**Further details of genome structure.** In addition to the large transposon interrupting one of the flagellar biosynthetic operons (Fig. S1), another transposase gene \((Tmden\_1713)\) is located near a tRNA\(_{\text{Thr}}\) gene, adjacent to a hypothetical protein gene \((Tmden\_1712)\) whose 3’-end is 82% identical, at the nucleotide level, with two transposase genes located downstream \((Tmden\_1724\text{ and } Tmden\_1725)\). These two genes, and regions 5’ and 3’ of each (totaling 1302 bp apiece), are 100% identical to each other. This region also includes genes encoding a recombinase and phage integrase \((Tmden\_1723; \text{ Fig. S2})\); the presence of the phage integrase gene, identical repeats, and juxtaposition to a tRNA gene suggest that this portion of the genome may be a remnant of a degraded or partially excised prophage. Other potential tranposase genes are present \((Tmden\_0961; Tmden\_1698; Tmden\_1708)\), but have insufficient sequence similarity to known proteins for deducing their function convincingly.

Twelve phage integrase genes are present. As expected for phage genes, six of these are near tRNA genes \((Tmden\_0248, Tmden\_0779, Tmden\_0800, Tmden\_1618, Tmden\_1723; Tmden\_1743)\) which are common insertion sites for lysogenic phages (1), and three are flanked by clusters of genes encoding hypothetical proteins \((Tmden\_1247; Tmden\_1618; Tmden\_1633)\), which is consistent with the observation that many phage genes are unique and uncharacterized (2). Two phage integrase genes \((Tmden\_0938; Tmden\_0959)\) flank genes encoding a type I restriction modification system gene cluster and are part of a larger region \((\text{bp } 977850 – 1002764; 5 \text{ o’clock, Fig. 1})\), that have a negative G+C anomaly (31.1%). The remaining two are included in the large transposons described above \((Tmden\_0693 \text{ and } Tmden\_1590)\).
Another negative G+C anomaly, which also has a GC skew anomaly, is visible on the genome map at approximately 1 o’clock (bps 192095 – 210856; 30.5% G+C; Fig. 1). Although this region contains several genes encoding hypothetical proteins, as one might expect were it derived from a phage, it does not appear to include any transposase or integrase genes, nor does it include any repeated sequences that might suggest recent gene rearrangement in this region.

**Restriction-modification systems.** *S. denitrificans* has numerous restriction-modification (RM) systems encoded in its genome. Eleven DNA methyltransferase genes are present, and encode methyltransferases similar to those found in Type I (*Tmden_0697, Tmden_0942, Tmden_1594*), Type II (*Tmden_0121; Tmden_0129; Tmden_0130; Tmden_0478; Tmden_0537; Tmden_1565; Tmden_1839; Tmden_1855*) and Type III (*Tmden_1355*) restriction-modification systems. For 6 of these methyltransferases, genes encoding restriction enzymes are nearby (Type I: *Tmden_0700; Tmden_0948; Tmden_1597*; Type II: *Tmden_128; Tmden_1854*; Type III: *Tmden_1350*) and for two of them the genes appear to encode fused methylase/restriction enzymes (*Tmden_0478, Tmden_0537*). Based on genome sequence data, such large numbers of RM systems are not unusual for epsilonproteobacteria: *Helicobacter pylori* has 24 RM systems (3), *C. jejuni* has 10, and *W. succinogenes* has 5 (4, 5). If active, perhaps in *S. denitrificans* these systems provide a robust defense against the introduction of phage and other ‘non-native’ DNA into the genome.
REFERENCES


FIGURE LEGENDS.

FIG. S1. Two large identical transposons from the *Sulfurimonas denitrificans* genome. Numbers indicate the position of the regions, in nucleotides, with respect to the origin of replication, and the arrows indicate the presence of the inverted repeat sequences at each end: $\geq$ TGTCAATTACAA; $\leq$ TTGTAAATGACA.

FIG. S2. Map of a region from the *Sulfurimonas denitrificans* genome that includes a small repeated region. The duplicate copies of this repeat include the two adjacent transposase genes (shaded grey), while a third region with a high level of identity (82%) is included within a hypothetical gene upstream (also shaded grey). Numbers indicate the position of the regions, in nucleotides, with respect to the origin of replication.

FIG. S3. Phylogenetic relationships of NosZ from different bacteria and the archaeon *Pyrobaculum calidifontis*. Suden_1298 is part of a novel nos cluster previously identified in *Wolinella succinogenes* (63). All epsilonproteobacterial sequences have a C-terminal extension and contain a heme c-binding motif. The sequences from *Dechloromonas aromatica* and *Magentospirillum magnetotacticum* also have a (somewhat shorter) C-terminal extension, but are lacking a heme c-binding motif (63). Sequences were aligned using the program package MacVector. Neighbor-joining and Parsimony trees based on the predicted amino acid sequences were calculated using
PAUP 4.0b10. Bootstrap values (1,000 replicates) for the major nodes are given for the neighbor-joining (first value) and parsimony analyses (second value).
FIG. S2