



**Figure S2. Highly abundant retrotransposon/retrovirus-related ORFs in the metagenomic data from large size fractions.** ORFs predicted from the metagenomic assemblies were pooled together in this analysis if they originated from the same size fraction, irrespective of sampling depth (i.e., SUR or DCM). Five hundred CDD profiles that acquired the highest number of ORF assignments are shown on the left panels, while the forty best CDD profiles are shown on the right panels (with their CDD profile accession numbers). Red bars correspond to profiles related to retrotransposons/retroviruses, while black bars represent other CDD profiles. Sequence data from two whole genome amplified (WGA) samples (St TARA\_007/SUR or DCM/5-20  $\mu\text{m}$ ) are not included in this analysis.