

Supplementary Methods S1. Detailed statistical procedures.

The independent analytical unit in all statistical analyses was the sampling site. The dataset comprised 20 site-level observations: riverine sites ($n = 5$), channels ($n = 2$) and lakes ($n = 13$). Where more than one sequencing output was available for a riverine site, these were treated as technical sequencing replicates and combined before ecological inference. For phylum-level analyses, read counts from technical replicates were summed and then converted to relative abundances. For alpha-diversity indices, values obtained for technical replicates were averaged to provide one site-level value. Consequently, technical sequencing replicates were not treated as independent environmental observations.

The relationship between $\delta^2\text{H}$ and $\delta^{18}\text{O}$ in surface waters was estimated using ordinary least-squares regression. Uncertainty in the slope and fitted relationship was quantified using 20,000 non-parametric bootstrap resamples of sites. The global meteoric water line, GMWL, defined as $\delta^2\text{H} = 8 \times \delta^{18}\text{O} + 10$, was included only as a reference line (Craig, 1961). Because the study did not include a time-integrated precipitation isotope series, the fitted surface-water relationship was interpreted as an evaporation line rather than a local meteoric water line. This interpretation follows isotope-hydrology recommendations concerning meteoric-water-line fitting and evaporative deviations from meteoric input (Gat, 1996; Gibson et al., 2005; Crawford et al., 2014; Vystavna et al., 2021). Differences in d-excess among habitat types were evaluated using a permutation version of the Kruskal–Wallis rank statistic with 9,999 random reassignments of habitat labels.

Site-level alpha-diversity indices, including Chao1 richness, Shannon diversity and Simpson diversity, were summarised by habitat type. The richness metric was explicitly interpreted as Chao1 because the available source variable was MicroChao1 and was therefore not treated as directly observed ASV richness. Overall habitat differences were assessed separately for each alpha-diversity index using the permuted Kruskal–Wallis statistic with 9,999 random reassignments of habitat labels. Individual sites were retained in graphical summaries. Density distributions were not estimated for the channel group because this habitat type was represented by only two sites.

Bray–Curtis dissimilarities were calculated from site-level phylum relative-abundance profiles. These dissimilarities were visualised using classical principal coordinates analysis, PCoA. The percentages shown on the PCoA axes refer to the variation represented by positive eigenvalues. For presentation purposes, the sign of PCoA1 was oriented so that increasing PCoA1 values corresponded to increasing $\delta^{18}\text{O}$ values. This transformation does not alter inter-site distances or the underlying ordination structure. The association between PCoA1 and $\delta^{18}\text{O}$ was estimated by ordinary least-squares regression. A 20,000-resample non-parametric bootstrap confidence interval was calculated for the regression slope, and the significance of the correlation was assessed using a two-sided permutation test based on 9,999 permutations.

Habitat-associated differences in bacterioplankton composition were additionally tested using permutational multivariate analysis of variance, PERMANOVA, following Anderson (2001). Homogeneity of multivariate dispersion was assessed to distinguish differences in community

location among habitats from differences in within-habitat heterogeneity, following Anderson (2006).

For redundancy analysis, RDA, phylum relative abundances were Hellinger transformed by taking the square root before ordination, following recommendations for community-composition data analysed with linear methods (Legendre and Gallagher, 2001). Environmental predictors were centred and scaled prior to analysis. Dissolved organic carbon, DOC, was retained a priori as the mechanistically relevant indicator of organic-carbon availability. Because DOC and $\delta^{18}\text{O}$ were strongly correlated ($r = 0.933$), these variables were not included together in the same constrained model. $\delta^{18}\text{O}$ was therefore analysed separately, whereas the final RDA model included conductivity, DOC, total nitrogen, TN, and total phosphorus, TP. Collinearity among retained predictors was evaluated using variance-inflation factors, VIFs, and all retained variables had $\text{VIF} < 1.40$. Overall model significance was assessed with 999 permutations. Both raw and adjusted constrained R^2 values were calculated. To preserve figure legibility, only the five bacterial phyla with the longest fitted vectors were displayed in the RDA plot.

Exploratory associations between bacterial phyla and isotope variables were calculated using Pearson correlations based on site-level phylum relative abundances and $\delta^{18}\text{O}$, $\delta^2\text{H}$ and d-excess values ($n = 20$). Two-sided p-values were obtained from 9,999 permutations and adjusted across the displayed taxon–isotope matrix using the Benjamini–Hochberg false-discovery-rate procedure (Benjamini and Hochberg, 1995). Because phylum-level relative-abundance data are compositional and the study has an observational design, these correlations were interpreted as directional associations rather than independent taxon-specific effects.

The conceptual synthesis figure was constructed as a descriptive integration of observed habitat-level isotope, DOC and microbial patterns. Arrows in this figure indicate the analytical sequence linking hydrological-isotopic gradients, physicochemical background and bacterioplankton turnover, and should not be interpreted as experimentally demonstrated causal pathways.