

iScience, Volume 24

Supplemental information

Saline lakes on the Qinghai-Tibet Plateau

harbor unique viral assemblages mediating

microbial environmental adaption

Chengxiang Gu, Yantao Liang, Jiansen Li, Hongbing Shao, Yong Jiang, Xinhao Zhou, Chen Gao, Xianrong Li, Wenjing Zhang, Cui Guo, Hui He, Hualong Wang, Yeong Yik Sung, Wen Jye Mok, Li Lian Wong, Curtis A. Suttle, Andrew McMinn, Jiwei Tian, and Min Wang

Fig. S1 Histogram showing the proportion of lysogenic and lytic viruses. Related to Figure 1.

The histogram showed the proportion of putative lysogenic and lytic viruses in all predicted viruses of six viromes of three saline lakes on Qinghai-Tibet Plateau using VIBRANT.

Fig. S2 Abundant vOTUs in six viromes of three saline lakes on Qinghai-Tibet Plateau. Related to Figure 1.

A) 1,244 abundant vOTUs with cumulative relative abundance of 80 % in three lakes of six viromes of three saline lakes on Qinghai-Tibet Plateau

B) The most 40 abundant vOTUs in three lakes of six viromes

Fig. S3 Positive (blue) and negative (green) Pearson's correlation results comparing Shannon Index of viral community with different prokaryotic, geographical and physicochemical parameters. Related to Figure 1.

No significant correlation ($p < 0.05$) was found.

Fig. S4 The maximum likelihood phylogenetic trees of viral contigs based on the terminase large subunit domains of Caudoviruses (A) and major capsid protein of nucleocytoplasmic large DNA viruses (NCLDVs) (B) in six viromes of three saline lakes on Qinghai-Tibet Plateau. Related to Figure 1.

Tree annotations from inside to the outside: family of reference and source of viruses in viromes, relative abundance of new-found viruses.

Fig. S5 Functional analysis of viral contigs of six viromes of three saline lakes on Qinghai-Tibet Plateau. Related to Figure 5.

The coding sequence were compared with eggNOG database using eggNOG-mapper (--seed_ortholog_evalue 0.001 --seed_ortholog_score 60 --query-cover 20 --subject-cover 0).

Fig. S6 Ribosomal genes predicted by KEGG Automatic Annotation Server (KAAS). Related to Figure 5.

The coding sequence were piped through KAAS to detect ribosomal genes. Colors on the foursquare means source of genes. Red represents coding sequence from XQL and green represents sequence from QhL. If a ribosomal gene was detected in both XQL and QhL, red and green will fill the foursquare. Blank foursquares were not found in six viromes of three saline lakes on Qinghai-Tibet Plateau.

Fig. S7 Comparative genomics-based synteny (A) and genomic (B and C) analysis of two metagenomic assembled viral genomes named as XQL_135 (B) and the Lake Baikal phage Baikal-20-5-C28 (C). Related to Figure 3.

A) The star indicated the position of the novel CRISPR sequence. The blue part showed the GC content. Adaptive analysis based on the tetranucleotide frequency of XQL_135 (B) and the Lake Baikal phage Baikal-20-5-C28 (C). The lower value indicated the

larger difference of tetranucleotide frequency between the position and the total genome.

Fig. S8 The geographical position of other viromes in non-metric multidimensional scaling (NMDS) analysis. Related to Figure 6C. Color represent group of viromes.

Fig. S9 The heatmap of average nucleotide identity of genomes selected from superclade 1 of Fig. 3A. The mini phylogenetic tree was shown on the top of the heatmap and the clade was shown on the left.

Fig. S1

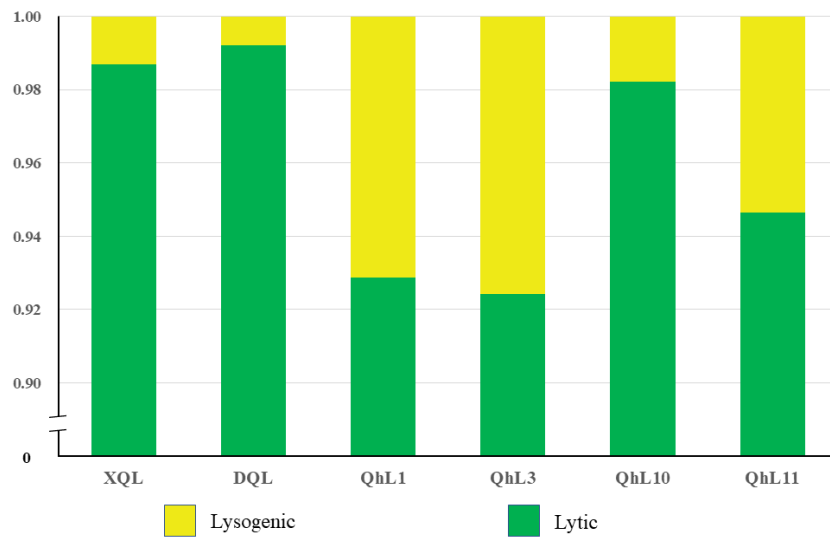


Fig. S2

A



B

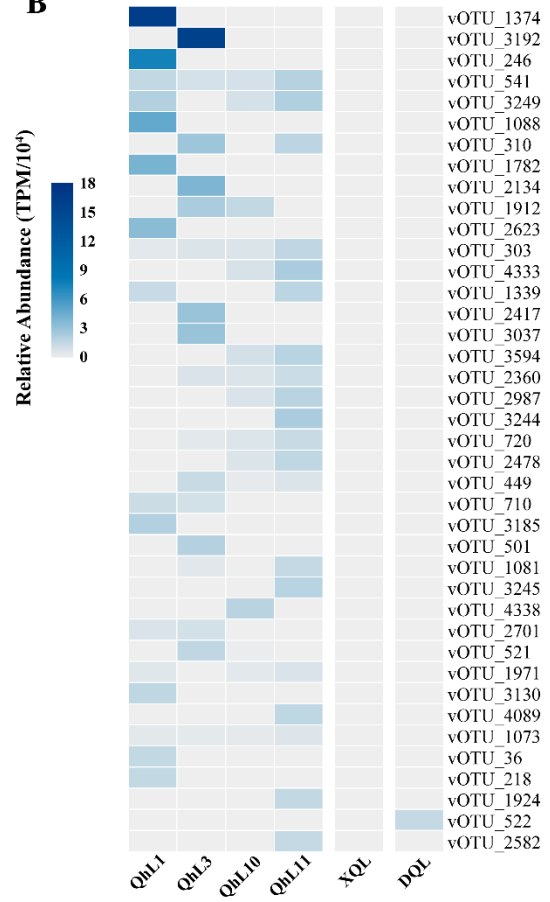


Fig. S3

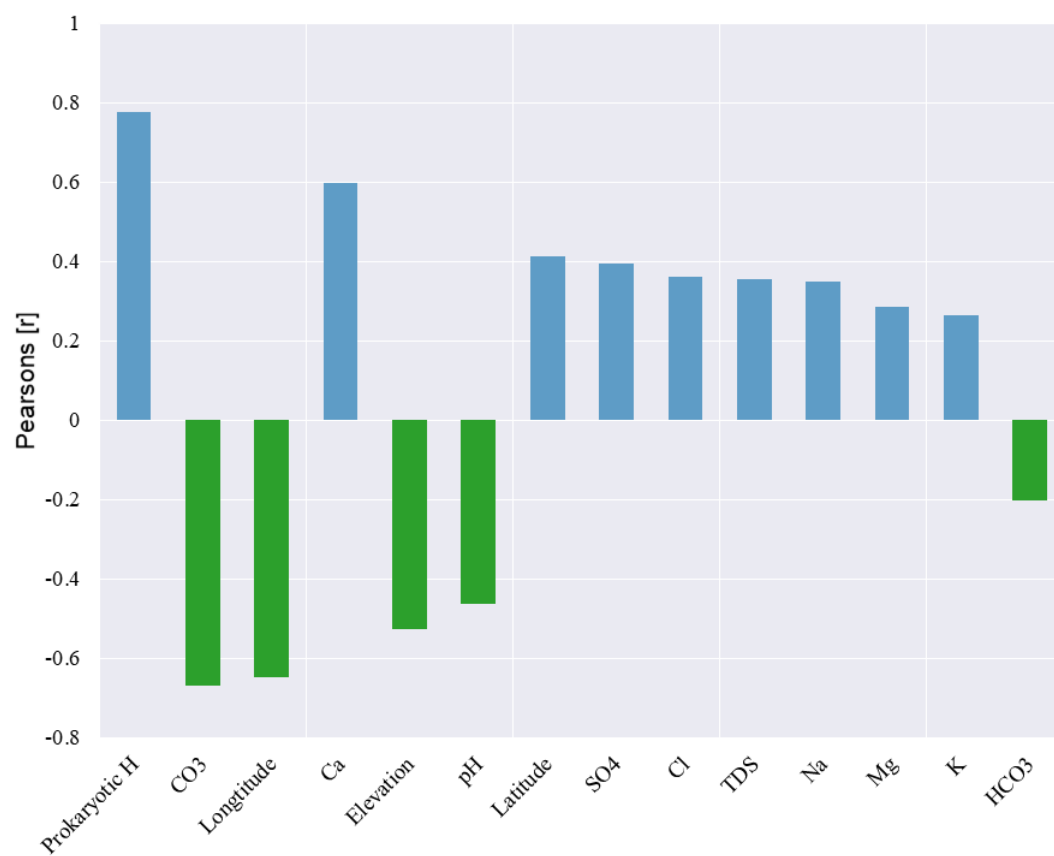


Fig. S4

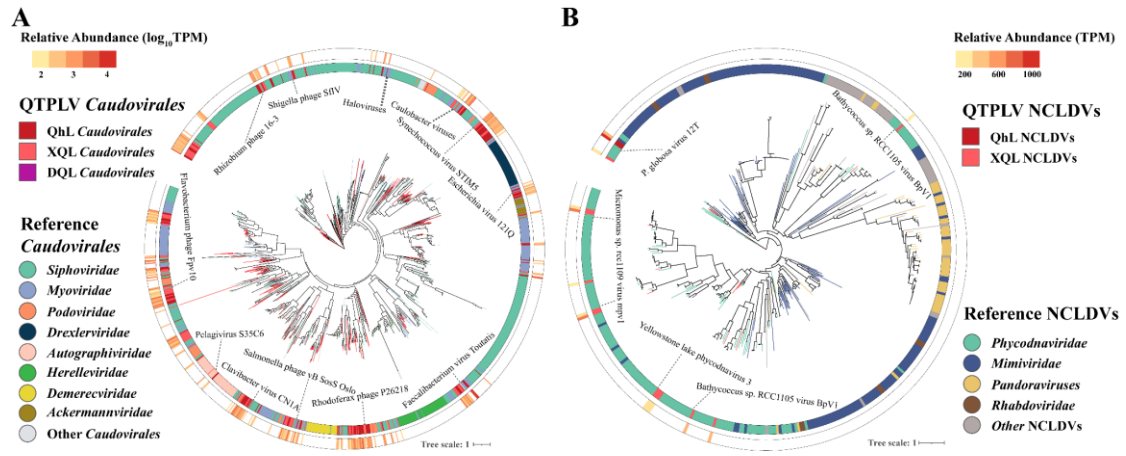


Fig. S5

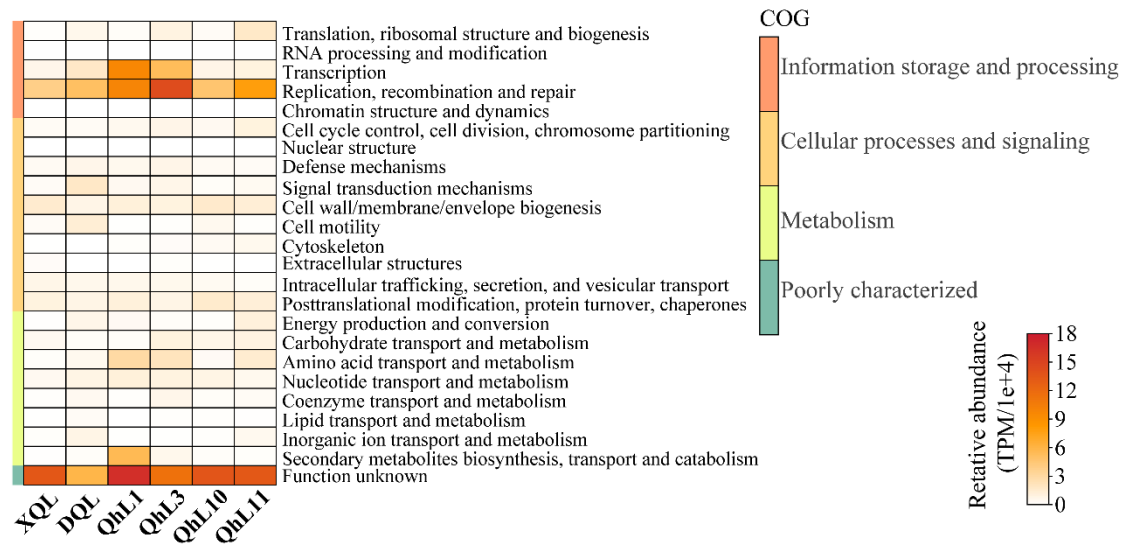


Fig. S7

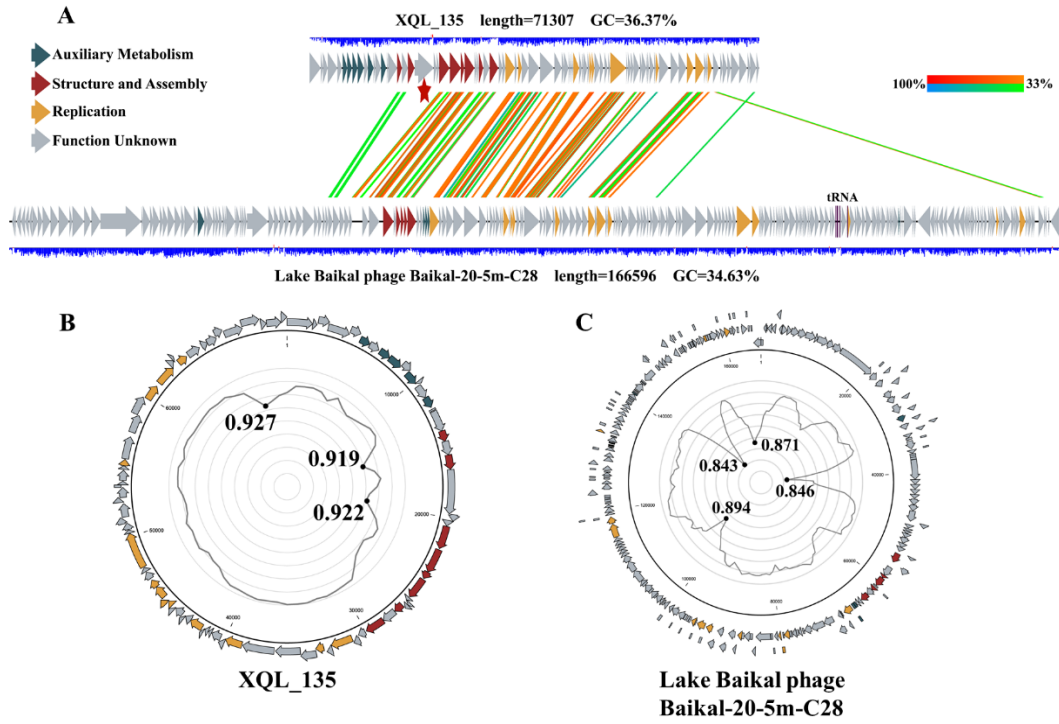


Fig. S8

