Succession within the prokaryotic communities during the VAHINE mesocosms experiment in the New Caledonia lagoon

**U. Pfreundt[1], F. Van Wambeke[2], M. Caffin[2], S. Bonnet[2, 3] and W. R. Hess[1]**

[1]{University of Freiburg, Faculty of Biology, Schaenzlestr. 1, D-79104 Freiburg, Germany}

[2]{Aix Marseille Université, CNRS/INSU, Université de Toulon, IRD, Mediterranean Institute of Oceanography (MIO) UM110, 13288, Marseille, France

[3] {Institut de Recherche pour le Développement, AMU/CNRS/INSU, Université de Toulon, Mediterranean Institute of Oceanography (MIO) UM 110, 13288, Marseille -Nouméa, France - New Caledonia}

Supplementary Figures S1 - S5

Supplementary Figure S1. a) Development of the Shannon (red) and Simpsons (black) diversity indices in the Nouméa lagoon and mesocosm M1 shows that these are very similar. Generalized linear models were fitted with glm() in R and the predicted values plotted as straight lines. b) Shannon evenness calculated using the RAM package for R and the function evenness(index="shannon"). c) OTU richness calculated as Chao1 estimate with estimateR() from the R package vegan.
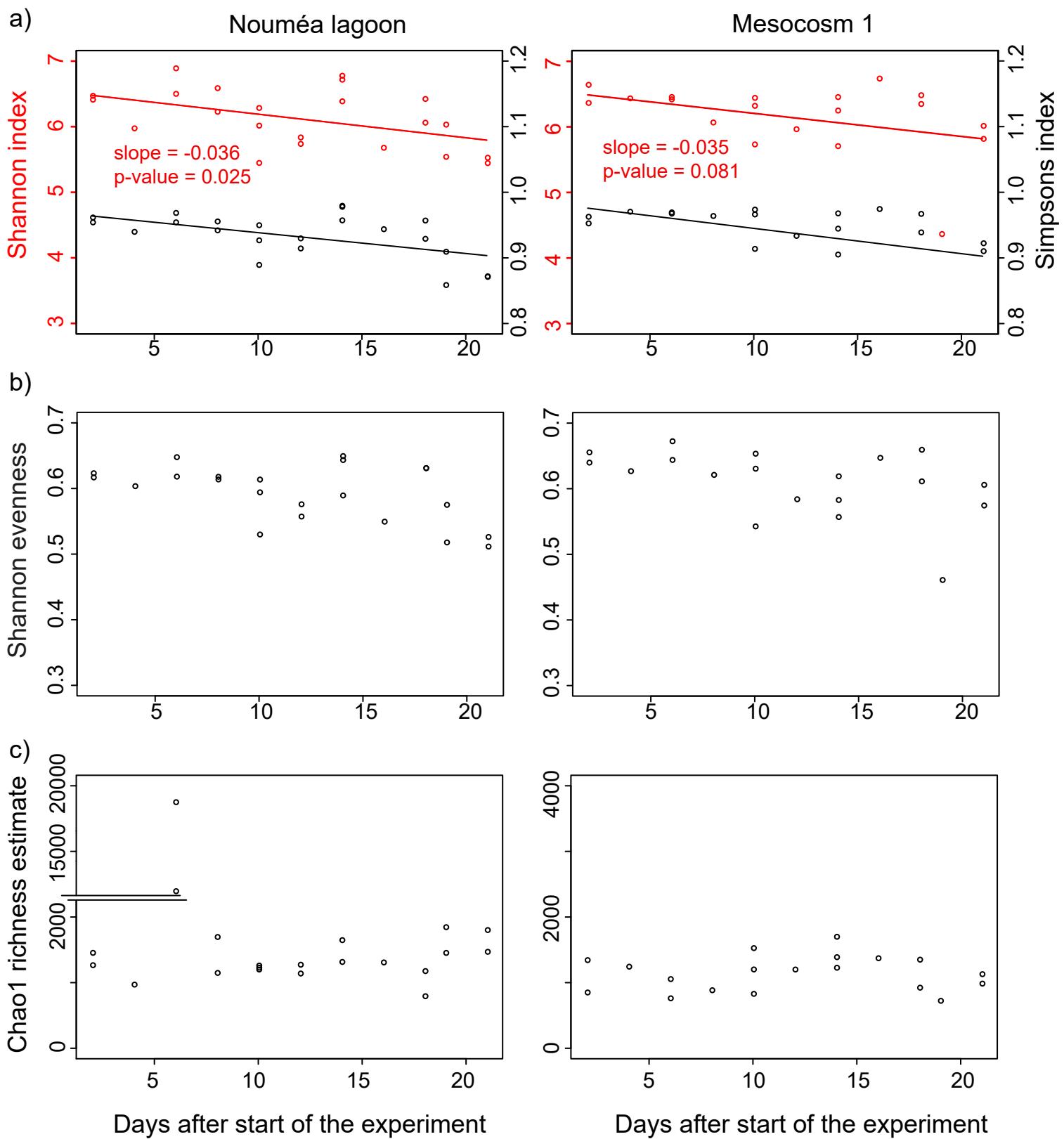
Supplementary Figure S2. For all investigated phylogenetic groups (Fig. S4, S5), the relative abundances in M1 and the Noumea lagoon were correlated by Spearman rank correlation. Note that these correlations are represented by the diagonals from top left to bottom right. All other positions in the heatmap represent correlation of one group in M1 versus another group in the lagoon. Values inside the heatmap are Spearman correlation coefficients, significant values are white and bold. a) Dominant classes >5 % relative abundance, b) classes 0.5-5 % rel. abundance, c-f) families >1 % rel. abundance within alpha- and gammaproteobacteria, Flavobacteriia, respectively, and genera within cyanobacteria.
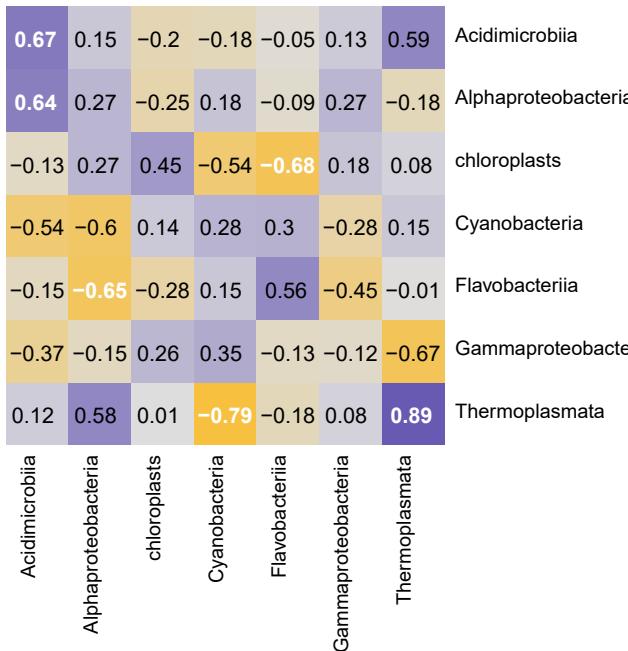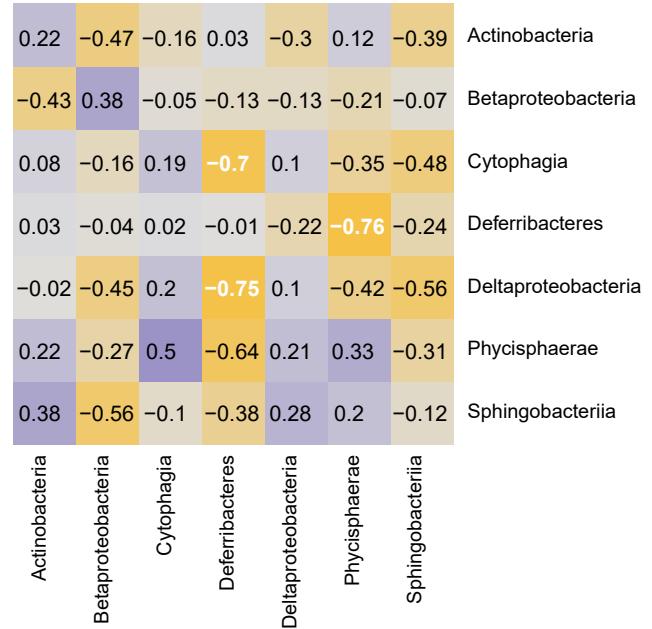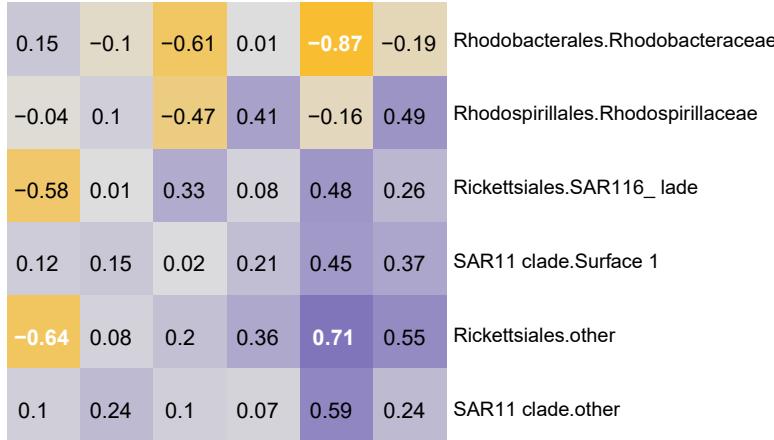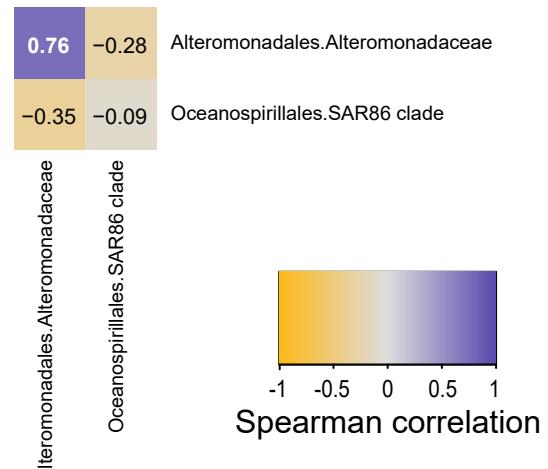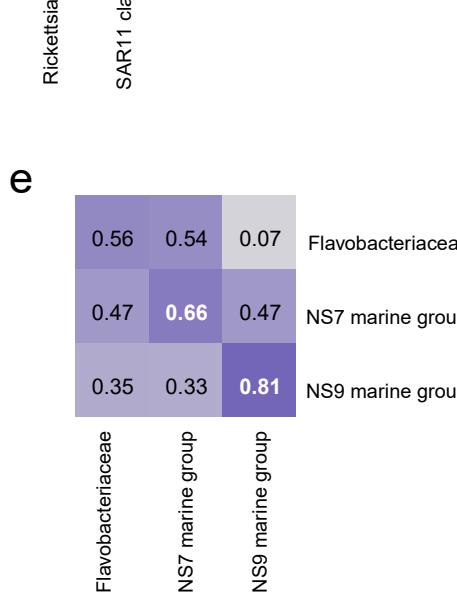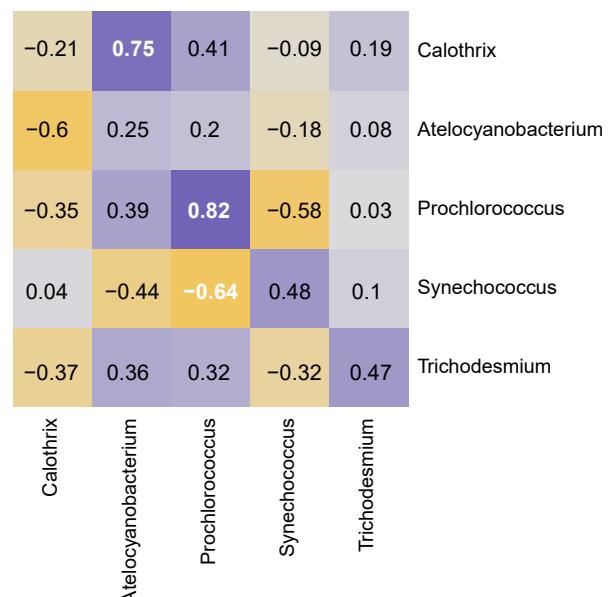
Supplementary Figure S3. *Defluviicoccus* sequences were downloaded with the SILVA browser from SSU version r123. Only sequences >1000 nt were selected, and from these, all sequences including `marine´,`deep-sea´, `soil´, or `symbiont´ were extracted plus 40 sequences that were not further classified (uncultured *Defluviicoccus*). These were aligned together with our *Defluviicoccus* OTUs using the SINA aligner (Pruesse et al, 2012), common gaps removed, and the alignment shortened to the length of the OTUs. The tree was built with Mr Bayes v3.2.5 (Ronquist et al, 2012), visualized in FigTree v1.4.2 (Rambaut & Drummond, 2009). The branches are weighted by posterior probability and the probablility of the basic nodes given in %. The summed up read count (normalized to sample total) is displayed together with the respective OTU ID. Our OTUs fall into a very distinct subcluster containing only marine *Defluviicoccus* 16S with one exception, which is annotated as coming from a soil sample.
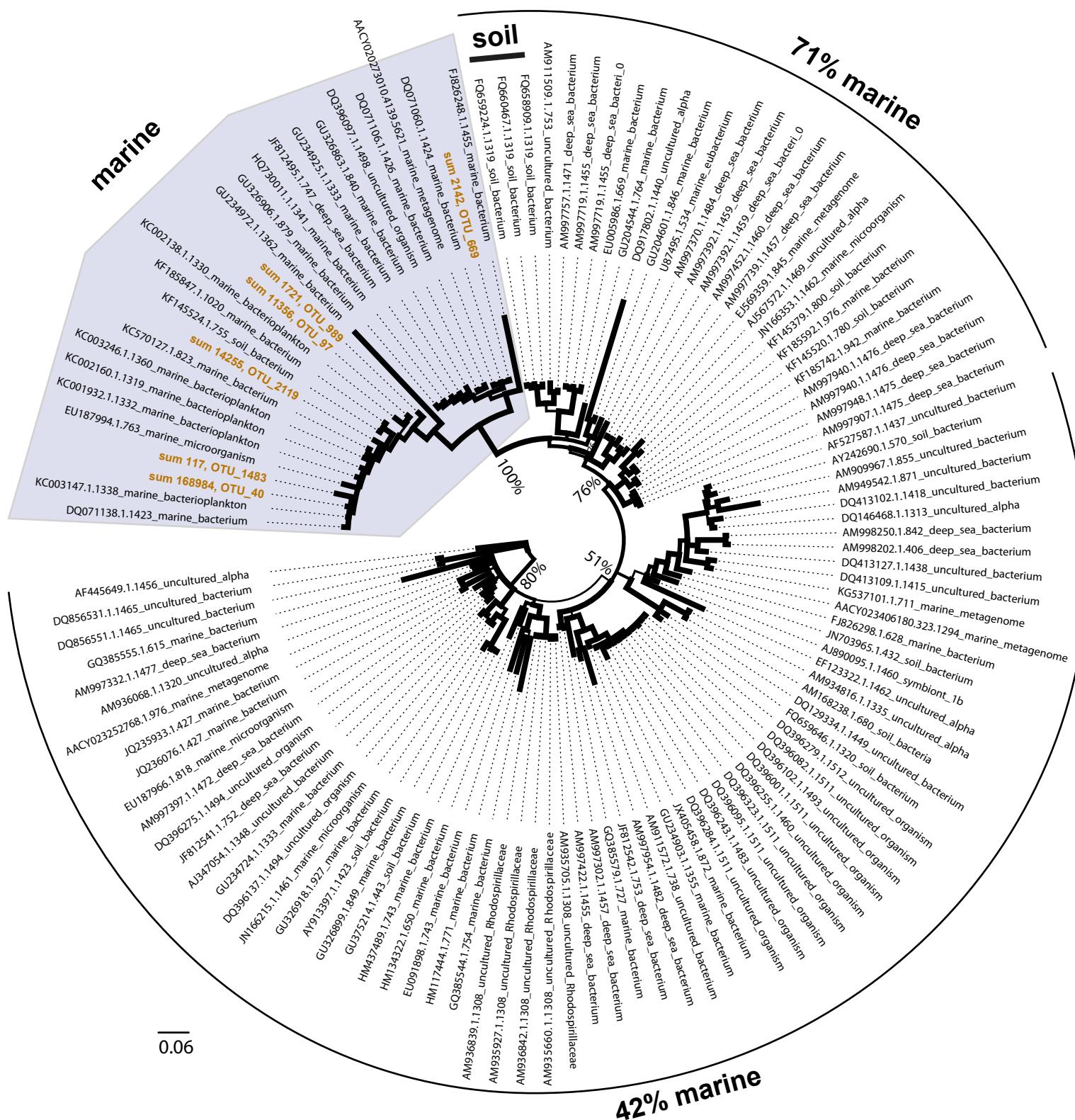
Supplementary Figures S4 and S5. These figures complement the inferred absolute abundances presented in the main figures 3 and 4 and show evolution of the 16S community with time for surface samples (1 m depth) inside mesocosm M1 and the Nouméa lagoon. All abundances are given as a percentage of total 16S reads. For better visualization, the classes with (a) a maximum abundance above 5 % and (b) a maximum abundance between 0.5–5 % were plotted separately. (c–e) Families with > 1% of total 16S reads (at their respective maximum) amongst the dominant non-photosynthetic classes Alphaproteobacteria, Gammaproteobacteria, and Flavobacteriia. (f) All genera > 0.05% of total 16S reads (at their respective maximum) within photosynthetic Cyanobacteria. Dashed lines were used to connect the data points for better visualization only.

Supplementary Figure S4: Mesocosm M1
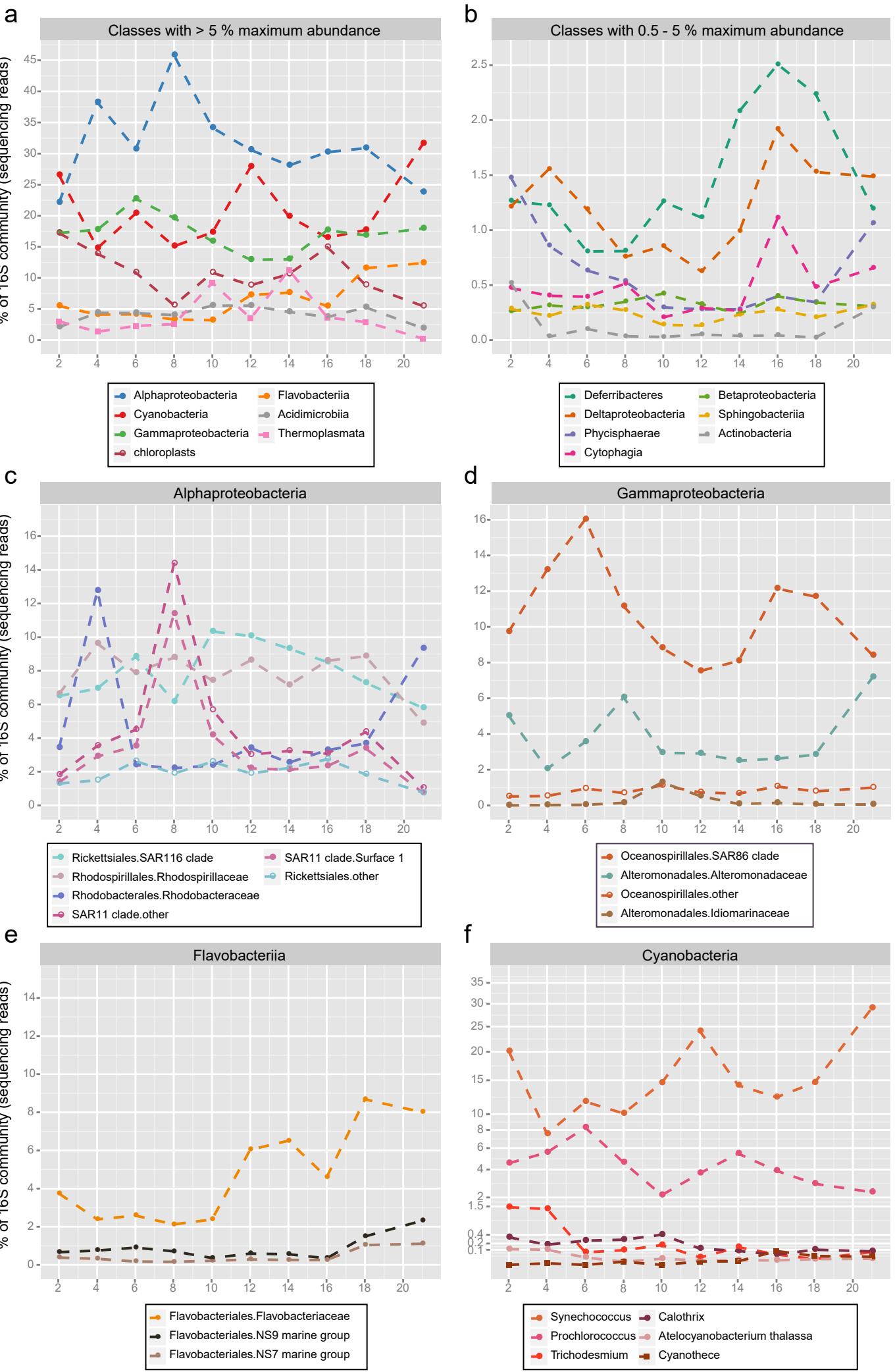
Supplementary Figure S5: Nouméa lagoon