but more robust phylogenies (Ludwig *et al.*, 2004). DsrAB trees were calculated as mentioned previously. 16S rRNA trees were calculated using maximum likelihood (RAxML), maximum parsimony (PHYLIP DNAPARS), and neighbor joining (PHYLIP NEIGHBOR) methods and branch lengths in the 16S rRNA consensus tree were adjusted using PHYLIP DNAML.

Uncultured family-level DsrAB lineages

By adapting previously established criteria (Pester *et al.*, 2012), environmental DsrAB sequences were assigned into uncultured family-level lineages. First, sequences sharing $\geq 64\%$ amino acid identity were clustered into groups using the furthest neighbor algorithm in mothur (Schloss *et al.*, 2009). This very conservative threshold was based on minimum intra-family amino acid sequence identities of known families of sulfate-reducing microorganisms (SRM), which range from 64% to 89%. Second, a $\geq 64\%$ DsrAB sequence identity cluster was only designated as an “uncultured DsrAB lineage” if it consisted of at least three sequences and formed a monophyletic group in the extended majority rule-based consensus tree with a bootstrap support of $>70\%$ in at least two treeing methods. CodeML (PAML 4.8) (Yang, 2007) was used to determine non-synonymous/synonymous substitution rate ratios for the uncultured DsrAB lineages.

Comparisons between *dsrAB* and 16S rRNA gene identities

Gene identity plots of *dsrAB* and 16S rRNA genes of organisms for which both genes are known were performed in R (R Development Core Team, 2008) using nucleic acid distance matrices calculated in ARB without an alignment filter. A nucleotide identity species-level threshold for bacterial type *dsrAB* was inferred from a *dsrAB*/16S rRNA gene identity plot of organisms with non-laterally acquired bacterial type *dsrAB*. A 99% nucleotide similarity was used as a threshold for approximate species level assignment on 16S rRNA level (Stackebrandt & Ebers, 2006).

SUPPLEMENTARY RESULTS AND DISCUSSION

The four families and root of the DsrAB tree

Since *dsrA* and *dsrB* originated from an ancient gene duplication event that likely preceded the divergence of the bacterial and archaeal domains of life (Dahl *et al.*, 1993; Wagner *et al.*, 1998), it is possible to determine the root of the DsrAB tree by paralogous rooting (Iwabe *et al.*, 1989). An alignment of DsrA to DsrB based on 163 conserved homologous amino acid positions supports a nearly bilaterally symmetrical tree in which DsrA and DsrB sequences form mirrored branches converging at the root (Supplementary Figure S4). The family of oxidative DsrAB sequences from sulfide-oxidizing bacteria (SOB) branches off after the split between bacterial type and archaeal type sequences and thus likely represents a secondary, functional diversification. Hence, the branching pattern suggests that ancestral DsrAB functioned in the reductive direction, a finding that is corroborated by geochemical data that suggests that sulfite reduction occurred very early in Earth's history and possibly predated the evolution of sulfate reduction (Skyring & Donnelly, 1982). It is not possible to conclusively determine the root of the DsrAB tree with respect to the fourth type of DsrAB family, constituted by the exceptional second DsrAB copy of *Moorella thermoacetica*, because its position differs in the DsrA and DsrB branches of the tree (Supplementary Figure S4). The basal placement of the *M. thermoacetica* DsrAB copy 2 subunits in the DsrA/DsrB paralog tree might be a treeing artifact caused by long-branch attraction (Bergsten, 2005) between the two subunits. We thus performed long-branch extraction (Siddall & Whiting, 1999) by repeating the tree calculation and by either omitting the DsrA or the DsrB subunit of *M. thermoacetica* DsrAB copy 2 in any one calculation. Erroneous phylogenetic placement of *M. thermoacetica* DsrAB copy 2 is unlikely because the phylogenetic position of the respective subunit in both test trees was identical to the position in the tree calculated with both subunits (data not shown). We thus conclude that the root of the DsrAB tree is either between *M. thermoacetica* DsrAB copy 2 and all other DsrAB sequences or between

the *M. thermoacetica* DsrAB copy 2/reductive archaeal type DsrAB branch and the oxidative/reductive bacterial type DsrAB branch.

DsrAB consensus phylogeny and members of main lineages

Reductive bacterial type DsrAB. The *Deltaproteobacteria* supercluster contains the majority of known bacterial type DsrAB sequences, including *dsrAB* from all *deltaproteobacterial* SRM, *Deltaproteobacteria*-related *dsrAB* from (mostly) thermophilic sulfate and/or sulfite-reducing members of the phyla *Firmicutes* (i.e., members of *Desulfotomaculum* subclusters Ib, Ic, Id, and Ie, *Ammonifex degensii*, *Candidatus Desulforudis audaxviator*, *Desulfoviregula thermocuniculi*, *Sporotomaculum hydroxybenzoicum*, and *Moorella thermoacetica dsrAB* copy 1) and *Thermodesulfobacteria* (*Thermodesulfobacterium* and *Thermodesulfatator*), uncultured DsrAB lineages 1 and 11 and other environmental sequences (Figure 1, Supplementary Figure S1). We have renamed the *Thermodesulfovibrio* supercluster into *Nitrospirae* supercluster because it now comprises *dsrAB* from two taxa within the phylum *Nitrospirae*, namely the genus *Thermodesulfovibrio* and *Candidatus Magnetobacterium casensis* (Lin *et al.*, 2014). The *Nitrospirae* supercluster additionally contains uncultured DsrAB lineage 10 and 13, and other environmental sequences. The environmental supercluster 1 does so far not contain *dsrAB* from any cultured organism and includes uncultured DsrAB lineages 8, 9, and 12, and other environmental *dsrAB* sequences. Most sequences from cultured organisms in the *Firmicutes* group belong to the phylum *Firmicutes* (i.e., the SRM genera *Desulfotomaculum*, *Desulfosporosinus*, and *Desulfurispora*, the non-SRM genera *Desulfitobacterium*, *Carboxydotherrmus*, *Desulfitibacter*, and *Thermanaeromonas* that are able to reduce other sulfur compounds, the syntroph *Pelotomaculum propionicum*, and members of the family *Sporomusaceae* (Yutin & Galperin, 2013), the acetogen *Acetonema longum* and the thiosulfate-utilizing *Thermosinus carboxydivorans*) yet these do not form a monophyletic DsrAB cluster. Instead, they are phylogenetically intermingled with members of the taxonomically uncertain genus *Thermodesulfobium* (Mori *et al.*, 2003), *dsrAB*-carrying members of the phyla *Actinobacteria*, *Aigarchaeota* and *Caldiserica*, uncultured DsrAB lineages 2 to 7, and other environmental *dsrAB* sequences (Figure 1, Supplementary Figure S1). Continuous genomic analysis of cultured microorganisms or single microbial cells has led to the discovery of *dsrAB* in members of phyla previously not known to possess these genes. Hence, DsrAB-coding genes were detected on amplified genomes of individual cells belonging to the bacterial phylum *Caldiserica* (JGI_0000059-M03, accession AQSQ01000030) and the archaeal candidate phylum *Aigarchaeota* (pSL4 archaeon JGI_0000106-J15, accession ASPF01000004) (Rinke *et al.*, 2013) (Figure 1). The phylum *Caldiserica* (formerly known as candidate phylum OP5) so far contains only one characterized species, *Caldisericum exile*, which uses thiosulfate, sulfite, and elemental

sulfur, but not sulfate, as electron acceptors, yet its genome (NC_017096) does not contain *dsrAB* (Mori *et al.*, 2009). The genome of the intestinal actinobacterium *Gordonibacter pamelaee* (NC_021021) (Würdemann *et al.*, 2009) also contains *dsrAB* that consistently cluster with the *Firmicutes* sister genera *Desulfosporosinus* and *Desulfitobacterium*. Importantly, *G. pamelaee* lacks further genes of the canonical dissimilatory sulfate reduction pathway in its genome and is thus likely not an SRM.

Oxidative bacterial type DsrAB. The majority of oxidative DsrAB sequences belong to the *Gammaproteobacteria* cluster that contains cultivated members of families *Chromatiaceae* (nine genera), *Ectothiorhodospiraceae* (genera *Alkalilimnicola*, *Halorhodospira*, and *Thioalkalivibrio*) and *Thiotrichaceae* (genus *Thiothrix*) (Figure 1, Supplementary Figure S2). Interestingly, DsrAB of *Thioalkalivibrio nitratreducens* does not fall into this *Gammaproteobacteria* cluster, in contrast to sequences of the other members of the genus, *T. sulfidiphilus*, *T. thiocyanodenitrificans*, and *T. paradoxus* [the genome sequence of *T. paradoxus*, NZ_AGFB01000003 was wrongly published as *T. thiooxyanoxidans* (D.Y. Sorokin, personal communication) and has since been removed]. The second largest oxidative DsrAB cluster, the *Alphaproteobacteria* cluster, is subdivided into two main branches, one harbors members of the genera *Magnetospirillum* and *Azospirillum* (family *Rhodospirillaceae*, order *Rhodospirillales*), and the other contains *Rhodomicrobium vannielii* (family *Hyphomicrobiaceae*, order *Rhizobiales*) and environmental sequences. The *Betaproteobacteria* cluster contains *dsrAB* from members of the genus *Thiobacillus*, from *Sulfuricella denitrificans* (both family *Hydrogenophilaceae*, order *Hydrogenophilales*), and from *Sideroxydans lithotrophicus* (family *Gallionellaceae*, order *Gallionellales*). The *Chlorobi* cluster contains *dsrAB* from members of the family *Chlorobiaceae*. *Magnetococcus marinus* (family *Magnetococcaceae*, order *Magnetococcales*) has been provisionally included in the *Alphaproteobacteria* (Bazylnski *et al.*, 2013), but its DsrAB is very dissimilar from other alphaproteobacterial DsrAB and instead clusters with DsrAB of *Chlorobi* and the *deltaproteobacterial* SAR324 clade.

Reductive archaeal type DsrAB. This type of DsrAB is present in three genera of hyperthermophilic *Crenarchaeota* (*Pyrobaculum*, *Vulcanisaeta* and *Caldivirga*). Each of these three genera represents a distinct monophyletic group in the archaeal DsrAB tree with *Caldivirga maquilingensis*, an organism capable of reducing sulfate, thiosulfate and sulfur (Itoh *et al.*, 1999), occupying the deepest branch (Figure 1, Supplementary Figure S3). *C. maquilingensis* is the only organism in which the neighboring genes *dsrA* and *dsrB* are arranged in different directions. *Vulcanisaeta distributa* (Itoh *et al.*, 2002) and *V. moutnovskia* (Prokofeva *et al.*, 2005) utilize elemental sulfur and thiosulfate for growth. However, both species might also be capable

of sulfate reduction (Itoh *et al.*, 2002) since genes for the complete canonical sulfate reduction pathway are present in their genomes (Mavromatis *et al.*, 2010; Gumerov *et al.*, 2011). All *Pyrobaculum* species seem to be capable of using thiosulfate as an electron acceptor (Völkl *et al.*, 1993; Molitor *et al.*, 1998; Huber *et al.*, 2000; Sako *et al.*, 2001; Amo *et al.*, 2002) and some were also shown to use elemental sulfur (*P. arsenaticum* (Huber *et al.*, 2000), *P. islandicum* (Molitor *et al.*, 1998), *P. neutrophilum* (Fischer *et al.*, 1983), *P. oguniense* (Sako *et al.*, 2001)) as a terminal electron acceptor whereas other species are inhibited by elemental sulfur (*P. aerophilum* (Völkl *et al.*, 1993), *P. calidifontis* (Amo *et al.*, 2002)). *P. islandicum* was also shown to utilize other sulfur compounds like sulfite, L-cysteine and oxidized glutathione (Molitor *et al.*, 1998). Interestingly, some but not all *Pyrobaculum* genomes harbor multiple *dsrAB* copies with a phylogeny that indicates a complex evolutionary history of *dsrAB* in members of this genus involving gene duplications, gene losses, and/or intra-genus LGT. Some *Pyrobaculum* species thus possess two (*P. aerophilum*, *P. calidifontis*) or three (*P. arsenaticum*, *P. oguniense*) *dsrAB* copies, all of which have intact reading frames and conserved siroheme binding sites. These multiple copies do not cluster together according to species affiliation (Supplementary Figure S3). No reports exist on function of specific copies of *dsrAB* in *Pyrobaculum* genomes and there is no obvious pattern regarding copy number and/or type and metabolic potential of *Pyrobaculum* species. Gene duplications can be a source of new protein functions because one (or both) of the now functionally redundant paralogs experience a period of relaxed selection and accelerated evolution (Ohno, 1970) in enzymes leading, for example, to different substrate affinities or specificities (Baani & Liesack, 2008). Presence of multiple *dsrAB* in some *Pyrobaculum* species is very unusual as *dsrAB* are mostly present as single-copy genes. The only other organisms known to possess more than one copy of *dsrAB* is *Moorella thermoacetica* with its two significantly different copies of *dsrAB*, and the SOB *Chlorobaculum tepidum*, which possesses two nearly identical *dsrAB* copies, but one has an authentic frameshift in *dsrB* and is likely not functional (Eisen *et al.*, 2002).

Moorella thermoacetica DsrAB copy 2. The unusual second *dsrAB* copy of *M. thermoacetica* is the only representative of this type of DsrAB. The two strains of *M. thermoacetica* for which genome sequences are available, strain ATCC 39073 (Pierce *et al.*, 2008) and strain Y72 (Tsukahara *et al.*, 2014), are the only bacteria known to possess two significantly different copies of *dsrAB*. The first copy (NC_007644, 1634923..1637803; BARR01000003, 15717..18335) is a bacterial DsrAB related to *Firmicutes* sequences within the *Deltaproteobacteria* supercluster, whereas the second copy (NC_007644, 1664993..1667213; BARR01000005, 9074..11294) is not closely related to other DsrAB sequences (Pierce *et al.*, 2008). *M. thermoacetica* can reduce

thiosulfate and dimethylsulfoxide (Drake & Daniel, 2004) but nothing specific is known about the individual functions of its DsrAB copies.

Putative lateral gene transfers of *dsrAB*

Consensus trees of corresponding DsrAB and 16S rRNA sequences were compared for topological incongruences as signs of LGT (Figures 2 and 3). Additionally, we plotted identity values of 16S rRNA and *dsrAB* genes of pairs of pure cultures and genomes against each other (Supplementary Figure S6). 16S rRNA gene and *dsrAB* identities are highly correlated, which indicates that these genes generally evolve in parallel at constant mutation rates. Considerable deviations from linear regression in such a plot thus indicate changes in the standard evolutionary mechanism. In comparisons of the major DsrAB families (Supplementary Figure S6), deviations towards the top left of the plot (*dsrAB* identity > 16S rRNA identity) is indicative of laterally acquired *dsrAB* (LA-*dsrAB*), while deviation towards the bottom right of the plot (*dsrAB* identity < 16S rRNA identity) might have multiple causes such as LGT, diversification of *dsrAB* through accelerated mutation or a combination of these processes with gene duplication/loss events.

DsrAB and 16S rRNA branching patterns are generally similar but show previously recognized and newly identified topological inconsistencies. Acquisition of *dsrAB* of a group of *Firmicutes* from deltaproteobacterial ancestors of the *Desulfatiglans anilini* (formerly *Desulfobacterium anilini*) (Suzuki et al., 2014)) lineage (Figure 2) (Klein et al., 2001; Zverlov et al., 2005) is confirmed by the *dsrAB*-16S rRNA gene identity plot. When compared with all organisms carrying “standard” vertically inherited bacterial *dsrAB*, non-LA-*dsrAB* *Firmicutes* cluster along the linear regression, while LA-*dsrAB* *Firmicutes* show slightly higher *dsrAB* gene identity (Supplementary Figure S6 B). Additionally, *dsrAB* of the latter organisms possess characteristic insertions that are typical for deltaproteobacterial *dsrAB* and provide further independent evidence of a deltaproteobacterial origin (Klein et al., 2001). It has been proposed that members of the phylum *Thermodesulfobacteria* (genera *Thermodesulfobacterium* and *Thermodesulfatator*) also have LA-*dsrAB* (Klein et al., 2001). *Thermodesulfobacteria* form a monophyletic branch within the *Deltaproteobacteria* cluster in the DsrAB tree (Figure 2) and possess the characteristic deltaproteobacterial *dsrAB* insertions, but, in contrast to *Firmicutes* with LA-*dsrAB*, show no deviations in the *dsrAB*-16S rRNA gene identity plot when compared to non-LA bacterial *dsrAB* (Supplementary Figure S6 C). Notably, it has recently been suggested that the phylogenetic inconsistency between 16S rRNA and DsrAB is due to incorrect placement of the phylum *Thermodesulfobacteria* in the 16S rRNA tree (Lang et al., 2013). Phylogenetic analysis of a set of 24 concatenated phylogenetic marker genes identified the *Thermodesulfobacteria* as a

sister-group of the deltaproteobacterial order *Desulfovibrionales* (Lang et al., 2013). However, *Thermodesulfobacteria* were clearly unrelated to *Deltaproteobacteria* in another whole genome tree based on 38 concatenated marker genes (Rinke et al., 2013). Because of these conflicting findings, it currently remains unresolved if *Thermodesulfobacteria* received *dsrAB* via LGT.

All evidence points to a bacterial origin of *dsrAB* in members of the archaeal genus *Archaeoglobus*. *Archaeoglobus* DsrAB branches unambiguously in the reductive bacterial type DsrAB family (Klein et al., 2001) (Figure 2). Furthermore, *Archaeoglobus* species (*Euryarcheota*) are more similar to bacteria on *dsrAB* level than to *Crenarchaeota* and thus deviate considerably from the linear regression of *dsrAB*-16S rRNA sequence identities, whereas *Crenarchaeota* do not (Supplementary Figure S6 C). This is consistent with paralogous rooting analysis and confirms an archaeal origin of *dsrAB* in *Crenarchaeota* members.

Analogous to *Archaeoglobus*, DsrAB in the *Aigarchaeota* member clearly belongs to the reductive bacterial type DsrAB family (Figures 1 and 2). Substantial differences in GC content between host genomes and acquired genes can be used to infer LGT events (Lawrence & Ochman, 1997) and a difference of more than 10% was previously used as an additional indication of LA-*dsrAB* (Klein et al., 2001). The GC content of the *dsrAB* sequence (44%) of the aigarchaeon is more than 10% different from the whole genome (38%), whereas the most closely related *dsrAB* sequences have GC contents between 51 and 55%, suggesting that *dsrAB* was relatively recently acquired from a donor with a higher genomic GC content and is currently undergoing evolutionary adaptation into the new genome (Lawrence & Ochman, 1997). The *Aigarchaeota* member also shows a deviation pattern in the *dsrAB*-16S rRNA gene identity plot that is similar to *Archaeoglobus*, with its *dsrAB* being much more similar to bacterial sequences than its 16S rRNA sequence (Supplementary Figure S6 C).

DsrAB of the actinobacterium *G. pamelaeae* forms a stable monophyletic group with DsrAB of the *Firmicutes* genera *Desulfosporosinus* and *Desulfitobacterium* (Figures 1 and 2), which suggests that *G. pamelaeae dsrAB* were laterally acquired from a *Firmicutes* member. However, there is no evidence that this potential LGT event occurred recently, since the GC content of *dsrAB* (65.3%, the highest value of any known organisms with a bacterial type DsrAB) is similar to the GC content of the whole genome (64%) of *G. pamelaeae*, but considerably higher than the GC content of *dsrAB* and genome sequences of *Desulfosporosinus* and *Desulfitobacterium* species (41-49%). Furthermore, *dsrAB* of *G. pamelaeae* lack a characteristic insertion at position 1053 that is shared by all *Desulfosporosinus* and *Desulfitobacterium* species. The *dsrAB*-16S rRNA gene identity plot does not suggest LGT (Supplementary Figure S6 C). So far *G. pamelaeae* is the only

member of the phylum *Actinobacteria* with *dsrAB* and there are no closely related environmental *dsrAB* sequences with a similar, high GC content that might hint at the presence of further actinobacterial *dsrAB* sequences in the database.

The branching patterns of the only known *dsrAB*-carrying member of the phylum *Caldiserica* do not suggest LGT of its *dsrAB* as it branches deeply in both the DsrAB tree and the 16S rRNA tree (Figure 2). However, the GC content is >10% different between *dsrAB* (44.1%) and genome (35%) and more similar to phylogenetically related *dsrAB* sequences, hinting to lateral acquisition. Furthermore, in the *dsrAB*-16S rRNA gene identity plot it shows considerable deviation from other bacterial *dsrAB* sequences (Supplementary Figure S6 C).

Primers targeting *dsrA* and *dsrB*

The very first *dsrAB*-targeted primers were designed on the basis of nucleotide sequence homologies of only two sequences, namely *dsrAB* from *Archaeoglobus fulgidus* and *Desulfovibrio vulgaris* (Karkhoff-Schweizer *et al.*, 1995). Shortly after, first versions of the now widely used DSR1F and DSR4R primers were published (Wagner *et al.*, 1998). They target highly conserved sequence regions of *dsrAB* and amplify a ~1.9kb fragment that covers approximately 85% of *dsrA*, 70% of *dsrB* and the intergenic spacer region between the two genes (Supplementary Figure S7). These degenerated primers were repeatedly updated by introducing additional primer variants (Loy *et al.*, 2004; Zverlov *et al.*, 2005; Pester *et al.*, 2010) and were modified (Kondo *et al.*, 2004; Suzuki *et al.*, 2005; Schmalenberger *et al.*, 2007; Loy *et al.*, 2009; Lenk *et al.*, 2011; Lever *et al.*, 2013) to adjust their coverage or improve PCR efficiency (Supplementary Tables S1 and S3). Also, several primers were developed that bind within the region amplified by DSR1F/DSR4R primer variants and were applied to amplify shorter fragments of either *dsrA* or *dsrB* for denaturing gradient gel electrophoresis (Geets *et al.*, 2005; Steger *et al.*, 2011), terminal restriction fragment length polymorphism analysis (Santillano *et al.*, 2010), quantitative real-time PCR (Kondo *et al.*, 2004; Ben-Dov *et al.*, 2007; Chin *et al.*, 2008; Gittel *et al.*, 2009; Pereyra *et al.*, 2010) or (nested) PCR (Dhillon *et al.*, 2003; Giloteaux *et al.*, 2010; Akob *et al.*, 2012; Lever *et al.*, 2013). Analogous to primers targeting the reductive bacterial type *dsrAB*, primers were also developed for oxidative bacterial type *dsrAB* of SOB (Loy *et al.*, 2009; Mori *et al.*, 2010; Lenk *et al.*, 2011; Luo *et al.*, 2011; Lever *et al.*, 2013) (Supplementary Tables S2 and S4). Primers for amplification of reductive, archaeal type *dsrAB* sequences or the second *dsrAB* copy of *M. thermoacetica* are not yet available.

References

Akob DM, Lee SH, Sheth M, Küsel K, Watson DB, Palumbo AV *et al.* (2012). Gene Expression Correlates with Process Rates Quantified for Sulfate- and Fe(III)-Reducing Bacteria in U(VI)-Contaminated Sediments. *Front Microbiol* **3**: 280.

Amo T, Paje ML, Inagaki A, Ezaki S, Atomi H, Imanaka T. (2002). *Pyrobaculum calidifontis* sp. nov., a novel hyperthermophilic archaeon that grows in atmospheric air. *Archaea* **1**: 113-121.

Baani M, Liesack W. (2008). Two isozymes of particulate methane monooxygenase with different methane oxidation kinetics are found in *Methylocystis* sp. strain SC2. *Proc Natl Acad Sci U S A* **105**: 10203-10208.

Bazylinski DA, Williams TJ, Lefevre CT, Berg RJ, Zhang CLL, Bowser SS *et al.* (2013). *Magnetococcus marinus* gen. nov., sp nov., a marine, magnetotactic bacterium that represents a novel lineage (Magnetococcaceae fam. nov., Magnetococcales ord. nov.) at the base of the Alphaproteobacteria. *Int J Syst Evol Microbiol* **63**: 801-808.

Ben-Dov E, Brenner A, Kushmaro A. (2007). Quantification of sulfate-reducing bacteria in industrial wastewater, by real-time polymerase chain reaction (PCR) using *dsrA* and *apsA* genes. *Microb Ecol* **54**: 439-451.

Benson DA, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. (2014). GenBank. *Nucleic Acids Res* **42**: D32-37.

Bergsten J. (2005). A review of long-branch attraction. *Cladistics* **21**: 163-193.

Chin KJ, Sharma ML, Russell LA, O'Neill KR, Lovley DR. (2008). Quantifying expression of a dissimilatory (bi) sulfite reductase gene in petroleum-contaminated marine Harbor Sediments. *Microb Ecol* **55**: 489-499.

Dahl C, Kredich NM, Deutzmann R, Trüper HG. (1993). Dissimilatory sulphite reductase from *Archaeoglobus fulgidus*: physico-chemical properties of the enzyme and cloning, sequencing and analysis of the reductase genes. *J Gen Microbiol* **139**: 1817-1828.

Dhillon A, Teske A, Dillon J, Stahl DA, Sogin ML. (2003). Molecular characterization of sulfate-reducing bacteria in the Guaymas Basin. *Appl Environ Microbiol* **69**: 2765-2772.

Drake HL, Daniel SL. (2004). Physiology of the thermophilic acetogen *Moorella thermoacetica* (vol 155, pg 422, 2004). *Res Microbiol* **155**: 868-883.

Eisen JA, Nelson KE, Paulsen IT, Heidelberg JF, Wu M, Dodson RJ *et al.* (2002). The complete genome sequence of *Chlorobium tepidum* TLS, a photosynthetic, anaerobic, green-sulfur bacterium. *Proc Natl Acad Sci U S A* **99**: 9509-9514.

Felsenstein J. (1989). PHYLIP-phylogeny inference package. *Cladistics* **5**: 163-166.

Fischer F, Zillig W, Stetter KO, Schreiber G. (1983). Chemolithoautotrophic metabolism of anaerobic extremely thermophilic archaeobacteria. *Nature* **301**: 511-513.

Fish JA, Chai B, Wang Q, Sun Y, Brown CT, Tiedje JM *et al.* (2013). FunGene: the functional gene pipeline and repository. *Front Microbiol* **4**: 291.

Geets J, Borremans B, Vangronsveld J, Diels L, van der Lelie D. (2005). Molecular monitoring of SRB community structure and dynamics in batch experiments to examine the applicability of in situ precipitation of heavy metals for groundwater remediation. *J Soils Sediments* **5**: 149-163.

Giloteaux L, Goni-Urriza M, Duran R. (2010). Nested PCR and new primers for analysis of sulfate-reducing bacteria in low-cell-biomass environments. *Appl Environ Microbiol* **76**: 2856-2865.

Gittel A, Sørensen KB, Skovhus TL, Ingvorsen K, Schramm A. (2009). Prokaryotic Community Structure and Sulfate Reducer Activity in Water from High-Temperature Oil Reservoirs with and without Nitrate Treatment. *Appl Environ Microbiol* **75**: 7086-7096.

Gumerov VM, Mardanov AV, Beletsky AV, Prokofeva MI, Bonch-Osmolovskaya EA, Ravin NV *et al.* (2011). Complete genome sequence of "Vulcanisaeta moutnovskia" strain 768-28, a novel member of the hyperthermophilic crenarchaeal genus Vulcanisaeta. *J Bacteriol* **193**: 2355-2356.

Huber R, Sacher M, Vollmann A, Huber H, Rose D. (2000). Respiration of arsenate and selenate by hyperthermophilic archaea. *Syst Appl Microbiol* **23**: 305-314.

Itoh T, Suzuki K, Sanchez PC, Nakase T. (1999). *Caldivirga maquilingensis* gen. nov., sp. nov., a new genus of rod-shaped crenarchaeote isolated from a hot spring in the Philippines. *Int J Syst Bacteriol* **49 Pt 3**: 1157-1163.

Itoh T, Suzuki K, Nakase T. (2002). *Vulcanisaeta distributa* gen. nov., sp. nov., and *Vulcanisaeta souniana* sp. nov., novel hyperthermophilic, rod-shaped crenarchaeotes isolated from hot springs in Japan. *Int J Syst Evol Microbiol* **52**: 1097-1104.

Iwabe N, Kuma K, Hasegawa M, Osawa S, Miyata T. (1989). Evolutionary relationship of archaeobacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc Natl Acad Sci U S A* **86**: 9355-9359.

Jones DT, Taylor WR, Thornton JM. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* **8**: 275-282.

Karkhoff-Schweizer RR, Huber DP, Voordouw G. (1995). Conservation of the genes for dissimilatory sulfite reductase from *Desulfovibrio vulgaris* and *Archaeoglobus fulgidus* allows their detection by PCR. *Appl Environ Microbiol* **61**: 290-296.

Klein M, Friedrich M, Roger AJ, Hugenholtz P, Fishbain S, Abicht H *et al.* (2001). Multiple lateral transfers of dissimilatory sulfite reductase genes between major lineages of sulfate-reducing prokaryotes. *J Bacteriol* **183**: 6028-6035.

Kondo R, Nedwell DB, Purdy KJ, Silva SD. (2004). Detection and enumeration of sulphate-reducing bacteria in estuarine sediments by competitive PCR. *Geomicrobiol J* **21**: 145-157.

Lang JM, Darling AE, Eisen JA. (2013). Phylogeny of bacterial and archaeal genomes using conserved genes: supertrees and supermatrices. *PloS one* **8**: e62510.

Lawrence JG, Ochman H. (1997). Amelioration of bacterial genomes: rates of change and exchange. *J Mol Evol* **44**: 383-397.

Lenk S, Arnds J, Zerjatke K, Musat N, Amann R, Mussmann M. (2011). Novel groups of Gammaproteobacteria catalyse sulfur oxidation and carbon fixation in a coastal, intertidal sediment. *Environ Microbiol* **13**: 758-774.

- Letunic I, Bork P. (2007). Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**: 127-128.
- Lever MA, Rouxel O, Alt JC, Shimizu N, Ono S, Coggon RM *et al.* (2013). Evidence for microbial carbon and sulfur cycling in deeply buried ridge flank basalt. *Science* **339**: 1305-1308.
- Lin W, Deng A, Wang Z, Li Y, Wen T, Wu LF *et al.* (2014). Genomic insights into the uncultured genus 'Candidatus Magnetobacterium' in the phylum Nitrospirae. *ISME J.*
- Loy A, Küsel K, Lehner A, Drake HL, Wagner M. (2004). Microarray and functional gene analyses of sulfate-reducing prokaryotes in low-sulfate, acidic fens reveal cooccurrence of recognized genera and novel lineages. *Appl Environ Microbiol* **70**: 6998-7009.
- Loy A, Duller S, Baranyi C, Musmann M, Ott J, Sharon I *et al.* (2009). Reverse dissimilatory sulfite reductase as phylogenetic marker for a subgroup of sulfur-oxidizing prokaryotes. *Environ Microbiol* **11**: 289-299.
- Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363-1371.
- Luo JF, Lin WT, Guo Y. (2011). Functional genes based analysis of sulfur-oxidizing bacteria community in sulfide removing bioreactor. *Appl Microbiol Biotechnol* **90**: 769-778.
- Markowitz VM, Chen IM, Chu K, Szeto E, Palaniappan K, Grechkin Y *et al.* (2012). IMG/M: the integrated metagenome data management and comparative analysis system. *Nucleic Acids Res* **40**: D123-129.
- Mavromatis K, Sikorski J, Pabst E, Teshima H, Lapidus A, Lucas S *et al.* (2010). Complete genome sequence of *Vulcanisaeta distributa* type strain (IC-017). *Stand Genomic Sci* **3**: 117-125.
- Molitor M, Dahl C, Molitor I, Schäfer U, Speich N, Huber R *et al.* (1998). A dissimilatory sirohaem-sulfite-reductase-type protein from the hyperthermophilic archaeon *Pyrobaculum islandicum*. *Microbiology* **144 (Pt 2)**: 529-541.
- Mori K, Kim H, Kakegawa T, Hanada S. (2003). A novel lineage of sulfate-reducing microorganisms: Thermodesulfobiaceae fam. nov., Thermodesulfobium narugense, gen. nov., sp. nov., a new thermophilic isolate from a hot spring. *Extremophiles* **7**: 283-290.
- Mori K, Yamaguchi K, Sakiyama Y, Urabe T, Suzuki K. (2009). *Caldisericum exile* gen. nov., sp. nov., an anaerobic, thermophilic, filamentous bacterium of a novel bacterial phylum, *Caldiserica* phyl. nov., originally called the candidate phylum OP5, and description of *Caldisericaceae* fam. nov., *Caldisericales* ord. nov. and *Caldisericia* classis nov. *Int J Syst Evol Microbiol* **59**: 2894-2898.
- Mori Y, Purdy KJ, Oakley BB, Kondo R. (2010). Comprehensive Detection of Phototrophic Sulfur Bacteria Using PCR Primers That Target Reverse Dissimilatory Sulfite Reductase Gene. *Microbes Environ* **25**: 190-196.
- Ohno S. (1970). *Evolution by gene duplication*. London: George Alien & Unwin Ltd. Berlin, Heidelberg and New York: Springer-Verlag.
- Pereyra LP, Hiibel SR, Riquelme MVP, Reardon KF, Pruden A. (2010). Detection and Quantification of Functional Genes of Cellulose-Degrading, Fermentative, and Sulfate-Reducing Bacteria and Methanogenic Archaea. *Appl Environ Microbiol* **76**: 2192-2202.

- Pester M, Bittner N, Deevong P, Wagner M, Loy A. (2010). A 'rare biosphere' microorganism contributes to sulfate reduction in a peatland. *ISME J* **4**: 1591-1602.
- Pester M, Knorr KH, Friedrich MW, Wagner M, Loy A. (2012). Sulfate-reducing microorganisms in wetlands - fameless actors in carbon cycling and climate change. *Front Microbiol* **3**: 72.
- Pierce E, Xie G, Barabote RD, Saunders E, Han CS, Detter JC *et al.* (2008). The complete genome sequence of *Moorella thermoacetica* (f. *Clostridium thermoaceticum*). *Environ Microbiol* **10**: 2550-2573.
- Prokofeva MI, Kublanov IV, Nercessian O, Tourova TP, Kolganova TV, Lebedinsky AV *et al.* (2005). Cultivated anaerobic acidophilic/acidotolerant thermophiles from terrestrial and deep-sea hydrothermal habitats. *Extremophiles* **9**: 437-448.
- R Development Core Team. (2008). R: A language and environment for statistical computing.: R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF *et al.* (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**: 431-437.
- Sako Y, Nunoura T, Uchida A. (2001). *Pyrobaculum oguniense* sp. nov., a novel facultatively aerobic and hyperthermophilic archaeon growing at up to 97 degrees C. *Int J Syst Evol Microbiol* **51**: 303-309.
- Santillano D, Boetius A, Ramette A. (2010). Improved dsrA-Based Terminal Restriction Fragment Length Polymorphism Analysis of Sulfate-Reducing Bacteria. *Appl Environ Microbiol* **76**: 5308-5311.
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al.* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537-7541.
- Schmalenberger A, Drake HL, Kusel K. (2007). High unique diversity of sulfate-reducing prokaryotes characterized in a depth gradient in an acidic fen. *Environ Microbiol* **9**: 1317-1328.
- Siddall ME, Whiting MF. (1999). Long-branch abstractions. *Cladistics* **15**: 9-24.
- Skyring GW, Donnelly TH. (1982). Precambrian Sulfur Isotopes and a Possible Role for Sulfite in the Evolution of Biological Sulfate Reduction. *Precambrian Res* **17**: 41-61.
- Stackebrandt E, Ebers J. (2006). Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today* **33**: 152.
- Stamatakis A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688-2690.
- Steger D, Wentrup C, Braunegger C, Deevong P, Hofer M, Richter A *et al.* (2011). Microorganisms with Novel Dissimilatory (Bi)Sulfite Reductase Genes Are Widespread and Part of the Core Microbiota in Low-Sulfate Peatlands. *Appl Environ Microbiol* **77**: 1231-1242.
- Sun S, Chen J, Li W, Altintas I, Lin A, Peltier S *et al.* (2011). Community cyberinfrastructure for Advanced Microbial Ecology Research and Analysis: the CAMERA resource. *Nucleic Acids Res* **39**: D546-551.

Suzuki D, Cui X, Li Z, Zhang C, Katayama A. (2014). Reclassification of *Desulfobacterium anilini* as *Desulfatiglans anilini* comb. nov. to *Desulfatiglans* gen. nov., and description of a 4-chlorophenol-degrading sulfate-reducing bacterium-*Desulfatiglans parachlorophenolica* sp. nov. *Int J Syst Evol Microbiol*: ijs-0.

Suzuki Y, Kelly SD, Kemner KM, Banfield JF. (2005). Direct microbial reduction and subsequent preservation of uranium in natural near-surface sediment. *Appl Environ Microbiol* **71**: 1790-1797.

Tsukahara K, Kita A, Nakashimada Y, Hoshino T, Murakami K. (2014). Genome-guided analysis of transformation efficiency and carbon dioxide assimilation by *Moorella thermoacetica* Y72. *Gene* **535**: 150-155.

Völkl P, Huber R, Drobner E, Rachel R, Burggraf S, Trincone A *et al.* (1993). *Pyrobaculum aerophilum* sp. nov., a novel nitrate-reducing hyperthermophilic archaeum. *Appl Environ Microbiol* **59**: 2918-2926.

Wagner M, Roger AJ, Flax JL, Brusseau GA, Stahl DA. (1998). Phylogeny of dissimilatory sulfite reductases supports an early origin of sulfate respiration. *J Bacteriol* **180**: 2975-2982.

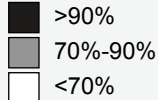
Würdemann D, Tindall BJ, Pukall R, Lunsdorf H, Strompl C, Namuth T *et al.* (2009). *Gordonibacter pamelaiae* gen. nov., sp. nov., a new member of the Coriobacteriaceae isolated from a patient with Crohn's disease, and reclassification of *Eggerthella hongkongensis* Lau *et al.* 2006 as *Paraeggerthella hongkongensis* gen. nov., comb. nov. *Int J Syst Evol Microbiol* **59**: 1405-1415.

Yang Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**: 1586-1591.

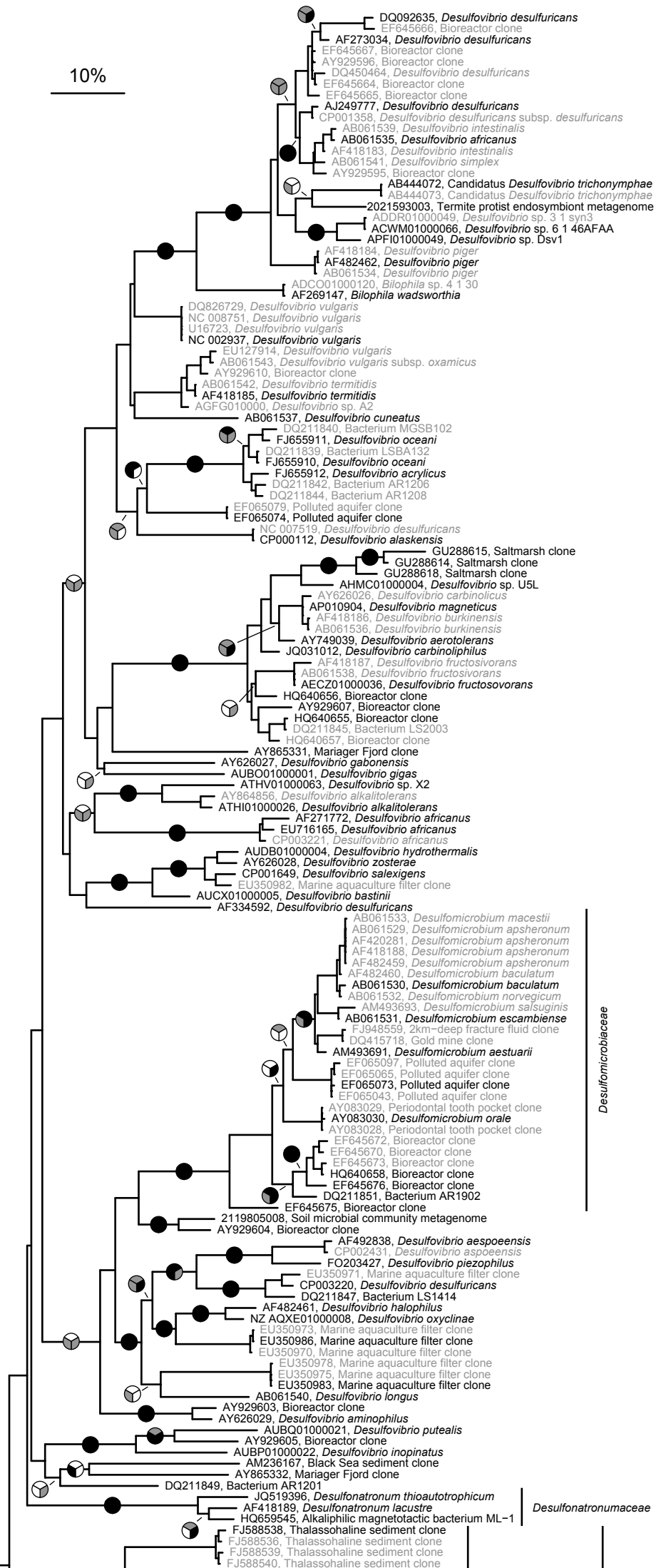
Yutin N, Galperin MY. (2013). A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environ Microbiol* **15**: 2631-2641.

Zverlov V, Klein M, Lückner S, Friedrich MW, Kellermann J, Stahl DA *et al.* (2005). Lateral gene transfer of dissimilatory (bi)sulfite reductase revisited. *J Bacteriol* **187**: 2203-2208.

Bootstraps:



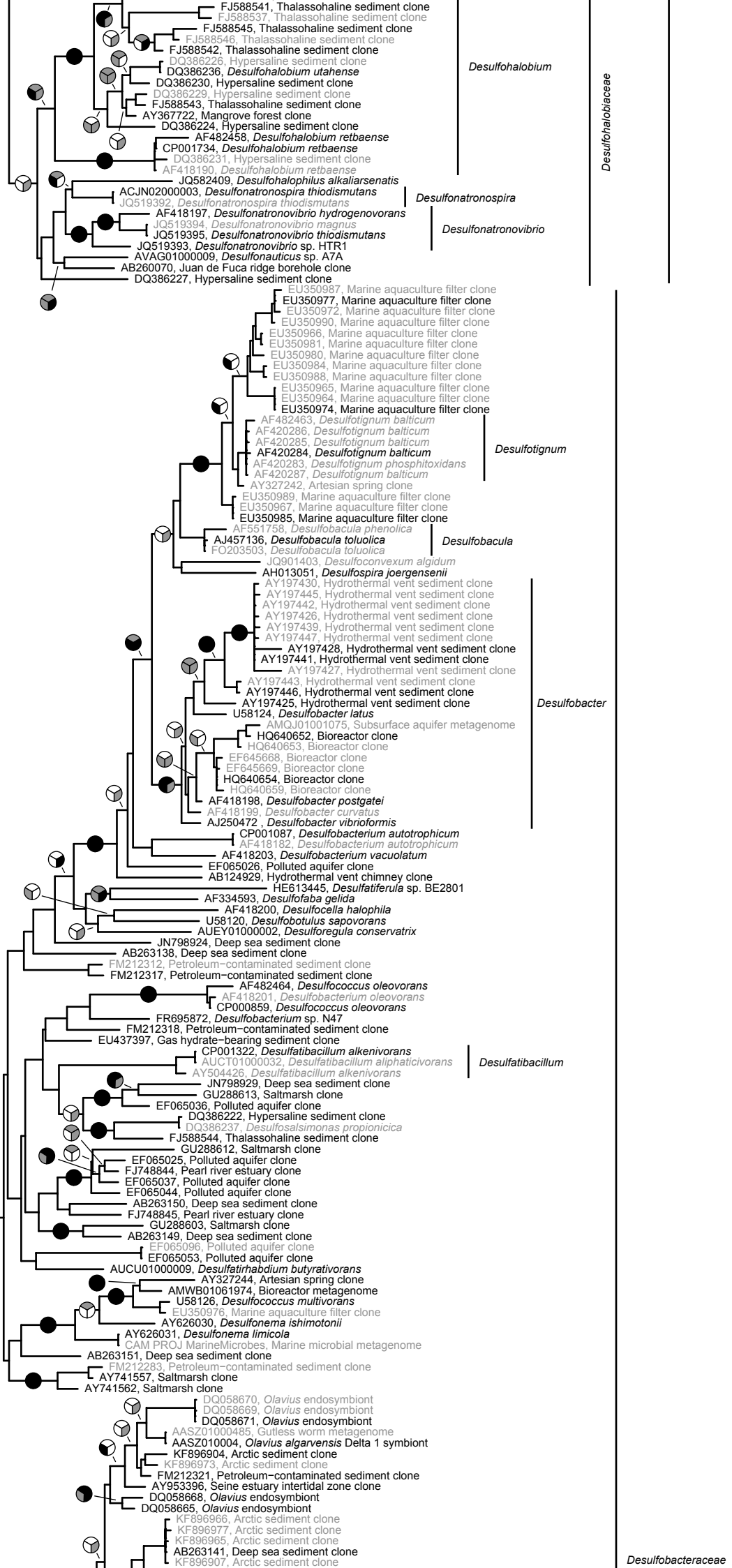
10%

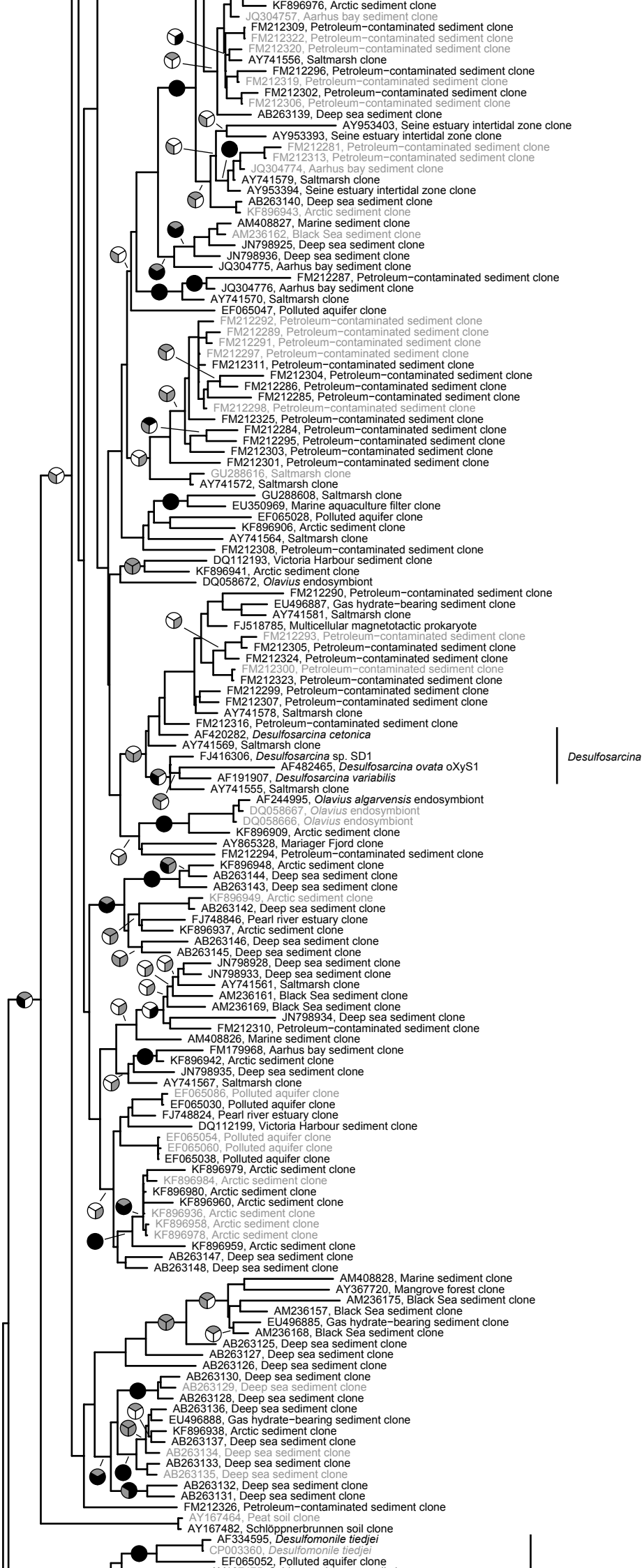


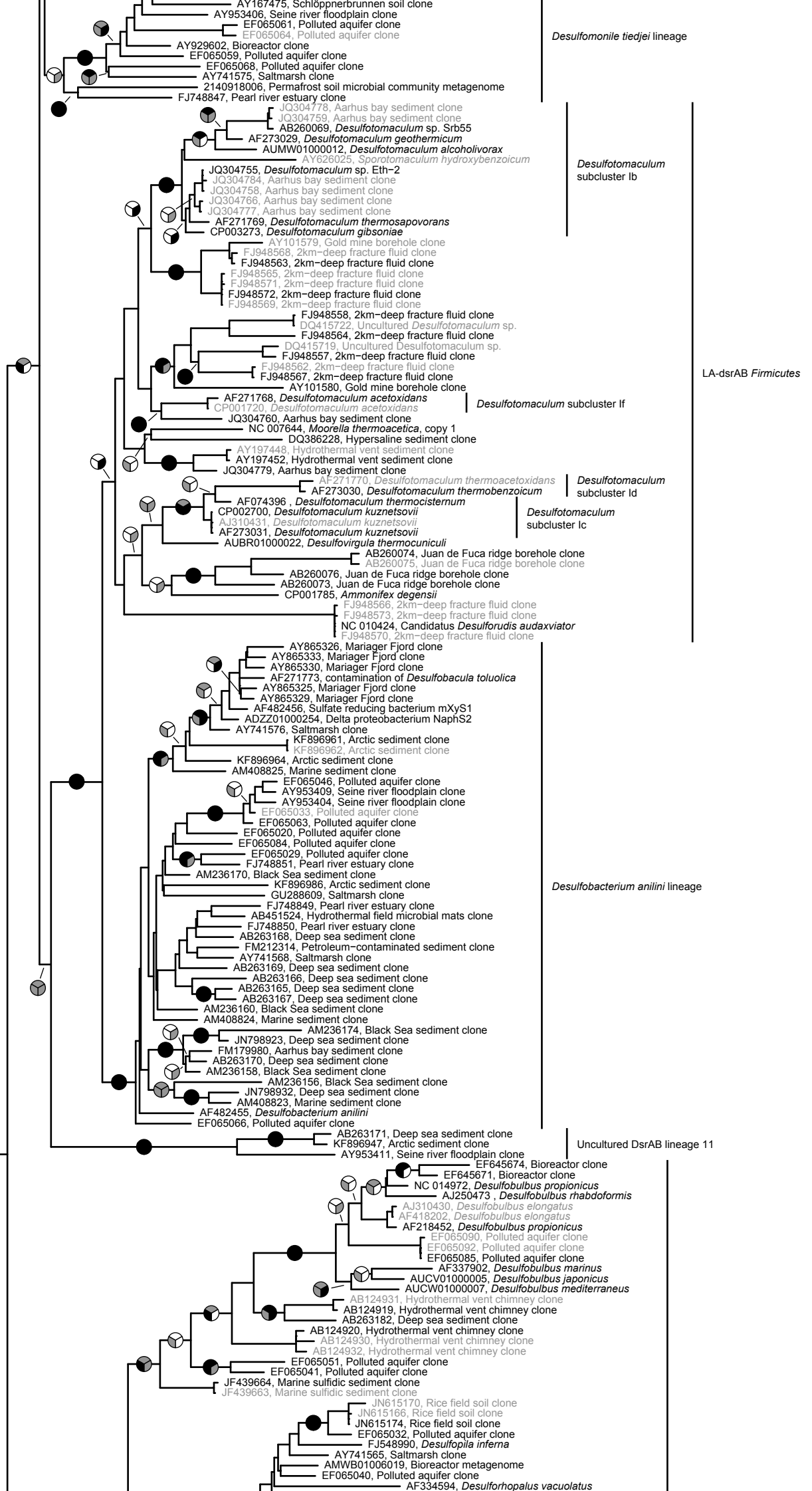
Desulfobacteriales

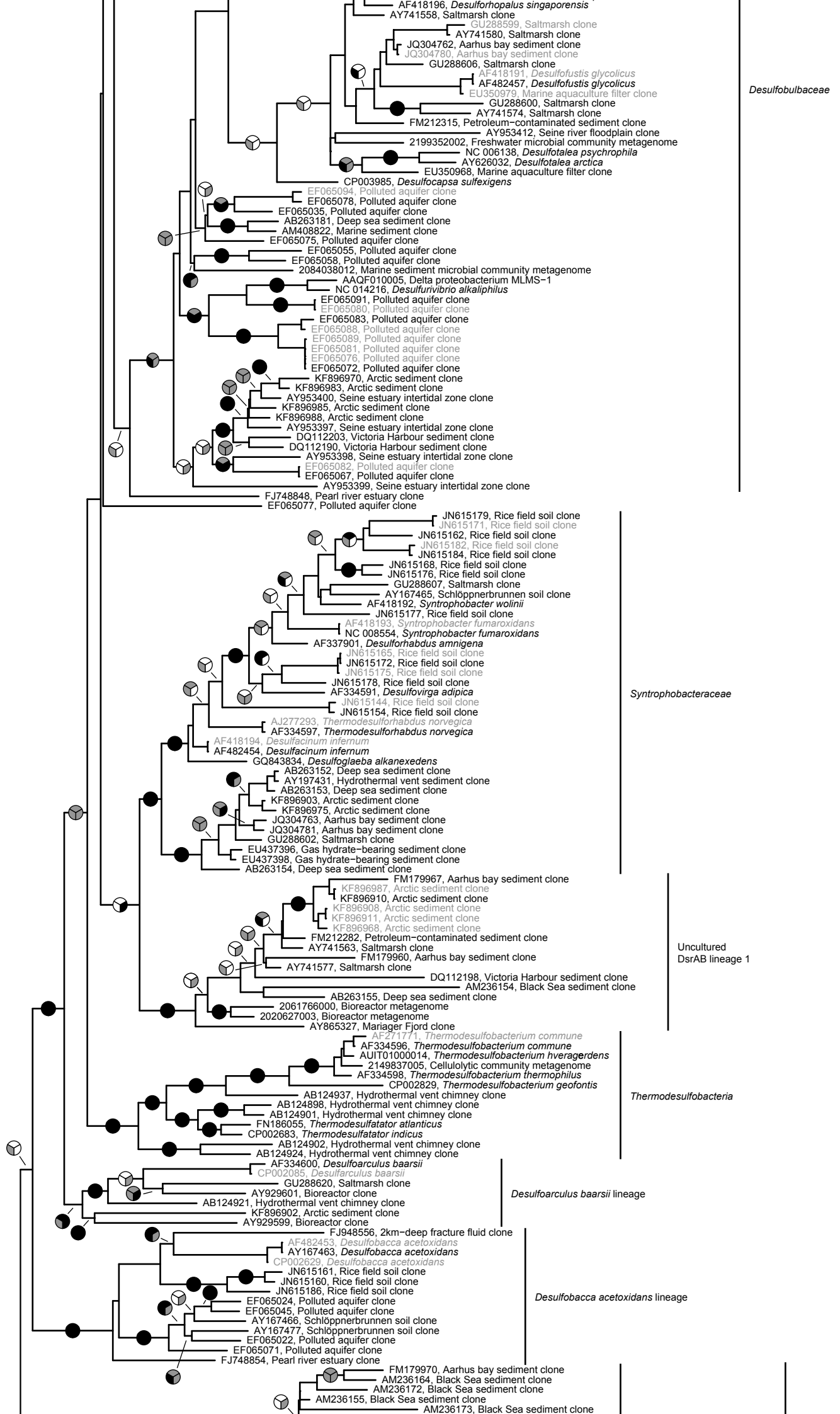
Desulfomicrobiaceae

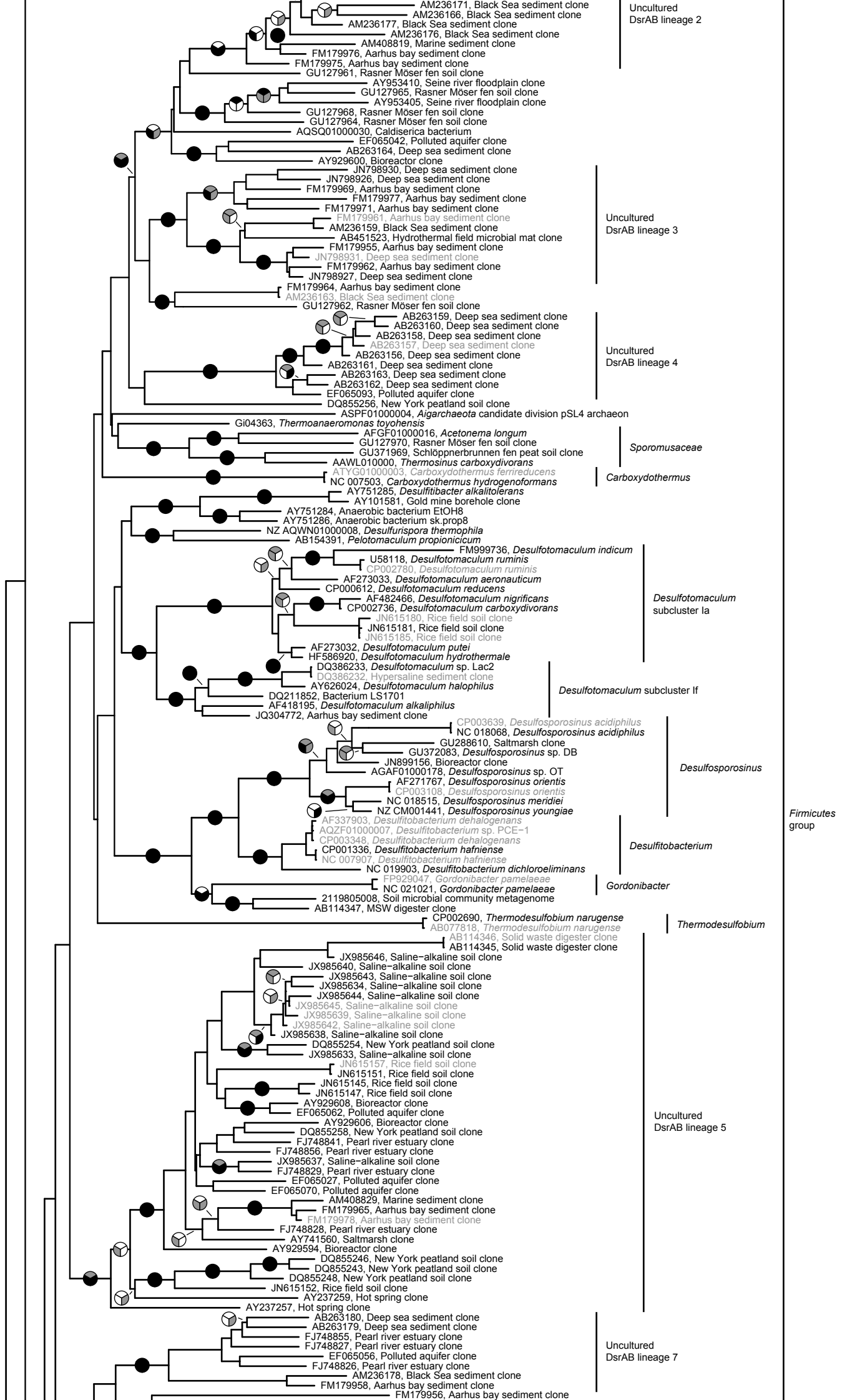
Desulfonatronumaceae

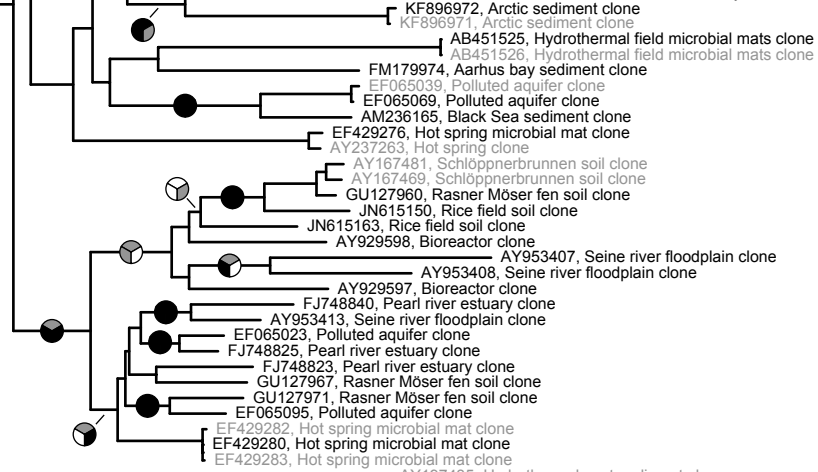




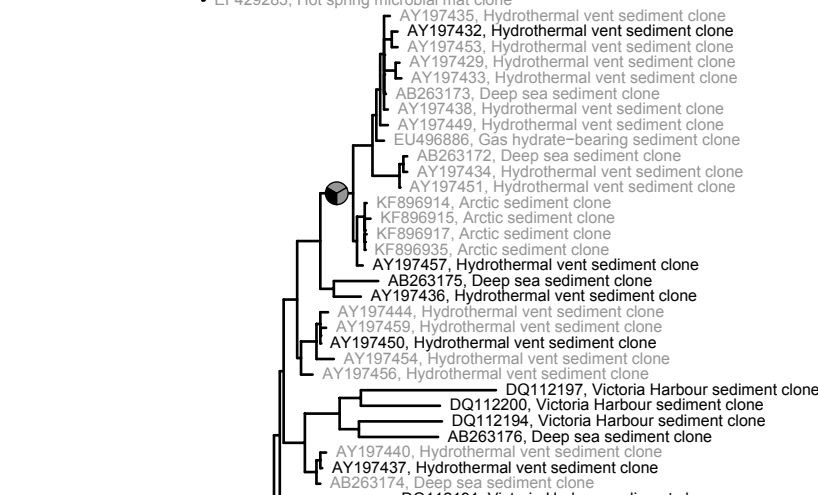




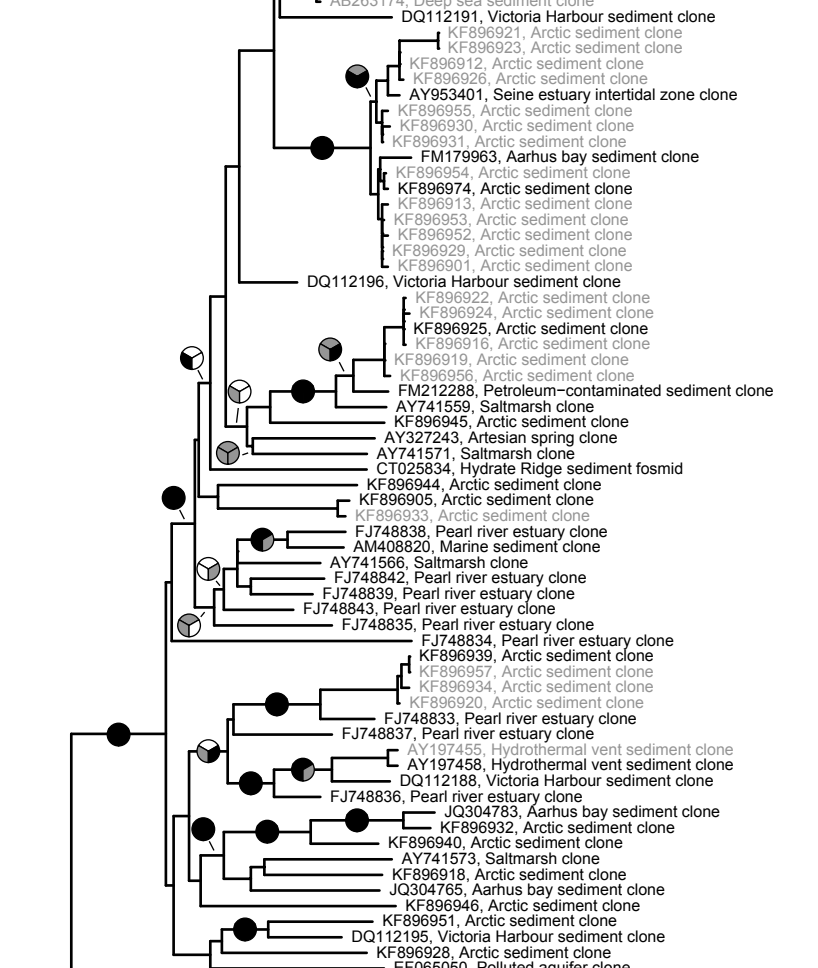




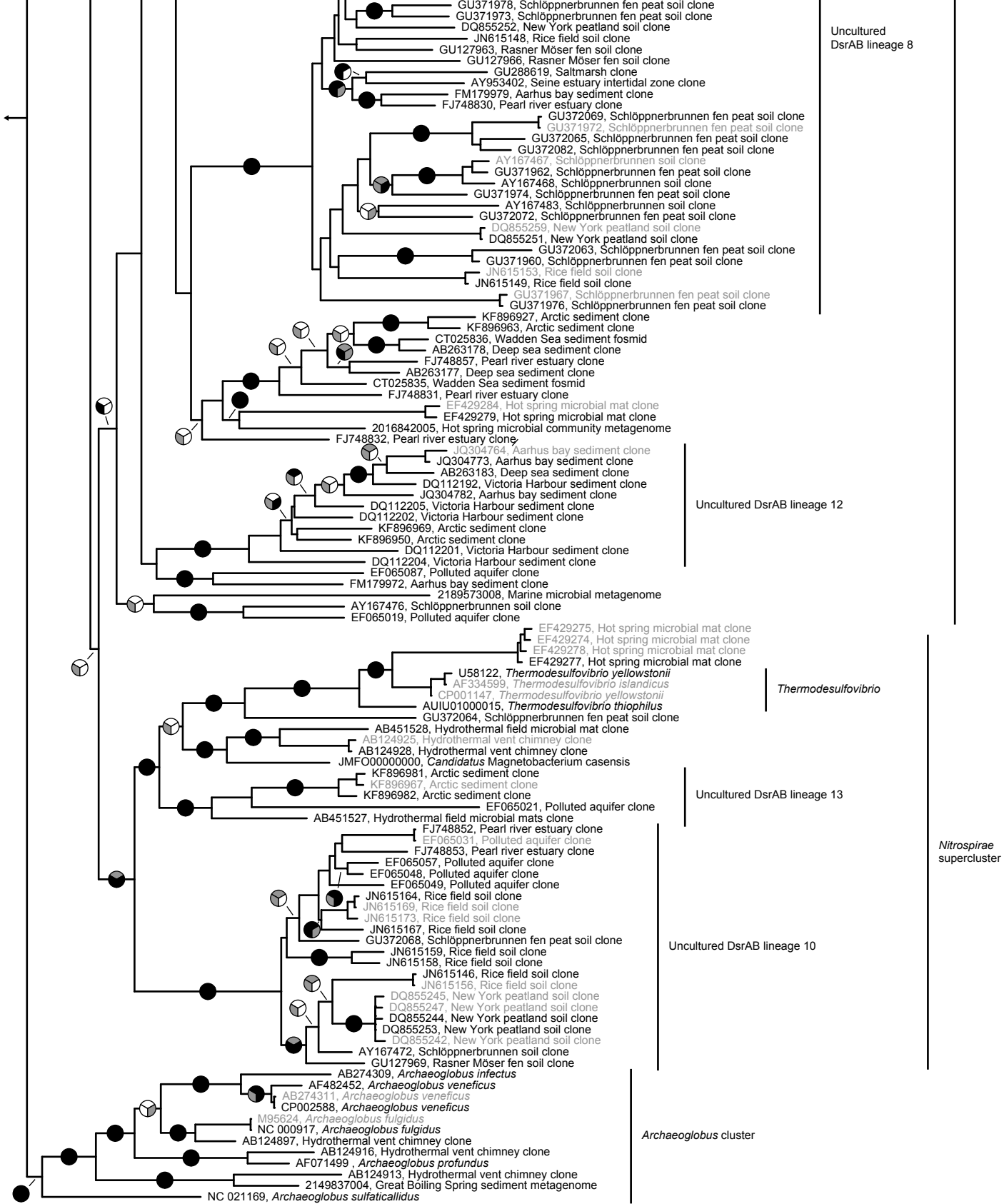
Uncultured
DsrAB lineage 6



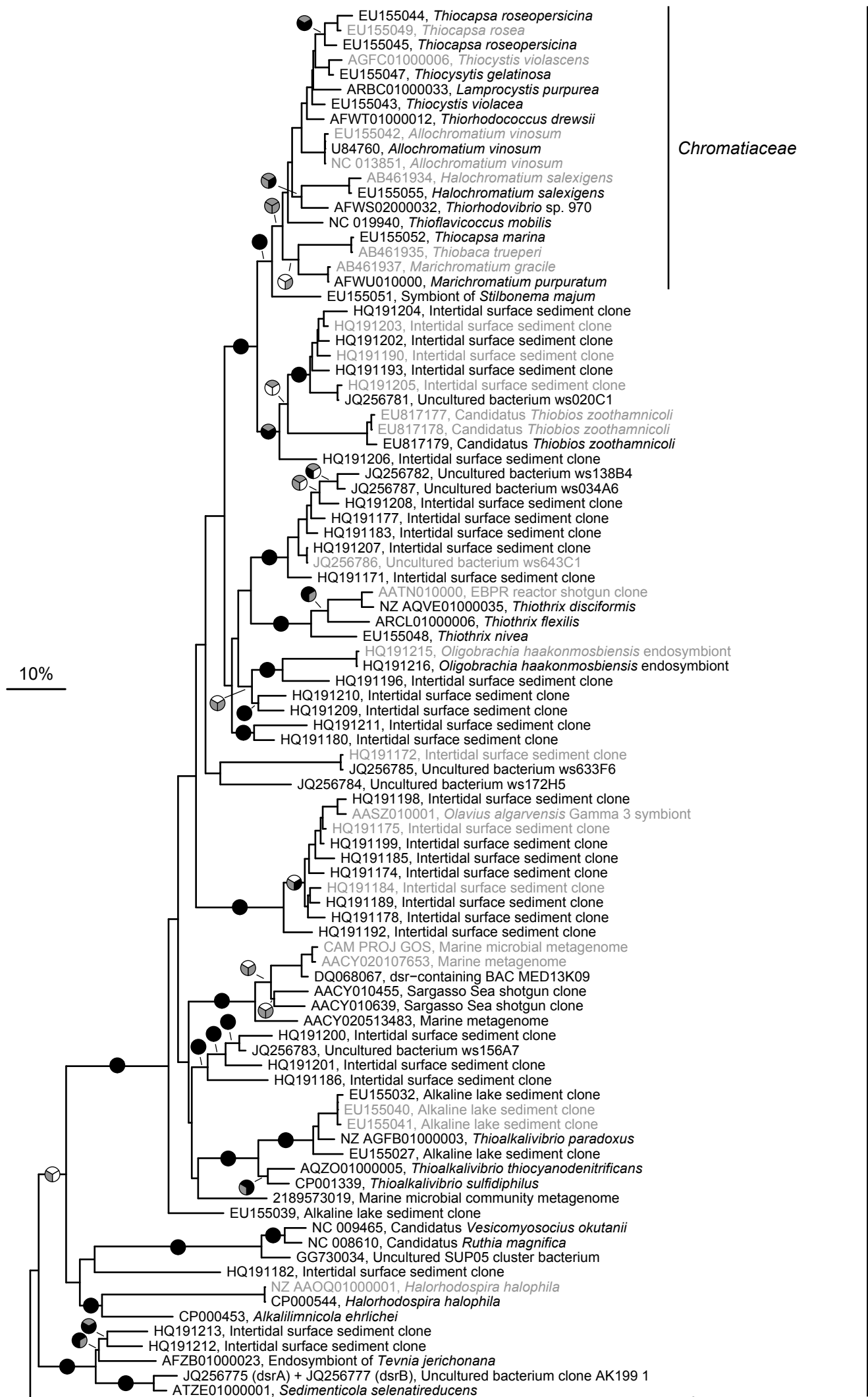
Uncultured
DsrAB lineage 9

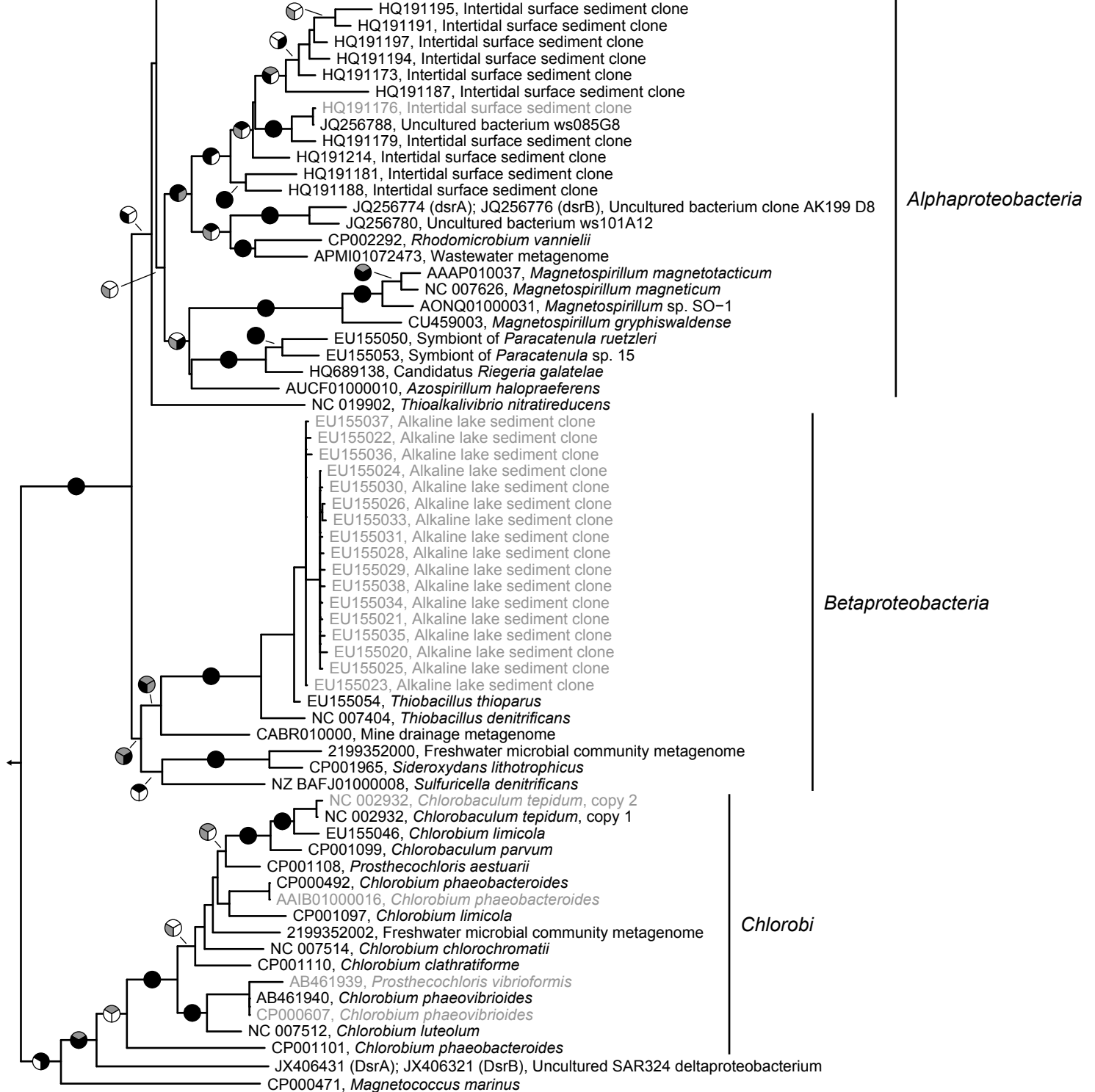


Environmental
supercluster 1

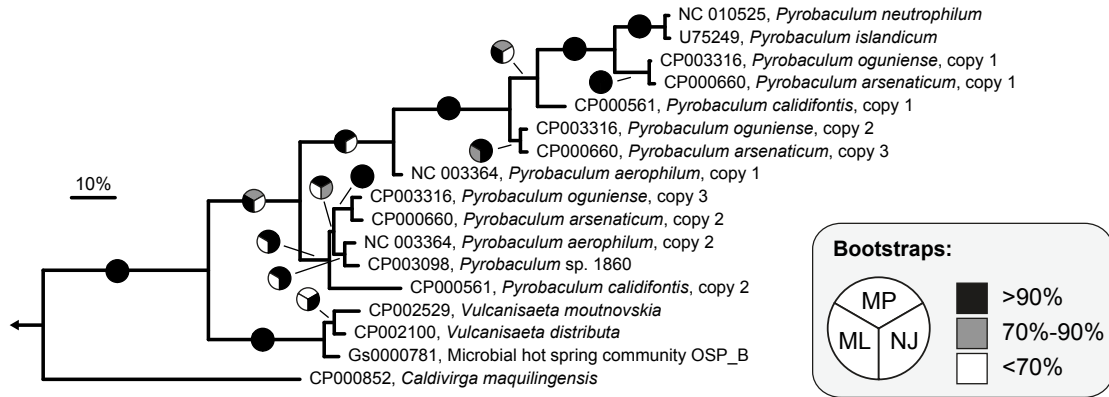


Supplementary Figure S1. Consensus phylogeny of reductive bacterial type DsrAB sequences. This figure shows an un-collapsed version of the consensus tree shown in Figure 1 depicting only reductive bacterial type DsrAB. Sequences in grey were subsequently added to the consensus tree without changing its topology. Scale bar indicates 10% sequence divergence.

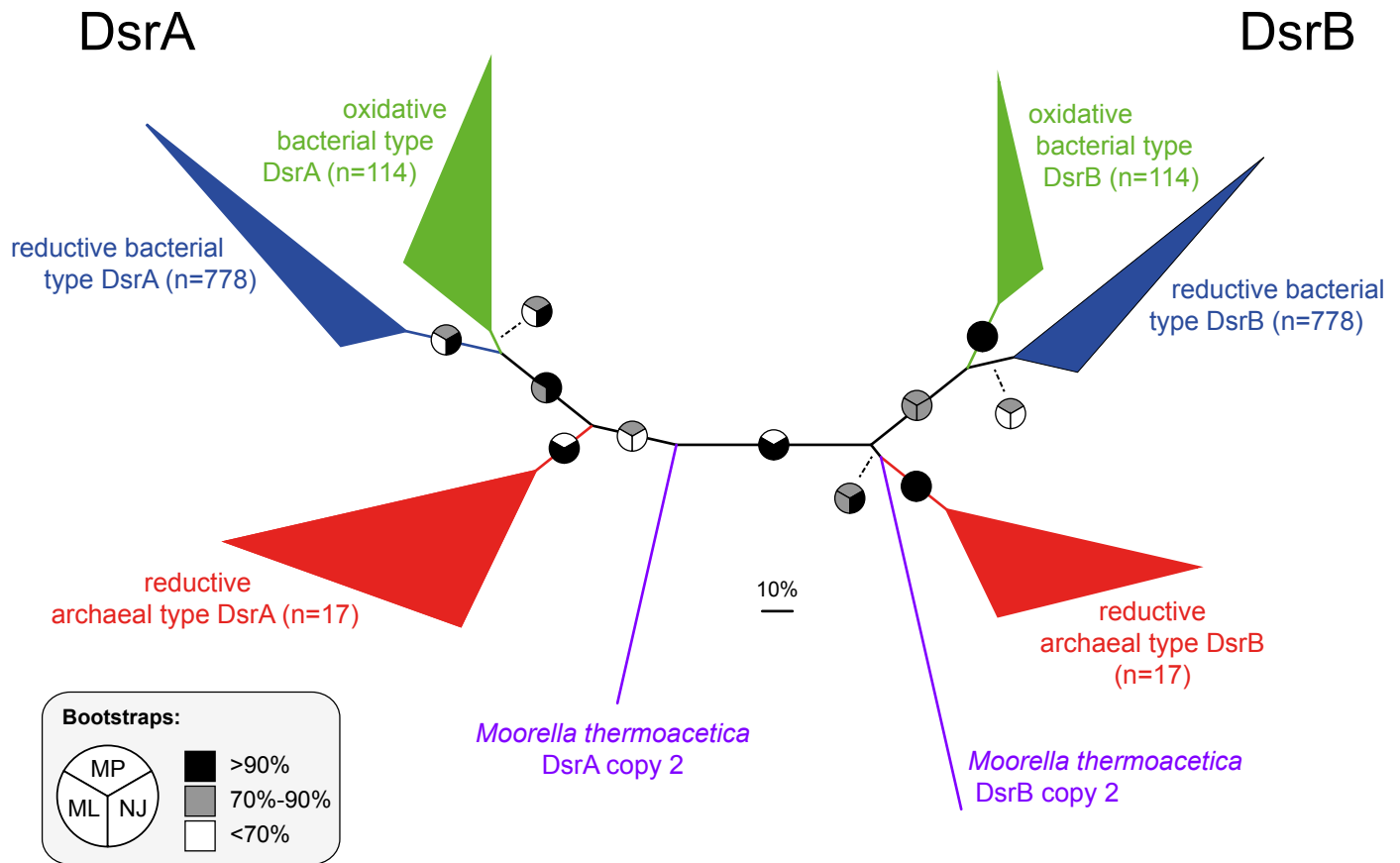




Supplementary Figure S2. Consensus phylogeny of oxidative bacterial type DsrAB sequences. Trees for reconstruction of the consensus tree (extended majority rule) were calculated using an amino acid alignment of 115 representative oxidative DsrAB sequences (clustered at 97% amino acid identity) and a filter covering 552 amino acid positions (omitting insertions/deletions). The tree was rooted with reductive bacterial type DsrAB sequences as outgroup. Sequences in grey (n=45) were subsequently added to the consensus tree without changing its topology. Scale bar indicates 10% sequence divergence. Bootstrap support (1000 re-samplings) is shown by split circles (top: maximum parsimony, bottom left: maximum likelihood, bottom right: neighbor joining) at the respective branches; with black, grey, and white indicating $\geq 90\%$, 70%-90%, and $< 70\%$ support, respectively.

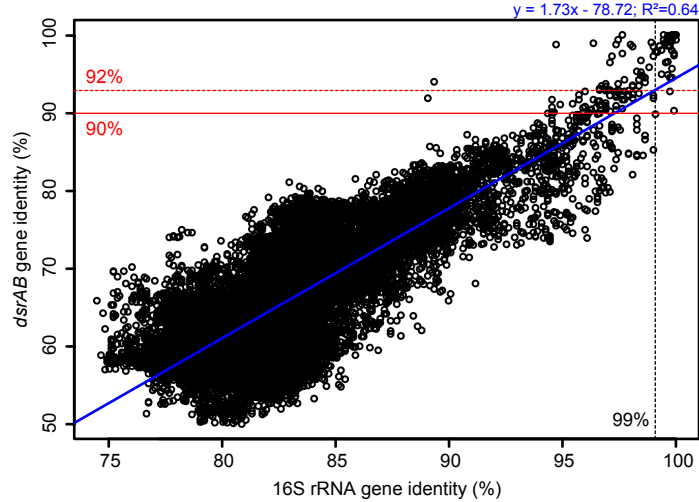


Supplementary Figure S3. Consensus phylogeny of reductive archaeal type DsrAB sequences. Trees for reconstruction of the consensus tree (extended majority rule) were calculated using an amino acid alignment of 17 archaeal type DsrAB sequences and a filter covering 629 amino acid positions (omitting insertions/deletions). The tree was rooted with bacterial DsrAB sequences as outgroup. Scale bar indicates 10% sequence divergence. Bootstrap support (1000 re-samplings) is shown by split circles (top: maximum parsimony, bottom left: maximum likelihood, bottom right: neighbor joining) at the respective branches; with black, grey, and white indicating $\geq 90\%$, 70%-90%, and $< 70\%$ support, respectively.

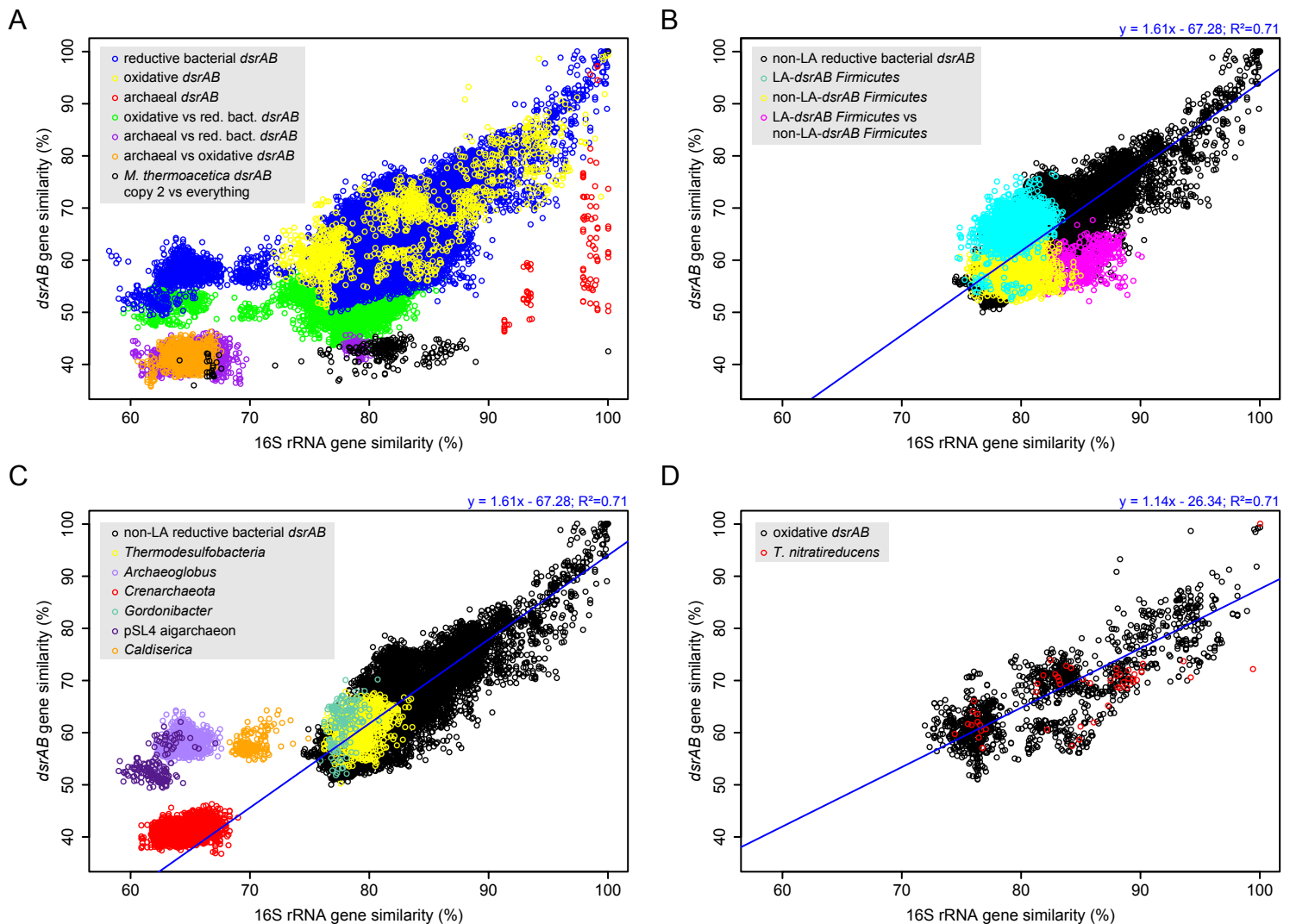


Supplementary Figure S4. Revealing the root in the DsrAB tree by paralogous rooting.

Trees for reconstruction of the consensus tree (extended majority rule) were calculated using an alignment of 910 DsrA to corresponding DsrB sequences and a filter covering 163 amino acid positions that are conserved and homologous between DsrA and DsrB subunits. Scale bar indicates 10% sequence divergence. Bootstrap support (100 re-samplings) is shown by split circles (top: maximum parsimony, bottom left: maximum likelihood, bottom right: neighbor joining) at the respective branches; with black, grey, and white indicating $\geq 90\%$, 70%-90%, and $< 70\%$ support, respectively. Clusters of the four, most basal DsrAB branches are collapsed and shown in different colors.

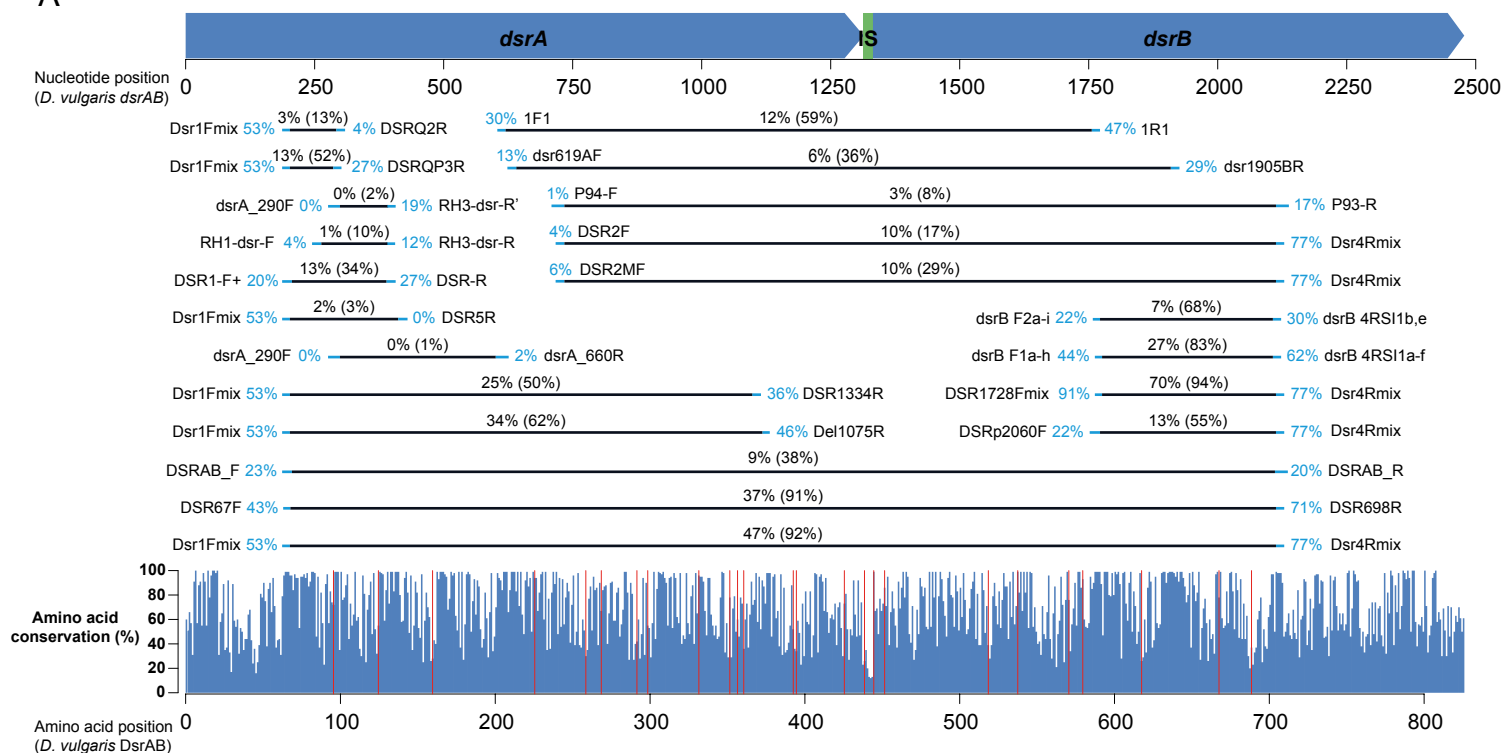


Supplementary Figure S5. Species-level threshold for *dsrAB* nucleotide identity inferred from a gene identity plot of non-laterally acquired bacterial *dsrAB*-carrying organisms using 99% nucleotide identity on 16S rRNA level as threshold for species level.

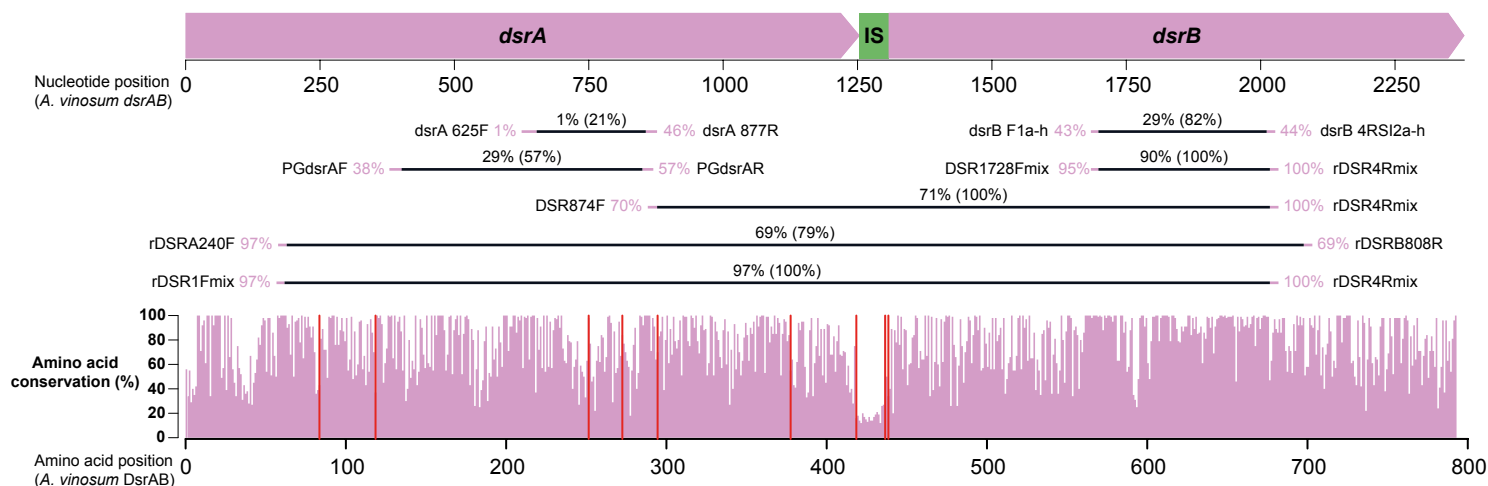


Supplementary Figure S6. Gene identity plot of *dsrAB* against 16S rRNA. Sequence identities of *dsrAB* and 16S rRNA for pairs of organisms for which both genes are known (mostly pure cultures and genomes) were plotted against each other. A) Pairwise comparisons between 259 corresponding *dsrAB*-16S rRNA pairs. Comparisons within and among the three main DsrAB enzyme families as well as comparisons with *M. thermoacetica dsrAB* copy 2 are highlighted in different colors. B) Comparisons of *Firmicutes* with laterally acquired and vertically inherited *dsrAB* with all organisms that have vertically inherited *dsrAB*. Linear regression of comparisons between organisms with non-laterally acquired reductive bacterial type *dsrAB* is shown for reference. C) Comparisons between organisms with laterally acquired *dsrAB* and organisms with vertically inherited *dsrAB*. Linear regression of comparisons between organisms with non-laterally acquired reductive bacterial type *dsrAB* is shown for reference. D) Comparisons of *T. nitratireducens* with reverse *dsrAB*-carrying organisms. Linear regression of comparisons between organisms with oxidative type *dsrAB* is shown for reference.

A



B



Supplementary Figure S7. Binding sites and *in silico* coverages of reductive (A) and oxidative (B) bacterial type *dsrAB*-targeted primers. Perfect-match coverage values are displayed for individual primers (blue or pink) and primer pairs (black), coverage values using one weighted mismatch are additionally indicated in parentheses for primer pairs. Displayed coverage values for primers binding at the target sites or outside the amplification region of (r)DSR1F/(r)DSR4R are based on full length datasets of reductive (n=115) and oxidative (n=62) bacterial type *dsrAB* sequences, whereas coverage values for primers binding within the region amplified by (r)DSR1F/(r)DSR4R are based on a core dataset of reductive (n=1110) and oxidative (n=159) bacterial type *dsrAB* sequences that completely cover this region. Bar charts show positional amino acid sequence variability (calculated in ARB using the ARB_PHYLO 'filter by base frequency' method and indicated in blue or pink bars) and regions with deletions/insertions (red bars) in DsrAB. Sequence numbering and primer binding positions correspond to *dsrAB*/DsrAB of *Desulfovibrio vulgaris* (NC_002937, 449880..452365) or *Allochromatium vinosum* (NC_013851, 1439735..1442113).

Supplementary Table S2

Name	Sequence (5'-3')	Target gene	Position ^d	Deg. ^e	full length <i>dsrAB</i> ^b (%)		core dataset <i>dsrAB</i> ^c (%)		Reference
					0 MM	1 wMM	0 MM	1 wMM	
rDSR1Fmix		<i>dsrA</i>	169-184	80	97	100	<i>n.a.</i>	<i>n.a.</i>	Loy <i>et al.</i> , 2009
rDSR1Fa	AARGGNTAYTGGGAARG	<i>dsrA</i>	169-184	32	69	98	<i>n.a.</i>	<i>n.a.</i>	Loy <i>et al.</i> , 2009
rDSR1Fb	TTYGGNTAYTGGGAARG	<i>dsrA</i>	169-184	32	0	69	<i>n.a.</i>	<i>n.a.</i>	Loy <i>et al.</i> , 2009
rDSR1Fc	ATGGGNTAYTGGGAARG	<i>dsrA</i>	169-184	16	27	71	<i>n.a.</i>	<i>n.a.</i>	Loy <i>et al.</i> , 2009
rDSRA240F	GGNTAYTGGGAARGGNGG	<i>dsrA</i>	172-188	64	97	100	<i>n.a.</i>	<i>n.a.</i>	Lenk <i>et al.</i> , 2011
PGdsrAF	CAYGBCAGACCGGBRAYATYATG	<i>dsrA</i>	379-402	144	39	56	38	61	Mori <i>et al.</i> , 2010
dsrA 625F	TTCAGTTCTCCGGCTGCSNAAAYGACTG	<i>dsrA</i>	625-653	16	2	26	1	21	Luo <i>et al.</i> , 2011
PGdsrAR	RCAGTGCATRCACKGHCACRCA	<i>dsrA</i>	850-870	48	60	81	57	89	Mori <i>et al.</i> , 2010
dsrA 877R	CGTTSANRCAGTGCAATGACGCG	<i>dsrA</i>	856-877	16	39	79	46	86	Luo <i>et al.</i> , 2011
DSR874F	TGYATGCAYTYGYTVAAYG	<i>dsrA</i>	859-877	96	71	100	70	100	Loy <i>et al.</i> , 2009
DSR1728Fmix		<i>dsrB</i>	1684-1698	77	90	100	95	99	Steger <i>et al.</i> , 2011
DSR1728FmixA	CAYACCCAGGGNTGG	<i>dsrB</i>	1684-1698	8	56	74	65	82	Steger <i>et al.</i> , 2011
DSR1728FmixB	CAYACBCAAGGNTGG	<i>dsrB</i>	1684-1698	24	18	97	14	98	Steger <i>et al.</i> , 2011
DSR1728FmixC	CATACDCAGGGHTGG	<i>dsrB</i>	1684-1698	9	6	27	6	28	Steger <i>et al.</i> , 2011
DSR1728FmixD	CACACDCAGGGNTGG	<i>dsrB</i>	1684-1698	12	10	63	9	68	Steger <i>et al.</i> , 2011
DSR1728FmixE	CATACHCAGGGNTAY	<i>dsrB</i>	1684-1698	24	0	74	0	82	Steger <i>et al.</i> , 2011
dsrB F1a-h		<i>dsrB</i>	1684-1698	8	47	90	43	89	Lever <i>et al.</i> , 2013
dsrB F1a	CACACCCAGGGCTGG	<i>dsrB</i>	1684-1698	1	35	63	32	70	Lever <i>et al.</i> , 2013
dsrB F1b	CATACTCAGGGCTGG	<i>dsrB</i>	1684-1698	1	0	18	1	13	Lever <i>et al.</i> , 2013
dsrB F1c	CATACCCAGGGCTGG	<i>dsrB</i>	1684-1698	1	10	53	9	61	Lever <i>et al.</i> , 2013
dsrB F1d	CACACTCAAGGTTGG	<i>dsrB</i>	1684-1698	1	0	21	1	11	Lever <i>et al.</i> , 2013
dsrB F1e	CACACACAGGGATGG	<i>dsrB</i>	1684-1698	1	0	11	0	13	Lever <i>et al.</i> , 2013
dsrB F1f	CACACGCAGGGATGG	<i>dsrB</i>	1684-1698	1	2	8	1	7	Lever <i>et al.</i> , 2013
dsrB F1g	CACACGCAGGGGTTGG	<i>dsrB</i>	1684-1698	1	0	6	0	5	Lever <i>et al.</i> , 2013
dsrB F1h	CATACGCAAGGTTGG	<i>dsrB</i>	1684-1698	1	0	8	0	5	Lever <i>et al.</i> , 2013
dsrB 4RSI2a-h		<i>dsrB</i>	2011-2027	1	44	90	<i>n.a.</i>	<i>n.a.</i>	Lever <i>et al.</i> , 2013
dsrB 4RSI2a	CAGGCGCCGAGCAGAT	<i>dsrB</i>	2011-2027	1	23	60	<i>n.a.</i>	<i>n.a.</i>	Lever <i>et al.</i> , 2013
dsrB 4RSI2b	CAGGCGCCGAGCAGCAC	<i>dsrB</i>	2011-2027	1	8	15	<i>n.a.</i>	<i>n.a.</i>	Lever <i>et al.</i> , 2013
dsrB 4RSI2c	CATGCTCCGAGCAGAT	<i>dsrB</i>	2011-2027	1	2	13	<i>n.a.</i>	<i>n.a.</i>	Lever <i>et al.</i> , 2013
dsrB 4RSI2d	CACGCGCCGCAAGCCAC	<i>dsrB</i>	2011-2027	1	3	3	<i>n.a.</i>	<i>n.a.</i>	Lever <i>et al.</i> , 2013
dsrB 4RSI2e	CATGCACACAACAAAT	<i>dsrB</i>	2011-2027	1	2	13	<i>n.a.</i>	<i>n.a.</i>	Lever <i>et al.</i> , 2013
dsrB 4RSI2f	CAGGCACACAGCAGAT	<i>dsrB</i>	2011-2027	1	0	18	<i>n.a.</i>	<i>n.a.</i>	Lever <i>et al.</i> , 2013
dsrB 4RSI2g	CAGGCTCCGAGCAGAT	<i>dsrB</i>	2011-2027	1	5	44	<i>n.a.</i>	<i>n.a.</i>	Lever <i>et al.</i> , 2013
dsrB 4RSI2h	CAGGCGCCGAGCAT	<i>dsrB</i>	2011-2027	1	2	42	<i>n.a.</i>	<i>n.a.</i>	Lever <i>et al.</i> , 2013
rDSR4Rmix		<i>dsrB</i>	2017-2033	96	100	100	<i>n.a.</i>	<i>n.a.</i>	Loy <i>et al.</i> , 2009
rDSR4Ra	CCRAARCAIGCNCCRCA	<i>dsrB</i>	2017-2033	32	23	50	<i>n.a.</i>	<i>n.a.</i>	Loy <i>et al.</i> , 2009
rDSR4Rb	GGRWARCAIGCNCCRCA	<i>dsrB</i>	2017-2033	64	77	100	<i>n.a.</i>	<i>n.a.</i>	Loy <i>et al.</i> , 2009
rDSRB808R	CCDCNACCCADATNGC	<i>dsrB</i>	2011-2027	144	69	79	<i>n.a.</i>	<i>n.a.</i>	Lenk <i>et al.</i> , 2011

^a Recommended primers are highlighted in gray.

^b Data indicate primer coverage of all full length oxidative bacterial type *dsrAB* sequences (n=62); 0 MM, no mismatches; 1 wMM, one weighted mismatch.

^c Data indicate primer coverage of core dataset oxidative bacterial type *dsrAB* sequences (n=159); 0 MM, no mismatches; 1 wMM, one weighted mismatch; *n.a.*, not applicable for primers binding at the target sites or outside the amplification region of rDSR1F/rDSR4R.

^d Position is relative to *Allochroamium vinosum dsrAB* (NC_013851, 1439735..1442113).

^e Degeneracy is given as the number of oligonucleotides that comprise the primer.

Supplementary Table S3

Primer pairs targeting reductive bacterial type <i>dsrAB</i> ^a				full length <i>dsrAB</i> ^b (%)		core dataset <i>dsrAB</i> ^c (%)		Phylogenetic clusters ^d (%)					Reference
Primers	Target gene	Position ^e	Length ^f	0 MM	1 wMM	0 MM	1 wMM	DpS	Fg	ES1	NsS	AgC	
DSR1Fmix/DSR1334R	<i>dsrA</i>	187-1115	929	25	50	n.a.	n.a.	39	0	0	0	0	Santillano <i>et al.</i> , 2010
DSR1Fmix/Del1075R	<i>dsrA</i>	187-1132	946	34	62	n.a.	n.a.	51	0	20	0	0	Gittel <i>et al.</i> , 2009
DSR1Fmix/DSR4Rmix	<i>dsrAB</i>	187-2129	1943	47	92	n.a.	n.a.	39	7	0	0	40	Pester <i>et al.</i> , 2010
DSRAB_F/DSRAB_R	<i>dsrAB</i>	187-2136	1950	9	38	n.a.	n.a.	13	0	0	0	0	Schmalenberger <i>et al.</i> , 2007
DSR1Fmix/DSRQP3R	<i>dsrA</i>	187-302	116	13	52	n.a.	n.a.	19	0	20	0	0	Akob <i>et al.</i> , 2012
DSR1Fmix/DSRQ2R	<i>dsrA</i>	187-309	123	3	13	n.a.	n.a.	5	0	0	0	0	Chin <i>et al.</i> , 2007
DSR1-F+/DSR-R	<i>dsrA</i>	187-407	221	13	34	n.a.	n.a.	20	0	0	0	0	Kondo <i>et al.</i> , 2004
DSR1Fmix/DSR5R	<i>dsrA</i>	187-430	244	2	3	n.a.	n.a.	3	0	0	0	0	Wagner <i>et al.</i> , 1998
DSR67F/DSR698R	<i>dsrAB</i>	189-2129	1941	37	91	n.a.	n.a.	45	18	20	0	40	Suzuki <i>et al.</i> , 2005
RH1-dsr-F/RH3-dsr-R	<i>dsrAB</i>	245-408	164	3	18	1	10	2	0	0	0	0	Ben-Dov <i>et al.</i> , 2007
dsr_290F/RH3-dsr-R'	<i>dsrAB</i>	276-408	133	0	1	0	2	0	0	0	0	0	Pereyra <i>et al.</i> , 2010
dsr_290F/dsrA_660R	<i>dsrAB</i>	276-626	351	0	0	0	1	0	0	0	0	0	Pereyra <i>et al.</i> , 2010
1F1/1R1	<i>dsrAB</i>	604-1778	1175	13	47	12	59	34	39	14	0	0	Dhillon <i>et al.</i> , 2003
dsr619AF/dsr1905BR	<i>dsrAB</i>	623-1926	1304	10	36	6	36	8	4	0	3	0	Giloteaux <i>et al.</i> , 2010
P94-F/P93-R	<i>dsrAB</i>	709-2138	1430	3	8	n.a.	n.a.	4	0	0	0	20	Karkhoff-Schweizer <i>et al.</i> , 1995
DSR2MF/DSR4Rmix	<i>dsrAB</i>	717-2129	1413	10	29	n.a.	n.a.	15	0	0	0	20	Akob <i>et al.</i> , 2012
DSR2F/DSR4Rmix	<i>dsrAB</i>	717-2129	1413	10	17	n.a.	n.a.	13	0	0	0	20	Wagner <i>et al.</i> , 1998
DSRp2060F/DSR4Rmix	<i>dsrB</i>	1752-2129	378	13	55	n.a.	n.a.	20	0	0	0	0	Geets <i>et al.</i> , 2005
dsrB F2a-i/4RSI1b,e	<i>dsrB</i>	1758-2123	366	7	68	n.a.	n.a.	11	0	0	0	0	Lever <i>et al.</i> , 2013
dsrB F1a-h/4RSI1a-f	<i>dsrB</i>	1762-2123	362	27	83	n.a.	n.a.	36	14	0	0	0	Lever <i>et al.</i> , 2013
DSR1728mix/DSR4Rmix	<i>dsrB</i>	1762-2129	368	70	94	n.a.	n.a.	72	64	60	0	100	Steger <i>et al.</i> , 2011

^a Coverage values for primer pairs published using DSR1F or DSR4R are calculated using the most up to date version of the primer mix. Recommended primer pairs are highlighted in gray.

^b Data indicate primer coverage of all full length reductive bacterial type *dsrAB* sequences (n=115); 0 MM, no mismatches; 1 wMM, one weighted mismatch.

^c Data indicate primer coverage of all core reductive bacterial type *dsrAB* sequences (n=1110); 0 MM, no mismatches; 1 wMM, one weighted mismatch; n.a., not applicable for primers binding at the target sites or outside the amplification region of DSR1F/DSR4R.

^d Primer coverage of full length/core dataset *dsrAB* sequences in higher taxonomic clusters. DpS, *Deltaproteobacteria* supercluster (n=75/709); Fg, *Firmicutes* group (n=28/180); ES1, Environmental supercluster 1 (n=5/170); NsS, Nitrospirae supercluster (n=2/39); AgC, *Archaeoglobus* cluster (n=5/12). The larger core dataset is used for primers binding within the DSR1F/DSR4R region.

^e Position is relative to *Desulfovibrio vulgaris* Hildenborough *dsrAB* (NC_002937, 449888..452365).

^f Expected length of the PCR amplicon.

Supplementary Table S4

Primer pairs targeting oxidative bacterial type <i>dsrAB</i> ^a				full length <i>dsrAB</i> ^b (%)		core dataset <i>dsrAB</i> ^c (%)		Reference
Name	Target gene	Position ^d	Length ^e	0 MM	1 wMM	0 MM	1 wMM	
rDSR1Fmix/rDSR4Rmix	<i>dsrAB</i>	169-2033	1865	97	100	n.a.	n.a.	Loy <i>et al.</i> , 2009
rDSRA240F/rDSRB808R	<i>dsrAB</i>	172-2027	1856	69	79	n.a.	n.a.	Lenk <i>et al.</i> , 2011
PGdsrAF/PGdsrAR	<i>dsrA</i>	379-870	492	32	53	29	57	Mori <i>et al.</i> , 2010
dsrA 625F/dsrA 877R	<i>dsrA</i>	625-877	253	2	26	1	21	Luo <i>et al.</i> , 2011
DSR874F/rDSR4Rmix	<i>dsrAB</i>	859-2033	1175	71	100	n.a.	n.a.	Loy <i>et al.</i> , 2009
dsrB F1a-h/4RSI2a-h	<i>dsrB</i>	1684-2027	344	29	82	n.a.	n.a.	Lever <i>et al.</i> , 2013
DSR1728mix/rDSR4Rmix	<i>dsrB</i>	1684-2033	350	90	100	n.a.	n.a.	Steger <i>et al.</i> , 2011

^a Recommended primer pairs are highlighted in gray.

^b Data indicate primer coverage of all full length reverse *dsrAB* sequences (n=62); 0 MM, no mismatches; 1 wMM, one weighted mismatch.

^c Data indicate primer coverage of core dataset reverse *dsrAB* sequences (n=159); 0 MM, no mismatches; 1 wMM, one weighted mismatch; n.a., not applicable for primers binding at the target sites or outside the amplification region of rDSR1F/rDSR4R.

^d Position is relative to *Allochrochromatium vinosum dsrAB* (NC_013851, 1439735..1442113).

^e Expected length of the PCR amplicon.