

Supplementary Data Legends

File name: Supplementary Data 1.

Description: Summary of 16S rRNA gene amplicon data from HTC17 and HTC18.

File name: Supplementary Data 2.

Description: Detailed information for isolate genomes reported in this study. The genome summary information originated from checkM (v1.1.2). In the “# contigs” column, (*) indicates the genome has been manually verified to be closed.

File name: Supplementary Data 3.

Description: Summary statistics for genomes used in this study, including new isolate genomes reported here, previously published isolate genomes, and high-quality single amplified genomes (SAGs) used in the extended SAR11 phylogeny. The genomes used in read-recruitment are indicated. In the “Accession” column, (*) indicates the accession is the IMG Genome ID instead of the NCBI Accession.

File name: Supplementary Data 4.

Description: Genes included in the SAR11_165 gene set.

File name: Supplementary Data 5.

Description: Final 95% gANI clusters including cluster number, final representative genome, and other genomes within each cluster.

File name: Supplementary Data 6.

Description: Average genome statistics for 24 monophyletic clusters within the *Pelagibacteraceae* as well as the Ic (*Mesopelagibacteraceae*), II (*Cosmipelagibacteraceae*), and III (*Fontibacteriaceae*) families.

File name: Supplementary Data 7.

Description: Studies used as a source for metagenomes.

File name: Supplementary Data 8.

Description: Summary of competitive read recruitment.

File name: Supplementary Data 9.

Description: Coverage values from competitive read recruitment.

File name: Supplementary Data 10.

Description: Detection values from competitive read recruitment.

File name: Supplementary Data 11.

Description: Number of metagenomes per bin included in Figure 2.

File name: Supplementary Data 12.

Description: Average detection across genome clusters of bins included in Figure 2.

File name: Supplementary Data 13.

Description: Summary of type genomes and classification hierarchy for the *Pelagibacterales*.

File name: Supplementary Data 14.

Description: Average coverage values for genomes HIMB1556 and GCA_902533445.1 across BATS, Station ALOHA, and KByT metagenomes included in Figure 3.

File name: Supplementary Data 15.

Description: Gene IDs for genomes HIMB1556 and GCA_902533445.1.

File name: Supplementary Data 16.

Description: Summary of RED group classifications and the GTDB-based classifications of genomes included in the pruned SAR11 phylogeny.