Rice paddy *Nitrospirae* encode and express genes related to sulfate respiration: proposal of the new genus *Candidatus Sulfobium*

Sarah Zecchin\(a,b\), Ralf C. Mueller\(a\), Jana Seifert\(c\), Ulrich Stingl\(d\), Karthik Anantharaman\(e\), Martin von Bergen\(f\), Lucia Cavalca\(b\), Michael Pester\(a,g\)

*contributed equally, #corresponding author

\(a\) Department of Biology, University of Konstanz, Konstanz, Germany; \(b\) Dipartimento di Scienze per gli Alimenti, la Nutrizione e l'Ambiente (DeFENS), Università degli Studi di Milano, Milano, Italy; \(c\) Institute of Animal Science, Hohenheim University, Stuttgart, Germany; \(d\) University of Florida, UF/IFAS, Department for Microbiology & Cell Science, Fort Lauderdale Research and Education Center, Davie, FL, USA; \(e\) Department of Earth and Planetary Science, University of California, Berkeley, CA, USA; \(f\) Helmholtz Centre for Environmental Research—UFZ, Department of Molecular Systems Biology, Leipzig, Germany; \(g\) Department Microorganisms, Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany

Short title: Rice paddy *Nitrospirae* encode sulfate respiration

Keywords: sulfate-reducing microorganisms, rice paddies, gypsum fertilization, *dsrAB* genes, *Nitrospirae*

Word count abstract: 250 words

Word count main text: 4960 words

#Corresponding author: Michael Pester, Department Microorganisms, Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures, D-38124 Braunschweig, Germany; phone: +49-531-2616-237; fax: +49-531-2616-418; e-mail: Michael.Pester@dsmz.de
Nitrospirae spp. distantly related to thermophilic, sulfate-reducing Thermodesulfovibrio species are regularly observed in environmental surveys of anoxic marine and freshwater habitats. Here, we present a metaproteogenomic analysis of Nitrospirae bacterium Nbg-4 as a representative of this clade. Its genome was assembled from replicated metagenomes of rice paddy soil that was used to grow rice in the presence and absence of gypsum (CaSO₄·2H₂O). Nbg-4 encoded the full pathway of dissimilatory sulfate reduction and showed expression thereof in gypsum-amended anoxic bulk soil as revealed by parallel metaproteomics. In addition, Nbg-4 encoded the full pathway of dissimilatory nitrate reduction to ammonia (DNRA) with expression of its first step being detected in bulk soil without gypsum amendment. The relative abundance of Nbg-4 was similar under both treatments indicating that it maintained stable populations while shifting its energy metabolism. Whether Nbg-4 is a strict sulfate reducer or can couple sulfur oxidation to DNRA by operating the pathway of dissimilatory sulfate reduction in reverse could not be resolved. Further genome reconstruction revealed the potential to utilize butyrate, formate, H₂, or acetate as electron donor, with the Wood-Ljungdahl pathway being expressed under both conditions. Comparison to publicly available Nitrospirae genome bins revealed the pathway for dissimilatory sulfate reduction also in related Nitrospirae recovered from groundwater. Subsequent phylogenomics showed that such microorganisms form a novel genus within the Nitrospirae, with Nbg-4 as a representative species. Based on the widespread occurrence of this novel genus, we propose for Nbg-4 the name Candidatus Sulfobium mesophilum, gen. nov., spec. nov.

Rice paddies are indispensable for food supply but are a major source of the greenhouse gas methane. If not counterbalanced by cryptic sulfur cycling, methane emission from rice paddy fields would be even higher. However, the microorganisms involved in this sulfur cycling are little understood. By using an environmental systems biology approach of Italian rice paddy soil, we
could retrieve the population genome of a novel member of the phylum *Nitrospirae*. This microorganism encoded the full pathway of dissimilatory sulfate reduction and expressed it in anoxic paddy soil under sulfate-enriched conditions. Phylogenomics and comparison to environmental surveys showed that such microorganisms are actually widespread in freshwater and marine environments. At the same time, they represent a yet undiscovered genus within the little explored *Nitrospirae*. Our results will be important to design enrichment strategies and postgenomic studies to fully understand the contribution of these novel *Nitrospirae* to the global sulfur cycle.

**Introduction**

Sulfate reducing microorganisms (SRM) are regularly observed in rice paddy fields (1-8). Despite the prevailing low sulfate concentrations in this habitat (lower μM-range, 9, 10), the rice rhizosphere and bulk soil are characterized by high sulfate reduction rates, which are comparable to marine surface sediments (11). This observation is explained by a cryptic sulfur cycle. Here, the small sulfate pool is rapidly reduced to sulfide but the latter also rapidly re-oxidized to sulfate thus keeping a highly active sulfur cycle running (10-13). This cryptic sulfur cycle can occur at oxic-anoxic interfaces such as rice roots but apparently runs also in the completely anoxic bulk soil (10). Under the latter conditions, reduced sulfur species may be re-oxidized with the help of iron minerals or redox-active parts of humic material such as quinone moieties as shown for other freshwater habitats (14-16).

The ability to perform dissimilatory sulfate reduction is most widespread among members of the *Deltaproteobacteria* and *Firmicutes* (17). Additional and exclusively thermophilic sulfate reducers are affiliated with the archaeal phyla *Euryarchaeota* and *Crenarchaeota* and the bacterial phyla *Thermodesulfobacteria* and *Nitrospirae* (17, 18). The only known SRM in the phylum *Nitrospirae* are bacteria belonging to the genus *Thermodesulfovibrio* (19-23). All described species of this genus are thermophilic with their common metabolic properties comprising the reduction of sulfate, thiosulfate and in some cases sulfite with a limited range of electron donors. These include pyruvate
and lactate, which are incompletely oxidized to acetate, or H₂ and formate in a background of acetate as auxiliary carbon source. The inability for autotrophic growth and the incomplete oxidation of organic substrates to acetate is a characteristic feature of this genus. Alternative electron acceptors used by Thermodesulfovibrio spp. are Fe(III) and in the case of Thermodesulfovibrio islandicus DSM 12570 nitrate (19-23).

In addition to the genus Thermodesulfovibrio, the phylum Nitrospirae currently encompasses the genera Nitrospira and Leptospirillum (24). Nitrospira spp. are known to have a versatile metabolism ranging from chemolithoautotrophic ammonia, nitrite, or hydrogen oxidation coupled to oxygen respiration to formate-driven nitrate respiration to nitrite (reviewed in 25). Leptospirillum spp. are described as iron oxidizers (24). A group of still uncultured Nitrospirae, which form a sister clade to the genus Thermodesulfovibrio, is represented by magnetotactic bacteria belonging to the putative genera Candidatus Magnetobacterium (26-28), Candidatus Thermomagnetovibrio (29), Candidatus Magnetooovum (30, 31), and Candidatus Magnetominusculus (32). These microorganisms are typically encountered at the oxic-anoxic interface of sediments but were also enriched from water of hot springs (33). The observation of sulfur-rich inclusions in the cells of Ca. Magnetobacterium bavaricum (27), Ca. Magnetooovum chiemensis (31), and Ca. Magnetooovum mohavensis (30), the presence of sulfur metabolism genes in the genomes of the former two species (31), and their predominant occurrence at oxic-anoxic interfaces led to the hypothesis that these microorganisms could be involved in sulfur oxidation (27, 31, 33).

All SRM have the canonical pathway of dissimilatory sulfate reduction in common, which is an intracellular process that involves an eight-electron reduction of sulfate to sulfide. This pathway proceeds through the enzymes sulfate adenylyltransferase (Sat), adenylyl phosphosulfate reductase (Apr), dissimilatory sulfite reductase (Dsr), and the sulfide-releasing DsrC (34). In addition, the complexes QmoAB(C) and DsrMK(JOP) are important in transferring reducing equivalents towards the pathway of sulfate reduction (35). The only known exception to this rule are ANME-archaea
that anaerobically oxidize methane by a yet unresolved mechanism of sulfate reduction to zero-valent sulfur (36).

The two different subunits of the heterotetrameric dissimilatory sulfite reductase Dsr are encoded by the paralogous genes dsrA and dsrB, which are frequently used as functional phylogenetic markers for SRM (37). The phylogeny of reductive bacterial-type DsrAB is subdivided into the Deltaproteobacteria, Firmicutes, Environmental, and Nitrospirae superclusters (37). DsrAB sequences affiliated with the Nitrospirae supercluster have been predominantly found in freshwater and soil environments and to a smaller extent in marine, industrial, or hot-temperature habitats (37). Intriguingly, they have also been detected in Italian (10) and Chinese (4, 8) rice paddy soils, but the detailed phylogenetic affiliation of these dsrAB-carrying microorganisms and their possible involvement in rice paddy sulfur cycling remained unclear.

Here, the draft genome of a novel and putatively sulfate reducing species belonging to the phylum Nitrospirae has been obtained from a metagenome survey of rice paddy soil. We present its metabolic potential and phylogeny as reconstructed from its genome and compare this to Nitrospirae genome bins recently recovered from metagenome studies of groundwater habitats. To support our conclusions, we present protein expression patterns of this novel Nitrospirae species as inferred by a metaproteome analysis of rice paddy soil.

Results

A Nitrospirae genome from rice paddy soil

We used a metagenomics approach to identify novel microorganisms involved in rice paddy sulfur cycling. For this purpose, replicated metagenomes (Table S1) were sequenced from bulk and rhizosphere soils of rice plants, which were grown to their late vegetative phase either in gypsum-amended (CaSO₄·2H₂O) or un-amended (control) soils. In addition, metagenomes from freshly flooded and unplanted soils were analyzed. Among the 159 population genome bins that could be
retrieved, *Nitrospirae* genome bin Nbg-4 was outstanding because it encoded *dsrAB*, was of high quality with ≤2% residual contamination, showed no strain heterogeneity, and had an estimated genome completeness of 75% (Table 1). The relative abundance of Nbg-4 was highest in the bulk soils (averaging 17 RPKM; reads per kilobase of scaffold per million reads) and roughly three times lower in rhizosphere soils (Figure 1). A two-way analysis of variance (ANOVA) showed that soil compartment had a significant effect on the relative abundance of Nbg-4 ($F_{2,14}=36.16$, $p<0.001$), while gypsum amendment ($F_{1,14}=0.17$, $p=0.69$) and the interaction of soil compartment and gypsum amendment ($F_{1,12}=0.03$, $p=0.87$) remained insignificant. To estimate the index of replication (iRep, 38) of Nbg-4, single reads of metagenomic replicates were combined per soil habitat to achieve sufficient coverage. This analysis indicated that roughly three quarters of the population were replicating their genome in freshly flooded soils, while roughly one third replicated its genome in bulk soils after 58-59 days of incubation irrespective of gypsum treatment (Table 1). For rhizosphere soils, the coverage was insufficient to perform an iRep analysis.

Reconstruction of a dissimilatory sulfur metabolism

The complete pathway for dissimilatory sulfate reduction was recovered in Nbg-4 (Figure 2). Besides genes encoding Sat and the β-subunit of Apr, which catalyze the activation of sulfate and its concomitant reduction to sulfite, respectively, also genes for DsrAB and DsrC, which reduce sulfite further to sulfide could be detected. *aprA* was probably missing because of an assembly break in the scaffold after *aprB* (typically *aprA* is downstream of *aprB*). In addition, genes encoding the QmoABC and DsrMK complexes were detected, which couple quinol reduction to electron transfer to AprAB and DsrC, respectively. *Thermodesulfovibrio* spp. possess in addition to the module DsrMK also the module DsrJOP, which form together the membrane-bound electron-transferring complex DsrMKJOP (23, 35). Since *dsrMK* were located at the end of one scaffold in Nbg-4 and another scaffold started with a long fragment of *dsrP*, it is likely that also Nbg-4 encodes a complete DsrMKJOP complex. In addition, the presence of *dsrD* directly adjacent to *dsrAB* was
detected. DsrD is a small protein of putative regulatory function present in all sulfate reducers (39) with sporadic encounters in genomes of sulfide and sulfur-oxidizing bacteria (40). In addition, dsrN and dsrT as typical genes of the dsr operon in sulfate reducers and sulfur-oxidizing green sulfur bacteria (39, 41) and hppA, which codes for a membrane-bound and proton-translocating pyrophosphatase to pull, e.g., the energy-demanding reaction of Sat, were detected (Table S2). With the exception of a membrane-bound sulfide-quinone oxidoreductase (Sqr) (Table S2), no genes that are essential for the oxidative sulfur metabolism of chemolithotrophic sulfur oxidizers (41-43) were detected.

All soil samples that were used for metagenome sequencing were also analyzed for their metaproteome. In bulk soil treated with gypsum, a search against Nbg-4 encoded proteins identified peptides specific for Sat and DsrA as two essential components of the first and last step of sulfate reduction, respectively (Figure 2). Peptides specific for DsrA of Nbg-4 were also detected in rhizosphere soil treated with gypsum but with lower intensity (label free quantification (LFQ) value in bulk soil: $1.20 \times 10^8$, LFQ in rhizosphere soil: $6.05 \times 10^4$). In contrast, no peptides matching Nbg-4 sulfur metabolism proteins were detected in control treatments without gypsum, neither in the bulk soil nor in the rhizosphere (Table S2). The fragmented recovery of proteins involved in dissimilatory sulfate reduction is certainly a result of the low coverage of the proteome of a single microbial population in the background of the whole soil metaproteome.

Based on the recovery of the dissimilatory sulfate reduction pathway in Nbg-4, NCBI’s sequence repositories were searched for additional dsrAB-carrying Nitrospirae genome bins of high assembly quality. This analysis identified fourteen additional bins recovered from metagenomes: three from aquifer sediments (44), nine from aquifer groundwater (44), and two from a deep subsurface water (45) (Table S3). In-depth analysis of four bins that represent the three additional habitat types revealed not only the presence of dsrAB but also of the complete dsr operon including dsrC, dsrD, dsrN, dsrT, and dsrMKJOP, which were all in synteny to the respective genes of Nbg-4 (Figure. 3).
Only *Nitrospirae* bacterium CG1-02-44-142 recovered from deep subsurface water had an inversion of *dsrC*, *dsrT*, and *dsrMKJOP* on its genome. Interestingly, also all other components of the dissimilatory sulfate reduction pathway including *sat*, *aprBA*, *qmoABC*, and *hppA* were encoded on these *Nitrospirae* genome bins, either completely or partially depending on the assembly breaks of the respective scaffolds (Table 2).

Nitrate reduction as an alternative respiratory metabolism

Nbg-4 also encoded a full set of genes necessary for dissimilatory nitrate reduction to ammonia (DNRA) (Figure 2). DNRA is employed by members of the genera *Thermodesulfovibrio*, *Desulfovibrio*, *Desulfobulbus*, *Desulfobacterium*, and *Desulfotomaculum* as alternative respiratory pathway in the absence of sulfate (39). The first step of DNRA is the reduction of nitrate to nitrite. To perform this step, Nbg-4 contains a periplasmic nitrate reductase NapA that forms a soluble complex with cytochrome c-containing NapB and couples electron transfer from the quinone pool by the membrane-associated quinol dehydrogenase module formed by NapGH (Table S2). In Nbg-4, the nap operon lacks a gene encoding NapC, which is a proposed electron-transferring, membrane-associated protein typically observed in DNRA-performing SRM. The lack of NapC resembles the situation in *Wolinella succinogenes* that also lacks this protein while being able to perform DNRA (46). The second step of DNRA employs a six-electron transfer to reduce nitrite to ammonia. In Nbg-4, this step might be catalyzed by the membrane-bound nitrite reductase complex formed by NrfA, a periplasmic nitrite reductase, and NrfH, a membrane-associated quinol dehydrogenase that delivers electrons to NrfA. Screening of the obtained metaproteomes for DNRA-related proteins of Nbg-4 identified peptides specific for NapA and NapG in bulk soils (LFQ: $1.35 \times 10^7$ and $4.87 \times 10^3$, respectively) without gypsum treatment. Lack of peptides specific for proteins involved in the second step of DNRA could again be due to the fragmented recovery of the metaproteome. However, a specific expression of the first DNRA step only without further conversion of the produced nitrite to ammonium cannot be excluded. No peptides of DNRA-
related proteins were detected in bulk soil treated with gypsum or in the rhizosphere samples, irrespective of gypsum treatment (Table S2).

The genetic potential for complete oxidation of organic matter to CO$_2$

The genome of Nbg-4 encoded the capacity for complete oxidation of acetate to CO$_2$. This included the acetate transporter ActP, activation of acetate to acetyl-CoA by an AMP-forming acetyl-CoA synthetase (AcsA) and the complete Wood-Ljungdahl pathway (Figure 2, Table S2). Peptides specific for several of these enzymes could be detected with higher signal intensities in the bulk soil (LFQ: 3.55-5.20 × 10$^{08}$) as in the rhizosphere (LFQ: 2.11 × 10$^{04}$ – 1.69 × 10$^{06}$), whereas gypsum treatment had no apparent effect (Table S2). The Wood-Ljungdahl pathway included at the end of its methyl branch a formate dehydrogenase, which provides Nbg-4 with the potential to utilize also formate as an electron donor. In addition, a periplasm-oriented, membrane-bound [NiFeSe] hydrogenase (HysLS) was detected, which connects to the quinone pool in the membrane (Figure 2). However, no peptides related to either one of these two enzyme complexes could be detected (Table S2). Furthermore, the potential for butyrate degradation via a $\beta$-oxidation was encoded. With the exception of the activation step of butyrate to butyryl-CoA, all genes encoding for the necessary enzymes were recovered (Figure 2). Peptides that match Nbg-4 enzymes involved in butyrate degradation were detected in rhizosphere but not in bulk soil metaproteomes (LFQ: 4.18 × 10$^{04}$ – 2.66 × 10$^{06}$; Table S2).

In addition to the H$^+$-translocating quinol reductase complexes mentioned above for the sulfate and nitrate reduction pathways, coupling of electron transfer to energy conservation could be mediated in Nbg-4 by an electron-bifurcating Fd:NADP oxidoreductase (NfnAB), a H$^+$/Na$^+$-pumping Rnf complex (RnfCDGEAB), and a NADH-quinone oxidoreductase (respiratory complex I, NuoABCDEFGHIJKLMN) (35). In addition, the full set of genes encoding the ATP synthase was identified (AtpABCDEFHI) (Figure 2). Peptides specific for each of these Nbg-4 enzyme complexes were identified in the various bulk and rhizosphere soil metaproteomes (Table S2).
indicating their active role in electron transfer and energy conservation.

Phylogenetic affiliation of the *Nitrospirae* genome bin Nbg-4

A phylogenomic maximum-likelihood tree placed Nbg-4 and eight of the fourteen *dsrAB*-carrying *Nitrospirae* bacteria recovered in other studies (Table 2) in a stable cluster that branched off between *Thermodesulfovibrio* spp. and magnetotactic *Nitrospirae*. Two additional *dsrAB*-carrying *Nitrospirae* bacteria (GWA2-46-11 and GWB2-47-37) formed a sister branch to the Nbg-4 containing cluster and were more closely related to *Thermodesulfovibrio* species (Figure 4a). The remaining four *dsrAB*-carrying *Nitrospirae* bacteria branched off more basely within the phylum *Nitrospirae* forming two separate lineages with no clear affiliation to previously isolated species (Figure 4a).

The same branching pattern was recovered when analyzing deduced DsrAB sequences. Here, the well separated Nbg-4 containing cluster was most closely related to uncultured DsrAB family-level lineage 13 as defined by A. L. Müller et al. (37). Both clusters shared a common origin branching off between *Thermodesulfovibrio* species and magnetotactic *Nitrospirae* (Figure 4b). As in the phylogenomics approach, *Nitrospirae* bacteria GWA2-46-11 and GWB2-47-37 formed a stable sister branch that was more closely related to *Thermodesulfovibrio* species. Interestingly, the DsrAB of *Nitrospirae* bacterium RBG-13-39-12 and CG2-30-41-42, which were the closest relatives to Nbg-4 in the phylogenomics approach, did not fall into the *Nitrospirae* supercluster but were most closely related to uncultured DsrAB family-level lineage 11, which belongs to the Deltaproteobacteria supercluster (Fig. S1). This indicates lateral gene transfer of *dsrAB* within the phylum *Nitrospirae*, which is further supported by the DsrAB phylogeny of the basely branching *Nitrospirae* bacterium RBG-16-64-22. Here, the respective DsrAB sequences were clearly affiliated to the oxidative bacterial-type DsrAB having the alphaproteobacterium *Magnetococcus marinus* and *Chlorobi* spp. as closest relatives (Fig. S1). In contrast, DsrAB of *Nitrospirae* bacteria that formed the second basely branching lineage in the phylogenomics approach were also clustering
basely in the DsrAB *Nitrospirae* supercluster and clustered within or as closest relatives to uncultured DsrAB family-level lineage 10 (Fig. 4B).

In a third approach, the phylogenetic position of the partial 23S rRNA gene of Nbg-4 was inferred when placed into a full-length 23S rRNA gene tree of cultured and uncultured members of the phylum *Nitrospirae*. Also here, Nbg-4 branched off between stable clusters related to *Thermodesulfovibrio* species and magnetotactic *Nitrospirae* (Figure 4c), thus corroborating the phylogenetic placement of the other two approaches.

In parallel, a genome-wide average nucleotide identity (gANI) and average amino acid identity (gAAI) analysis was performed (47-49). The gANI analysis revealed that all *Nitrospirae* genomes used for the phylogenomic tree reconstruction were less similar than 70% to the genome of Nbg-4 (Table S4). Since this is well below the proposed value of 96.5% to group bacterial strains into the same species (48), Nbg-4 represents a novel species. The gAAI analysis mainly mirrored the phylogenomic tree reconstruction. Here, all genomes within the Nbg-4 containing cluster as well as the sister branch that encompasses *dsrAB*-carrying *Nitrospirae* bacteria GWA2-46-11 and GWB2-47-37 shared identities between 55 and 100% (Table S5). At the same time, these genomes shared less than 55% identity to representatives of other genera within the *Nitrospirae*. In addition, the two basely branching lineages of *dsrAB*-carrying *Nitrospirae* genome bins represented either by *Nitrospirae* bacterium RBG-16-64-22 or *Nitrospirae* bacteria GWC2-57-13, GWD2-57-8, and GWD2-57-9 shared less than 55% gAAI identity to *Nitrospirae* spp. outside of their respective lineage. At the same time, the later three *dsrAB*-carrying *Nitrospirae* bacteria shared among themselves gAAI identities of 62-99% (Table S5). Since 55% gAAI is the lower boundary that is currently recommended to group bacterial strains into the same genus (47), Nbg-4 and the additional uncultured *dsrAB*-carrying *Nitrospirae* bacteria listed in Table S3 form three independent genera.
**Discussion**

Members of the phylum *Nitrospirae*, which form a stable clade between thermophilic *Thermodesulfovibrio* spp. and magnetotactic *Nitrospirae*, are regularly observed in 16S rRNA gene- and *dsrAB*-based surveys of anoxic freshwater and marine environments of moderate temperature. These environments include marine (37) and estuarine (50) sediments, groundwater (44, 51), lake sediment (52), wetland soil (53), an anoxic bioreactor (54), and rice paddy fields (10, 55, 56). Also in rice paddy soil analyzed in this study, eight species-level operational taxonomic units (OTUs) of such *Nitrospirae* were observed previously by 16S rRNA gene-based amplicon sequencing (Fig. S2, 7). Here, we present a detailed genome analysis of *Nitrospirae* bacterium Nbg-4 as a representative of this clade and analyzed its protein expression profile under sulfate-enriched and sulfate-depleted conditions in planted rice paddy microcosms.

Nbg-4 encoded the complete pathway for dissimilatory sulfate reduction (Figure 2). Indeed, there are several lines of evidence that this newly discovered member of the *Nitrospirae* could represent an active sulfate reducer in rice paddy soil. From a genomic perspective, Nbg-4 carries not only all genes necessary for sulfate reduction but also genes of unknown function that are typically found in SRM such as *dsrD*, *dsrN* and *dsrT* (39). The same *dsr* operon organization (Figure 3) as well as the presence of all sulfate reduction-related genes (Table 2) were observed in the genomes of the other *dsrAB*-carrying *Nitrospirae* bacteria that form a stable phylogenetic lineage with Nbg-4 (Figure 4).

From a phylogenetic perspective, DsrAB of Nbg-4 and related *Nitrospirae* bacteria were clearly affiliated to the branch of reductively operating DsrAB of bacterial origin, which are phylogenetically separated from oxidatively operating DsrAB of bacterial origin (37). Most importantly, peptides clearly belonging to enzymes involved in sulfate reduction were preferentially detected for Nbg-4 in gypsum-treated bulk soil, i.e. under completely anoxic and sulfate-enriched conditions. On the contrary, under sulfate-depleted conditions in control bulk soil, peptides clearly belonging to the enzyme complex involved in the first step of DNRA were detected. From pure
culture SRM capable of DNRA, it is known that sulfate is preferentially respired even in the presence of the thermodynamically more favorable electron acceptor nitrate and that expression of DNRA-related enzymes is only induced in the absence of sulfate, which acts as repressor (57).

Nevertheless, an involvement of Nbg-4 and related *dsrAB*-carrying *Nitrospirae* in anaerobic sulfur oxidation cannot be ruled out. For example, a study conducted in parallel to our’s reported recently on the enrichment of a novel *Nitrospirae* species in an anoxic bioreactor that operated under simultaneous sulfide, methane, and ammonium consumption at the expense of nitrate (54). This novel *Nitrospirae* species closely resembled genomic and phylogenetic features of Nbg-4 with sulfide oxidation coupled to DNRA being one of several explanations that lead to its enrichment (54). Also, dense cell suspensions of the SRMs *Desulfovibrio desulfuricans* and *Desulfobulbus propionicus* are capable of coupling sulfide oxidation to nitrate reduction (58) and $S^0$ oxidation to electron transfer to a graphite electrode (59), respectively. In addition, *Desulfurivibrio alkaliphilus* was recently shown to grow by sulfide oxidation coupled to DNRA while encoding and transcribing DsrAB affiliated to the phylogenetic branch of reductively operating sulfite reductases (40). *D. alkaliphilus* encoded and also expressed all other genes of the canonical pathways of sulfate reduction while oxidizing sulfide coupled to DNRA. At the same time, it lacked all typical sulfur metabolism genes of chemolithotrophic sulfur oxidizers with the exception of a membrane-bound sulfide-quinone oxidoreductase (Sqr). This led to the proposal that the canonical pathway of sulfate reduction could act in reverse when coupled to Sqr (40). Interestingly, Nbg-4 encoded Sqr as well, which showed a moderate similarity (54% amino acid identity) to Sqr of *D. alkaliphilus*. However, Sqr of Nbg-4 could not be identified to be expressed in the analyzed rice paddy metaproteomes (Table S2). The overall picture is further complicated by the phylogenetic placement of Nbg-4 and related *dsrAB*-carrying *Nitrospirae* between the genus *Thermodesulfovibrio*, which contains exclusively sulfate-reducing species, and magnetotactic *dsrAB*-carrying *Nitrospirae*, which are proposed to be capable of sulfur oxidation. Since genes encoding the biosynthesis of magnetosomes were not detected in the largely recovered genome of Nbg-4, and it was significantly more abundant
in the completely anoxic bulk soil (Figure 1), a lifestyle comparable to magnetotactic *Nitrospirae*
can be most likely excluded.

In a preceding study, exclusively members of the Deltaproteobacteria (*Syntrophobacter, Desulfovibrio*, unclassified Desulfobulbaceae, and unclassified Desulfobacteraceae species) were identified to respond by population increase towards higher sulfate availability in rice paddy soil (7). The current study utilized soil from exactly the same experiment and identified Nbg-4 as an additional potential SRM. Nbg-4 did clearly not respond by changes in population size towards sulfate availability (Figure 1) but most likely by a switch in energy metabolism, i.e., from nitrate reduction under sulfate-depleted conditions to sulfate reduction under sulfate-enriched conditions (see above). This interpretation is supported by porewater sulfate turnover in microcosms incubated in parallel to the ones analyzed in this study (7), where sulfate concentrations steadily declined from 2.6 to 0.5 mM throughout the incubation period in gypsum-amended bulk soil but were below the detection limit in unamended bulk soil. Together, both studies reveal that rice paddy SRM may follow different ecological strategies, either by activity response coupled to growth (Deltaproteobacteria) or by switching the energy metabolism to maintain a stable population (Nbg-4). Interestingly, species-level OTUs obtained in the previous study and which fall into a phylogenetic lineage resembling the Nbg-4 cluster (Fig. S2), constituted relative population sizes of up to 0.2% of the overall bacterial community in bulk soil irrespective of gypsum treatment (re-analyzed from 7). This is clearly above the currently recognized threshold of the so-called “rare biosphere” (<0.1%) but below the threshold of dominating species (> 1%) (60, 61). As such, these novel *Nitrospirae* constitute moderately abundant members of the bacterial bulk soil community. This is in accordance to a study of three different Chinese rice paddy soils, where comparable population sizes were recorded (55).

Nbg-4 and related *dsrAB*-carrying *Nitrospirae*, which were all recovered from groundwater systems, clearly formed a separate lineage within the *Nitrospirae*. This was supported by three
independent phylogeny inference approaches as based on highly conserved marker genes, the *dsrAB* genes, and the 23S rRNA gene (Figure 4). Further indirect evidence was provided by the same branching pattern of 16S rRNA genes affiliated to the phylum *Nitrospirae* and recovered from the same microcosms (Fig. S2). In accordance with the performed gAAI analysis, Nbg-4 and related *dsrAB*-carrying *Nitrospirae* that form this separate lineage, constitute a newly discovered genus (Table S5). In addition, Nbg-4 represents a clearly distinct species in comparison to all members within this novel genus based on the performed gANI analysis (Table S4).

**Description of a new Candidatus genus and species**

Based on its distinct potential physiology, separation into an own phylogenetic lineage, and predominant occurrence in habitats of moderate temperature, the following name is proposed for Nbg-4: *Candidatus Sulfobium mesophilum* [etymology: *Sulfobium* gen. nov. (Sul.fo'bi.um. L. n. *sulfur* sulfur; Gr. n. *bios* life; N.L. neut. n. *Sulfobium* a living entity metabolizing sulfur compounds), *S. mesophilum* sp. nov. (me.so'phi.lum. Gr. adj. *mesos* middle; Gr. adj. *philos* friend, loving; N.L. neut. n. *mesophilum*, loving medium temperatures)]. *Candidatus* *S. mesophilum* encodes the complete pathways for dissimilatory sulfate reduction and nitrate reduction to ammonia. Based on its genome, it is able to utilize butyrate, acetate, formate, and molecular hydrogen as electron donors. In evidence of a complete Wood-Ljungdahl pathway, *Candidatus S. mesophilum* possesses the metabolic potential to oxidize organic matter completely to CO₂.

**Materials and methods**

**Rice paddy microcosms**

Soil from planted rice paddy microcosms described in S. Wörner et al. (7) was analyzed. In brief, microcosms were sampled destructively after 58-59 days of a greenhouse incubation (late vegetative phase of rice plants) to obtain rhizosphere and bulk soil of microcosms treated without (control) and with gypsum (0.15% (w/w) CaSO₄·2H₂O). In addition, freshly flooded soil was
incubated for three days in the absence of a rice seedling and denoted as T₀. As such, the experimental setup resulted in five different soil habitats: bulk soil with and without gypsum addition, rhizosphere soil with and without gypsum addition, and freshly flooded soil. Sampling from the different soil compartments and DNA extraction based on beat beating and phenol-chloroform extraction were as described in S. Wörner et al. (7).

**Metagenome sequencing, assembly, and binning**

Rhizosphere- and bulk soil-derived DNA extracts were obtained from four separate microcosms per treatment (gypsum and control). In addition, three DNA samples were obtained from freshly flooded soil. For each replicate, 2 µg of DNA were used for metagenomic library preparation and paired-end sequencing (2 × 100 bp) on an Illumina HiSeq 2000 platform at the King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. Raw reads were processed in the CLC Genomics Workbench 5.5.1 (CLC bio, Aarhus, Denmark) using only paired-end reads >50 bp with ≤1 ambiguity and a quality score ≥0.03 (corresponds to 99% accuracy). De novo assembly of pooled reads per habitat type was done in CLC using a k-mer size of 41 (determined as optimal in preliminary tests). Contigs with <2000 bp were discarded. Scaffolds containing 16S rRNA genes, 23S rRNA genes, or dsrAB were identified by a blastn search (62) against the respective SILVA reference databases v.123 (63) or a dsrAB reference database (37). Coverage of scaffolds was determined in CLC using 100% identity over the full length of quality trimmed reads. This was done for each sequenced replicate separately for statistical analysis and in addition using pooled replicates per habitat type for genome binning.

Genome binning was performed according to M. Albertsen et al. (64) using the gypsum and control treatment as differential coverage conditions (Fig. S3). From the 159 obtained genome bins, a dsrAB-carrying *Nitrospirae* bin assembled from gypsum-treated bulk soil was selected for further refinement (Figure S1). First, quality-trimmed reads that mapped to the *Nitrospirae* bin as well as to taxonomically unaffiliated scaffolds of similar coverage were re-assembled in CLC and binned as
outlined above. Thereafter, obtained scaffolds were co-assembled with quality-trimmed reads of the first step using SPAdes (65). Binning resulted in the genome bin Nbg-4 (Nitrospirae genome bin from bulk soil treated with gypsum). Using this procedure, the genome of Nbg-4 could be extended from 1.15 Mbp with 57 out of 107 queried essential single-copy genes (ESG) to 2.77 Mbp that covered 92 ESGs, with 91 of these ESGs being present as one copy. Assembly refinement of a 23S rRNA gene fragment encoded at the end of one Nbg-4 scaffold is described in Supplementary Material. Completeness, contamination and strain heterogeneity of Nbg-4 were evaluated using CheckM (66). To assess its relative abundance in the different soil habitats, quality-trimmed reads of sequenced soil replicates were mapped with a similarity threshold of 100% over the complete read to the Nbg-4 scaffolds using CLC. Mapped reads were normalized to RPKM values (reads per kilobase of scaffold per million reads).

**Annotation and additional analyses**

The MicroScope platform was used for automatic annotation (67, 68). Annotation refinement was done as follows: proteins with an amino acid identity ≥40% (over ≥80% of the sequence) to a SwissProt entry (69) were annotated as homologous to proteins with a known function. Proteins with an amino acid identity ≥25% (over ≥80% of the sequence) to a SwissProt or TrEMBL (69) entry were annotated as putative homologs of the respective database entries.

Genome-wide average nucleotide identity (ANI, 49) and average amino acid identity (AAI, 47) comparisons were performed using the web service of the Konstantinidis laboratory at the Georgia Institute of Technology, GA, USA (env-omics.ce.gatech.edu). The index of replication (iRep) was calculated using the iRep software (38). SAM files needed as input for iRep were created using bowtie2 (70).

To estimate the effect of soil habitat, gypsum treatment and the interaction thereof on the relative abundance of the Nitrospirae genome bin, a two-way ANOVA was performed based on RPKM values of its longest scaffold (106,945 bp) in the different replicated metagenomes. This was done
using the base package of the program R, version 3.1.1 (71). Assumptions of variance homogeneity and normality were tested using Levene’s test in the R package lawstat (72). Significant differences between differently treated soil habitat types were inferred using Tukey’s test of honest significant difference.

Metaproteomics of rice paddy soils

Total proteins were extracted from the same replicated soil samples as used for metagenome sequencing. Protein extraction and subsequent in-gel tryptic digestion followed the procedure outlined in R. Starke et al. (73). Briefly, 2 g of soil was used for a phenol extraction procedure with a subsequent ammonium acetate precipitation. Tryptic peptides were analyzed using a UPLC-LTQ Orbitrap Velos LC-MS/MS (74). Peptide searches were performed using the MaxQuant algorithm with the following parameters: tryptic cleavage with maximum two missed cleavages, a peptide tolerance threshold of ±10 ppm and an MS/MS tolerance threshold of ±0.5 Da, and carbamido methylation at cysteines as static and oxidation of methionines as variable modifications. As sample specific database, the Nbg-4 genome was used. According to currently accepted practice in metaproteomics (75, 76), proteins were considered as identified with at least one unique peptide with high confidence (false discovery rate-corrected p-value <0.01). To check for false positive assignments, all metaproteome replicates were also searched against the complete bacterial protein database of NCBI (08/2017).

Phylogenetic analysis

Additional Nitrospirae genome bins carrying dsrAB were identified using a blast search (62) against NCBI’s sequence repositories (77). Only Nitrospirae genome bins with a completeness above 70% and a contamination below 7% according to CheckM (66) were considered for further analysis. The phylogenetic affiliation of Nbg-4 and public dsrAB-carrying Nitrospirae genome bins was inferred using a phylogenomics approach based on 43 conserved marker genes with largely congruent phylogenetic histories as defined by D. H. Parks et al. (66) as well as using dsrAB and 23S rRNA
genes as phylogenetic markers. Respective maximum likelihood trees were calculated using RAxML v8.2.9 (78) as implemented on the CIPRES webserver (79, www.phylo.org). Details are provided in Supplementary Material.

Sequence information

All sequences are available in the Short Read Archive of NCBI under bioproject number PRJNA391190. The draft genome of Nbg-4 has been deposited in EMBL under the study accession number PRJEB21584. The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (80) with the dataset identifier PXD007817.

Funding information

This research was financed by the German Research Foundation (DFG, PE 2147/1-1 to MP) and the European Union (FP7-People-2013-CIG, Grant No PCIG14-GA-2013-630188 to MP). Furthermore, this research was supported by the PhD School in Food Systems from the University of Milano as well as by an ERASMUS+ placement studentship, both awarded to SZ. Funding for US was provided through baseline funds from KAUST and through the USDA National Institute of Food and Agriculture, Hatch project FLA-FTL-005631. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication. The authors declare no conflict of interests.

Acknowledgements

We are grateful to Prof. Dr. Bernhard Schink and Dr. Nicolai Müller for helpful discussions and support in naming the novel Candidatus genus and species. We also express our gratitude to Alexander Loy and Bela Hausmann for the continuous exchange of ideas. A pre-print of this manuscript was deposited at bioRxiv (www.biorxiv.org) under the doi: https://doi.org/10.1101/196774.


bacterium *Thermodesulfovibrio yellowstonii*, gen. nov. and sp. nov.: its phylogenetic relationship to *Thermodesulfobacterium commune* and their origins deep within the bacterial domain. Archives of Microbiology 161:62-69.


350:1541-1545.


stable isotope probing of rRNA. Environmental Microbiology 9:223-237.


27


77. NCBI-Resource-Coordinators. 2017. Database resources of the National Center for


Figures

**Figure 1.** Average relative abundances (± one standard deviation) of *Nitrospirae* bacterium Nbg-4 in the differently treated soil habitats as inferred from the RPKM values (reads per kilobase of scaffold per million reads) of its longest scaffold. Significant differences are indicated by different letters and were inferred by a two-way ANOVA and a post-hoc Tukey test (p<0.001).

**Figure 2.** Schematic view of reconstructed energy metabolism pathways in *Nitrospirae* bacterium Nbg-4. Expression of proteins in bulk soil treated with gypsum as revealed by metaproteomics is color-indicated. Protein expression in other soil habitats and treatments is given in Table S2.

**Figure 3.** Organization and synteny of the *dsr* operon in *Nitrospirae* bacterium Nbg-4 in comparison to other *dsrAB*-carrying members of the phylum *Nitrospirae*. In addition, typical representatives of known sulfate-reducing microorganisms within the *Deltaproteobacteria* (*Desulfovibrio vulgaris* Hildenborough), *Firmicutes* (*Desulfosporosinus meridiei*), and *Archaea* (*Archaeoglobus fulgidus*) are shown.

**Figure 4.** Phylogeny of *Nitrospirae* bacterium Nbg-4 (in bold) and related *dsrAB*-carrying *Nitrospirae* bacteria recovered from metagenomes of groundwater systems (44, 45). Uncultured *dsrAB*-carrying *Nitrospirae* bacteria that form separate genera as inferred by the genome-wide AAI approach are color coded. Maximum likelihood trees were inferred using the RAxML algorithm (78) and (A) a concatenated alignment of 43 essential proteins (66), (B) deduced DsrAB sequences, and (C) the 23S rRNA gene. The partially recovered 23S rRNA gene of Nbg-4 was added to a RAxML tree of almost full-length 23S rRNA genes using the Quick add parsimony tool as implemented in ARB (81) without changing the tree topology. This is indicated by the dashed branch leading to Nbg-4 in this tree. Bootstrap support is indicated by closed (≥90%) and open (≥70%) circles at the respective branching points. The scale bar indicates 10 or 5% estimated sequence divergence, respectively.
Tables

Table 1. Characteristics of the obtained draft genome of *Nitrospirae* bacterium Nbg-4.

<table>
<thead>
<tr>
<th>Genome feature</th>
<th>Nitrospirae bacterium Nbg-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome size (Mbp)</td>
<td>2.77</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>49</td>
</tr>
<tr>
<td>Number of scaffolds</td>
<td>151</td>
</tr>
<tr>
<td>Number of CDS</td>
<td>2855</td>
</tr>
<tr>
<td>Average CDS length (bp)</td>
<td>855</td>
</tr>
<tr>
<td>Protein coding density (%)</td>
<td>87</td>
</tr>
<tr>
<td>Number of rRNA genes</td>
<td>1</td>
</tr>
<tr>
<td>Number of tRNA genes</td>
<td>21</td>
</tr>
<tr>
<td><strong>CheckM analysis</strong></td>
<td></td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>75.5</td>
</tr>
<tr>
<td>Contamination (%)</td>
<td>2.0</td>
</tr>
<tr>
<td>Strain heterogeneity (%)</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>iRep analysis</strong></td>
<td></td>
</tr>
<tr>
<td>iRep initial soil</td>
<td>1.73</td>
</tr>
<tr>
<td>iRep bulk soil without gypsum</td>
<td>1.34</td>
</tr>
<tr>
<td>iRep bulk soil with gypsum</td>
<td>1.31</td>
</tr>
</tbody>
</table>
Table 2. Locus tag of genes involved in dissimilatory sulfate reduction in *Nitrospirae* bacterium 748
related *dsrAB* carrying *Nitrospirae* recovered from groundwater metagenomes (44, 45), and 749
*Thermodesulfovibrio yellowstonii*.

<table>
<thead>
<tr>
<th>Genome</th>
<th><strong>dsrA</strong></th>
<th><strong>dsrB</strong></th>
<th><strong>dsrD</strong></th>
<th><strong>dsrN</strong></th>
<th><strong>dsrC</strong></th>
<th><strong>dsrT</strong></th>
<th><strong>dsrM</strong></th>
<th><strong>dsrK</strong></th>
<th><strong>dsrI</strong></th>
<th><strong>dsrO</strong></th>
<th><strong>dsrP</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Nitrospirae</em> bacterium Nbg-4 (NBG4)</td>
<td>480011</td>
<td>480010</td>
<td>480009</td>
<td>480008</td>
<td>480005</td>
<td>480003</td>
<td>480002</td>
<td>480001</td>
<td>-</td>
<td>-</td>
<td>1080008</td>
</tr>
<tr>
<td><em>Nitrospirae</em> bacterium GWF2-44-13 (A2X54)</td>
<td>05135</td>
<td>05130</td>
<td>05125</td>
<td>05120</td>
<td>00165</td>
<td>00170</td>
<td>00175</td>
<td>00180</td>
<td>00185</td>
<td>00190</td>
<td>00195</td>
</tr>
<tr>
<td><em>Nitrospirae</em> bacterium CG1-02-44-142 (AUJ60)</td>
<td>04265</td>
<td>04260</td>
<td>04255</td>
<td>09835</td>
<td>04175</td>
<td>04180</td>
<td>04185</td>
<td>04190</td>
<td>04195</td>
<td>04200</td>
<td>0425</td>
</tr>
<tr>
<td><em>Nitrospirae</em> bacterium GWB2-47-37 (A2X55)</td>
<td>01500</td>
<td>01495</td>
<td>01490</td>
<td>01485</td>
<td>01475</td>
<td>01470</td>
<td>01465</td>
<td>01460</td>
<td>01455</td>
<td>01450</td>
<td>01445</td>
</tr>
<tr>
<td><em>Nitrospirae</em> bacterium RBG-13-39-12 (A2Y97)</td>
<td>05480</td>
<td>05485</td>
<td>-</td>
<td>05450</td>
<td>05445</td>
<td>05440</td>
<td>05435</td>
<td>05430</td>
<td>05425</td>
<td>05420</td>
<td>05415</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genome</th>
<th><strong>aprA</strong></th>
<th><strong>aprB</strong></th>
<th><strong>sat</strong></th>
<th><strong>hppA</strong></th>
<th><strong>qmoA</strong></th>
<th><strong>qmoB</strong></th>
<th><strong>qmoC</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Nitrospirae</em> bacterium Nbg-4</td>
<td>-</td>
<td>690001</td>
<td>690002</td>
<td>30083</td>
<td>30087</td>
<td>30086</td>
<td>30085</td>
</tr>
<tr>
<td><em>Nitrospirae</em> bacterium GWF2-44-13 (A2X54)</td>
<td>02100</td>
<td>02105</td>
<td>02110</td>
<td>02080</td>
<td>02095</td>
<td>02090</td>
<td>02085</td>
</tr>
<tr>
<td><em>Nitrospirae</em> bacterium CG1-02-44-142 (AUJ60)</td>
<td>-</td>
<td>-</td>
<td>03990</td>
<td>08585</td>
<td>08565</td>
<td>08570</td>
<td>08575</td>
</tr>
<tr>
<td><em>Nitrospirae</em> bacterium GWB2-47-37 (A2X55)</td>
<td>02795</td>
<td>02800</td>
<td>02805</td>
<td>02770</td>
<td>02790</td>
<td>02785</td>
<td>02780</td>
</tr>
<tr>
<td><em>Nitrospirae</em> bacterium RBG-13-39-12 (A2Y97)</td>
<td>02630</td>
<td>02635</td>
<td>02645</td>
<td>02470</td>
<td>02455</td>
<td>02460</td>
<td>02465</td>
</tr>
<tr>
<td><em>Thermodesulfovibrio yellowstonii</em> (THEYE)</td>
<td>A1832</td>
<td>A1833</td>
<td>A1835</td>
<td>-</td>
<td>A1831</td>
<td>A1830</td>
<td>A1829</td>
</tr>
</tbody>
</table>
Supplementary figure legends

**Figure S1.** Phylogeny of deduced DsrAB sequences of *Nitrospirae* bacterium Nbg-4 and related *dsrAB*-carrying *Nitrospirae* bacteria recovered from metagenomes of groundwater systems (44, 45). A maximum likelihood tree were inferred using the RAxML algorithm (78). Bootstrap support is indicated by closed (≥90%) and open (≥70%) circles at the respective branching points. *Nitrospirae* bacteria with *dsrAB* that underwent horizontal gene transfer are marked with an asterisk. The scale bar indicates 10% estimated sequence divergence.

**Figure S2.** Maximum likelihood 16S rRNA gene tree showing the phylogenetic position of species-level OTUs (97% nucleic acid sequence identity) affiliated to the phylum *Nitrospirae*, which were obtained in a previous study (7) using the same rice paddy soil samples as analyzed in the current study. The tree was reconstructed using the RAxML algorithm (78) as implemented in ARB (81) using 1,222 unambiguously aligned nucleotide positions and a 50% conservation filter for the domain Bacteria. The representative 454 amplicon sequences were added to the tree by using ARB’s Parsimony Interactive tool as indicated by the dashed branch. Solid circles indicate ≥90% and open circles ≥70% bootstrap support (1000 replications). The bar represents 10% inferred sequence divergence.

**Figure S3.** Schematic overview of the bioinformatics workflow to obtain the high quality draft genome of *Nitrospirae* bacterium Nbg-4.

Supplementary table legends

**Table S1.** Key characteristics of sequenced metagenomes.

**Table S2.** Annotation and locus of genes involved in energy and biosynthesis metabolism in *Nitrospirae* bacterium Nbg-4. Expression of respective genes as proteins is indicated in the metaproteomes of the respective analyzed soil replicates.
Table S3. Main characteristics of members of the phylum *Nitrospirae*.

Table S4. Genome-wide average nucleotide identity (gANI) of *Nitrospirae* bacterium Ngb-4 in comparison to other members of the phylum *Nitrospirae*.

Table S5. Genome-wide average amino acid identity (gAAI) of *Nitrospirae* bacterium Ngb-4 in comparison to other members of the phylum *Nitrospirae*.
Table 1. Characteristics of the obtained draft genome of *Nitrospirae* bacterium Nbg-4.

<table>
<thead>
<tr>
<th>Genome feature</th>
<th>Nitrospirae bacterium Nbg-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome size (Mbp)</td>
<td>2.77</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>49</td>
</tr>
<tr>
<td>Number of scaffolds</td>
<td>151</td>
</tr>
<tr>
<td>Number of CDS</td>
<td>2855</td>
</tr>
<tr>
<td>Average CDS length (bp)</td>
<td>855</td>
</tr>
<tr>
<td>Protein coding density (%)</td>
<td>87</td>
</tr>
<tr>
<td>Number of rRNA genes</td>
<td>1</td>
</tr>
<tr>
<td>Number of tRNA genes</td>
<td>21</td>
</tr>
<tr>
<td><strong>CheckM analysis</strong></td>
<td></td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>75.5</td>
</tr>
<tr>
<td>Contamination (%)</td>
<td>2.0</td>
</tr>
<tr>
<td>Strain heterogeneity (%)</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>iRep analysis</strong></td>
<td></td>
</tr>
<tr>
<td>iRep initial soil</td>
<td>1.73</td>
</tr>
<tr>
<td>iRep bulk soil without gypsum</td>
<td>1.34</td>
</tr>
<tr>
<td>iRep bulk soil with gypsum</td>
<td>1.31</td>
</tr>
</tbody>
</table>
Table 2. Locus tag of genes involved in dissimilatory sulfate reduction in *Nitrospira* bacterium *Nbg-4*, related *dsrAB*-carrying *Nitrospira* recovered from groundwater metagenomes (44, 45), and *Thermodesulfovibrio yellowstonii*.

<table>
<thead>
<tr>
<th>Genome</th>
<th>dsrA</th>
<th>dsrB</th>
<th>dsrD</th>
<th>dsrN</th>
<th>dsrC</th>
<th>dsrT</th>
<th>dsrM</th>
<th>dsrK</th>
<th>dsrJ</th>
<th>dsrO</th>
<th>dsrP</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Nitrospira</em> bacterium <em>Nbg-4</em> (NBG4)</td>
<td>480011</td>
<td>480010</td>
<td>480009</td>
<td>480008</td>
<td>480003</td>
<td>480002</td>
<td>480001</td>
<td>-</td>
<td>-</td>
<td>1080008</td>
<td></td>
</tr>
<tr>
<td><em>Nitrospira</em> bacterium GWF2-44-13 (A2X54)</td>
<td>05135</td>
<td>05130</td>
<td>05125</td>
<td>05120</td>
<td>00165</td>
<td>00170</td>
<td>00175</td>
<td>00180</td>
<td>00185</td>
<td>00190</td>
<td>00195</td>
</tr>
<tr>
<td><em>Nitrospira</em> bacterium CG1-02-44-142 (AUJ60)</td>
<td>04265</td>
<td>04260</td>
<td>04255</td>
<td>09835</td>
<td>04175</td>
<td>04180</td>
<td>04185</td>
<td>04190</td>
<td>04195</td>
<td>04200</td>
<td>0425</td>
</tr>
<tr>
<td><em>Nitrospira</em> bacterium GWB2-47-37 (A2X55)</td>
<td>01500</td>
<td>01495</td>
<td>01490</td>
<td>01485</td>
<td>01475</td>
<td>01470</td>
<td>01465</td>
<td>01460</td>
<td>01455</td>
<td>01450</td>
<td>01445</td>
</tr>
<tr>
<td><em>Nitrospira</em> bacterium RBG-13-39-12 (A2Y97)</td>
<td>05490</td>
<td>05485</td>
<td>-</td>
<td>05450</td>
<td>05445</td>
<td>05440</td>
<td>05435</td>
<td>05430</td>
<td>05425</td>
<td>05420</td>
<td>05415</td>
</tr>
<tr>
<td></td>
<td>aprA</td>
<td>aprB</td>
<td>sat</td>
<td>hppA</td>
<td>qmoA</td>
<td>qmoB</td>
<td>qmoC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Nitrospira</em> bacterium <em>Nbg-4</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Nitrospira</em> bacterium GWF2-44-13 (A2X54)</td>
<td>02100</td>
<td>02105</td>
<td>02110</td>
<td>02080</td>
<td>02095</td>
<td>02090</td>
<td>02085</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Nitrospira</em> bacterium CG1-02-44-142 (AUJ60)</td>
<td></td>
<td></td>
<td></td>
<td>03990</td>
<td>08585</td>
<td>08565</td>
<td>08570</td>
<td>08575</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Nitrospira</em> bacterium GWB2-47-37 (A2X55)</td>
<td>02795</td>
<td>02800</td>
<td>02805</td>
<td>02770</td>
<td>02790</td>
<td>02785</td>
<td>02780</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Nitrospira</em> bacterium RBG-13-39-12 (A2Y97)</td>
<td>02630</td>
<td>02635</td>
<td>02645</td>
<td>02470</td>
<td>02455</td>
<td>02460</td>
<td>02465</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Thermodesulfovibrio yellowstonii</em> (THEYE)</td>
<td>A1832</td>
<td>A1833</td>
<td>A1835</td>
<td></td>
<td>A1831</td>
<td>A1830</td>
<td>A1829</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>