**Disturbance reverses classic island biogeography predictions in river-like landscapes**

**Authors:** Eric Harvey, Isabelle Gounand, Emanuel A. Fronhofer, and Florian Altermatt

Supporting Information

Appendix A - Robustness 2

Appendix B – Positive mortality-patch size relationships 4

Appendix C - video analysis 5

Supporting Figures 10

Figure S1. Robustness to dispersal modes 10

Figure S2. Robustness to dispersal directionality 11

Figure S3. Sensitivity to dispersal rate and directionality 12

Figure S4. Temporal dynamics of alpha diversity (model) 13

Figure S5. Patch size – mortality rate relationships 14

Figure S6. Diversity pattern with mortality and dispersal rates (model) 15

Figure S7. Patch size – mortality rate positive relationships 16

Figure S8. Diversity patterns with positive patch size–mortality relationships 17

Figure S9. River-like landscape (experiment) 18

Figure S10. Temporal dynamics of alpha diversity (experiment) 19

Figure S11. Community structure over time (model) 20

Supplementary Tables 21

Table S1. Variables and parameters of the model 21

Table S2. Effect of patch size on protist alpha diversity (experiment) 22

# Appendix A - Robustness

We tested the robustness of the mechanism producing the reversed diversity pattern in dendritic networks to changes in dispersal mode and dispersal directionality. For that, we looked at the difference in extinction times between large and small patches (Fig. 3b) under contrasting scenarios.

In the main text we scaled the dispersal rate per edge, , with the size of the patch () from which the dispersal origins (emigration), and we considered equal flows going upstream and downstream (strict bi-directionality), to mimic the experimental settings. To assess the generality of the reversed diversity pattern, we ran supplementary simulations in which the dispersal rate per edge did not depend on patch size, and was either fixed per edge, or per patch (, with the dispersal rate per patch and the number of edges linked to the patch of origin of dispersers), which could mimic passive or active dispersal of organisms, respectively. Moreover we contrasted the initial bi-directional scenario with a strongly directionally biased scenario, where dispersal goes 99% downward and 1% upward to account for the effect of flow drift commonly found in rivers. To implement this, we simply multiplied dispersal rate per edge by a factor and , for downstream and upstream flows of dispersal, respectively. We set in the strictly bi-directional scenario and in the strongly directionally biased scenario. To make simulations comparable, we homogenized the mean dispersal per edge among dispersal mode scenarios. Based on the same settings than in figure 3, with (0.05 multiplied by 2 to account for the factor of directionality, ) in the size-scaled dispersal scenario, an equivalent is obtained for in the constant dispersal per edge scenario, and in the constant dispersal per patch. We ran simulations of these different scenarios for the 5 different networks and with the same parameters other than dispersal than in figure 3. Results show that the reversed diversity pattern is not qualitatively affected by dispersal mode or directionality (Figs S1 and S2): even if strong directionality reduces the amplitude of the diversity patterns (compare Figs S1 and S2), competitive exclusion still occurs later in small than in large patches with disturbance (Figs S1 and S2, bottom panels).

In addition, we ran a more extensive sensitivity analysis on flow directionality. We crossed 5 levels of dispersal () with 3 levels of directionality (), with patch-size dependent mortality, for the case of dispersal scaled to the volume (i.e., the dispersal mode detailed in the main text). For each dispersal x directionality combination we ran fifty simulations (5 landscape networks x 10 community matrices). Figure S3 summarizes the variation in the reversed pattern strength displayed in these 750 additional simulations. We find that the reversed pattern is robust to dispersal directionality over a wide range of parameter settings. Only very high dispersal combined with very high dispersal directionality weakens the pattern.

# Appendix B – Positive mortality-patch size relationships

We tested the occurrence of diversity patterns for positive mortality rate-patch size relationships, that is, with higher mortality rates in large downstream than in small upstream patches. We explored similar parameters than for negative slopes (Figs 4, S5, S6), that is 6 levels of dispersal with and 21 mortality rate combinations. We used all combinations with in the largest and smallest patches, with local mortality in the smallest patches being lower or equal to in the largest patches (Fig. S7). For each parameter combination we run 50 simulations (i.e., 5 replicate landscapes x 10 replicate communities) resulting in a total of 6,300 additional simulations. Figure S8 shows that increasing positive slopes strengthen the expected classic diversity patterns, until a threshold where diversity pattern shifts toward the reversed diversity pattern. Actually, the reversed pattern appears when the mortality rate is maximal and limits the persistence of species in downstream patches. Along with the patterns observed with negative patch size-mortality rate slopes, dispersal weakens the diversity patterns by homogenizing species presence across the landscape. In particular, it could shift reversed patterns to classical ones at high dispersal, especially when mortality rates in small patches is low or equal to those in large patches (Fig. S8).

# Appendix C - video analysis

Video analysis was conducted based on well-established protocols and methods (1–3). We here only describe the specific details of the analytical/measurement settings, and respective deviations of the already published protocols.

Video Records

At each sampling date, a constant volume of 34.4 µL sampled from each patch was placed under a dissecting microscope (Leica M205 C) with a 16-fold magnification, connected to a camera (Hamamatsu Orca Flash 4) and a computer, and recorded in a 5 s video. We also recorded videos of each species initial population in monoculture to serve as reference to recognize species in mixtures.

General principle of the analysis

We analysed these videos using a customized version of the R-package bemovi (1), and an image processing software (ImageJ, National Institute of Health, USA) to extract the number of moving organisms per video frame along with a suite of different morphological and movement traits for each organism (e.g., shape, size, speed). We then used these traits to filter out background movement noise (e.g., particles from the medium) and identify species in a mixture.

Getting abundance and trait data

To first detect the moving particles and filter out the background noise we used the functions locate\_and\_measure\_particles(), link\_particles() and filter\_data() of the R-package bemovi, with the following parameters: min\_size = 5, max\_size = 1000, linkrange = 2, disp = 20, net\_filter = 20, duration\_filter = 0.5, detect\_filter = 0.1, median\_step\_filter = 3. From this first step we got the number of moving particles per sample and the movement and morphological traits for each of them.

Principle of species identification analysis

To identify species in mixtures we followed the general ideas presented in (3). We used a Support Vector Machine algorithm with the function svm() in the R-package “e1071” (4). The training data set was made of the trait data set of the initial monocultures of each species, where the species of each individual is known. A model is built by providing this training data set to SVM. The model is finally run on the set of species in mixture to predict the most likely identity of each organism in the mixtures, by comparing its traits to the set of traits of reference for which the species identity is known. SVM algorithm has the advantage over RandomForest algorithm (used in (3)) to be more insensitive towards imbalanced numbers of individuals per species, which was the case here, notably with far lower densities in large than in small species populations.

Trait addition

Initial analyses of the trait distributions from monocultures showed partial significant overlaps between the smallest species (*Tetrahymena* and *Chilomonas*) and among large ones (*Blepharisma*, *Paramecium* and *Cephalodella*). To limit such misclassifications in the species identification step, we first added new traits to the set provided by the bemovi package by calculating different metrics to describe the distribution of these traits measured over the 125 frames of each video, that is the kurtosis, the skewedness, and the quartiles 25 and 75% for each trait, with the R-package “moments” (5). We furthermore (i) combined different models, which training data sets where optimized to limit specific sources of confusion, and (ii) we ran two steps of species identification, the first one to well identify the large species, the second one to focus on discriminating small species:

(i) Model combination

The principle of model combination was to run different SVM models on the species mixture, and then, for each individual in the mixture, we looked at the probability associated with the species prediction of each model and we recorded the prediction with the highest associated probability (highest confidence). The models included whether all the individuals of the monocultures or only individuals with typical traits that could discriminate one species from another with which it is commonly confounded. For instance, we built a model to discriminate well among *Paramecium*, *Cephalodella* and *Spirostomum* individuals. We looked at the trait distributions to select traits with minimum overlaps and we selected the individuals with non-overlapping values for these traits to integrate in the model. We choose thresholds values in order to integrate the maximum number of individuals while excluding outliers for these traits. This method allowed us to obtain an acceptable range of error rates for most species (<12%), which were similar to error rates obtained via visual identification among similar species with trained operators (based on our experience).

(ii) Two-step analysis

However, no model addition that we tested could improve at the same time *Tetrahymena* and *Chilomonas* error rates without increasing the error rates of other species. To reach acceptable error rates for the small species, we therefore ran a separate analysis while keeping the predictions made by the model combination above for the other species. For that, we built a new training dataset with all the individuals in monoculture but the species other than *Tetrahymena* and *Chilomonas* were grouped and identified as “Other”. Again, we combined this complete model with two “partial” models, which included only typical individuals with the most discriminating traits (i.e. excluding large *Chilomonas* individuals or small *Tetrahymena* individuals with trait similar to *Chilomonas*). We ran these three models on the individuals in mixture previously identified as *Tetrahymena* or *Chilomonas* and kept the predictions that were associated with the highest probability.

Overall, this resulted in an accuracy of species identification comparable to analogue analyses (3). Note that from the second step of the identification, about 0.7% of the individuals where identified as “Other” (large species initially identified as *Tetrahymena* or *Chilomonas*). For the sake of simplicity, we integrated these individuals in our analysis as a new species called “Other”, which corresponded to intermediate morphotypes in terms of morphology and movement behaviour. Completely omitting these individuals in the analyses did not significantly change the qualitative/quantitative results, indicating the robustness of our approach/findings.

References

1. Pennekamp F, Schtickzelle N, Petchey OL (2015) BEMOVI, software for extracting behavior and morphology from videos, illustrated with analyses of microbes. *Ecol Evol* 5(13):2584–2595.

2. Altermatt F, et al. (2015) Big answers from small worlds: a user’s guide for protist microcosms as a model system in ecology and evolution. *Methods Ecol Evol* 6(2):218–231.

3. Pennekamp F, et al. (2017) Dynamic species classification of microorganisms across time, abiotic and biotic environments - a sliding window approach. *bioRxiv*:105395.

4. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F (2015) *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Fomerly: E1071).* (TU Wien) Available at: https://CRAN.R-project.org/package=e1071.

5. Komsta L, Novomestky F (2015) *moments: Moments, cumulants, skewness, kurtosis and related tests.* Available at: https://CRAN.R-project.org/package=moments.

# Supporting Figures

Macintosh HD:Users:eric:Documents:Research:1.Projects:Eawag:Projects:8.Blue-Green:5.Redaction:1.Turnover:4.PNAS:Figures_BlueGreen_20180213:FigureS1_20180213.pdf

## Figure S1. Robustness to dispersal modes

Time at extinction for each non-dominant species in the smallest and the largest patches, with disturbance (lower panels) or without disturbance (upper panels), for different dispersal modes (columns). Dispersal rate per edge is either (**a**) depending on patch volume, like in the experiment (same panel than Fig.3b), (**b**) constant across the landscape, or (**c**) depending on patch number of edges (constant dispersal per patch). Each boxplot represents the distribution of values for one patch size across the 5 replicate landscapes. Parameters are the same as in Fig.3, except the dispersal, which is set to conserve the same mean dispersal per edge across the among dispersal modes: , and in (a), (b), and (c) respectively, with an unbiased directionality: 50% of dispersal goes in upstream and 50% in downstream direction (see methods in Appendix A).

Macintosh HD:Users:eric:Documents:Research:1.Projects:Eawag:Projects:8.Blue-Green:5.Redaction:1.Turnover:4.PNAS:Figures_BlueGreen_20180213:FigureS2_20180213.pdf

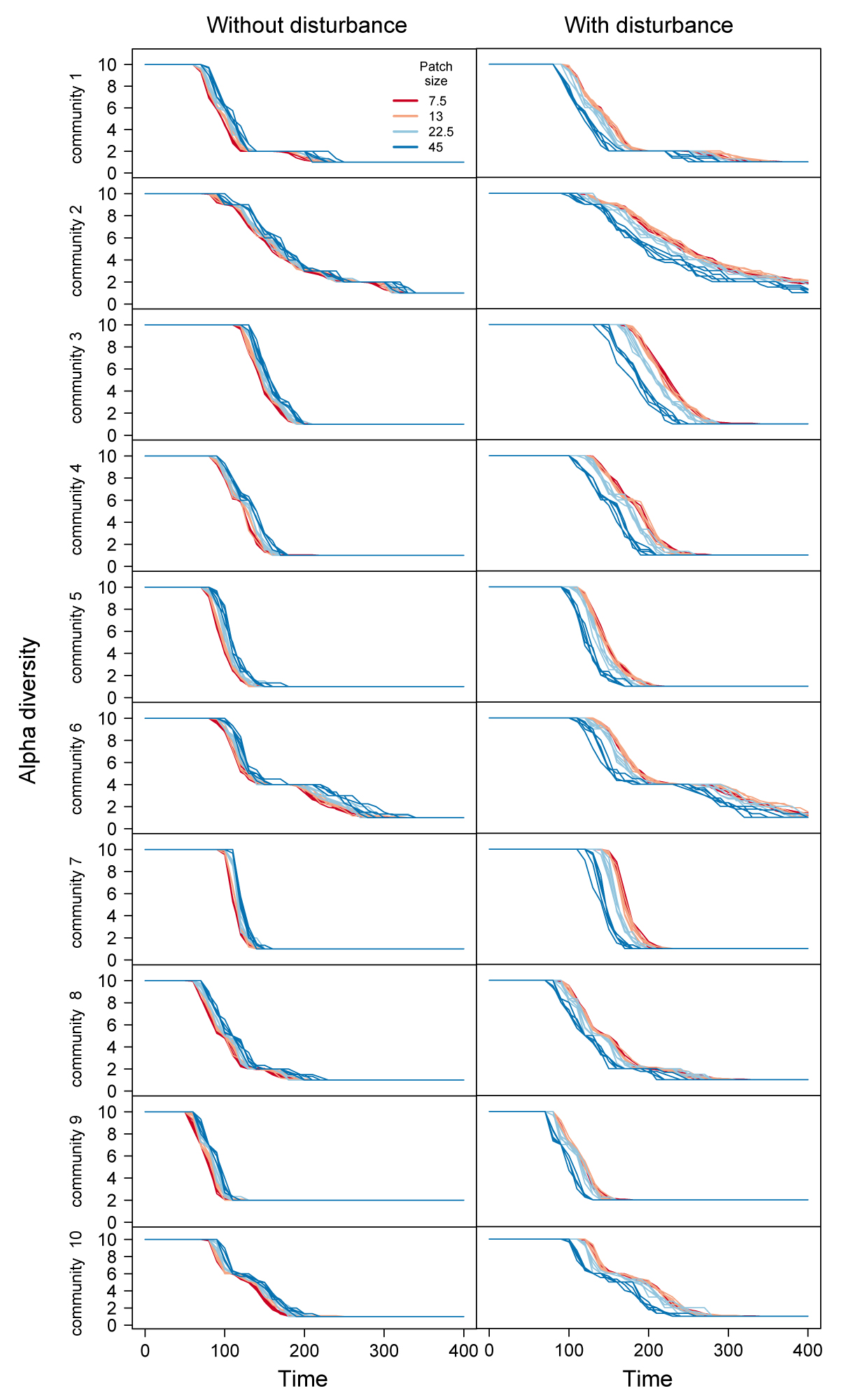
## Figure S2. Robustness to dispersal directionality

This is the same figure as Fig.S1, but with an additional directional bias of dispersal: dispersal goes 1% upstream versus 99% downstream (see methods in Appendix A). Values are times at extinction for each non-dominant species in the smallest and the largest patches, with disturbance (lower panels) or without disturbance (upper panels), for different dispersal modes (columns). Dispersal rate per edge is either (**a**) depending on patch volume, like in the experiment (same panel than Fig.3b), (**b**) constant across the landscape, or (**c**) depending on patch number of edges (constant dispersal per patch). Each boxplot represents the distribution of values for one patch size across the 5 replicate landscapes. Parameters are the same as in Fig.3, except the dispersal, which is set to conserve the same mean dispersal per edge across the among dispersal modes: , and in (a), (b), and (c) respectively, with an unbiased directionality: 50% of dispersal goes in upstream and 50% in downstream direction (see methods in Appendix A).

**Macintosh HD:Users:eric:Documents:Research:1.Projects:Eawag:Projects:8.Blue-Green:5.Redaction:1.Turnover:4.PNAS:Figures_BlueGreen_20180213:FigureS3_20180213.pdf**

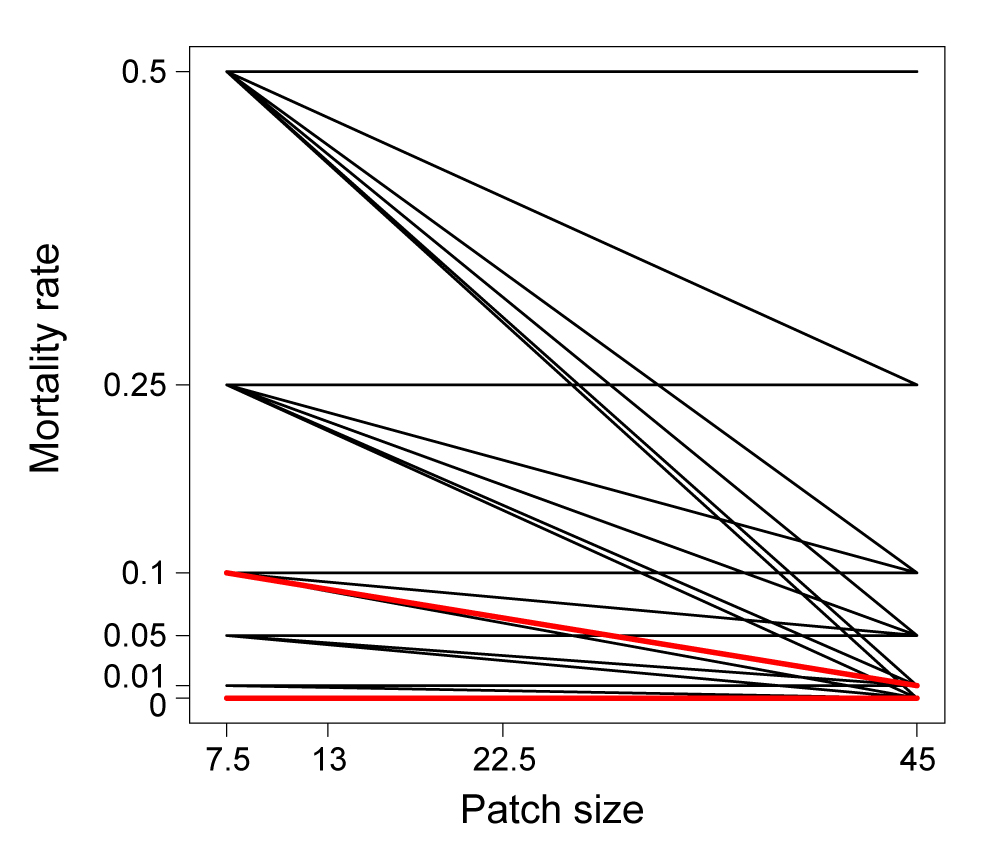
## Figure S3. Sensitivity to dispersal rate and directionality

The figure illustrates the parameter space within which the diversity pattern is found as a function of the magnitude of dispersal and the proportion of dispersal going downstream (increasing directionality), with patch-size dependent mortality (mortality rates are 0.1, 0.0868, 0.064 and 0.01 respectively from smaller to larger patches). D–U corresponds to the maximum observed difference, along the temporal dynamics, in mean richness between largest and smallest patches. When D–U is positive (green colours) the classical pattern emerges over the course of the dynamics, with higher diversity in large than in small patches, while when the difference is negative (pink colours) it is the reversed pattern, with higher diversity in small (i.e., headwater) rather than in large patches. Each square correspond to an average over fifty simulations (5 landscapes x 10 community matrices).



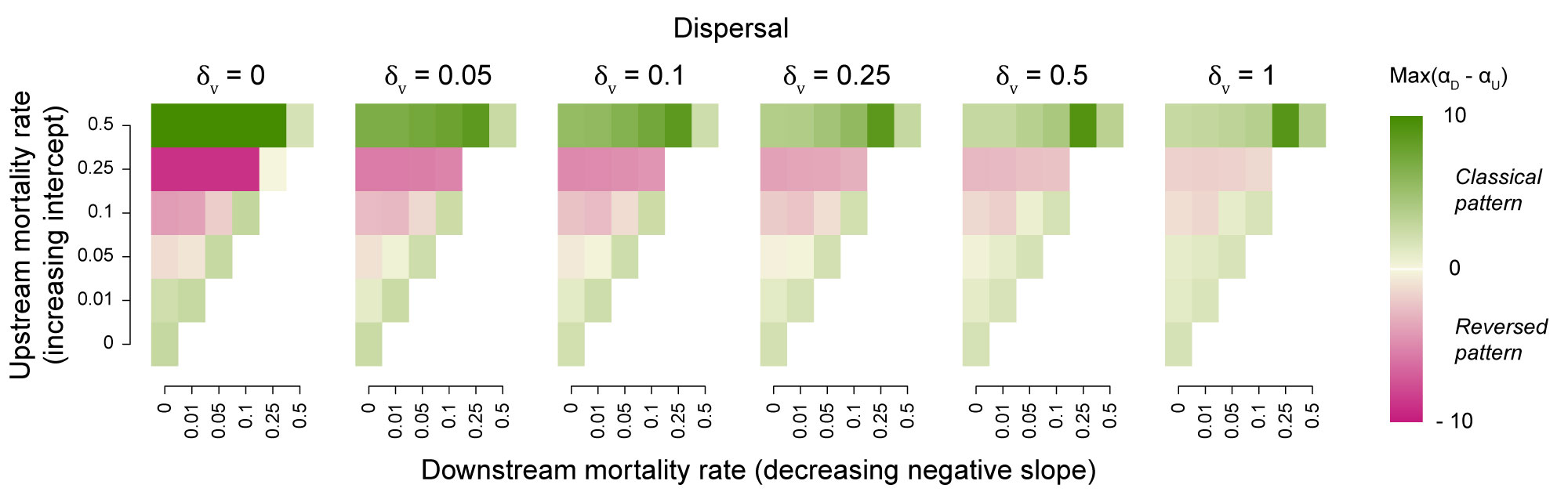
## Figure S4. Temporal dynamics of alpha diversity (model)

Species richness in local patches in dendritic, river-like networks for different communities (species sets) in rows, and with or without disturbance in columns. Each line corresponds to the alpha diversity averaged over all the patch of a given size (colours) in one landscape. Small patches are in red and large ones in blue. There are five lines per colour, one for each of the five landscapes. Mortality rates in the case with disturbance (right column) are 0.1, 0.0868, 0.064 and 0.01 for the 4 patch sizes with increasing sizes, and dispersal volume .

****

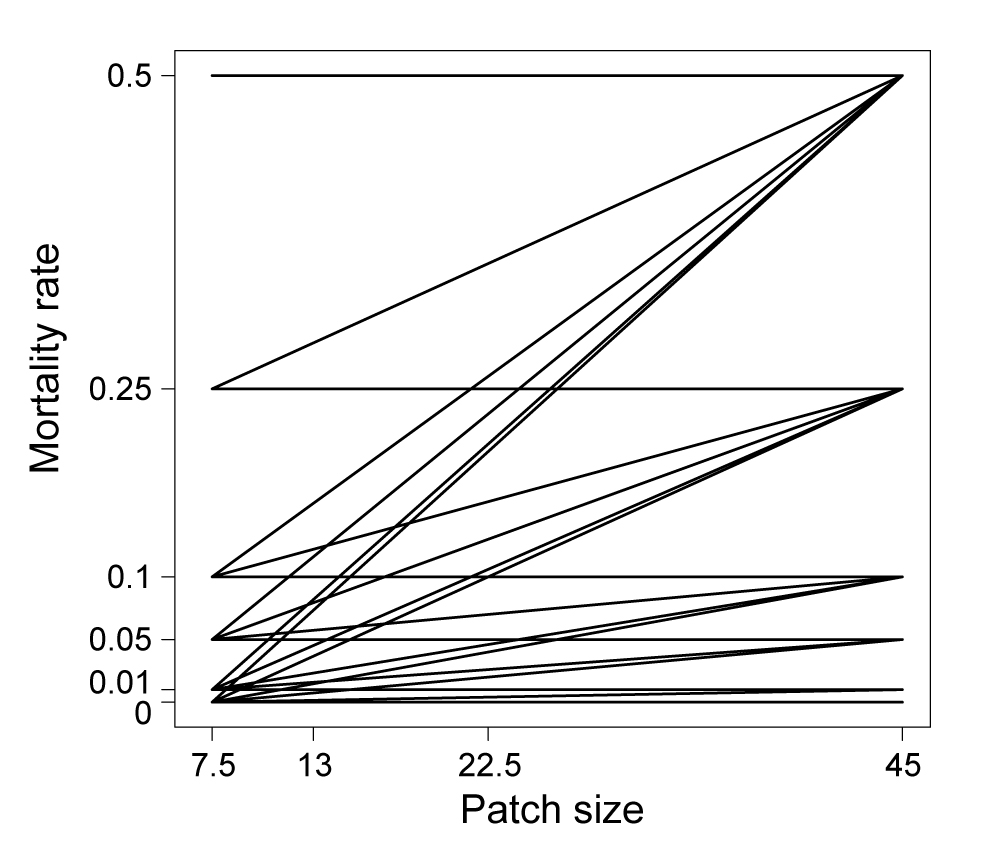
## Figure S5. Patch size – mortality rate relationships

Relationships between patch size and mortality rate used in the model to produce Figs. 4 and S6. The red bold lines show the combinations used to produce Figs. 1-3. In our model, patch size can be expressed in any dimension of volume (L3). To match our experiment, the unit would be mL. Dimension of mortality rates is T-1.



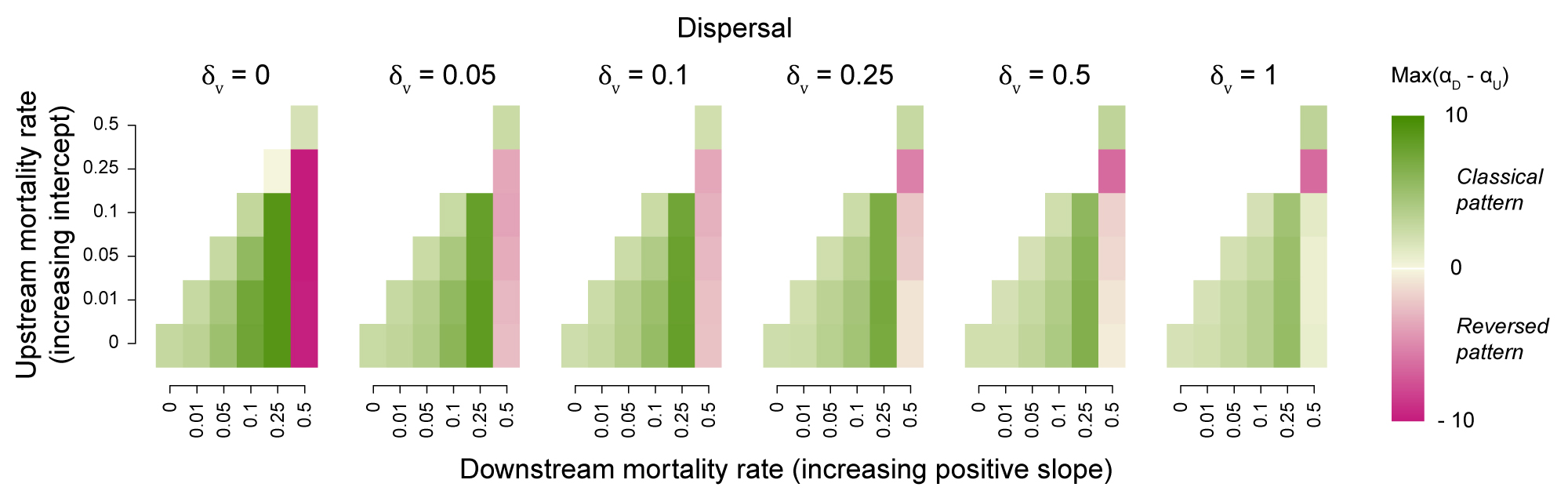
## Figure S6. Diversity pattern with mortality and dispersal rates (model)

Spatial diversity pattern shift with mortality and dispersal rates (, the volume dispersed, related to dispersal rate as: , with the volume of the patch = patch size). Colours represent the maximum difference of richness between downstream (large) and upstream (small) patches. When this difference is positive (green colours) the classical pattern emerges over the course of the dynamics, with higher diversity in large than in small patches, while when the difference is negative (pink colours) it is the reversed pattern, with higher diversity in small than in large patches. Each case represents an average over 50 replicate simulations (5 different landscapes x 10 different communities). Mortality rate increase in small upstream patches raises the intercept of the negative mortality rate-patch size relationship, while mortality rate increase in large dowstream patches, reduces its slope.

****

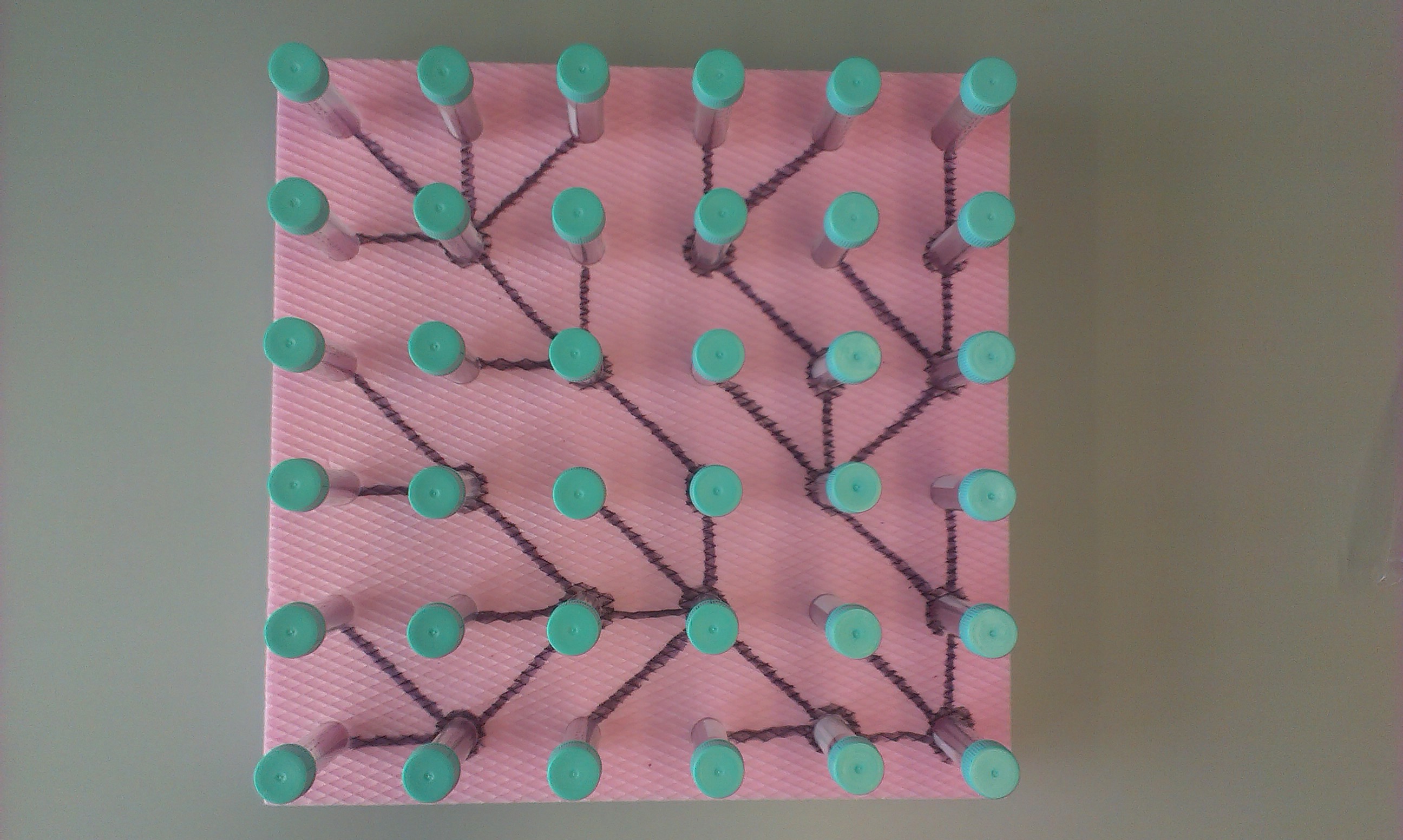
## Figure S7. Patch size – mortality rate positive relationships

Relationships between patch size and mortality rate used in the model to produce Figs. S11-12. In our model, patch size can be expressed in any dimension of volume (L3). To match our experiment, the unit would be mL. Dimension of mortality rates is T-1.



## Figure S8. Diversity patterns with positive patch size–mortality relationships (model)

Spatial diversity pattern shift with mortality and dispersal rates (, the volume dispersed, related to dispersal rate as: , with the volume of the patch = patch size). Colours represent the maximum difference of richness between downstream (large) and upstream (small) patches. When this difference is positive (green colours) the classical pattern emerges over the course of the dynamics, with higher diversity in large than in small patches, while when the difference is negative (pink colours) it is the reversed pattern, with higher diversity in small than in large patches. Each case represents an average over 50 replicate simulations (5 different landscapes x 10 different communities). Mortality rate increase in small upstream patches raises the intercept of the positive mortality rate-patch size relationship, while mortality rate increase in large dowstream patches, increases its slope.



## Figure S9. River-like landscape (experiment)

One realization of our four experimental dendritic, river-like landscapes.

****

## Figure S10. Temporal dynamics of alpha diversity (experiment)

Diversity distribution in dendritic, river-like landscapes, with disturbance, over experimental days. The figure illustrates how the effect of patch size on protist species richness changed over time as the effect of mortality on ecological dynamics unfolded. Each point is the mean species richness ± SE for patch size 7.5 (N=85), 13 (N=34), 22.5 (14) and 45 (11) over experimental days. See Table S1 for linear mixed effect results.

Macintosh HD:Users:eric:Documents:Research:1.Projects:Eawag:Projects:8.Blue-Green:5.Redaction:1.Turnover:5.ProcB:PROT_NET_COMP.pdf

## Figure S11. Community structure over time.

Changes in relative abundances over time (day 7 and 29) for each landscape replicate (A to D). Numbers indicate patch size in volume for each patch.

# Supplementary Tables

## Table S1. Variables and parameters of the model

M, L, T refers to mass, length and time dimensions respectively.

|  |  |  |  |
| --- | --- | --- | --- |
| Symbol | Definition | Dimension | Values used in main figures |
|  | Abundance of species i’s population in patch *x* (= population size)\* | M | - |
|  | Intrinsic growth rate of species *i*  (all species have the same ) | T-1 | 0.25 |
|  | per capita effect of species *j* on species *i* | ø | Negative values from a gaussian distribution  (,) |
|  | Carrying capacity of species *i* in patch *x* (all species have the same ) | M | 2000 \* |
|  | Dispersal rate per edge from patch *x* | T-1 |  |
|  | Dispersal volume per unit of time and per edge | L3 T-1 | 0, 0.01,0.05, 0.1, 0,25,0.5,1 |
|  | Dispersal rate per patch | T-1 | - |
|  | Volume of patch *x* (= patch size) | L3 | 7.5, 13, 22.5, 45 |
|  | Mortality rate in patch *x*  (all species have the same ) | T-1 | 0, 0.01, 0.05, 0.1, 0.25, 0.5 (values used for the largest and smallest patches) |
|  | Magnitude of stochasticity | ø | 0.25 |
|  | Random factor for the stochasticity term | ø | from a gaussian distribution  (,) |

\*Note that abundance and carrying capacities can take decimal values, because it is not an individual-based model. It can be viewed as the biomass corresponding to population abundance, with a mass dimension, M.

## Table S2. Effect of patch size on protist alpha diversity (experiment)

Effect of patch size on protist species richness over continuous time or at day 29. For the linear mixed effect model over time, we added landscape replicates and time as nested random factors to control for temporal pseudo-replications issues. For the LME model at day 20 we added landscape replicates as a random factor. The models were fitted by maximizing the restricted log-likelihood.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Estimate | SD | DF | t-value | p-value |
|  | Effect of patch size over time | | | | |
| Intercept | 8.48 | 0.15 | 565.00 | 55.19 | 0.0000 |
| Size13 | -0.25 | 0.28 | 565.00 | -0.88 | 0.3772 |
| Size22.5 | 0.11 | 0.40 | 565.00 | 0.28 | 0.7770 |
| Size45 | 0.59 | 0.44 | 565.00 | 1.33 | 0.1834 |
| day | -0.13 | 0.01 | 565.00 | -16.39 | 0.0000 |
| Size13:day | 0.01 | 0.01 | 565.00 | 1.02 | 0.3105 |
| Size22.5:day | -0.04 | 0.02 | 565.00 | -1.99 | 0.0474 |
| Size45:day | -0.08 | 0.02 | 565.00 | -3.42 | 0.0007 |
|  | Effect of patch size on last day (day 29) | | | | |
| Intercept | 4.72 | 0.15 | 137.00 | 30.72 | 0.0000 |
| Size13 | 0.14 | 0.29 | 137.00 | 0.47 | 0.6384 |
| Size22.5 | -1.36 | 0.41 | 137.00 | -3.34 | 0.0011 |
| Size45 | -2.17 | 0.45 | 137.00 | -4.80 | 0.0000 |