SUPPLEMENTAL MATERIAL

2   **Further details of genome structure.** In addition to the large transposon

3   interrupting one of the flagellar biosynthetic operons (Fig. S1), another transposase gene

4   (*Tmden_1713*) is located near a tRNA^{Thr} gene, adjacent to a hypothetical protein gene

5   (*Tmden_1712*) whose 3'-end is 82% identical, at the nucleotide level, with two

6   transposase genes located downstream (*Tmden_1724* and *Tmden_1725*). These two

7   genes, and regions 5' and 3' of each (totaling 1302 bp apiece), are 100% identical to each

8   other. This region also includes genes encoding a recombinase and phage integrase

9   (*Tmden_1723;* Fig. S2); the presence of the phage integrase gene, identical repeats, and

10   juxtaposition to a tRNA gene suggest that this portion of the genome may be a remnant of

11   a degraded or partially excised prophage. Other potential tranposase genes are present

12   (*Tmden_0961; Tmden_1698; Tmden_1708)*, but have insufficient sequence similarity to

13   known proteins for deducing their function convincingly.

14   Twelve phage integrase genes are present. As expected for phage genes, six of

15   these are near tRNA genes (*Tmden_0248, Tmden_0779, Tmden_0800, Tmden_1618,*

16   *Tmden_1723; Tmden_1743)* which are common insertion sites for lysogenic phages (1),

17   and three are flanked by clusters of genes encoding hypothetical proteins *(Tmden_1247;*

18   *Tmden_1618; Tmden_1633)*, which is consistent with the observation that many phage

19   genes are unique and uncharacterized (2). Two phage integrase genes *(Tmden_0938;*

20   *Tmden_0959)* flank genes encoding a type I restriction modification system gene cluster

21   and are part of a larger region (bp 977850 – 1002764; 5 o'clock, Fig. 1), that have a

22   negative G+C anomaly (31.1%). The remaining two are included in the large transposons

23   described above (*Tmden_0693* and *Tmden_1590).*

24   Another negative G+C anomaly, which also has a GC skew anomaly, is visible on

25   the genome map at approximately 1 o'clock (bps 192095 – 210856; 30.5% G+C; Fig. 1).

26   Although this region contains several genes encoding hypothetical proteins, as one might

27   expect were it derived from a phage, it does not appear to include any transposase or

28   integrase genes, nor does it include any repeated sequences that might suggest recent

29   gene rearrangement in this region.

30   **Restriction-modification systems.** *S. denitrificans* has numerous restriction-

31   modification (RM) systems encoded in its genome. Eleven DNA methyltransferase genes

32   are present, and encode methyltransferases similar to those found in Type I

33   (*Tmden_0697, Tmden_0942, Tmden_1594*), Type II (*Tmden_0121; Tmden_0129;*

34   *Tmden_0130; Tmden_0478; Tmden_0537; Tmden_1565; Tmden_1839; Tmden_1855*)

35   and Type III (*Tmden_1355*) restriction-modification systems. For 6 of these

36   methyltransferases, genes encoding restriction enzymes are nearby (Type 1:

37   *Tmden_0700; Tmden_0948; Tmden_1597;* Type II: *Tmden_128; Tmden_1854;* Type III:

38   *Tmden_1350)* and for two of them the genes appear to encode fused methylase/restriction

39   enzymes (*Tmden_0478, Tmden_0537*)(5). Based on genome sequence data, such large

40   numbers of RM systems are not unusual for epsilonproteobacteria: *Helicobacter pylori*

41   has 24 RM systems (3), *C, jejuni* has 10, and *W. succinogenes* has 5 (4, 5). If active,

42   perhaps in *S. denitrificans* these systems provide a robust defense against the introduction

43   of phage and other 'non-native' DNA into the genome.

44

REFERENCES

1.   **Canchaya, C., G. Fournous, and H. Brussow.** 2004. The impact of prophages on bacterial chromosomes. Molec Microbiol **53:**9-18.

2.   **Hambly, E., and C. A. Suttle.** 2005. The viriosphere, diversity, and genetic exchange within phage communities. Curr Opin Microbiol **8:**444-450.

3.   **Kong, H., L. Lin, N. Porter, S. Stickel, D. Byrd, J. Posfai, and R. j. Roberts.** 2000. Functional analysis of putative restriction-modification system genes in the *Helicobacter pylori* J99 genome. Nucl Acids Res **28:**3216-3223.

4.   **Roberts, R. J., T. Vincze, J. Posfai, and D. Macelis.** 2007. REBASE--enzymes and genes for DNA restriction and modification. Nucl Acids Res **35:**D269-D270

5.   **Roberts, R. J., T. Vincze, J. Posfai, and D. Macelis.** 2005. REBASE Restriction enzymes and DNA methyltransferases. Nucl Acids Res **33:**D230-232.

60
61 **FIGURE LEGENDS.**
62
63   FIG. S1. Two large identical transposons from the *Sulfurimonas denitrificans*

64 genome. Numbers indicate the position of the regions, in nucleotides, with respect to the

65 origin of replication, and the arrows indicate the presence of the inverted repeat

66 sequences at each end: $>$ = TGTCATTTACAA; $<$ = TTGTAAATGACA.

67

68   FIG. S2. Map of a region from the *Sulfurimonas denitrificans* genome that

69 includes a small repeated region. The duplicate copies of this repeat include the two

70 adjacent transposase genes (shaded grey), while a third region with a high level of

71 identity (82%) is included within a hypothetical gene upstream (also shaded grey).

72 Numbers indicate the position of the regions, in nucleotides, with respect to the origin of

73 replication.

74

75
76   FIG. S3. Phylogenetic relationships of NosZ from different bacteria and the

77 archaeon *Pyrobaculum calidifontis*. Suden_1298 is part of a novel nos cluster previously

78 identified in *Wolinella succinogenes* (63). All epsilonproteobacterial sequences have a C-

79 terminal extension and contain a heme *c*-binding motif. The sequences from

80 *Dechloromonas aromatica* and *Magentospirillum magnetotacticum* also have a

81 (somewhat shorter) C-terminal extension, but are lacking a heme *c*-binding motif (63)

82 Sequences were aligned using the program package MacVector. Neighbor-joining and

83 Parsimony trees based on the predicted amino acid sequences were calculated using

84   PAUP 4.0b10. Bootstrap values (1,000 replicates) for the major nodes are given for the

85   neighbor-joining (first value) and parsimony analyses (second value).
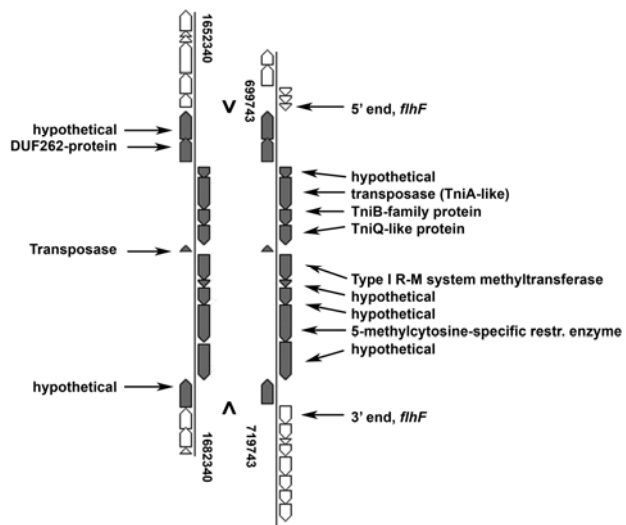
86

87

88

89

90

hypothetical
DUF262-protein

5' end, *flhF*

hypothetical
transposase (TniA-like)
TniB-family protein
TniQ-like protein

Transposase

Type I R-M system methyltransferase
hypothetical
hypothetical
5-methylcytosine-specific restr. enzyme
hypothetical

hypothetical

3' end, *flhF*

165 2340
699 743
168 2340
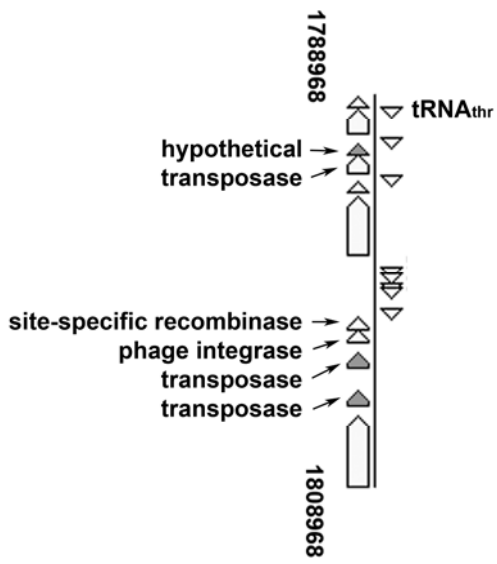719 743

91

92

93

94

95

96

97   FIG. S1

98

99

100

101

102

103

104

105

106

107

108

109   FIG. S2