

**Supplementary Data:**

**Supplementary Data 1a.** Metagenomic samples overview

**Supplementary Data 1b.** Summary of sample number

**Supplementary Data 2.** Virus overview (vOTU  $\geq 5$  Kb): quality assessment, lifestyle

**Supplementary Data 3.** Virus taxonomy (vOTU  $\geq 10$  Kb) assesment based on vConTACT3 and geNomad

**Supplementary Data 4a.** Normalized relative expression of viruses

**Supplementary Data 4b.** Scaling factor for metatranscriptome normalization

**Supplementary Data 4c.** Metatranscriptome statistical comparison. Data were analysed using two-sided Kruskal–Wallis tests, followed by two-sided Dunn or Wilcoxon pairwise tests with Bonferroni correction for multiple comparisons

**Supplementary Data 5a.** Normalized relative abundance table (vOTU  $\geq 5$  Kb)

**Supplementary Data 5b.** Normalized relative abundance table (vOTU  $\geq 10$  Kb)

**Supplementary Data 5c.** Normalized relative abundance table (1275 MAGs)

**Supplementary Data 5d.** Normalized relative abundance table (MAGs with virus links)

**Supplementary Data 5e.** Summary of read-mapping

**Supplementary Data 6.** Virus diversity assesment (vOTU  $\geq 5$  Kb)

**Supplementary Data 7.** Groundwater environmental data

**Supplementary Data 8a.** Virus-host link based on iPhop analysis and 1,274 MAGs

**Supplementary Data 8b.** The metagenomic-assembled genome (MAGs,  $n = 1,275$ ) used in virus-host analysis

**Supplementary Data 8c.** MAG-virus links based on iPhop analysis

**Supplementary Data 8d.** The metagenomic-assembled genome (MAGs,  $n = 1,275$ ) metabolic capability

**Supplementary Data 9a.** Virus-host ratio (VHR)

**Supplementary Data 9b.** VHR comparison analysis. Data were analyse using PERMANOVA, Mann Whitney U, and correlation tests (Spearman or Pearson) were all conducted as two-sided

**Supplementary Data 10a.** Identified CRISPR-Cas spacers in the groundwater MAGs

**Supplementary Data 10b.** Spacer blast analysis, one mismatch allowed

**Supplementary Data 10c.** Spacer blast analysis, zero mismatches allowed

**Supplementary Data 10d.** Spacer blast analysis, 100% matched

**Supplementary Data 10e.** IMG/Vr spacers blast (e-value=1e-05; hit=10)

**Supplementary Data 11.** Co-occurrence analysis

**Supplementary Data 12.** Auxiliary metabolic genes (AMGs) list

**Supplementary Data 13.** Metabolic reconstruction for host pathways and AMG

—