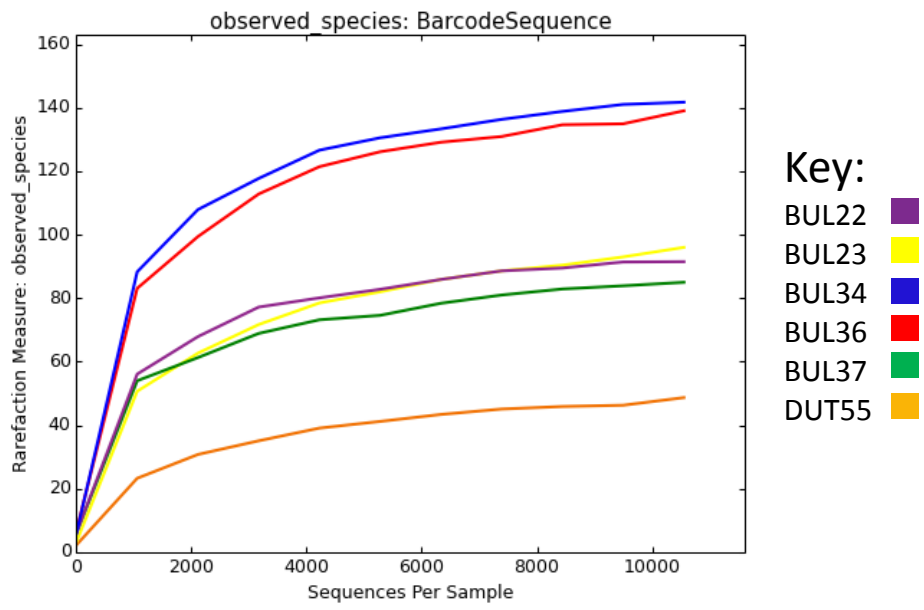
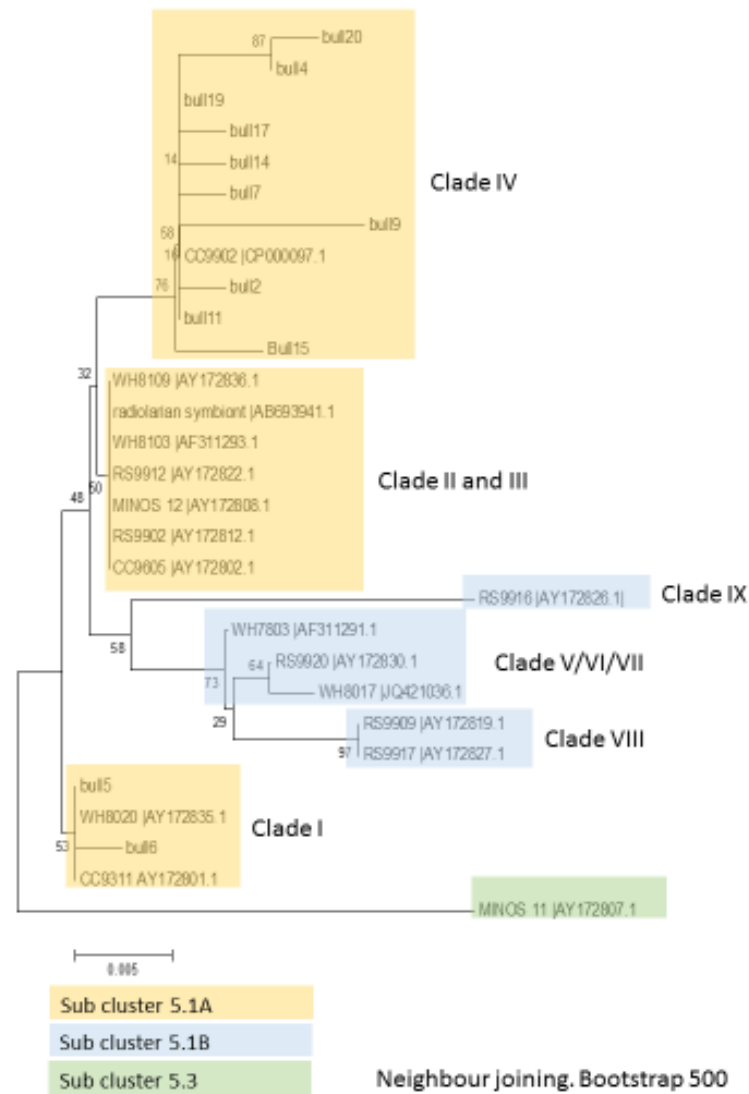
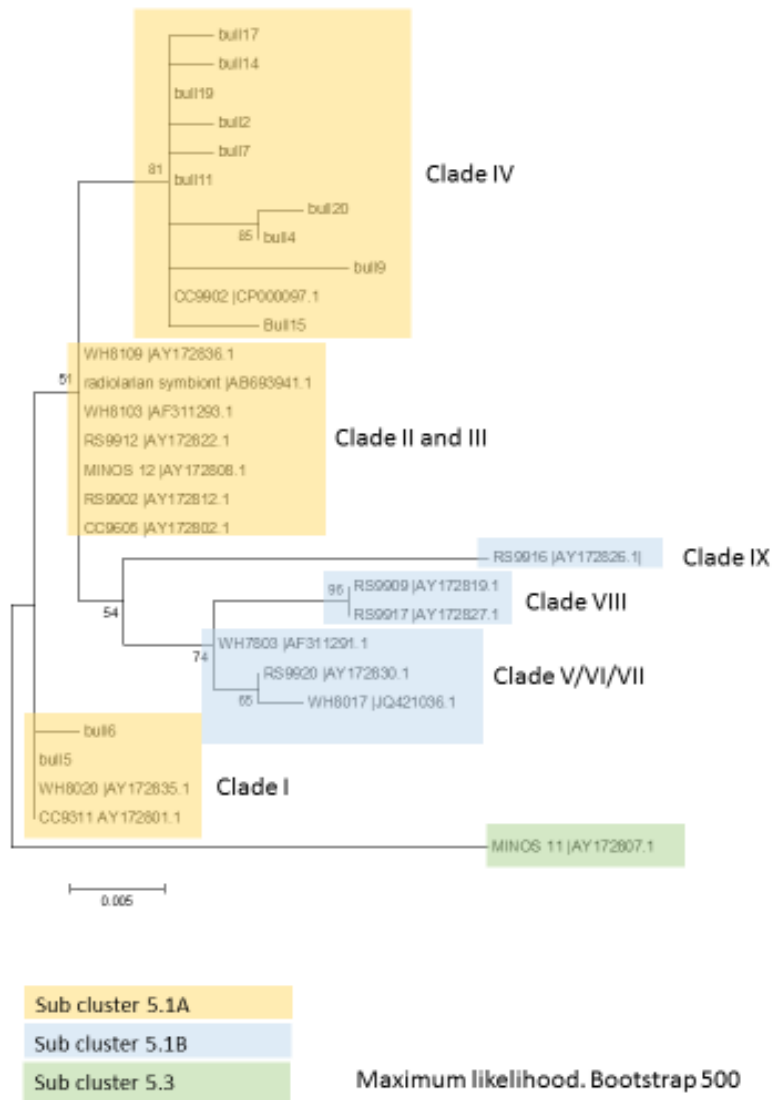


Supplementary Figure S1. OTU richness curves for each specimen generated using the alpha-rarefaction Observed Species metric in QIIME. The reduction in the gradient of the curves for each individual specimen with increasing sequencing effort demonstrates that the sequencing depth was sufficient to capture the full bacterial diversity present.

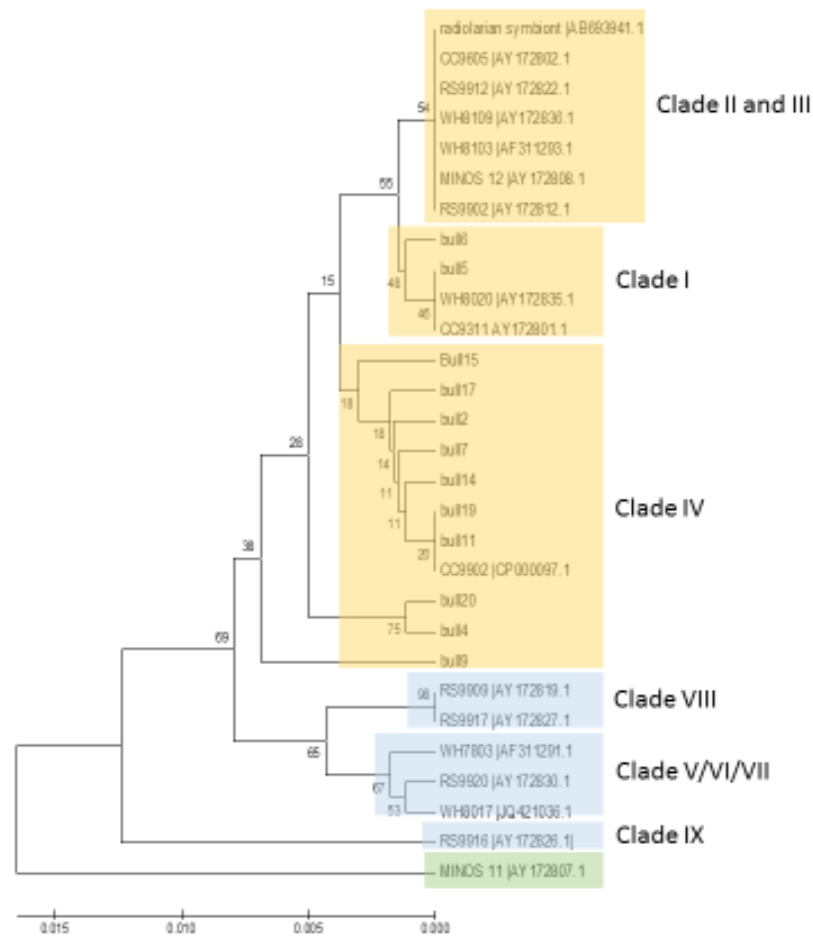
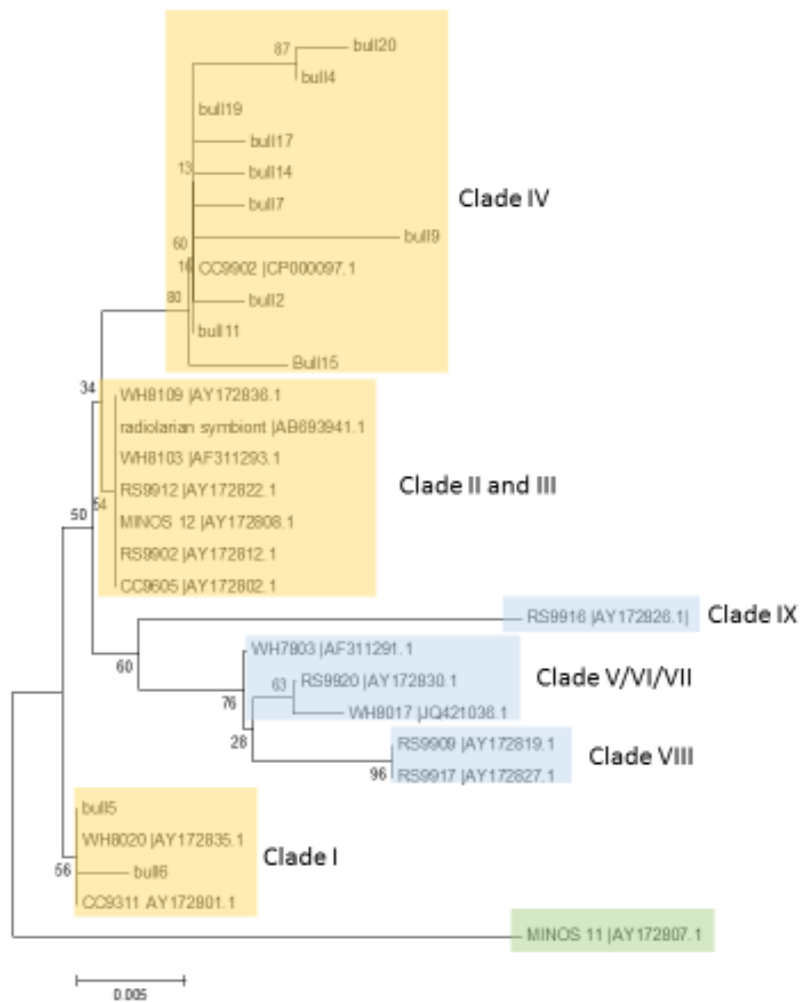
Supplementary Figure S2a-e. Phylogenetic trees (maximum likelihood Fig. S2a; neighbour-joining Fig. S2b; minimum evolution Fig. S2c; UPGMA Fig. S2d; maximum parsimony Fig. S2e) with 500 bootstrap resamplings of 425 bp fragments of *Synechococcus* 16S rRNA gene clone sequences (identified by the prefix 'bull' followed by a number) generated during this study and reference *Synechococcus* 16S rRNA gene sequences retrieved from the GenBank database (NCBI). A member (MINOS 11) of the open ocean sub-cluster 5.3 is used as the out group in all cases.



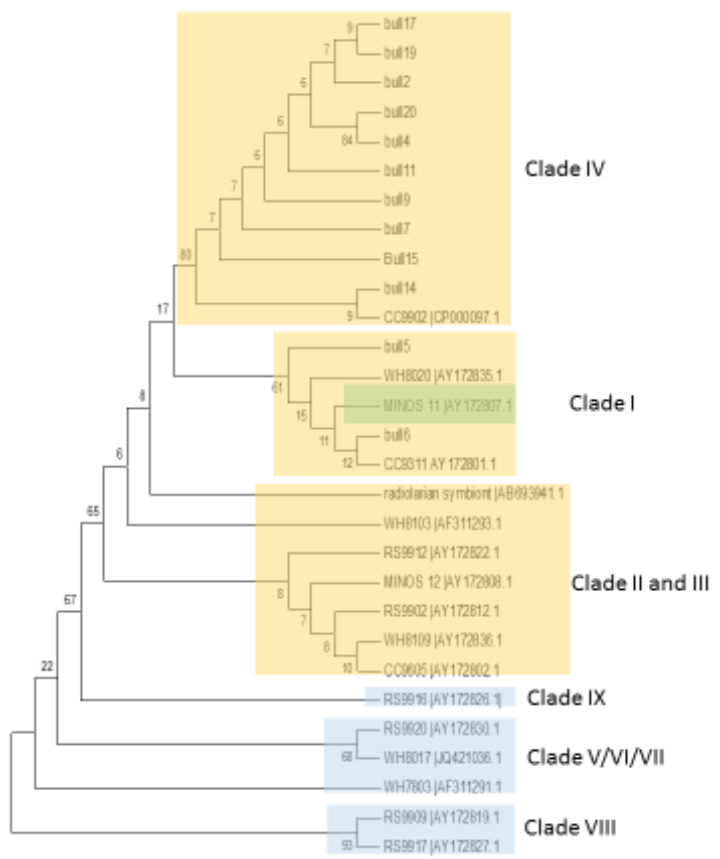
FigS1



FigS2a-b



FigS2c-d



Maximum parsimony. Bootstrap 500

FigS2e

Supplementary Table S1. The numbers of sequences and OTUs generated in individual specimens for both closed reference picking and de novo picking after removal of contaminants (Sect. 2.6.1). Note that total numbers of OTUs are not a sum of OTUs across all specimens as many OTUs are found in more than one specimen.

Morphospecies	Sample ID	Number of sequences assigned after closed reference picking	Number of OTUs in closed reference picking	Number of sequences assigned after <i>de novo</i> picking	Number of OTUs <i>de novo</i> picking
<i>G. bulloides</i>	BUL22	10,551	85	8,786	52
<i>G. bulloides</i>	BUL23	21,474	98	22,164	46
<i>G. bulloides</i>	BUL34	15,748	146	24,700	125
<i>G. bulloides</i>	BUL36	31,448	149	44,901	139
<i>G. bulloides</i>	BUL37	15,773	105	16,067	100
<i>N. dutertrei</i>	DUT59	119,093	74	119,052	92
Total sequences/OTUs		214,087	248	238,433	243