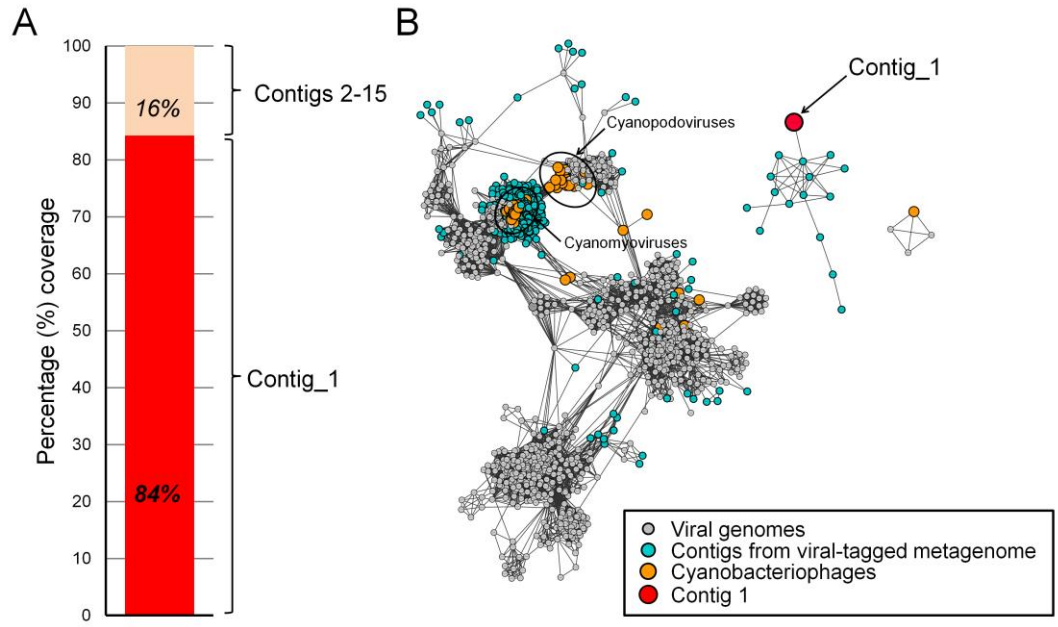

Supplementary Figure. Identification and characterization of lysate-associated viral genomic sequences. (A) The sequence coverage of each contig (%) from viral lysates was calculated by Bowtie 2, SAMtools and BamM. Note that contig_1 has the highest coverage with a majority (~85%) of the reads being assigned to it. (B) Left, a weighted gene-sharing network was built to characterize the relationships of contig_1 to other viruses, in which nodes represent viral genomes/contigs and edges represent the similarity between genomes/contigs based on the number of shared gene (i.e., homologous protein clusters, PCs). As the reference, a total of 2,010 bacterial and archaeal virus genomes (VirRefSeq v75) as well as Synechococcus-specific viral contigs retrieved from viral-tagging (VT) data (Deng et al., 2014) were used. For clarity, only 44,948 edges from 1,442 viral genomes/contigs are shown. Right, matrix view shows a list of clustered proteins (PCs) of contig_1, which are limited to those of cyanophages. A red-colored cell indicates the presence of homologs across the contig_1, cyanophages, and/or *Synechococcus*-specific contigs from VT data.

Deng L, Ignacio-Espinoza JC, Gregory AC, Poulos BT, Weitz JS, Hugenholtz P, Sullivan MB. 2014. Viral tagging reveals discrete populations in Synechococcus viral genome sequence space. Nature. 513(7517):242-245.