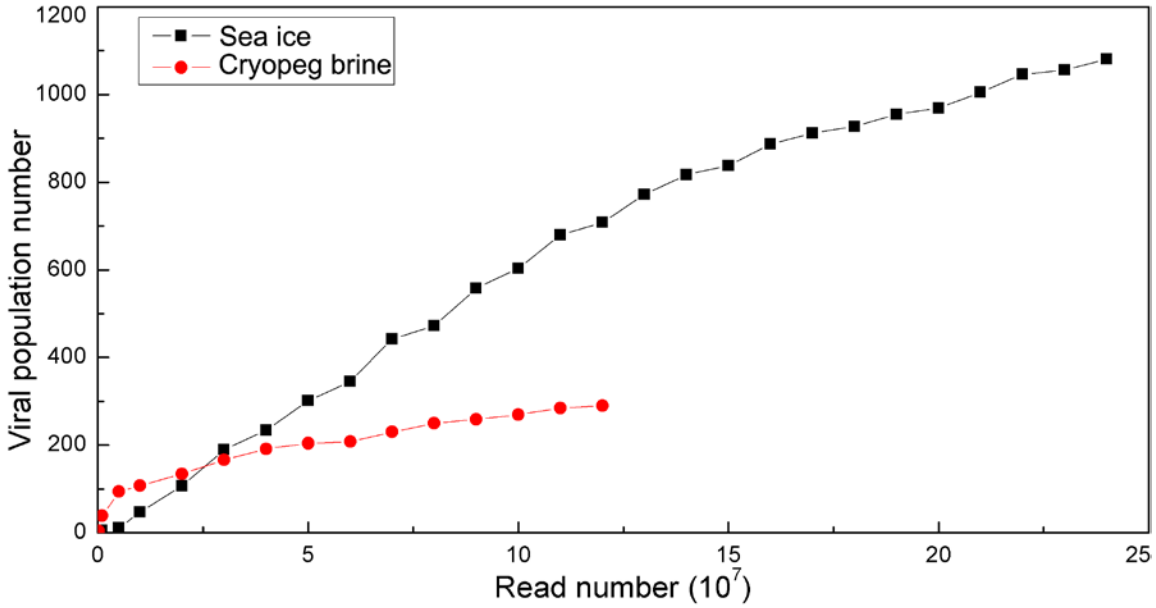
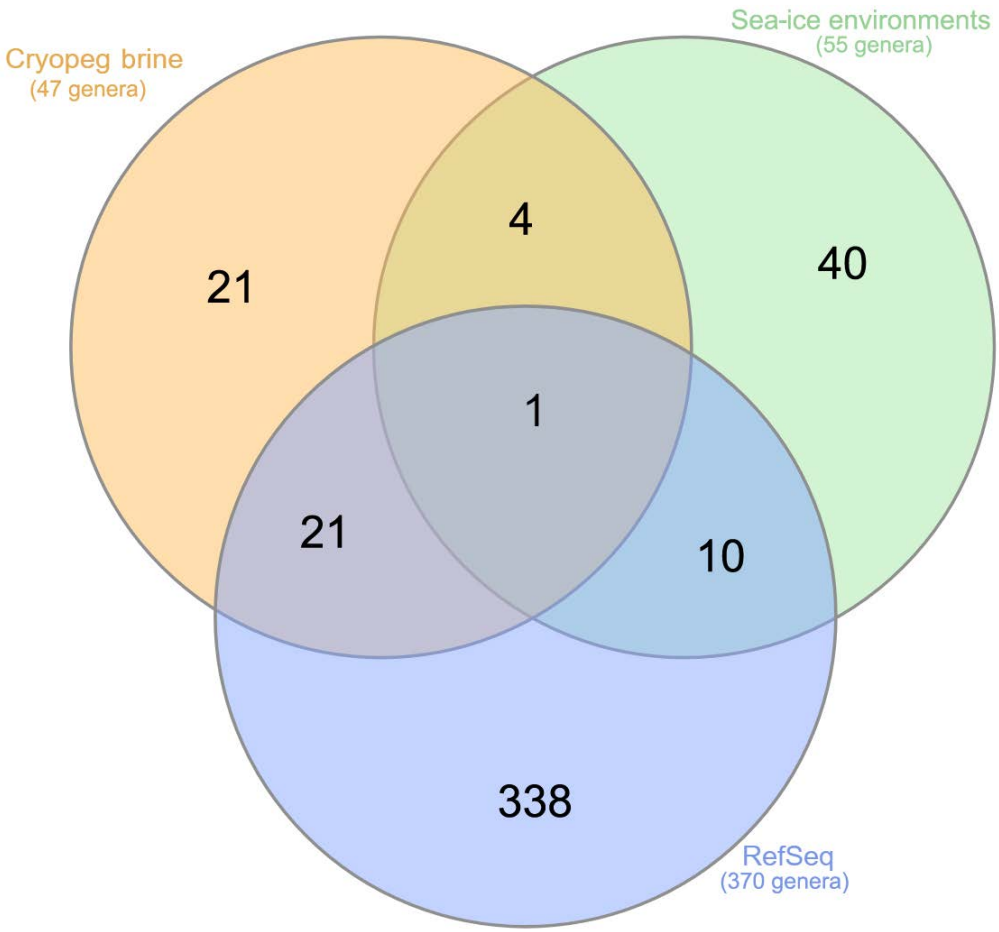
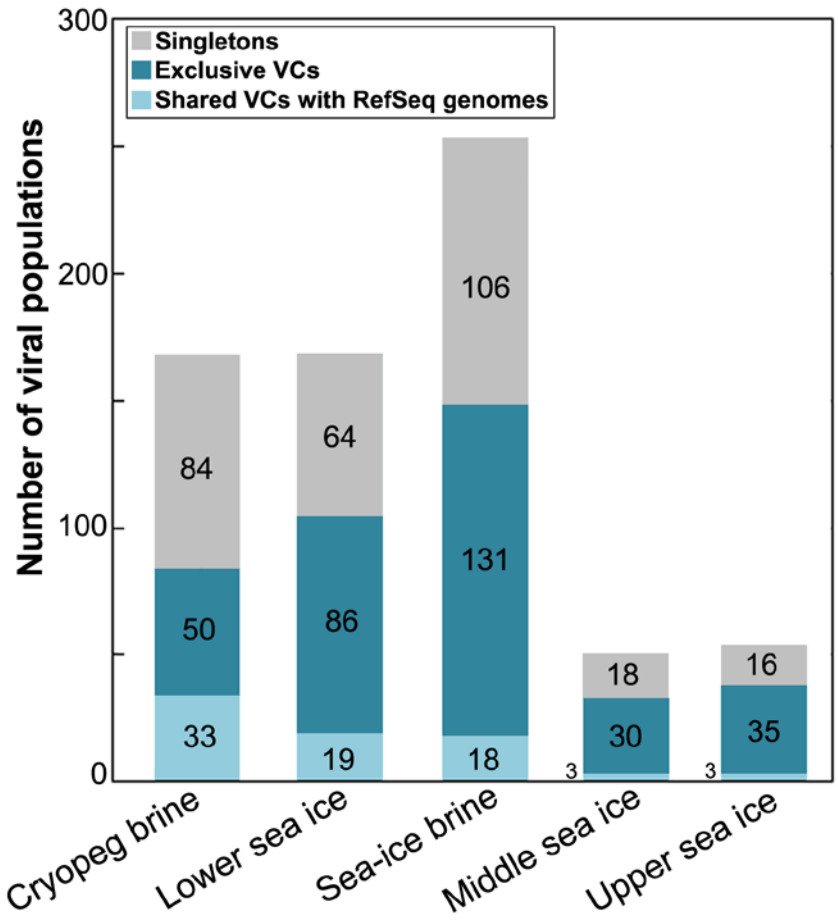


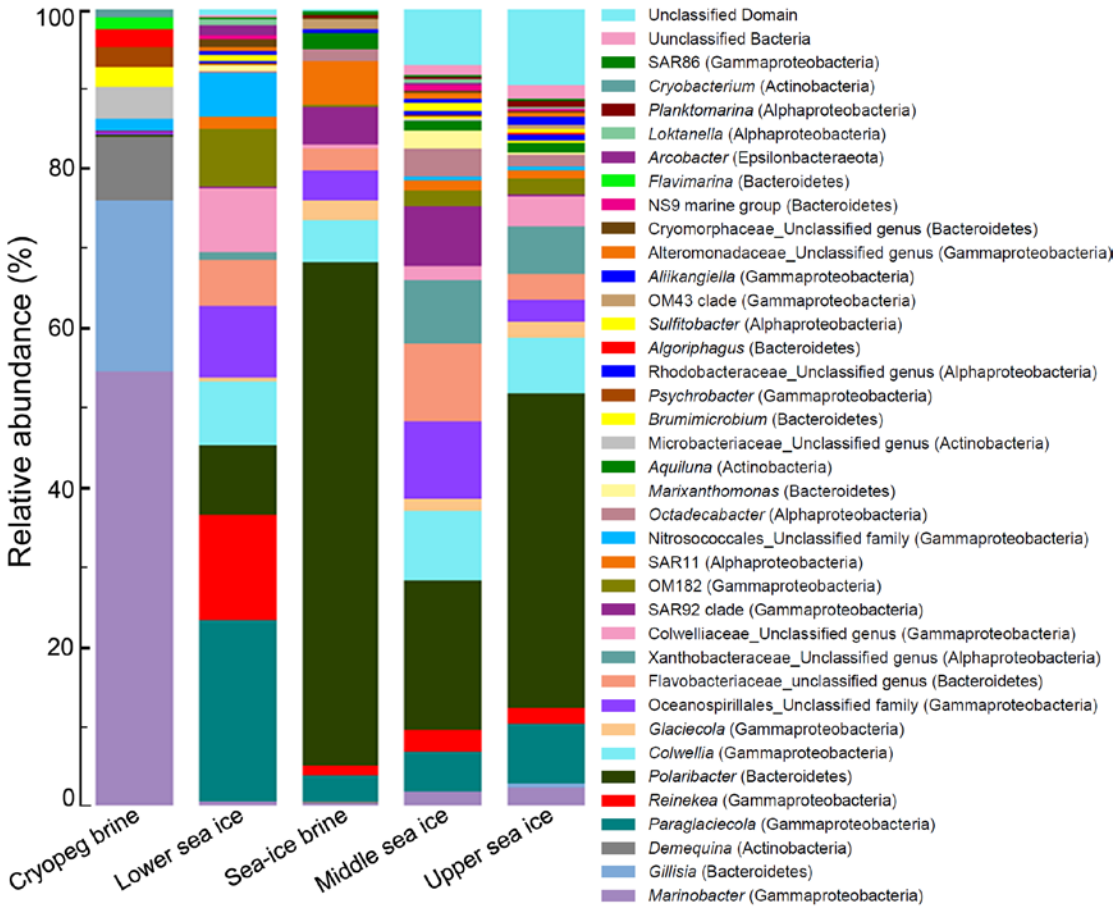
Supplementary Figure 1. Rarefaction curves of sea ice and cryopeg brine samples by viral population numbers. Samples collected from the same type of environment (sea ice) were combined.



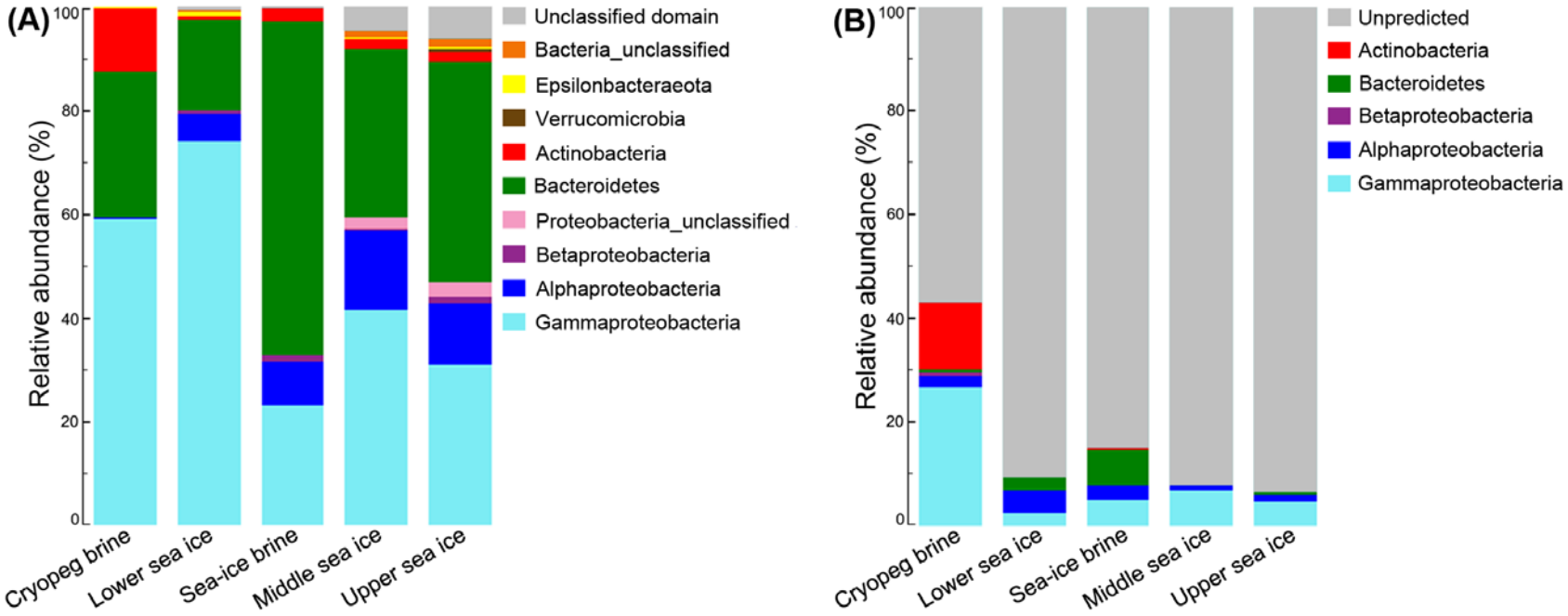
Supplementary Figure 2. Numbers of shared and specific viral genera from cryopeg and sea-ice environments and RefSeq database. Each viral genus represents a viral cluster. Data from sea-ice samples were combined to compare cryopeg and sea-ice habitats.



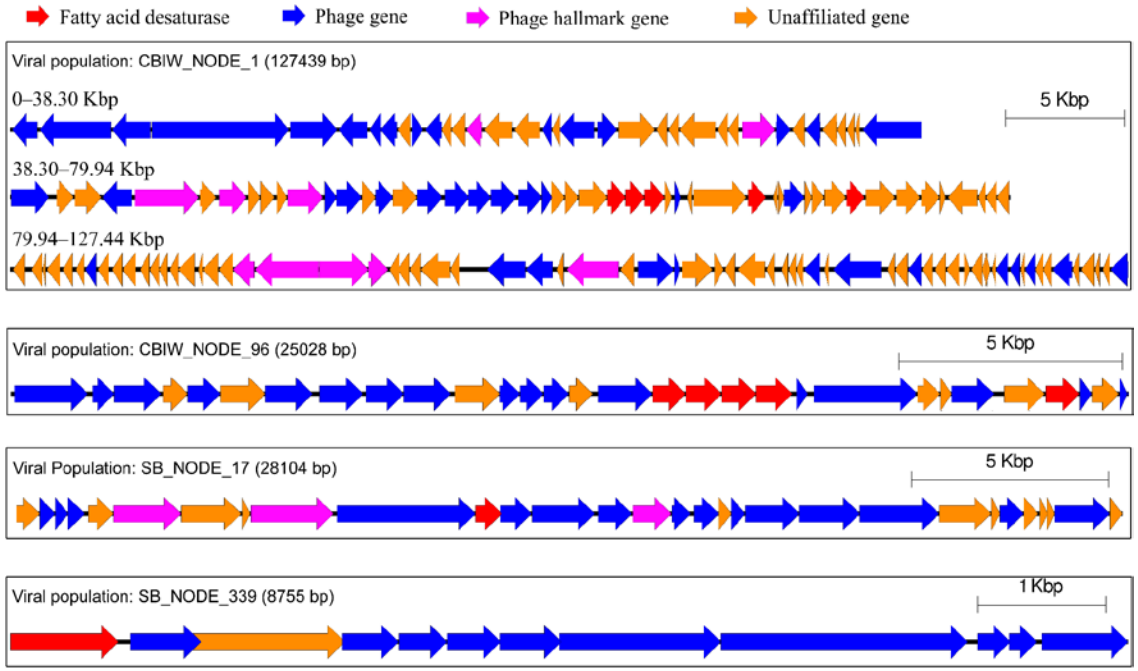
Supplementary Figure 3. Relationship between Utqiaġvik viruses and known viruses of RefSeq database. The bars show the numbers of viral populations detected from the field samples. The quality-controlled reads of each sample were mapped to these population contigs to count the presence of populations in each sample (see Methods). These populations were classified into three groups of VCs: singletons (gray) that had no close relatives, exclusive VCs (aqua) that comprised field populations only, and shared VCs with RefSeq genomes (sky blue) that comprised field populations with taxonomy assignments and genomes from RefSeq.



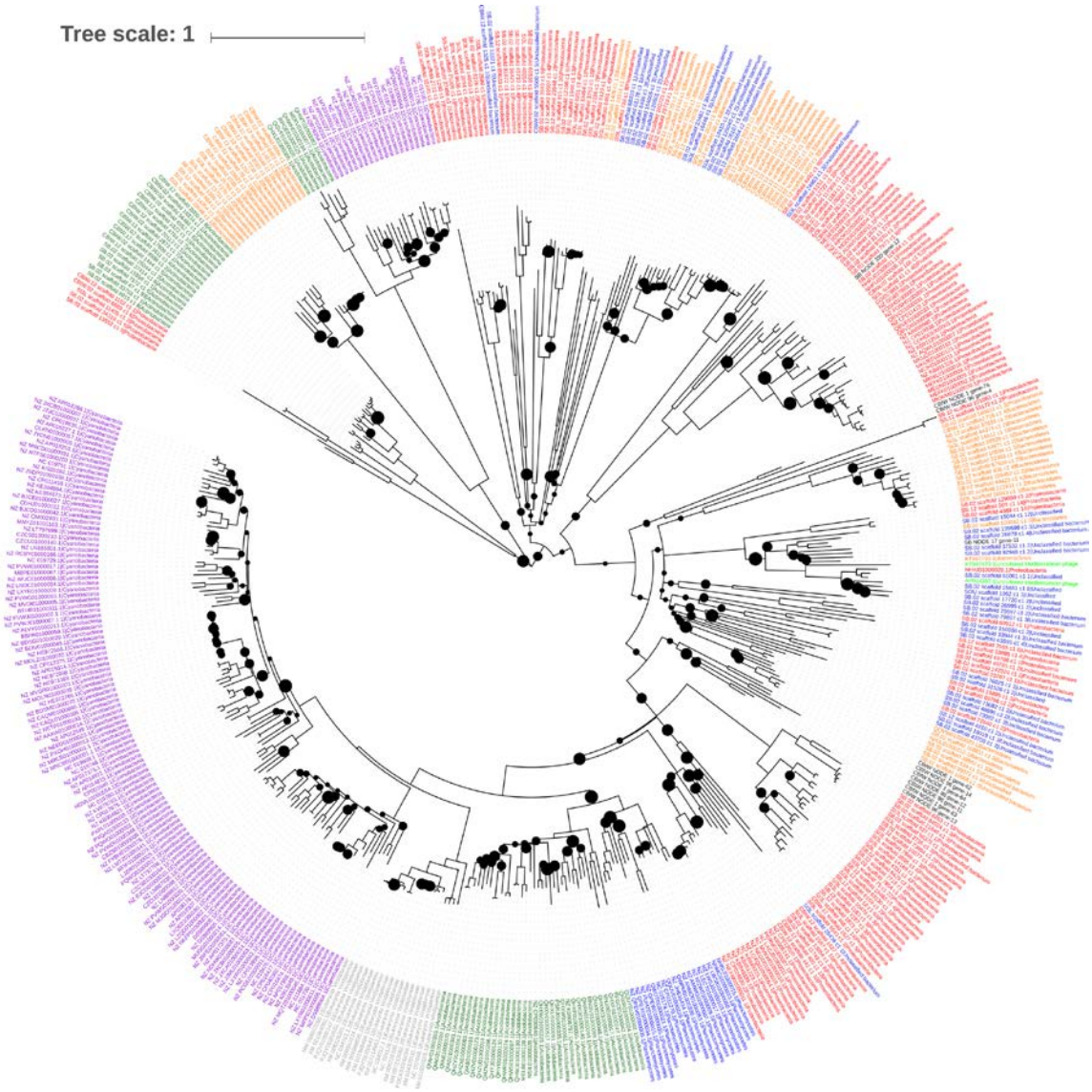
Supplementary Figure 4. Microbial community structure of the most abundant genera for the five field samples. Only genera with relative abundances > 0.5% in at least one sample were selected; these accounted for 92.3–97.4% of each community. The phylum (or class within Proteobacteria) for each genus is given in parentheses.



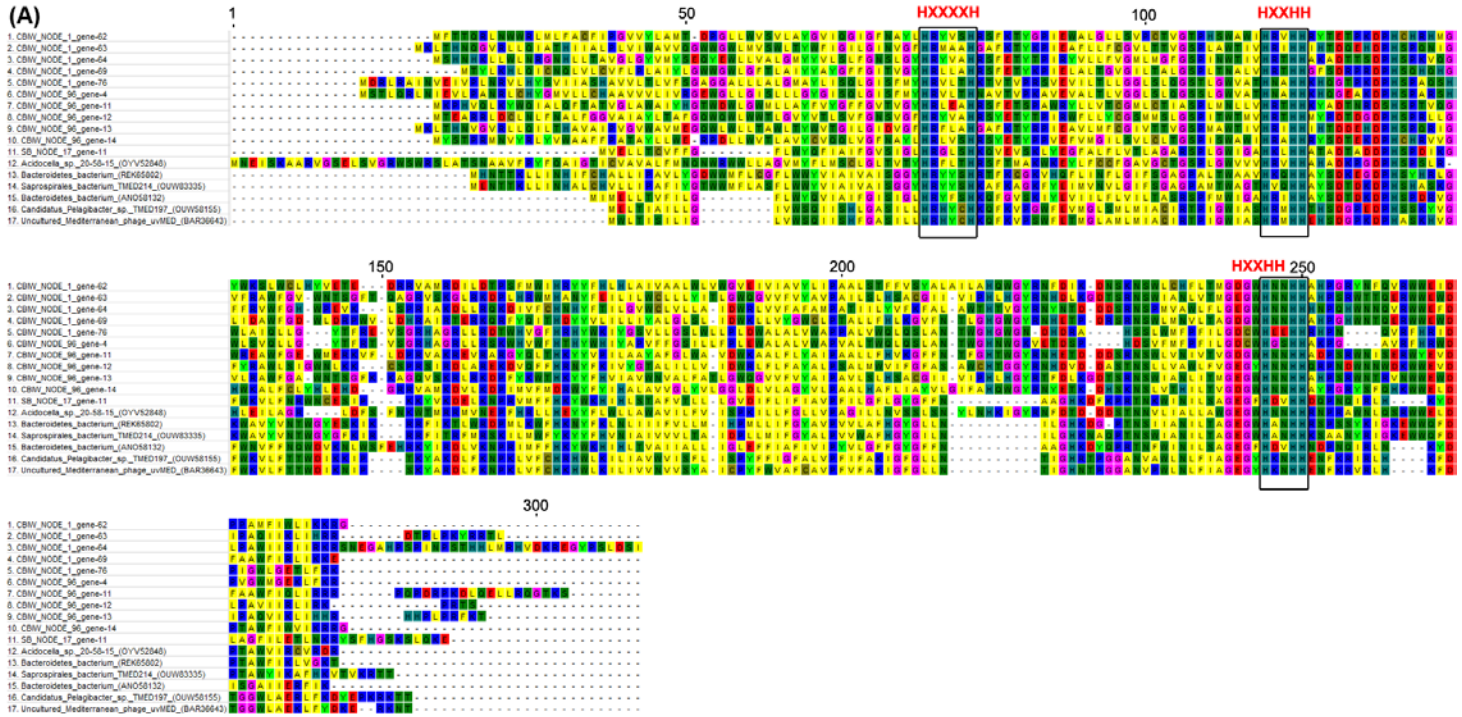
Supplementary Figure 5. Virus–host linkages and microbial profiles at the phylum (or class within Proteobacteria) level. (A) Microbial community structures are shown as the seven most abundant phyla (and four classes within the Proteobacteria, > 0.1% of relative abundance in at least one sample), which constituted > 99.6% of each community**. (B)** Summed relative abundances of viral populations associated with their hosts as predicted by VirHostMatcher.

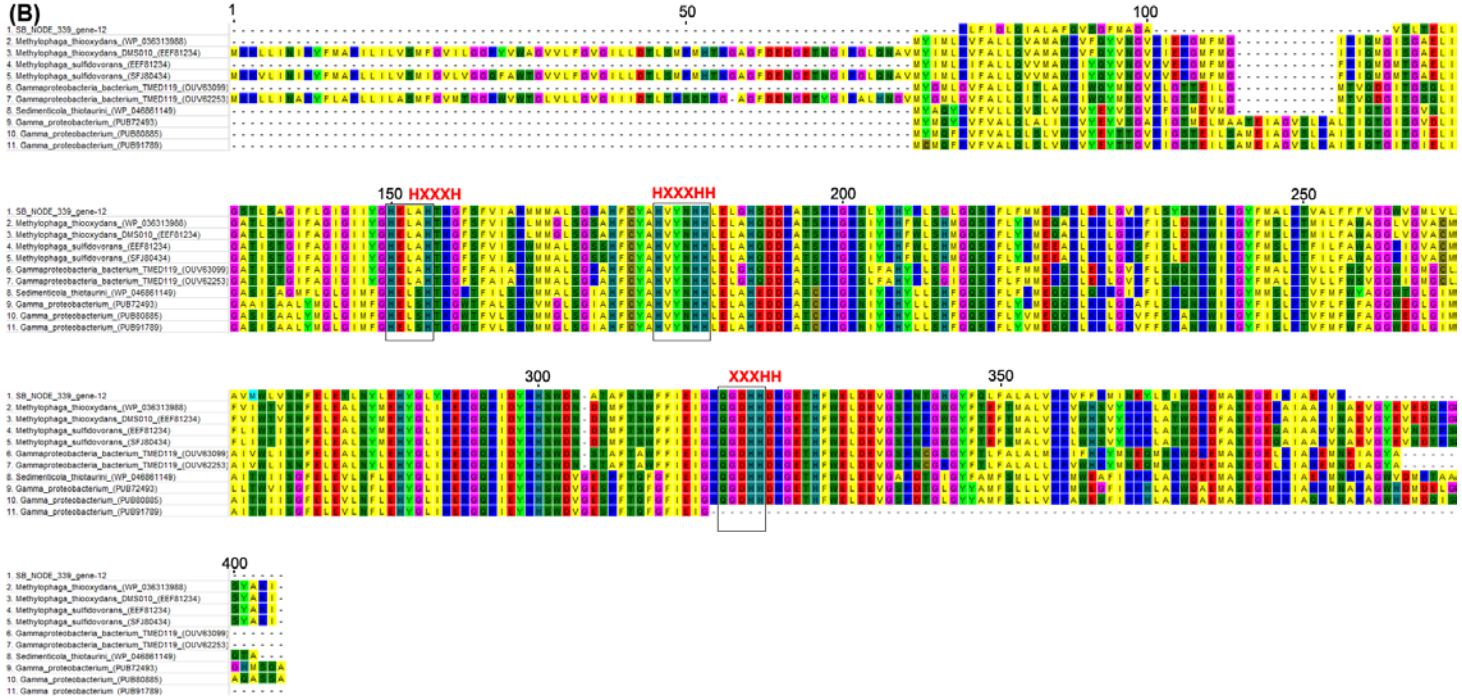


Supplementary Figure 6. Genomic maps of four viral populations containing fatty acid desaturase (*FAD*) genes. Genes were marked by four different colors to illustrate *FAD* genes (red), phage genes (blue), phage hallmark genes (purple), and unaffiliated genes (orange). The latter three groups of genes were classified by comparing their predicted protein sequences to those of viral genes in the extended RefSeqABVir database by VirSorter v1.0.3 in virome decontamination mode. Genes were marked as “phage genes” if they matched to the genes of viruses in RefSeqABVir database. Genes were considered “unaffiliated” if they had no hit to a sequence in RefSeqABVir.



Supplementary Figure 7. Phylogenetic analysis of *vFAD* and microbial *FAD* genes. A Neighbor Joining tree was constructed using nucleotides sequences of 11 *vFAD*, 263 *FAD* from top 40 hits in the NCBI nr database (Dec 2018), and 217 microbial *FAD* from microbial metagenomes. The clades are colored based on phylum group with the vFAD sequences colored in black. The scale bar indicates a distance of 1. Node supports values indicate bootstrap values that are greater than 70% unless indicated.


(A)
HXXXXH
HXXHH
HXXHH
1. CBW_NODE_1_gene-62
2. CBW_NODE_1_gene-63
3. CBW_NODE_1_gene-64
4. CBW_NODE_1_gene-69
5. CBW_NODE_1_gene-76
6. CBW_NODE_96_gene-4
7. CBW_NODE_96_gene-11
8. CBW_NODE_96_gene-12
9. CBW_NODE_96_gene-13
10. CBW_NODE_96_gene-14
11. SB_NODE_17_gene-11
12. Acidocella_sp._20-58-15_(OYV52848)
13. Bacteroidetes_bacterium_(REK65802)
14. Saprospirales_bacterium_TMED214_(OUW83335)
15. Bacteroidetes_bacterium_(ANO58132)
16. Candidatus_Pelagibacter_sp._TMED197_(OUW58155)
17. Uncultured_Mediterranean_phage_uvMED_(BAR36643)



Supplementary Figure 8. Multiple alignments of vFAD and FAD protein sequences. (A) Alignment of 11 *vFAD* (No. 1 to 11) and 6 *FAD* that encode three conserved histidine-cluster motifs including a HXXXXH and two HXXHH. **(B)** Alignment of *vFAD* SB_NODE_339_gene-12 (No. 1) and 10 *FAD* that encode a HXXXH, a HXXXHH, and a XXXHH motif as the conserved histidine-cluster motifs. The protein sequences were aligned using CLUSTAL X (gap opening penalty 1.0, gap extension penalty 0.2) version 2.0. Conserved histidine-cluster motifs are indicated by black boxes and red notes (e.g., HXXXXH). 'H' indicates Histidine and 'X' indicates any amino acid.