**Supplemental information**

Temperature drives local contributions to beta diversity in mountain streams: Stochastic and deterministic processes

Jianjun Wang1,2,10, \*, Pierre Legendre3, Janne Soininen2, Chih-Fu Yeh1, Emily Graham4, James Stegen4, Emilio O. Casamayor5, Jizhong Zhou6,7,8, Ji Shen1, Feiyan Pan9

1 State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China

2 Department of Geosciences and Geography, University of Helsinki, Helsinki FIN-00014, Finland

3 University of Chinese Academy of Sciences, Beijing 100049, China

4 Département de sciences biologiques, Université de Montréal, C.P. 6128, succursale Centre-Ville, Montréal, Québec H3C 3J7, Canada

5 Pacific Northwest National Laboratory, 902 Battelle Boulevard, P.O. Box 999, Richland, Washington 99352, USA

6 Integrative Freshwater Ecology Group, Center for Advanced Studies of Blanes-Spanish Council for Research CEAB-CSIC, Blanes E-17300, Spain

7 Institute for Environmental Genomics, Department of Microbiology and Plant Biology, and School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, Oklahoma, USA

8 State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China

9 Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, California 94270, USA

10 Jiangsu Key Laboratory for Molecular and Medical Biotechnology, Nanjing Normal University, Nanjing 210023, China

**\*** Correspondence: Jianjun Wang, jjwang@niglas.ac.cn

**Running title**: Local contributions to beta diversity in streams

**Table S1**. Brief information of the six mountains examined.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Mountains | Total samples | Latitude | Longitude | Elevation | Mean annual temperature (oC) | Total Phophorus (ppm) | Country |
| Balggesvarri | 19 | 69.378-69.385 | 20.296-20.398 | 18-771 | -3.3-0.7 | 0.104-0.457 | Norway |
| Pyrenees | 19 | 42.419-42.589 | 0.735-0.962 | 850-2500 | 1.0-9.7 | 0.202-0.488 | Spain |
| Haba | 20 | 27.361-27.382 | 100.076-100.186 | 1817-3943 | 4.2-15.8 | 0.088-1.350 | China |
| Laojun | 26 | 26.630-26.850 | 99.710-99.940 | 1828-4045 | 4.3-15.2 | 0.077-2.477 | China |
| Meili | 17 | 28.354-28.391 | 98.760-98.855 | 2211-3510 | 5.5-12.7 | 0.438-1.267 | China |
| Yulong | 24 | 27.139-27.170 | 100.220-100.431 | 1486-3316 | 7.3-18.1 | 0.377-1.240 | China |

../../../../../../../Documents/OneDriveBusiness/All.stream/Results_Biofilm_Beta_Legendre/Fig.%20Beta.div_Podani_Sørensen_replacement_richness_Propo

**Figure S1**. The proportions of replacement and richness difference components to total beta diversity.Partitioning of beta diversity into replacement (green) and richness difference (yellow) components is based on Podani’s Sørensendissimilarity of quantitative (upper panels) and qualitative (lower panels) community data for bacteria, diatoms and macroinvertebrates.



**Figure S2**. The correlations of the components in species replacement or richness difference (or nestedness) for Podani and Baselga family. Partitioning of beta diversity into species replacement (Repl.LCBD) and richness difference or nestedness (RichDiff.LCBD) components is based on Podani’s or Baselga’s Sørensendissimilarity of qualitative (upper panels) and quantitative (lower panels) and community data for bacteria, diatoms and macroinvertebrates. We did not distinguish the values for the three taxonomic groups because the positive relationships were obviously strong.



**Figure S3**. The brief framework for data analyses. Generality: The generality in the elevational patterns in local contributions to beta diversity (LCBD). Thresholds: The thresholds for non-detectable U-shaped elevational patterns in LCBD. Experimental species removal: We investigated to what extent the LCBD elevational pattern would be affected by generalist and specialist species by removing the species with large and small niche breadth (i.e. species elevational range size), respectively. Such experimental removal of generalists or specialists could identify the threshold for non-detectable U-shaped elevational patterns in LCBD. Ecological processes: Stochastic and deterministic processes. Details are shown in the Material and Methods section of the main text.

../../../../../../../Documents/OneDriveBusiness/All.stream/0_Biofilm_Codes_Beta_Bacteria/Legendre/R_2_Legendre/Figs/Fig%20Diversity_Elevation_3.taxa_Abundance-elevaion

**Figure S4**. Simulated coenoclines along the elevational gradient of Haba Mountain. We used the computed niche parameters (e.g., niche optimum, tolerance and height) of the species in observed communities for coenocline simulation along the elevational gradient. The maximum abundance for each species, that is niche height, were obtained from the observed communities. For each species, we estimated species’ weighted elevational optima by averaging the multiplications of species’ elevation values and abundances across all samples, and calculated species’ environmental tolerance with its elevational range size. If the elevational range size was zero (which would indicate that the species occurred in only one site), we replaced the elevational range size by half of the shortest elevation difference among sites in each mountain. Then we simulated coenoclines with the obtained weighted elevational optima and elevational range size for each species along an elevational gradient with the Gaussian response model in R package *coenocliner* (V0.1-0). We did not simulate species counts or occurrences with error from the parameterised species response curves, and thus the values of the parameterised response curve evaluated at the gradient locations were returned. The coenocline simulations were carried out for each taxonomic group. For better visualization of bacteria, only 300 randomly selected species abundance distributions are shown (colour lines in each graph). Quite similar to Haba Mountains, the simulated coenoclines for the other five mountains were not shown. Because the simulated LCBD with current parameter settings resulted in the expected U-shaped patterns (Figure S5), we did not calculate the parameters with modelling the observed species response curves or bootstrap resampling.



**Figure S5**. Elevational LCBD indices along elevational gradients. To support the observed elevational patterns of LCBD, we modelled the LCBD with Mid-Domain Effect (MDE) based on observed communities. The trends along elevations for LCBD were fitted with significant (*P* < 0.05) quadratic models. For better visualization, we showed only the results of bacteria, diatoms and macroinvertebrates on Haba Mountain. Similar U-shaped elevational patterns in LCBD were found for the other mountains.

../../../../../../../Documents/OneDriveBusiness/All.stream/Results_Biofilm_Beta_Legendre/Fig%20Diversity_Elevation_3.taxa_NULL_LCBD.null_locations.

**Figure S6**. LCBD indices along elevational gradients.To support the observed elevational patterns of LCBD, we calculated the LCBD of simulated coenoclines based on observed communities (as the examples shown in Figure S4). Solid dots indicate the significant (*P* ≤ 0.05) LCBD with permutation test. The trends along elevations for LCBD were fitted with linear and quadratic models, and the better model, shown as a solid red line, was selected based on the lower value of Akaike’s information criterion. For all panels, the quadratic model was significant (*P* < 0.05) and was selected as a better model. The values of adjusted R2 are shown for the significant regressions. The upper, middle and lower panels are for bacteria, diatoms and macroinvertebrates, respectively.

Suppl_R/Results_Fixed_Settings%20Figure%20fixed.optimum_fixed.tolerance_XX=Tolerance_Replacement.pdf

**Figure S7**. The effects of species tolerance and replacement on the patterns of local contributions in beta diversity (LCBD). We simulated the coenoclines for a community of 20 species at 30 equally spaced locations along a hypothetical spatial gradient. Each species’ abundance was set within a range of 0 - 20 individuals. For each simulation for a fixed niche breadth, the niche optima of 20 species was an arithmetic progression from 0 to 30, and the niche breadth of all 20 species was a fixed value selected from an arithmetic progression from 0.2 to 2.6 by a constant quantity of 0.05, which indicates the changes in species replacement. We then calculated the LCBD for each simulation. The trends of LCBD along the spatial gradient were fitted with quadratic models, and the trends were quantified with the coefficient of determination (R2). We also partitioned total beta diversity into replacement and species difference based on the replacement and richness difference matrices, respectively, with Sørensen dissimilarity in the Podani family of indices for both quantitative communities, and the percentage of species replacement was calculated as the ratio of species replacement diversity and total beta diversity. The relationships between adjusted R2 and species tolerance (A) or species replacement percentage (B) are shown for each simulation; non-significant quadratic models are indicated in red. Our results show that the U-shaped pattern in LCBD disappeared after a species replacement percentage of 89.6%, where the species niche breadth was 0.5 and a new species disappeared after six spatial steps. It should note that the simulation settings above may affect the results, but it happens for the general phenomenon of species replacement thresholds for significant U-shaped patterns in LCBD. More details of the coenocline simulation and beta diversity partitioning can be found in Supplemental material S1.

../../../../../../../Documents/OneDriveBusiness/All.stream/Results_Biofilm_Beta_Legendre/Fig%20Diversity_comm.richness_Total.LCBD_locations.face%5

**Figure S8**. The relationships between local contributions to beta diversity (LCBD) and species richness. LCBDs were calculated for each mountain based on abundance species matrix with Hellinger transformation. These relationships were fitted with linear and quadratic models, the significance of which (*P* ≤ 0.05) are shown with blue and red lines (solid or dotted). The better model, shown as solid line, was selected based on the lower value of Akaike’s information criterion. The values of adjusted R2 were shown for each significant regression. Non-significant trends for both models are shown with dotted gray lines. The upper, middle and lower panels are for bacteria, diatoms and macroinvertebrates, respectively. For better visualization, we standardized the species richness for each mountain to have mean = 0 and sd = 1.

../../../../../../../Documents/OneDriveBusiness/All.stream/Results_Biofilm_Beta_Legendre/Fig%20Qualitative_repl.LCBD_locations.

**Figure S9**. Replacement LCBD (REPLLCBD) along elevational gradients. These trends along elevations were fitted with linear and quadratic models, the significance of which (*P* ≤ 0.05) are shown with blue and red lines (solid or dotted). The better model, shown as solid line, was selected based on the lower value of Akaike’s information criterion. The values of adjusted R2 were shown for each significant regression. Non-significant trends for both models are shown with dotted gray lines. The upper, middle and lower panels are for bacteria, diatoms and macroinvertebrates, respectively.

../../../../../../../Documents/OneDriveBusiness/All.stream/Results_Biofilm_Beta_Legendre/Fig%20Qualitative_rich.LCBD_locations.

**Figure S10**. Richness difference LCBD (RichDiffLCBD) along elevational gradients.These trends along elevations were fitted with linear and quadratic models, the significance of which (*P* ≤ 0.05) are shown with blue and red lines (solid or dotted). The better model, shown as solid line, was selected based on the lower value of Akaike’s information criterion. The values of adjusted R2 were shown for each significant regression. Non-significant trends for both models are shown with dotted gray lines. The upper, middle and lower panels are for bacteria, diatoms and macroinvertebrates, respectively.



**Figure S11**. The environmental factors related to **replacement** **LCBD**. For all taxa groups, the explanatory factors were identified with information theoretic approach, and the significance was tested by asking whether the sum of the Akaike weights of each predictor variable was significantly greater than the summed Akaike weights obtained from a null distribution of sites. (A) For each mountain or taxa group, non-significant variables were shown in gray circles, and the significant variables were in green and light green for *P* < 0.01 and *P* < 0.05, respectively. (B) The frequency of environmental variables which significantly (*P* < 0.05) explained LCBD for each taxa group. MAT: mean annual temperature. MAT.squared: squared MAT. Depth: streamwater depth. Width: stream width. Velocity: current velocity. TP: total phosphorus. cDOM: chromophoric dissolved organic matter.



**Figure S12**. The environmental factors related to **richness differences** **LCBD**. For all taxa groups, the explanatory factors were identified with information theoretic approach, and the significance was tested by asking whether the sum of the Akaike weights of each predictor variable was significantly greater than the summed Akaike weights obtained from a null distribution of sites. (A) For each mountain or taxa group, non-significant variables were shown in gray circles, and the significant variables were in green and light green for *P* < 0.01 and *P* < 0.05, respectively. (B) The frequency of environmental variables which significantly (*P* < 0.05) explained LCBD for each taxa group. MAT: mean annual temperature. MAT.squared: squared MAT. Depth: streamwater depth. Width: stream width. Velocity: current velocity. TP: total phosphorus. cDOM: chromophoric dissolved organic matter.

../../../../../../../Documents/OneDriveBusiness/All.stream/Results_Biofilm_Beta_Legendre/Figure%20S%20R2-Environments_DeleteTwoSamples_RawR2_NoSampleNumbers_Final_Onl

**Figure S13**. The regressions between explanatory variables and beta diversity-elevation relationships. The LCBD-elevation relationships (LCBDer) are showing in Figure 1. We quantified the LCBDer of each mountain with the coefficient of determination (R2) as effect size within the meta-analysis context, and used mean annual temperature and heterogeneity in environmental variables of each mountain as explanatory variables. The explanatory variables were: (1) heterogeneity in morphology (stream width, shading, substratum size, depth and current velocity), (2) heterogeneity in chemistry (pH and conductivity), (3) heterogeneity in nutrients (TP and cDOM), (4) standard deviation (SD) in mean annual temperature (MAT) and (5) the mean value of MAT for each mountain. We modelled the relationships between LCBDer R2 and environmental variables with beta regression with the logit link, and reported the Pseudo R2 for the significant (*P* ≤ 0.05, in red lines) and marginally significant (P ≤ 0.1, in blue line) regressions.

../../../../../../../Documents/OneDriveBusiness/All.stream/Results_Biofilm_Beta_Legendre/RangeSizeCutoffs/Fig%20LCBD-Elevation_R2_Rare.

**Figure S14**. The effects of elevational range size of specialist species on LCBD-elevation relationships (LCBDer). The beta diversity was measured with local contributions to beta diversity (LCBD). The species having smaller elevational range size than specific cutoffs (x-axes) were removed from the community matrix and then LCBD were recalculated for each mountain and taxonomic group. The relationship of LCBD and elevation for each cutoff was quantified with quadratic model and the significant (P < 0.05) and non-significant (P > 0.05) relationships were shown with filled and open circles for adjusted R2, respectively.

../../../../../../../Documents/OneDriveBusiness/All.stream/Results_Biofilm_Beta_Legendre/RangeSizeCutoffs/Fig%20Processes-Elevation_Determin

**Figure S15**. The effects of species elevational range size on deterministic processes. The proportions of deterministic processes were quantified with Raup-Crick metric of beta diversity as the ratio between the sum of all positive pairwise tests (i.e., |RC| > 0.95) and the total number of possible pairwise comparisons. All species having larger elevational range size than specific cutoffs (x-axes) were removed from the observed community matrix and then the proportions of deterministic processes were calculated for the sub-communities of each mountain and taxonomic group.

../../../../../../../Documents/OneDriveBusiness/All.stream/Results_Biofilm_Beta_Legendre/Fig%20Diversity_comm.richness_Total.pa.LCBD_locations.face%5BFinal%5D_Bi

**Figure S16**. The relationships between local contributions to beta diversity (LCBD) and species richness. LCBDs were calculated for each mountain based on binary species matrix with Hellinger transformation. These relationships were fitted with linear and quadratic models, the significances of which (*P* ≤ 0.05) are shown with blue and red lines (solid or dotted). The better model, shown as solid line, was selected based on the lower value of Akaike’s information criterion. The values of R2 were shown for each significant regression. Non-significant trends for both models are shown with dotted gray lines. The upper, middle and lower panels are for bacteria, diatoms and macroinvertebrates, respectively. For better visualization, we standardized the species richness to have mean = 0 and sd = 1.

**Supplemental material S1**

# readme ####

# Script - Construct coenoclines with function coenocline() of {coenocliner}

# and analyse them using beta.div() to obtain LCBD indices.

# Supplemental material S1 for

# Wang et al, Submitted manuscript

# Coenoclines on mountainsides:

# Temperature drives local contributions to beta diversity along elevation gradients

# Variables included

# 1, species tolerance: fixed.

# 2, species optimum: fixed for each species.

# 3, Maximum abundance for each species: fixed as 30.

# 4, Sites = 30, species number = 20.

# ===============================================================

rm(list=ls())

library(coenocliner) # contains function: coenocline()

library(adespatial) # contains function: beta.div(), beta.div.comp()

library(vegan)

# local working directory

wd = "xx/dir"

setwd(wd)

#======== # fixed tolerance

# settings

# species replacement according to its tolerance

tol.range.list = seq(from=0.2, to=2.6, by = 0.05)

tol.range <- tol.range.list[1]

# go through tol.range

for (tol.range in tol.range.list) {

## simulate community along a spaital gradient

Sites = 1:30 # 30 sites along the transect

sp = 20 # 20 species

# 1, opt

opt <- seq(from=1, to=30, length=sp) # Optimum for each species

# 2, tol, height

tol <- rep(tol.range, sp) # Tolerance for each species

# 3, height

h <- rep(20, sp) # Maximum abundance for each species

# coenocline simulation

set.seed(1)

Sp.abund.1 <- coenocline(Sites, responseModel = "gaussian",

params = cbind(opt = opt, tol = tol, h = h),

countModel = "poisson", expectation = TRUE)

head(Sp.abund.1); dim(Sp.abund.1)

# LCBD calculation

res.beta.1 <- beta.div(Sp.abund.1)

summary(res.beta.1)

res.beta.1$LCBD

# model the LCBD-sites

lm.mod <- lm(res.beta.1$LCBD ~ Sites + I(Sites^2))

lm.summary <- summary(lm.mod)

# calculate the replacement ####

# LCBD partitioning

beta.div.Sor.ab = beta.div.comp(Sp.abund.1, coef="S", quant=TRUE) # compute dist; with abundance

beta.div.Sor.pa = beta.div.comp(Sp.abund.1, coef="S", quant=FALSE) # compute dist; without abundance

beta.div.OUT = data.frame(rbind(c(beta.div.Sor.ab$part,

beta.div.Sor.pa$part)))

colnames(beta.div.OUT) = c(paste(c("BDtotal", "Repl", "Rich", "Repl.perc", "Rich.perc"), "ab", sep="\_"),

paste(c("BDtotal", "Repl", "Rich", "Repl.perc", "Rich.perc"), "pa", sep="\_"))

#

mod.col.tmp = data.frame( Sites = 30,

sp.num = 20,

tol.range = tol.range,

model.p.value = summary(lm.mod)$coefficients[2,4],

r.squared = lm.summary$r.squared,

adj.r.squared = lm.summary$adj.r.squared,

beta.div.OUT

)

# collector

if (tol.range == tol.range.list[1]) {

mod.col = mod.col.tmp } else {

mod.col = rbind(mod.col, mod.col.tmp) }

} # end of tolerance loop

# subset

mod.col.nonsig <- subset(mod.col, model.p.value > 0.05)

mod.col.sig <- subset(mod.col, model.p.value <= 0.05)

# check the thresholds of U-shaped pattern

print(

rbind(mod.col.nonsig[nrow(mod.col.nonsig), 1:11],

mod.col.sig[1, 1:11])

)

#========

# plot the figures

#========

# final plot

# plot R2 of U-shaped patterns

pdf("Results\_Supplemental material S1.pdf", width=6, height=4)

par(mfrow=c(1,2))

par(mar=c(2, 3.5, 0, 0), # subplot margin

oma=c(1.1, 0, 0.5, 0.5)) # global margin

plot(mod.col$tol.range, mod.col$adj.r.squared, xlab="", ylab="")

points(mod.col.nonsig$tol.range, mod.col.nonsig$adj.r.squared, col="red")

mtext(text="Species tolerance", line = 2, side = 1)

mtext(text=expression(Adjusted~R^2), line = 2, side = 2)

text(0.25, 0.96, "A")

plot(mod.col$Repl.perc\_ab, mod.col$adj.r.squared, xlab="", ylab="")

points(mod.col.nonsig$Repl.perc\_ab, mod.col.nonsig$adj.r.squared, col="red")

mtext(text="Replacement (%)", line = 2, side = 1)

mtext(text=expression(Adjusted~R^2), line = 2, side = 2)

text(x = mod.col.sig[1, ]$Repl.perc\_ab-0.11,

y = mod.col.sig[1, ]$adj.r.squared, labels = round(mod.col.sig[1, 10], 3))

text(0.38, 0.96, "B")

dev.off()

# end