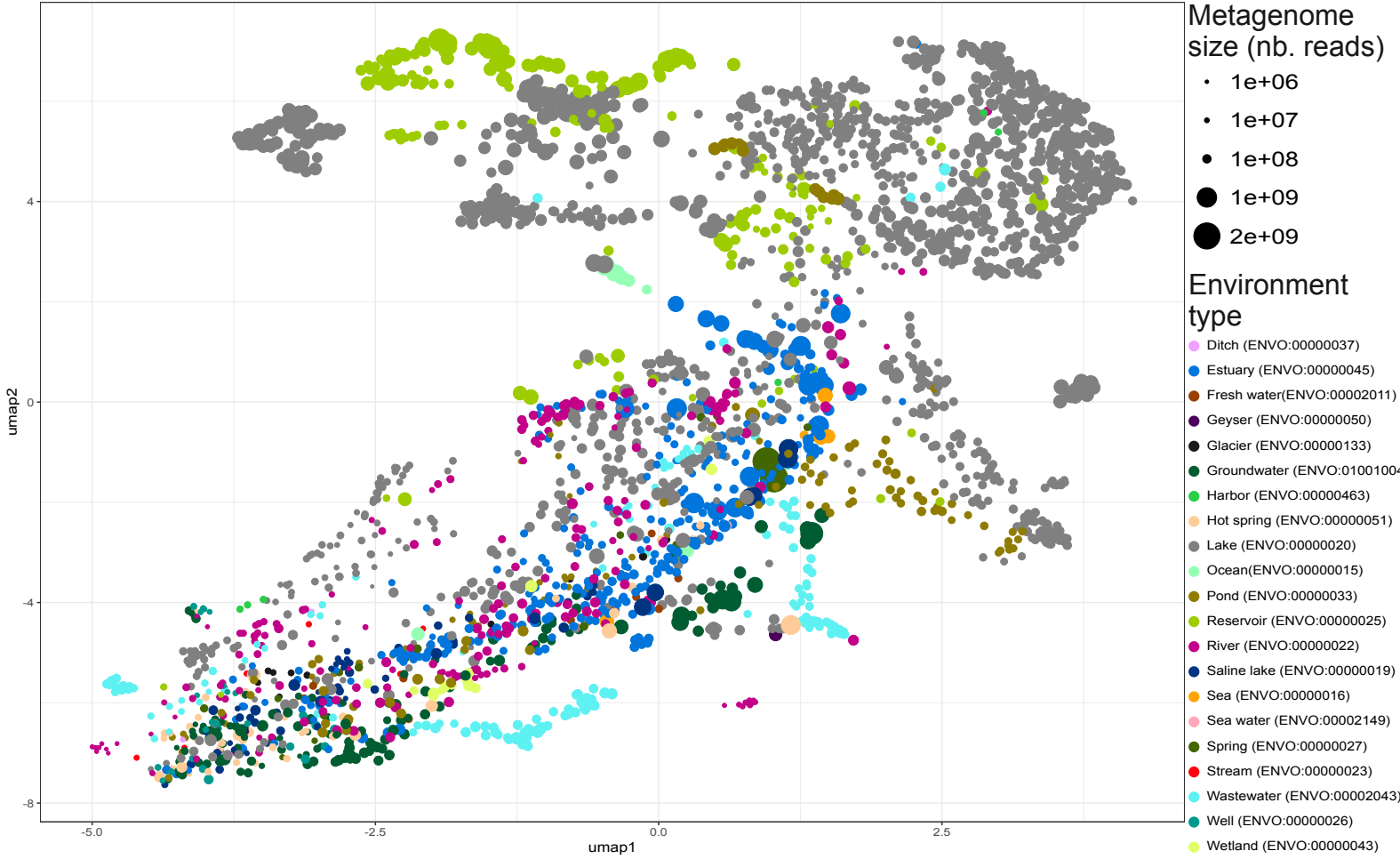
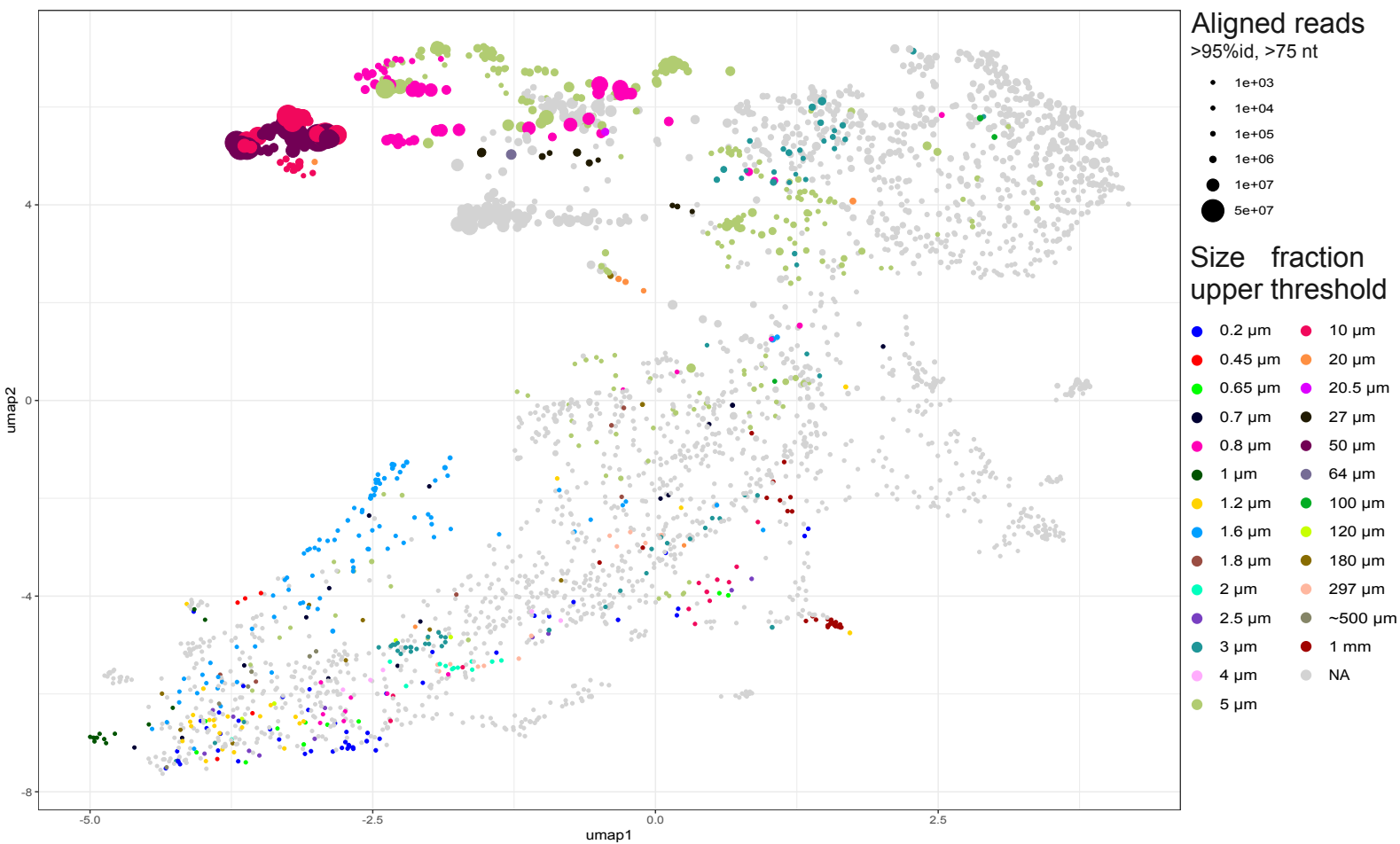
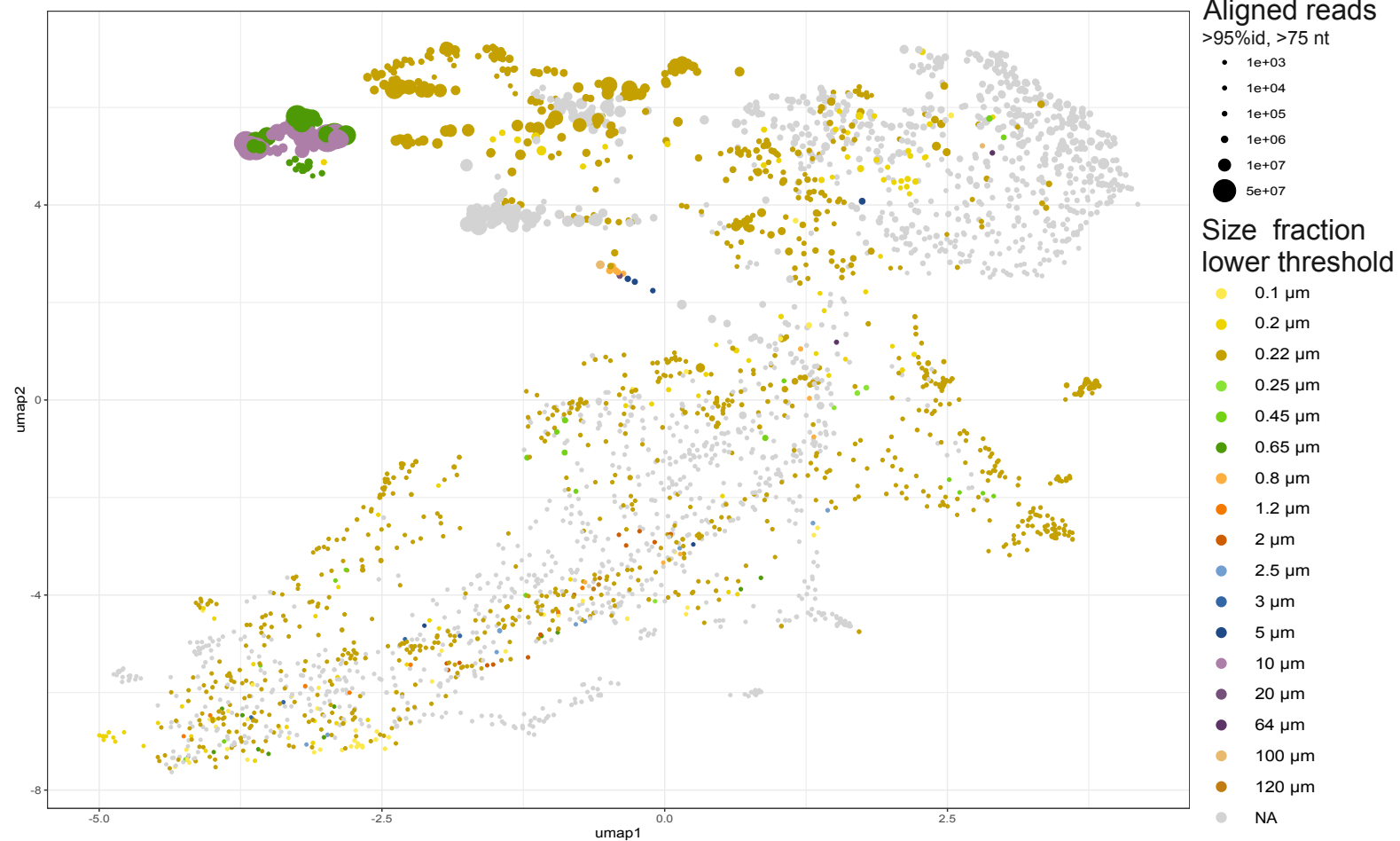


UMAP based on CLR-transformed raw read count of metagenomes mapped on MAGs and SAGs

Parameters : 25 neighbors, minimal distance 0.25





Supplementary Figure 1: UMAP dimension reduction of CLR-transformed raw read count from metagenomes mapped on MAGs and SAGs.

All panels represent the same UMAP, with different associated metadata. The UMAP was constructed with 25 neighbors and a minimal distance of 0.25, based on the CLR-transformed read count, with the UMAP *R* package¹. a, clusters of metagenomes based on the Aitchison distance (Euclidean distance on CLR-transformed data). b, the size (number of reads) and the environment type of each metagenomic dataset. c, the lower threshold of the size fraction (*i.e.* the filter size on which the sample was collected), and the number of read that aligned of the MAGs and SAGs. d, the upper threshold of the size fraction.

1. Konopka, T. Umap: Uniform Manifold Approximation and Projection. (2023).