**Functional and taxonomic diversity of a marine planktonic community**

An in depth description of the planktonic community at one site.

Effect of sequencing depth on community diversity



Fig. X. Relationship between prokaryotic community diversity (Shannon index) and sequencing depth (sequences per liter) for different size fractions. Communities were described by metabarcoding targeting the 16S rRNA gene. The 3-20 fraction had a significant linear regression. The horizontal bars represent the mean Shannon value.

For the 0.22 μm fraction, our study yielded 75 633 to 744 000 sequences per sample, and for the 0.2-3 μm fraction, 753 661 to 1 691 449 sequences. The 3-20 μm fractions had 1 159 442 to 2 747 157 sequences and the 20 μm had 713 112 to 1 824 136 sequences per sample (Sup Table X).

In order to compare the effect of sequencing depth on prokaryotic community diversity, we normalized the number of sequences per volume of filtered water. For both small fractions, an increase in sequencing depth did not increase significantly the diversity of the communities (stat results lm)(Fig. X). For the 0.22-3 μm fraction, however, the lowest Shannon value was measured with the lowest number of sequences (1 053 seq L-1) and the average Shannon value was reached with ca. 12 500 sequences L-1 . For the 0.22 μm filters, the sequencing effort was high (> 750 000 seq L-1) for all samples. For the larger size fractions, community diversity increased significantly with sequencing depth for the 3-20 μm size fraction (p<0.001) and average Shannon values were obtained at sequencing depth of ca. 12 500 sequences L-1. For the 20 μm fraction, diversity did not increase with sequencing depth and there was a high variability of Shannon values. High values were obtained with sequencing depth as low as 1650 sequence L-1 (Fig. X).



Fig. X. Relationship between prokaryotic community richness (number of ASVs) and filtered water volume. Communities were described by metabarcoding targeting the 16S rRNA gene. The horizontal bars represent the mean ASV value.

Their were no linear relationship between volume and community richness for all size fractions. Community richness varied, however, within all size fractions. For the small sizes, the 0.22 had highest variability within a same volume while the 0.22-3 um showed lowest variation (fig. X) . For the larger size fractions, volumes of 100L of filtered seawater showed highest variability.

Community diversity



Fig. X Diversity of prokaryotic communities described by 16S rRNA genes obtained from metabarcoding (metab) or metagenomics (metag) from different size fractions: 0.22μm (0.2), 0.22 to 3 μm (0.22-3), 3-20 (3-20) and 20 μm (20).

In the small size fraction (<0.22μm), for the prokaryotes there was no significant differences between the diversity found in the whole fraction versus the prefiltered one (t-test, p=0.3) (Fig. X). The estimated richness of the small size fraction was 1131 ASVs (Chao1 estimator). The richness measured with the mitags extracted from the metagenomes was much higher and reached an average of 3318 ASVs. For the large size fraction (>0.3 μm), the diversity of the samples was more variable for the 20 μm fraction for both the metabarcoding and the mitag approaches (Fig. 1). The average richness was 1297 and 1104 in the metabarcoding approach for the 3-20 and 20 fractions respectively, and 2137 and 1238 in the mitag approach for the for the 3-20 and 20 fractions respectively.



Fig. X. Diversity estimate (chao-1) of prokaryote and eukaryote communities in the small (<02um) and large size fractions (>3um). Communities were described by 16S rRNA and 18S rRNA miTags extracted from metagenomes.

For the prokaryotes, the small size fraction community diversity was higher than the large size fraction community diversity (Fig. X) (t test, p<0.001). For the eukaryotes, the large size fraction had highest diversity (p<0.001). Small prokaryote size fraction had higher diversity than eukaryotes and large eukaryotes had higher diversity than large prokaryotes (p<0.001) (Fig. X).

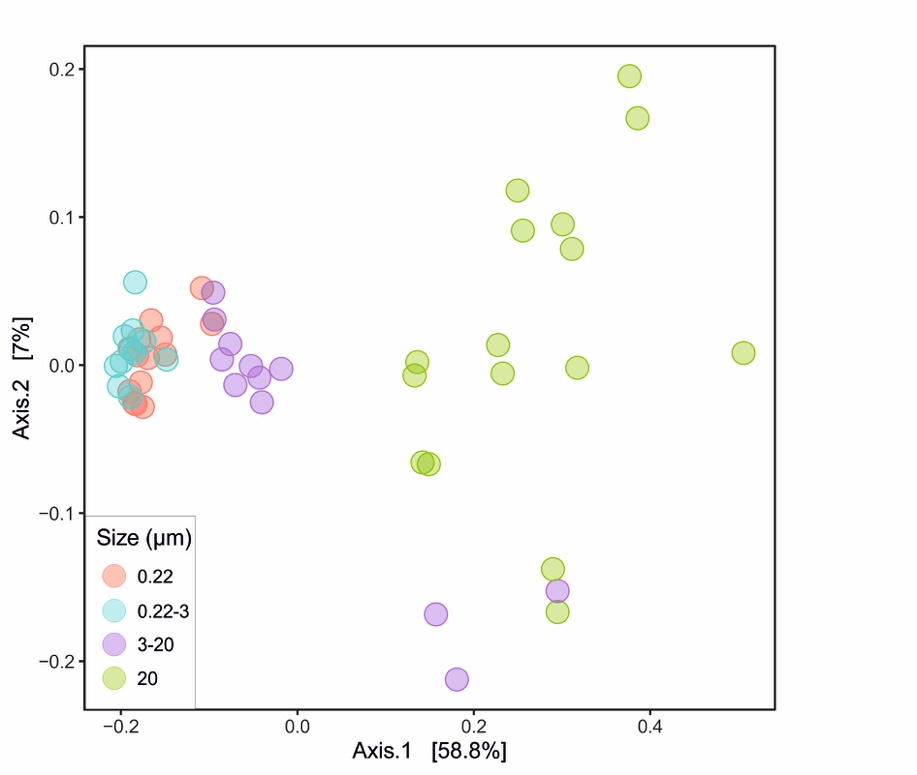


Fig. MDS ordination base on Bray Curtis dissimilarity comparing the prokaryotic community composition of samples from different size fractions. Communities are describes by metabarcoding targeting the 16S rRNA gene.

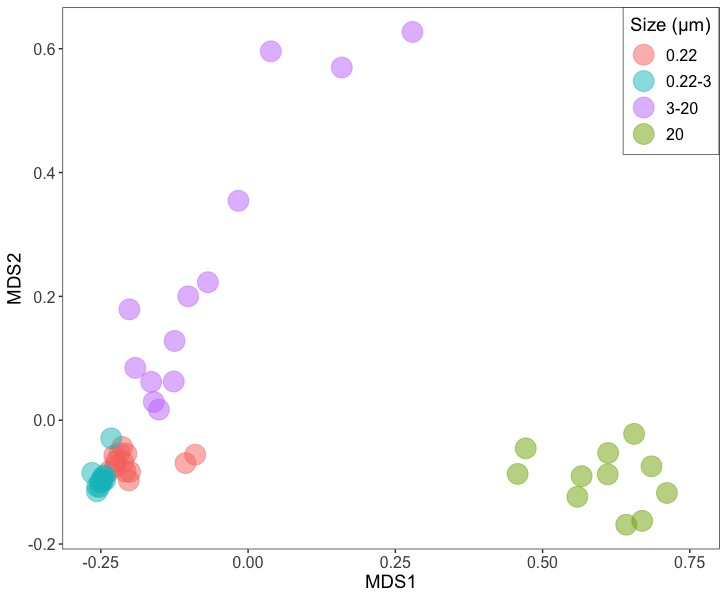


Fig. MDS ordination base on Bray Curtis dissimilarity comparing the entire metagenomes Kmer (21) composition for samples from different size fractions.

The comparison of community composition showed a clear separation by size fraction (Fig. X). The 0.22 and 0.22-3 fractions were similar, separated from the 3-20 and the 20 μm fraction. The 0.22-3 communities were less heterogenous in their composition followed by the 0.22 fraction. The largest size fraction (20 μm) was the most heterogeneous (Fig X).

Table Adonis (permanova)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| pairs | Df | SumsOfSqs | F.Model | R2 | p.value | p.adjusted |
| S02vsS20 | 1 | 2.60242755 | 45.62883 | 0.61971418 | 0.001 | 0.006 |
| S02vsS023 | 1 | 0.06240553 | 2.479973 | 0.09733027 | 0.087 | 0.522 |
| S02vsS320 | 1 | 0.43097374 | 8.099587 | 0.25232683 | 0.001 | 0.006 |
| S20vsS023 | 1 | 2.68133295 | 49.903199 | 0.66623588 | 0.001 | 0.006 |
| S20vsS320 | 1 | 1.20807119 | 15.385492 | 0.37176052 | 0.001 | 0.006 |
| S023vsS320 | 1 | 0.50057558 | 10.272539 | 0.32848434 | 0.001 | 0.006 |



Fig. Betadispersion of the prokaryotic community composition obtained on different size fractions and described by metabarcoding the 16S rRNA gene.

The 02-3 um size fraction communities were the less dissimilar, followed by the 02 fraction. The large size fraction were more variable.

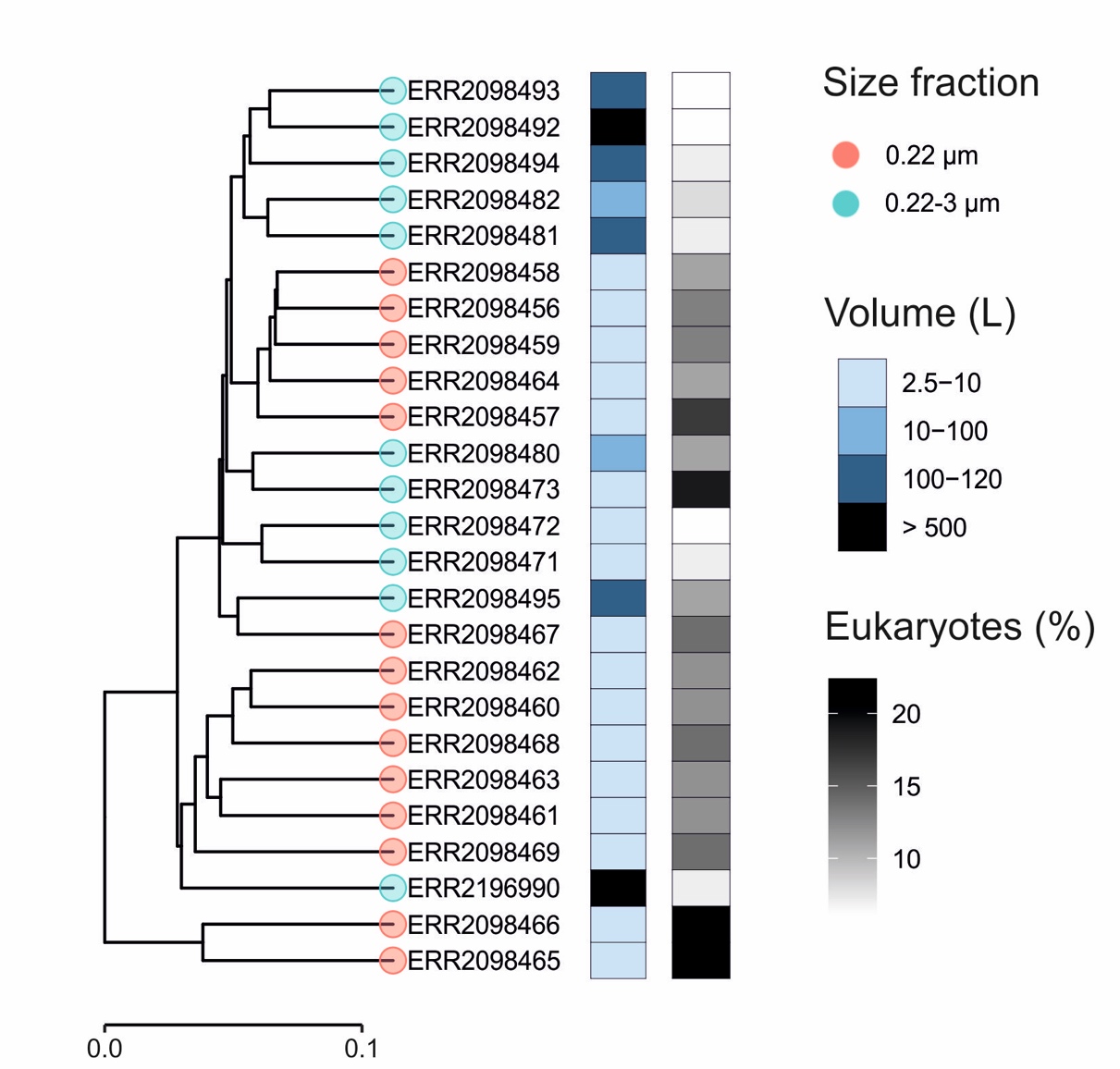


Fig.X. Bray Curtis dissimilarity between prokaryotic communities filtered on 0.22 and 0.22-3 μm filters and described by metabarcoding the 16S rRNA gene. The comparison was done after removing chloroplastic eukaryote sequences from the dataset. The proportion of eukaryote sequence present in the original dataset is shown in grey. The volume filtered is shown in blue.

When focusing on the small fraction (<0.22μm), the comparison between communities showed that the 0.22 μm fraction was split between 3 clusters. The 2 most distinct samples had the highest proportion of Eukaryotic sequences in their original sequence pool. The 0.22-3 μm fractions samples were also separated in 2 clusters. One contained samples obtained from volumes >10L and the other had samples from smaller volumes (Fig. X).

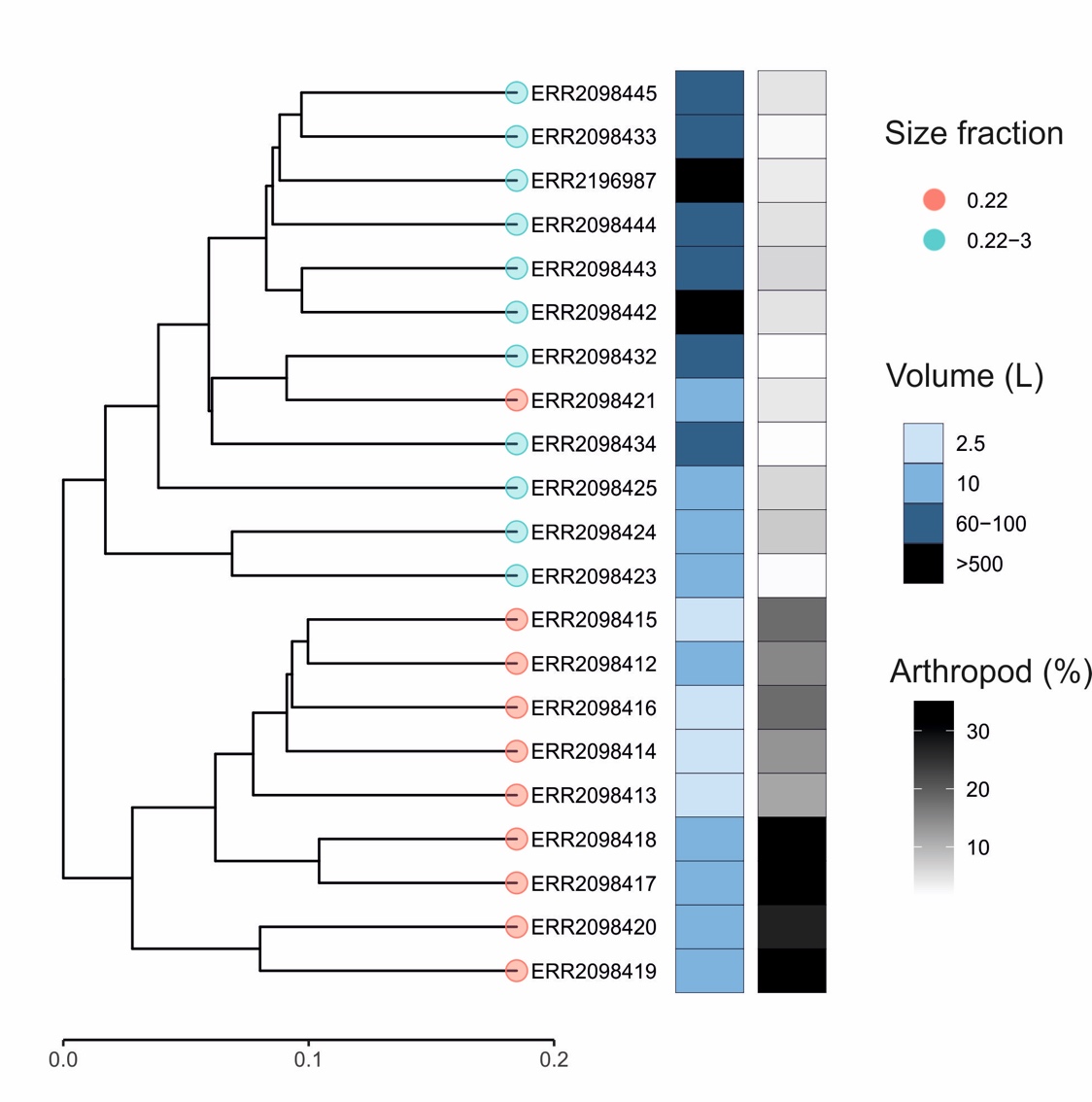


Fig.X. Bray Curtis dissimilarity between eukaryotic communities originating from 0.22 and 0.22-3 μm filters and described by metabarcoding the 18S rRNA gene. The proportion of Arthropod sequences present in the original dataset is shown in grey.

The similarity analysis on the small fraction of Eukaryotes showed that the samples separated in 2 main clusters, one containing all 0.22-3 um communities and one containing all, but one, 0.22 um samples (Fig. X). The 0.22 fraction had higher proportions of Arthropod sequences.