**Evaluation of sequential filtration and centrifugation to capture environmental DNA and survey microbial eukaryotic communities in aquatic environments**

**Running Title:**

Capture methods to survey protist diversity

**Authors:**

Ariane Atteia1\*, Béatrice Bec2, Camille Gianaroli1, Ophélie Serais1, Isaure Quétel1, Franck Lagarde1, Angélique Gobet1\*

1MARBEC, Univ Montpellier, CNRS, Ifremer, IRD, Sète, France

2MARBEC, Univ Montpellier, CNRS, Ifremer, IRD, Montpellier, France

**Supplementary data (excel file: 231010\_PICOTHAU\_Methods\_Supp.xlsx, with one table per sheet)**

**Supplementary table 1**. Total number of sequences retrieved after each step of the dada2 pipeline: input, total number of sequences retrieved after sequencing; filtered, total number of sequences retrieved after sequence filtration and trimming; denoisedF, total number of sequences retrieved after removing sequencing errors for the forward sequences; denoisedR, total number of sequences retrieved after removing sequencing errors for the reverse sequences; nonchim, total number of sequences retrieved after removing chimeras.

**Supplementary table 2**. Abundance table of the divisions identified in the surface waters of the Thau lagoon using the two capture methods. This table was used to produce figure 1.

**Supplementary table 3.** Abundance table representing the genera of the five most abundant divisions identified in the surface waters of the Thau lagoon using the two capture methods. This table was used to produce **figure 2 and supplementary figure 2**.

**Supplementary table 4.** Pairwise PERMANOVA comparing the sample dispersion for the two filtrated fractions and centrifugation.



**Supplementary figure 1. Taxonomic composition at the genus level of the eukaryotic community using the two methods for two of the five most abundant divisions. Metazoa (A) and Cryptophyta (B).** Taxonomic composition of each sample was surveyed on the dataset normalized by the number of sequences in each sample using median sequencing depth. For each division, genera were represented by their sequence relative abundance in a given size-fraction or in the pellet. Genera were ordered by sequence abundance within each class level. Genera with a lower number of sequences were grouped as “Other” for each class and division level. The list of “Other” genera can be found in the Supplementary table 3. Genera identified as NA were not taxonomically identified in the database PR2 and genera identified as \_XX were identified in the PR2 database but not yet described at this taxonomic level in the database. Por., Porifera; Cni., Cnidaria.

****

**Supplementary figure 2. Alpha-diversity measures for the different eDNA capture methods.** Observed richness, the Shannon index and the Simpson index were calculated for all samples and according to each method. An ANOVA followed by a post-hoc Tukey test indicated significant mean differences between the observed diversity in the 0.2-3 µm fraction and centrifugation while the other diversity measures did not show any significant differences between the methods.