## **Bacterial abundances and relationships between samples.**

******

#### **FIG S3: Relative abundance of the 20 most abundant bacterial orders in seawater samples. A.** Seawater samples collected in New Caledonia. **B.** Seawater samples collectedin French Polynesia.Each sample match to a given conditions corresponding to the sampling time: beginning of the experimentation (T0), 1 month later (T1); season: warm season (WS), cold season (CS). Orders written in grey are common in both sites.

The 20 most abundant bacterial orders across all the seawater by site (NC and FP) were displayed in the figure S3. The most abundant orders of the seawater samples explained more than 97 % of the bacterial diversity in both sites, thus, less than 3 % was explained by Other, *i.e.* lower abundant orders. The water samples microbial communities in NC were significantly different from the seawater samples ones in FP (based on the ASVs counts, Pairwise Wilcoxon, *p*-value 3.26x10-14). This inter-site variability can be observed in the order profiles through differential relative abundances (Fig. 2). There was higher abundance of *Synechococcales* in FP conditions compared to the NC conditions. The variability was also observed with orders specifically with higher abundances in both sites per examples with the *Arenicellales*, the *SBR*1031, or the *Thalassobaculales* were more abundant in NC (Fig. S3A) and with the *Rhizobiales*, the *Burkholderiales* or *Phycisphaerales* in FP (Fig. S3B). Despite this variability, there were 11 orders common among the 20 most abundant (in grey in Fig. S3A-B). Moreover, there were 100 ASVs constituting the core microbiota of the seawater samples in both sites (ASVs present in all the seawater samples, Table S10). There were also 183 ASVs specific of the seawater samples in NC and 165 ASVs specific of the seawater samples in FP. As each seawater condition were mono-sample, no further statistics tests were performed.

**Proportion of chimera identified with consensus or polled methods:**

Using the consensus method 30 722 bimeras were identified out of 44 356 input sequences while with the pooled method, 36 020 bimeras were detected.

As the consensus method is better when the study uses many samples: “For chimera removal, we have found that the "consensus" chimera removal method works better on large studies” (<https://benjjneb.github.io/dada2/bigdata.html>), we kept results from this method of chimera removal.