Supplementary Material

Spatial heterogeneity of food webs in a river–lake ecotone under flow regulation – a case study in northern China

Wei Yang, Xianting Fu, Xiaoxiao Li, Baoshan Cui, and Xinan Yin

Text S1. Samples for stable isotope analysis

Water samples that contained detritus (1 L), phytoplankton (5 L), and zooplankton (10 L) were collected from the surface water (to a depth of 100 cm) with an organic glass collector. Subsequently, phytoplankton and zooplankton samples were collected with a triple-layered plankton net with mesh sizes of 64 μm and 250 μm, respectively. The detritus and plankton samples were filtered through pre-combusted Whatman GF/F glass fiber filters.

Leaves samples of submerged macrophytes were collected with a sickle with an attached collection chamber and washed with distilled water to remove epiphytes, and then samples were preserved in zip-lock bags at 4℃ in the field until analysis. For zoobenthos, we used a Van Veen grab with a mouth area of 38 cm × 21 cm to excavate the substrate sludge, which we then washed through a 0.425 mm-mesh size filter to collect the organisms, and preserved these benthic samples in 75% ethanol. For the meiobenthos, we collected three samples for each species we encountered. For the macrozoobenthos, we preserved three muscle samples for each species.

For the fish community, we used multi-mesh gillnets with mesh sizes ranging from 5 to 55 mm and an overall size of 1.5 m × 30 m and 3 m × 30 m (Mao et al., 2014), which we installed beside a ground cage. The ground cage had a mesh size of 5 mm and was 25 m long, and was partitioned into 20 sections with 10-cm openings at the front and back for fish to enter the trap. All fish sampling started in the late afternoon (approx. 18:00 h) and ended the following morning (approx. 06:00 h), for a total of 12 h. The duration was chosen to limit the number of fish caught per net. All fish individuals were measured and weighed. We chose three individuals randomly and used their muscle tissues for the stable isotope analysis for each species. The remaining fish were released unharmed.

All stable isotope samples were kept in a 4℃ ice-chilled box until they could be oven-

dried at 60℃ for 48 h before analysis. To estimate the $\delta^{13}C$ and $\delta^{15}N$ content, each dried sample was ground into a homogenous powder using a ball mill. Subsamples were encapsulated in ultra-pure tin capsules and analyzed on an Elementar Vario Micro-Cube elemental analyzer (Flash EA1112; Thermo Scientific, Monza, Italy) coupled to a Continuous Flow Stable Isotope Ratio Mass Spectrometer (Delta V Advantage, Thermo Scientific, Dreieich, Germany) (Careddu et al., 2015; Hansen et al., 2018). The elemental analyzer and spectrometer were recalibrated after each five-sample run following the manufacturer's directions. All samples were analyzed twice. The obtained carbon (C) and nitrogen (N) stable isotope ratios (^{13}C:^{12}C and ^{15}N:^{14}N) were expressed as delta units, which represent the deviations (‰) from international standards: Vienna Pee Dee Belemnite (VPDB) for C and atmospheric nitrogen for N) according to the following equations:

$$\delta^{15}N = \left(\frac{^{15}N/\,^{14}N_{sample}}{^{15}N/\,^{14}N_{atmosphere}} - 1 \right) \times 1000 \quad \text{(S1)}$$

$$\delta^{13}C = \left(\frac{^{13}C/\,^{12}C_{sample}}{^{13}C/\,^{12}C_{VPDB}} - 1 \right) \times 1000 \quad \text{(S2)}$$

References

Careddu, G., Costantini, M. L., Calizza, E., Carlino, P., Bentivoglio, F., Orlandi, L., and Rossi, L. (2015). Effects of terrestrial input on macrobenthic food webs of coastal sea are detected by stable isotope analysis in Gaeta Gulf. Estuar. Coast. Shelf. Sci., 154, 158-168. http://dx.doi.org/10.1016/j.ecss.2015.01.013.

Hansen, A. G., Gardner, J. R., Connelly, K. A., Polacek, M. and Beauchamp, D. A. (2018). Trophic compression of lake food webs under hydrologic disturbance. Ecosphere, 9(6), e02304. http://dx.doi.org/10.1002/ecs2.2304.

Mao, Z. G., Gu, X. H., Zeng, Q. F., Gu, X. K. and Li, X. G., Wang, Y. P. (2014). Production sources and food web of a macrophyte-dominated region in Lake Taihu, based on gut contents and stable isotope analyses. J. Gt. Lakes. Res., 40(3), 656-665. http://dx.doi.org/10.1016/j.jglr.2014.05.005.

Text S2. Trophic level of consumers

The trophic level of each consumer was estimated using the formula proposed by Vander Zanden and Rasmussen (1999):

$$TL = \lambda + \left(\delta^{15}N_{consumer} - \delta^{15}N_{base}\right)/\delta^{15}N_{TEF} \quad \text{(S3)}$$

where *TL* represents the trophic level of each consumer, with the trophic level of basal food sources defined as 1; $\delta^{15}N_{consumer}$ represents the stable nitrogen isotope ratio for consumers; and $\delta^{15}N_{base}$ represents the stable nitrogen isotope for primary producers (λ = 1) or primary consumers (λ = 2) in the food web. In this study, we chose zooplankton as the baseline organisms. $\delta^{15}N_{TEF}$ represents the trophic enrichment factor for nitrogen.

References

Vander Zanden, M.J. and Rasmussen, J.B. (1999). Primary consumer $\delta^{13}C$ and $\delta^{15}N$ and the trophic position of aquatic consumers. Ecology, 80(4), 1395-1404. http://orcid.org/0000-0002-8122-1416.

Text S3. Meanings of the topological properties

Species richness (*S*) represents the number of species or functional groups present in the food-web model. The number of trophic links (*L*) represents the total number of interactions between organisms. The proportion of top species represents species with no predators. The proportions of intermediate and basal species represent consumer species preyed upon and species with predators or consumers but with no prey, respectively. The link density (*L/S*) represents the number of interaction links per species. Connectance (L/S^2) represents the number of links that exist divided by the total number of possible trophic links. Omnivores represent species that consume other organisms at two or more trophic levels. Herbivores transform plant biomass and transfer energy to the upper trophic levels of food webs. The proportions of links among the three categories of functional groups show predatory relationships within food webs, which represent energy pathways that begin with prey species and end with predators, i.e., between the top and intermediate, top and basal, intermediate and intermediate, and intermediate and basal functional group. Again, these four proportions are not independent as their sum equals 1. The number and proportion of these links are calculated from the trophic species' food-web matrixes. Links indicate predatory relations within a food web and represent energy pathways that begin with prey species and end with predators. A food chain is a linked path from a species to a basal species. Measure for mean chain length reflects an average of the different food chains across all the functional groups in each food web and captures the complexity of the food web. In addition, we defined a chain as a distinct path from a basal to a non-basal species, and chain length as the number of links in the chain (our food webs did not contain cycles). The number of chains (*NC*) was the total number of paths linking basal species with non-basal species. Mean chain length was computed as the total number of links in these paths divided by their number *NC*. The generality and vulnerability properties were the mean number of preys per consumer and the mean number of consumers per prey, respectively. Williams and Martinez (2000) introduced a measure of the variability for Generality and Vulnerability, the standard deviation of normalized generality (GenSD), and normalized vulnerability (VulSD). The proportion of omnivory, generality and vulnerability belong to feeding strategy metrics, and they refer to

the dietary niche properties of species.

Text S4. The related algorithms of statistical analysis

(1) One-way ANOVA and least significance difference (LSD) method

Analysis of Variance (ANOVA) is used to determine whether or not there is a statistically significant difference between the means of three or more independent samples (Fisher, 1918; Fisher and Mackenzie, 1923; Fisher, 1992). One-way ANOVA is used to investigate whether different levels of a control variable have a significant effect on the observed variable. The relevant hypothesis is as follows:

H_0: The means are equal for each sample

H_1: At least one of the means is different from the others

The statistic F is derived by comparing two-point estimates (i.e., between-classes variance (*MSB*) and within-classes variance (*MSE*)) of the assumed common variance:

$$MSB = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n}\left(\bar{x}_j-\bar{x}\right)^2}{k\text{-}1} \tag{S4}$$

$$MSE = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n}\left(x_{ij}-\bar{x}_j\right)^2}{n\text{-}k} \tag{S5}$$

$$F = \frac{MSB}{MSE} \tag{S6}$$

where $\bar{x}_j$ represents the mean of the jth sample; x_{ij} represents the ith value of the jth sample; $\bar{x}$ represents the total sample mean; k represents the number of samples, and n represents the sample size of the jth sample.

If the population means of the k samples are not equal, the between-classes variance (*MSB*) will be greater than the within-classes variance (*MSE*). We can reject the null hypothesis (H_0) when the value of the statistic F is greater than the critical value, which is determined by the given significance level (α) and the degrees of freedom. Thus, the rejection domain of the statistic F is $F > F_\alpha$ (k-1, n-k) when the given significance level is α.

The rejection of H_0 by one-way ANOVA can only indicate that the population means of multiple samples are not equal or not all equal. We can perform the least significance difference (*LSD*) test (Fisher, 1935) to identify the populations whose means are statistically different. The basic idea of the test is to compare the populations taken in pairs. In general, the standard deviation of the difference between the mean of sample i and the mean of sample j is equal to:

$$\sqrt{S_I^2 \frac{1}{n_i} + \frac{1}{n_j}} \tag{S7}$$

where S_I^2 is the estimation of the variance inside the samples:

$$S_I^2 = \frac{SCI}{N\text{-}k} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(X_{ij} - \overline{X}_{i.}\right)^2}{N\text{-}k} \tag{S8}$$

where *SCI* represents the sum of squares inside the samples, *N* represents the total number of observations, *k* represents the number of samples, X_{ij} represents the *j*th observation of sample *i*, $\overline{X}_{i.}$ represents the mean of sample *i*, n_i represents the number of observations in sample *i*, and n_j represents the number of observations in sample *j*.

The ratio follows the Student distribution with *N-k* degrees of freedom:

$$t_{obs} = \frac{\overline{X}_{i.} - \overline{X}_{j.}}{\sqrt{S_I^2 \frac{1}{n_i} + \frac{1}{n_j}}} \tag{S9}$$

The difference between a pair of means is significant when:

$$\left|\overline{X}_{i.} - \overline{X}_{j.}\right| \geq \sqrt{S_I^2 \frac{1}{n_i} + \frac{1}{n_j}}\, t_{(N\text{-}k,\ 1-\frac{\alpha}{2})} \tag{S10}$$

where $t_{(N\text{-}k,\ 1-\frac{\alpha}{2})}$ denoting the value of a Student variate with *N-k* degrees of freedom for a significance level set to α. The right-hand side of this equation is called the *LSD*.

(2) Kruskal-Wallis *H* test and Mann-Whitney *U* test

The Kruskal-Wallis *H* test from the bank of classical statistics tests is a well-known nonparametric alternative to the one-way ANOVA (Kruskal and Wallis, 1952). The statistic *H* tests the null hypothesis that the samples all come from identical populations. The rank test presented here requires that all the observations be ranked together, and the sum of the ranks obtained for each sample. The test statistic to be computed if there are no ties (that is, if no two observations are equal) is:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{C} \frac{R_i^2}{n_i} - 3(N+1) \tag{S11}$$

where *C* represents the number of samples; n_i represents the number of observations in the *i*th sample; $N = \sum n_i$ represents the number of observations in all samples combined; and R_i represents the sum of the ranks in the *i*th sample.

If there are ties, each observation is given the mean of the ranks for which it is tied. The

statistic H value calculated by the above formula is too small, and a general expression of whether or not there are ties, assuming that such ties as occur are given mean ranks: and the corrected H value can be calculated according to the following formula:

$$H = \frac{\frac{12}{N(N+1)}\sum_{i=1}^{C}\frac{R_i^2}{n_i}-3(N+1)}{1-\frac{\sum T}{N^3-N}} \quad \text{(S12)}$$

where the summation is over all groups of ties and $T= (t\text{-}1)t(t+1)=t^3\text{-}t$ for each group of ties, t being the number of tied observations in the group.

If the samples come from identical continuous populations and the n_i are not too small, H is distributed as $\chi^2(C\text{-}1)$, permitting use of the readily available tables of χ^2. Large values of H (i.e., $H > \chi^2_{\alpha}(C\text{-}1)$) lead to rejection of the null hypothesis.

The Mann-Whitney U test was further used for multiple comparisons of variables (Mann and Whitney, 1947). They discuss the test in terms of a statistic U which, as they point out, is equivalent to Wilcoxon's sum of ranks (Wilcoxon, 1945). When all observations from both samples are arranged in order, they count for each observation in one sample. The sum of these counts for the sample is called U:

$$U_1=R_1-\frac{n_1(n_1+1)}{2} \quad \text{(S13)}$$

$$U_2=R_2-\frac{n_2(n_2+1)}{2} \quad \text{(S14)}$$

where n_1 and n_2 are the sample sizes, and R_1 and R_2 are the sum of the ranks for the first and second samples, respectively. The minimum value of U_1 and U_2 is compared with the significance test U_α (see the Mann-Whitney table for specific values). If $U_{min} < U_\alpha$, the null hypothesis is rejected, indicating a significant difference between the two samples.

References

Fisher, R. A. (1918). The Correlation between Relatives on the Supposition of Mendelian Inheritance. Transactions of the Royal Society of Edinburgh, 52(2), 399-433. https://doi.org/10.1017/S0080456800012163.

Fisher, R. A. and Mackenzie, W. A. (1923). Studies in crop variation. II. The manurial response of different potato varieties. The Journal of Agricultural Science, 13(3), 311-320. https://doi.org/10.1017/S0021859600003592.

Fisher, R. A. (1935). The Design of Experiments. Oliver and Boyd, Edinburgh.

Fisher, R. A. (1992). Statistical Methods for Research Workers. In: Kotz, S., Johnson, N.L. (eds) Breakthroughs in Statistics. Springer Series in Statistics. Springer, New York, NY. https://doi.org/10.1007/978-1-4612-4380-9_6.

Kruskal, W. H. and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. Journal of the American statistical Association, 47(26), 583-621. https://doi.org/10.1080/01621459.1952.10483441.Mann, H. B. and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. The Annals of Mathematical Statistics, 18(1), 50-60. https://doi.org/10.1214/aoms/1177730491.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biometrics, 1(6), 80-83. https://doi.org/10.2307/3001968.

Figure S1. Contributions of the main food sources to the food webs in the four study areas.

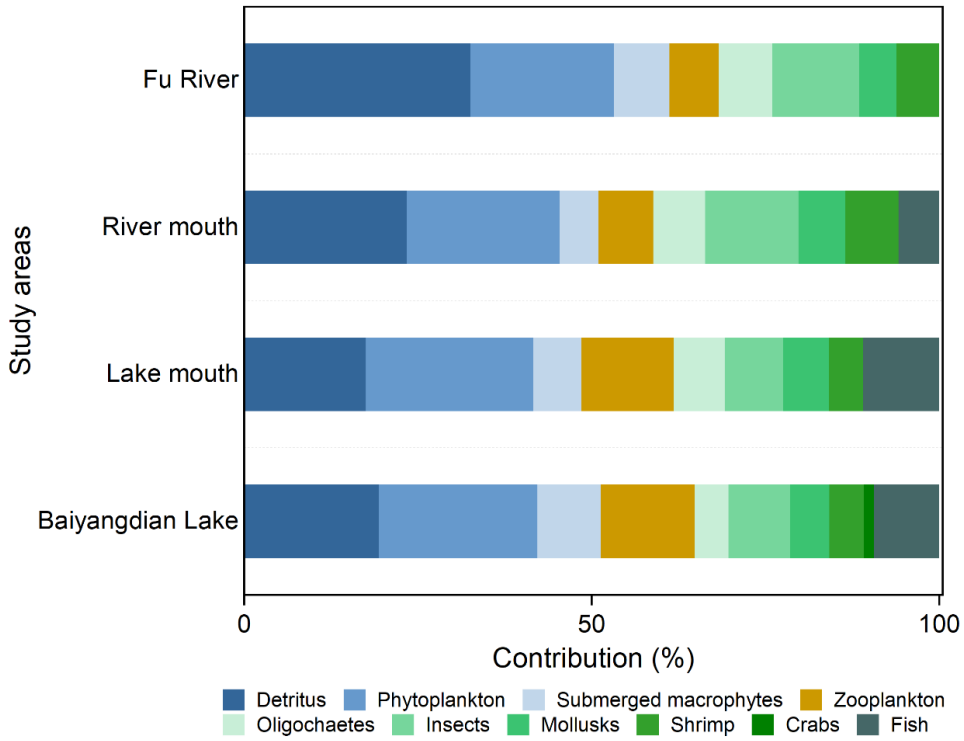

Figure S2. Topology structure of food webs at: (a) Fu River, (b) River mouth, (c) Lake mouth, and (d) Baiyangdian Lake.

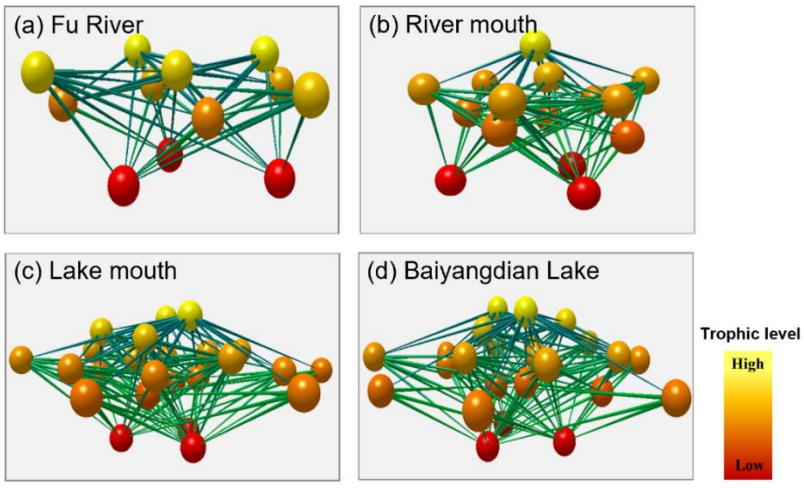

Table S1. Multiple comparisons of $\delta^{13}C$ and $\delta^{15}N$ of basal food sources between each ecosystem type

Ecosystem type 1-Ecosystem type 2	Adj. P-value				
	Submerged macrophytes $\delta^{15}N$	Detritus $\delta^{13}C$	Detritus $\delta^{15}N$	Phytoplankton $\delta^{13}C$	Phytoplankton $\delta^{15}N$
Baiyangdian Lake-Fu River	0.032	0.000	1.000	0.001	1.000
Baiyangdian Lake-River mouth	0.003	0.000	0.825	0.013	0.839
Baiyangdian Lake-Lake mouth	0.001	0.650	0.002	1.000	0.002
Fu River-River mouth	1.000	1.000	1.000	1.000	1.000
Fu River-Lake mouth	0.902	0.180	0.024	0.028	0.204
River mouth-Lake mouth	1.000	0.640	0.319	0.321	0.234

Table S2. Topological properties of the food web structure at each study area

Topological property	Fu River	River mouth	Lake mouth	Baiyangdian Lake
Species properties				
Species richness (*S*)	12	15	24	25
Number of trophic links (*L*)	46	72	140	152
Proportion of top level species	0.333	0.067	0.083	0.080
Proportion of intermediate level species	0.417	0.733	0.792	0.800
Proportion of basal level species	0.250	0.200	0.125	0.120
Proportion of herbivory	0.167	0.133	0.083	0.080
Proportion of omnivory	0.583	0.667	0.792	0.800
Link properties (complexity)				
Link density	3.833	4.800	5.833	6.080
Connectance ($C = L/S^2$)	0.319	0.320	0.243	0.243
Proportion of links between				
Top and intermediate levels	0.326	0.139	0.221	0.217
Top and basal levels	0.261	0.000	0.000	0.000
Intermediate and intermediate levels	0.130	0.444	0.421	0.434
Intermediate and basal levels	0.283	0.417	0.357	0.349
Chain properties				
Mean food chain length	2.514	3.450	3.451	3.556
Consumer–prey asymmetries				
Generality (*Gen*)	5.111	6.000	6.667	6.909
Vulnerability (*Vul*)	5.750	5.143	6.364	6.609
standard deviation of normalized generality (*GenSD*)	0.793	0.705	0.728	0.734
standard deviation of normalized vulnerability (*VulSD*)	0.896	0.823	1.058	1.067