

Figure S1. Complete protein sequence alignment of 6 proteorhodopsin and 6 FARhodopsin sequences. Sequences come from 6 SAGs (two of each taxon) containing both rhodopsins in their genomes. Dark and light grey regions represent conserved residues within the alignment.

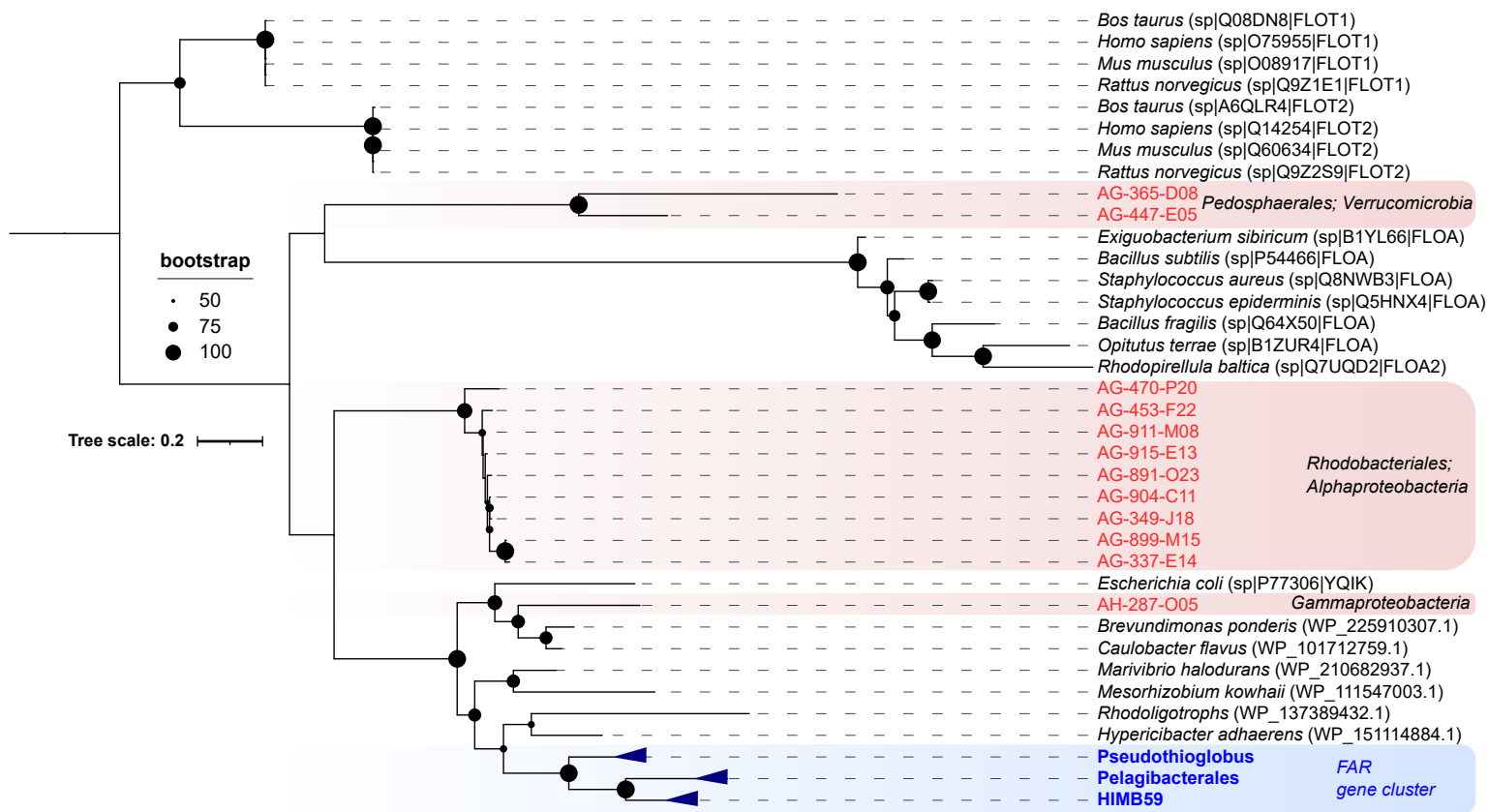


Figure S2. Maximum-likelihood phylogenetic tree of flotillin genes retrieved from marine SAGs. Reference flotillin sequences and their accession numbers within brackets are colored in black. Sequences in red represent sequences from marine SAGs that were not found near a FARhodopsin. Sequences in blue correspond to the flotillins of the three taxa coding for the tandem flotillin-FARhodopsin. Sequences in this branch were condensed for simplicity.

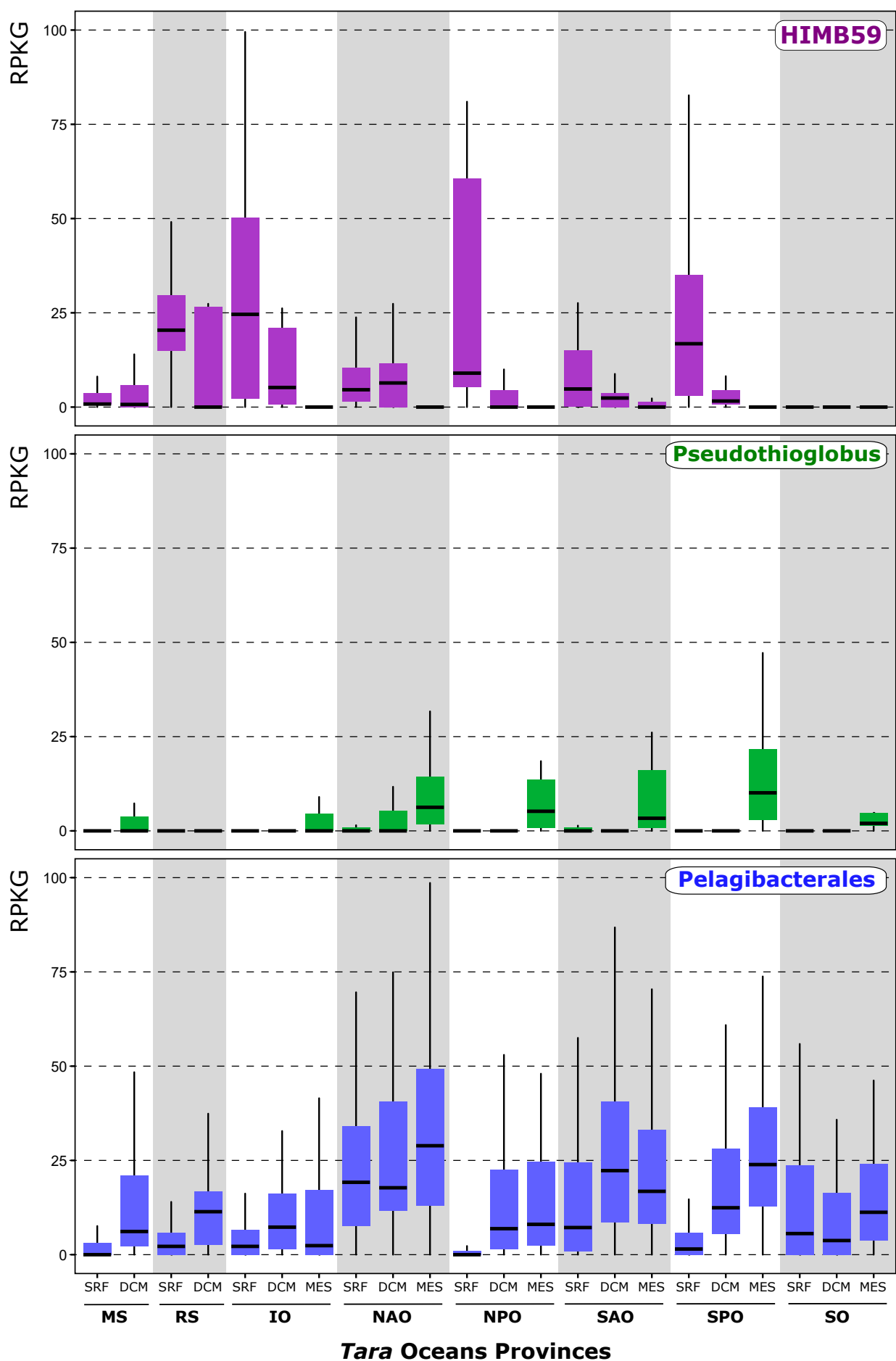


Figure S3. Relative abundance (measured in RPKG) of the FAR gene cluster in Tara Ocean metagenomes. Each province has been divided by depth in surface (SRF), deep chlorophyll maximum (DCM) and mesopelagic (MES) regions. The bars have been colored according to the microbes (blue, Pelagibacteriales; purple, HIMB59 and green, Pseudothioglobus). MS -- Mediterranean Sea; RS -- Red Sea; IO -- Indian Ocean; NAO and SAO -- North and South Atlantic Ocean; NPO and SPO -- North and South Pacific Ocean; SO -- Southern Ocean.