Appendix to « The evolution of dispersal and its consequences on α and β-diversity in a heterogeneous metacommunity »

Fabien Laroche, Philippe Jarne, Thomas Perrot & Francois Massol

Table of contents

Appendix 1: Metacommunities harbouring similar mean and variance but varying asymmetry of the carrying capacity distribution – examples using two values of carrying capacity

We define a metacommunity with two values of carrying capacity using three parameters:

- K_1 and K_2 the carrying capacities;

- π_1 the proportion of communities that have carrying capacity K_1.

We introduce the mean carrying capacity of metacommunities:

$$\bar{K} = \pi_1 K_1 + (1 - \pi_1)K_2$$

The squared coefficient of variation of carrying capacity γ_2 verifies:

$$\gamma_2 = \pi_1(1 - \pi_1)\left(\frac{K_2 - K_1}{\bar{K}}\right)^2$$

The standardized skewness of carrying capacity γ_3 verifies:

$$\gamma_3 = \frac{2\pi_1 - 1}{\sqrt{\pi_1(1 - \pi_1)}}$$

We aim at building landscapes with different standardized skewness (γ_3) while keeping squared coefficient of variation (γ_2) and $\bar{K}$ constant. As γ_3 only depends on π_1, controlling γ_3 is equivalent to controlling π_1. We may obtain different values of γ_3 using π_1 and then adjust values of K_1 and K_2 to obtain the same values for γ_2 and $\bar{K}$. The difficulty essentially lies with the fact that K_1 and K_2 must be non-negative integers. In the following, we assume, without loss of generality, that $K_2 \geq K_1$. Because we want to measure Simpson diversity in communities, we also impose $K_1 \geq 2$. We look for values of $\bar{K}$ and π_1 such that any integer value of K_1 such as $2 \leq K_1 \leq \bar{K}$ can generate valid metacommunities. K_2 can be expressed as a function of K_1, $\bar{K}$ and π_1:

$$K_2 = \frac{\bar{K} - \pi_1 K_1}{1 - \pi_1}$$

K_2 must be an integer for any K_1 such as $2 \leq K_1 \leq \bar{K}$, which is equivalent to imposing that $(\bar{K} - 2\pi_1)/(1 - \pi_1)$ and $\pi_1/(1 - \pi_1)$ are integers. $\pi_1/(1 - \pi_1)$ being a non-zero integer means that there exists a strictly positive integer k such as $\pi_1 = k/(k+1)$. Then:

$$\frac{\bar{K} - 2\pi_1}{1 - \pi_1} = \bar{K}(k+1) - 2k$$

$\bar{K}$ has to then be an integer. For a given integer value of $\bar{K} \geq 2$, we have a set of possible metacommunities:

$$(K_1, K_2, \pi_1) \in \mathcal{A} = \left\{ l, \bar{K} + k(\bar{K} - l), \frac{k}{k+1}, 2 \leq l \leq \bar{K}, 1 \leq k \right\}$$

We now seek to extract from $\mathcal{A}$ metacommunities having the same variation coefficient of carrying capacities. A metacommunity in $\mathcal{A}$ with parameters k and l has the coefficient of variation

$$\gamma_2 = k \left(1 - \frac{l}{\bar{K}} \right)^2$$

We then consider a metacommunity 0 with minimal $k_0 = 1$ and $l_0 = 2$ and we seek metacommunities in $\mathcal{A}$ with the same γ_2. These metacommunities verify:

$$k \left(1 - \frac{l}{\bar{K}} \right)^2 = \left(1 - \frac{2}{\bar{K}} \right)^2$$

Thus:

$$k = \left(\frac{\bar{K} - 2}{\bar{K} - l} \right)^2$$

k has to be a squared integer so that we introduce q such as $k = q^2$. (1) implies that q verifies:

$$q(\bar{K} - l) = \bar{K} - 2$$

so that q divides $\bar{K} - 2$. l can be define from q:

$$l = \frac{\bar{K}(q-1) + 2}{q}$$

We thus can provide a set of metacommunities with two only distinct values of carrying capacity such that $\bar{K}$ is constant for all the metacommunities and the coefficient of variation of all the metacommunities is also constant and verifies $\gamma_2 = (1 - 2/\bar{K})^2$. This set writes:

$$\mathcal{B} = \left\{ \left(\frac{\overline{K}(q-1)+2}{q}, \overline{K} + q(\overline{K} - 2), \frac{q^2}{q^2+1} \right), q \in \mathcal{D}_{\overline{K}-2} \right\}$$

where $\mathcal{D}_{\overline{K}-2}$ is the set of positive integers that divide $\overline{K} - 2$. This construction can easily be generalized to provide sets of metacommunities with only two distinct values of carrying capacity such that $\overline{K}$ is constant for all the metacommunities and the coefficient of variation of all the metacommunities is also constant and verifies $\gamma_2 = (1 - l/\overline{K})^2$ with $l \geq 2$ although we focus here on $l = 2$. Setting $\overline{K} = 8$ yields following possible values of q: 1, 2, 3, 6. Because, increasing q increases γ_3, the set $\mathcal{B}$ with $\overline{K} = 8$ allows generating metacommunities where the evolution of dispersal is predicted to undergo either evolutionary stable equilibrium or branching (Massol *et al.* 2011). We could generate two metacommunities (Figure A1.1) corresponding to following values of (K_1, K_2, π_1):

Metacommunity	Q	k	K_1	K_2	π_1
1	1	1	2	14	0.5
2	6	36	7	44	0.972973

Figures

Figure A1.1: Evolutionary scenario as a function of (γ_2, γ_3) as predicted by Massol et al. (2010). $\overline{K}$ is set to 8 and three values of dispersal cost c are explored: 0, 0.1, 0.5. The black zone corresponds to values of (γ_2, γ_3) that are not attainable because communities' carrying capacity must be superior or equal to 1. Black continuous curve shows the frontier between dispersal ESS scenario (below the curve) and branching scenario (above the curve) in terms of (γ_2, γ_3) for $c = 0.1$ (value used in main text). The same curves (dashed lines) are reported for $c = 0$ and 0.5. Dots represent metacommunities generated by our method. The two metacommunities used in main text are tagged here with numbers. The corresponding parameters are indicated in Table 1 of main text.
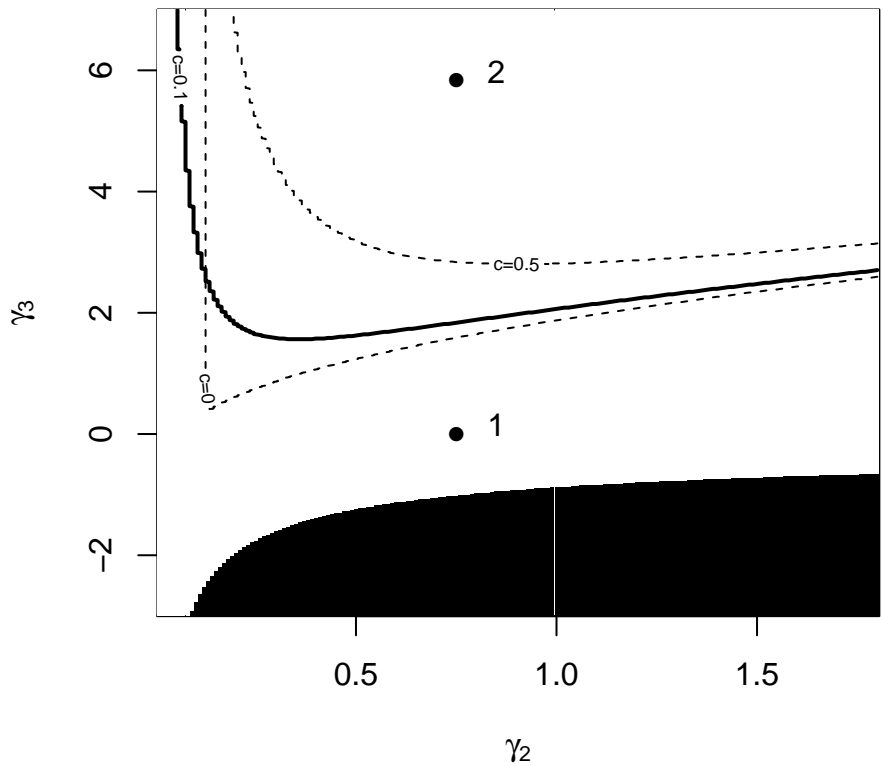
Appendix 2: Sensitivity of dispersal (as a function of time) to the distribution of carrying capacities (γ_2, γ_3) and dispersal cost (c) – a simulation exploration

The dispersal strategies obtained at evolutionary equilibrium depends on the simulation parameters. Figure A2.1 shows that a higher cost leads to a lower dispersal strategy in the ESS scenario as predicted by classic dispersal evolution theory. We also observed that higher cost leads to closer branches in the EB scenario.

Figures

Figure A2.1: Dispersal distribution through time in model M_1 for metacommunities 1 and 2 (see Table 1) for different values of dispersal cost c. Panels A and C (resp. panels B and D) present the symmetric (resp. asymmetric) community with $c = 0$ and 0.5 respectively. For each panel, one simulation was performed over 20000 generations (=20000*8880 cycles, see main text). Iconography and other parameter values are identical to Figure 1.
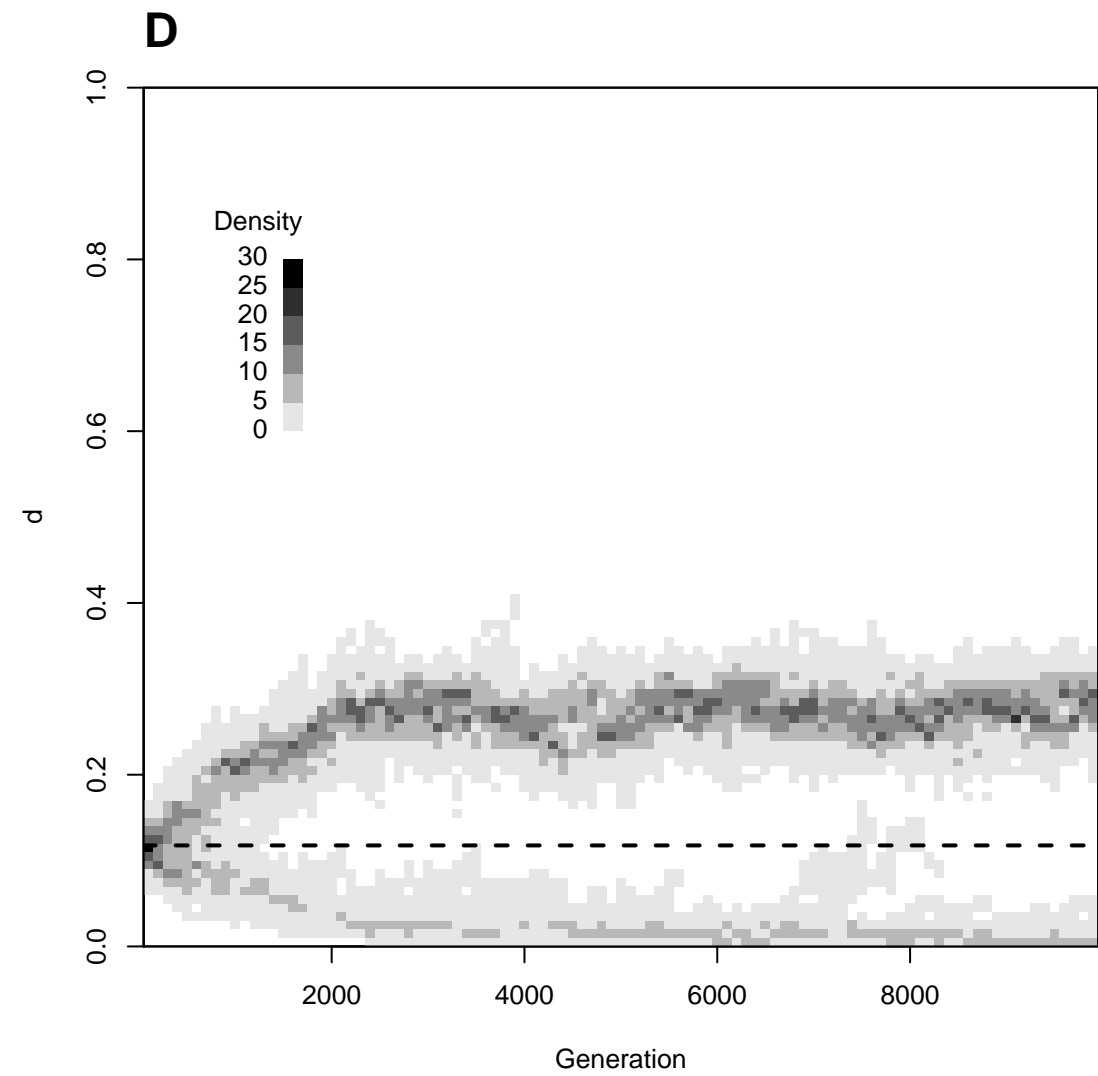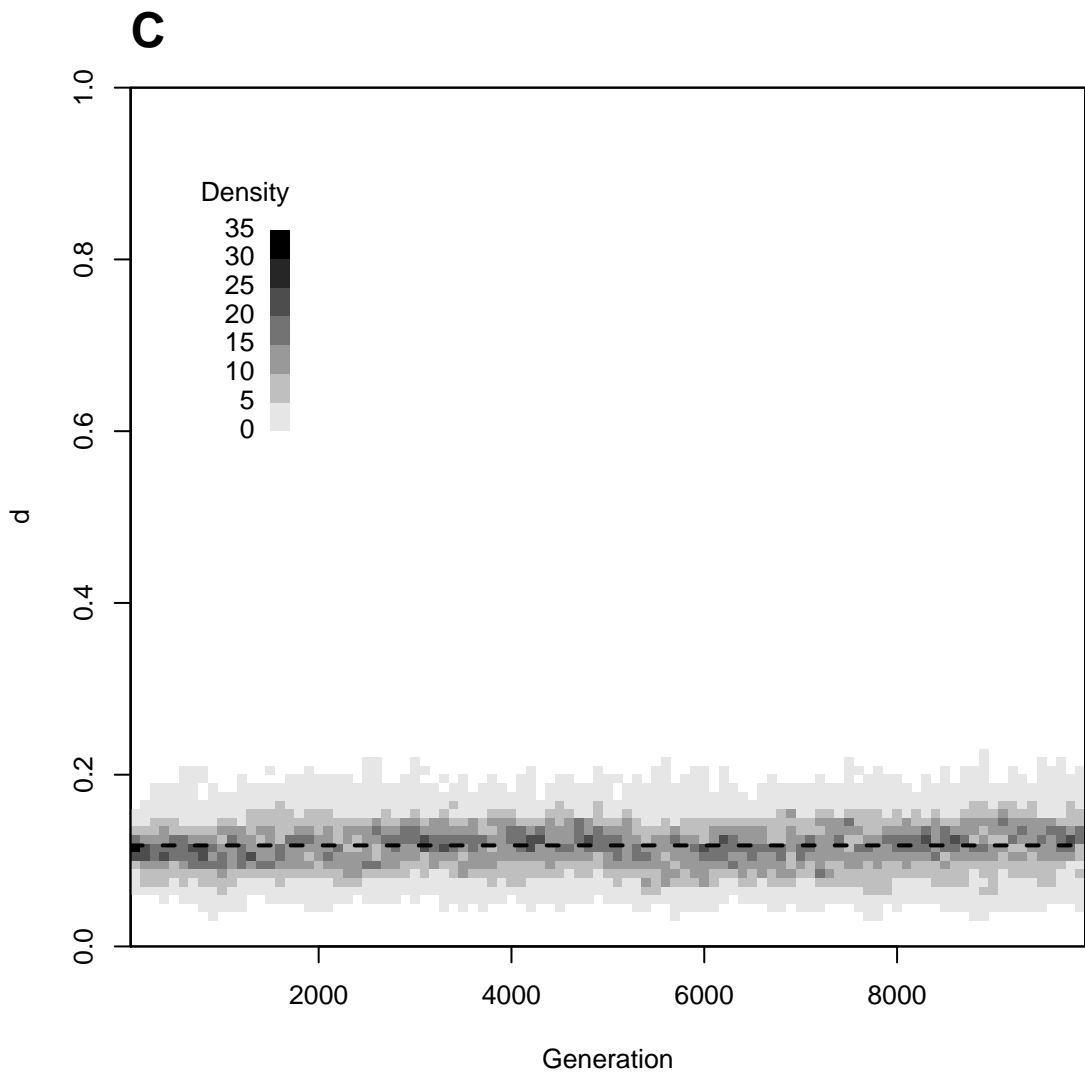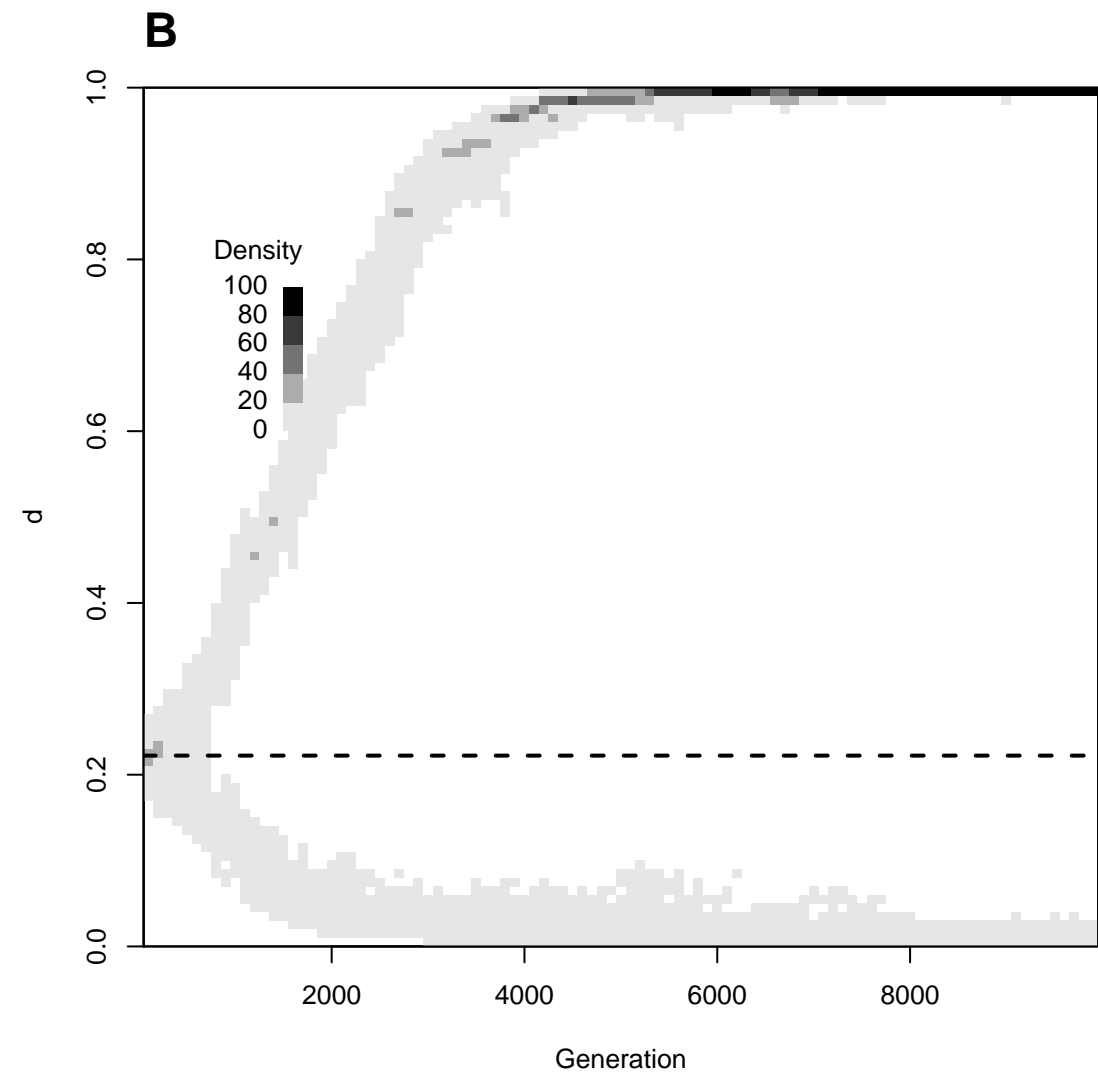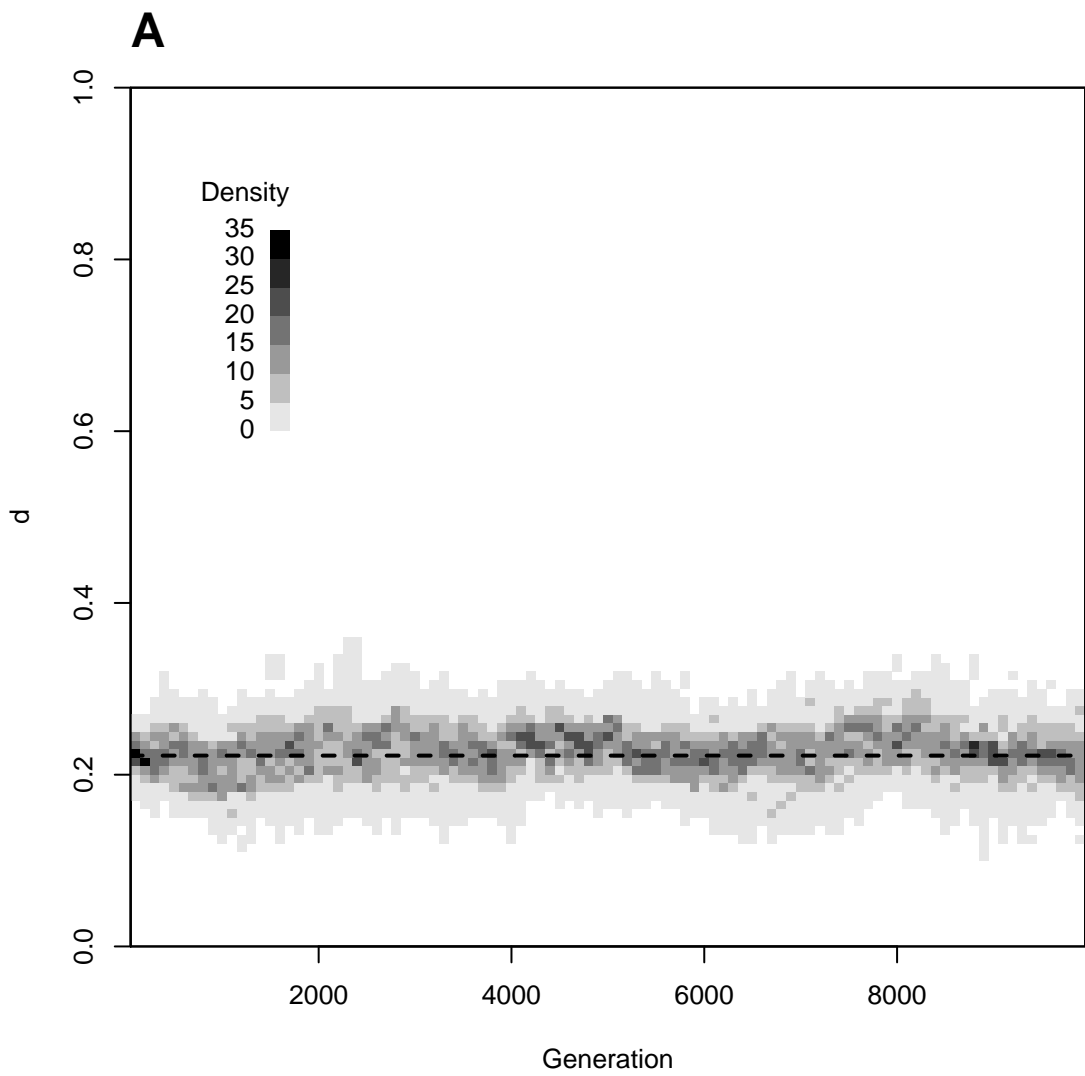
Figure A2.2: Dispersal values under the EB scenario for different values of the coefficient of variation (γ_2) and the asymmetry (γ_3) of the distribution of carrying capacities within the metacommunity. (A) $\gamma_3 = 2.27$; (B) $\gamma_3 = 2.73$; (C) $\gamma_3 = 4.13$. Blue (resp. grey) boxes represent the average distribution of the dispersal trait in the lower (resp. upper) branch over 12 replicated simulations.
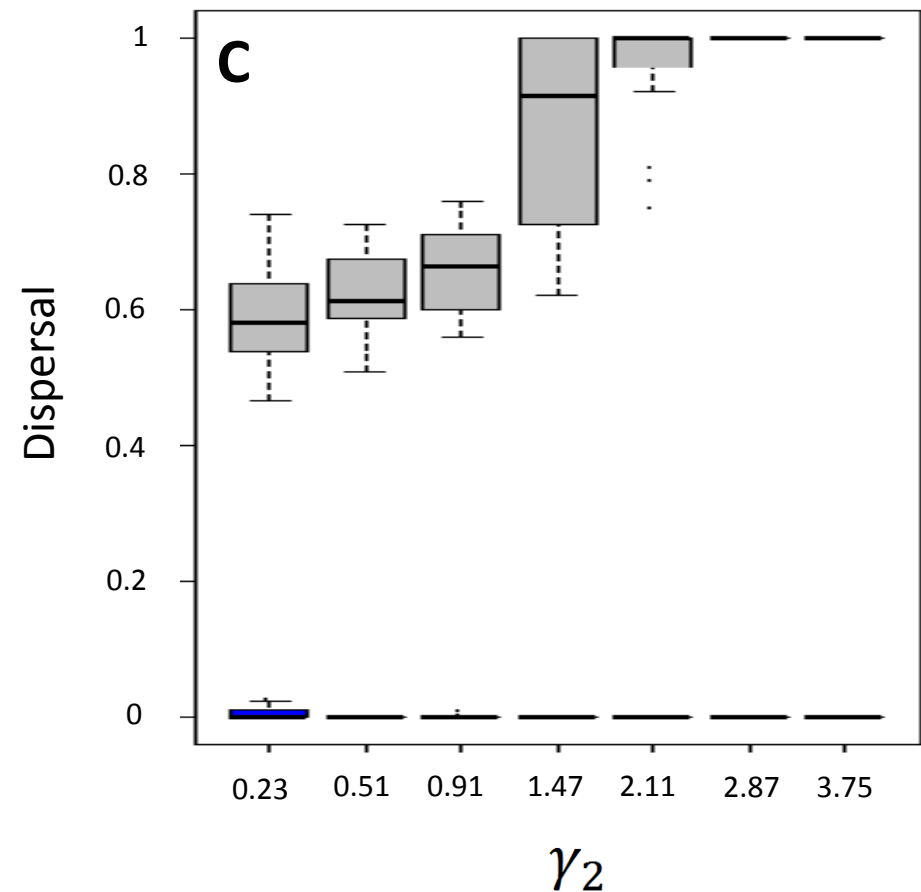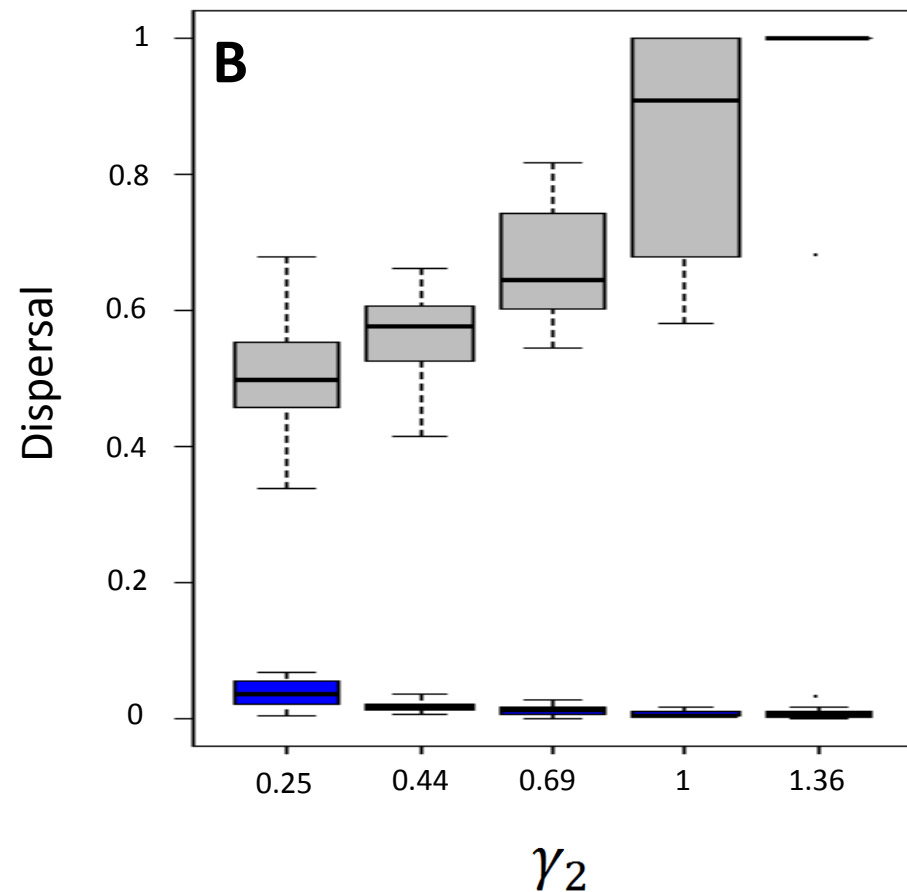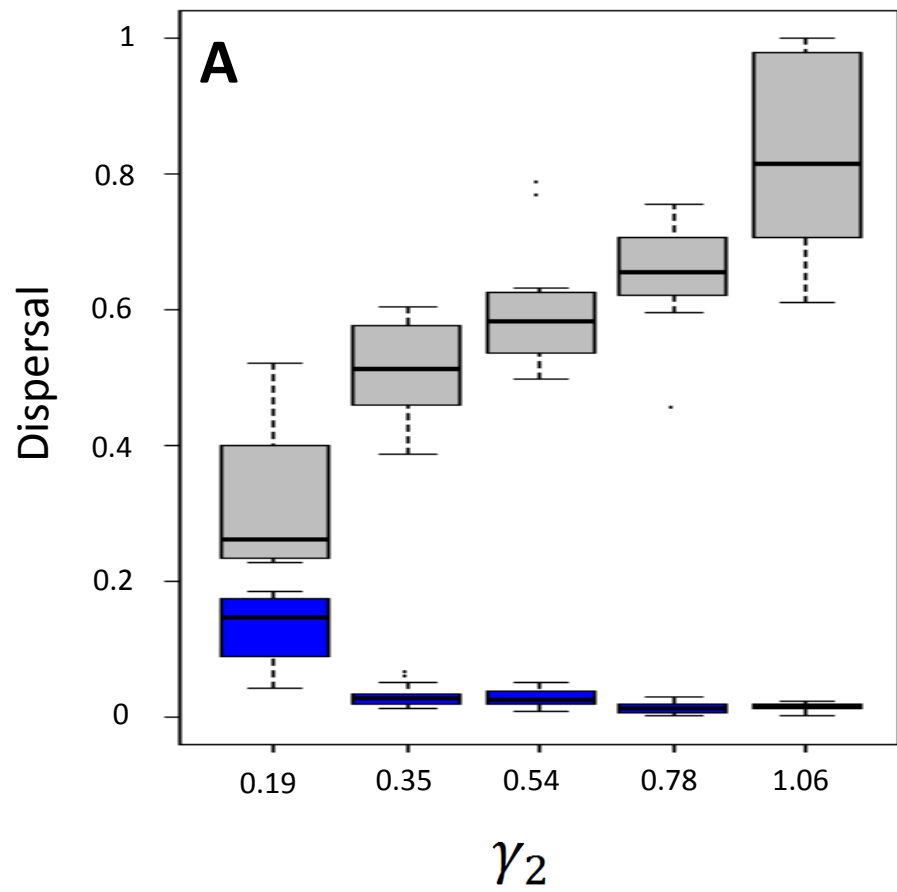
Appendix 3: The "capacity dispersal pattern" - theoretical derivation and robustness to dispersal cost

We explain why, in model M_1, the expected average dispersal value within a community decreases as the carrying capacity of the considered community increases - a feature that we call "carrying capacity-dispersal relationship". We do not consider the species label of individuals throughout this section.

We consider the limit case of model M_1 when the number of communities is infinite (*i.e.* $N \to +\infty$). We assume that:

- dispersal traits in the metacommunity have reached a stationary distribution at the metacommunity scale under which various trait values $\vec{D} = (d_1, d_2, \dots)$ have constant frequencies $\vec{F} = (f_1, f_2, \dots)$ with $\sum_i f_i = 1$.

- communities have very large carrying capacities (i.e. $K \to +\infty$ for all communities). This allows us to overlook the effects of demographic stochasticity within communities, which is not necessary to explain the pattern that we study here.

- $r = \frac{K}{\bar{K}}$ is finite for all the communities (no community is infinitely larger than the others), so that the individual contribution of any community to the metacommunity structure can always be neglected. This allows us to describe the model through a mainland-island dynamics in which all the communities receive migrants from a common regional pool with structure $(\vec{D}, \vec{F})$.

Because a community harbours an infinite number of individuals in this limit case, we describe its state through the local frequencies of each dispersal type, noted $\vec{x} = (x_1, x_2, \dots)$. In our asymptotic case in which there is no local demographic stochasticity, we look for the migration-selection equilibrium value $\overrightarrow{x_{eq}}$ which verifies:

$$\forall i, \beta_i(\overrightarrow{x_{eq}}) = \delta_i(\overrightarrow{x_{eq}}) \tag{3.1}$$

where $\beta_i(\vec{x})$ (resp. $\delta_i(\vec{x})$) is the rate of increase (resp. decrease) of the local proportion of dispersal type i. We derive their expression in our asymptotic case:

$$\begin{cases} \beta_i(\vec{x}) = (1-x_i)\dfrac{x_i(1-d_i)+f_i d_i\frac{(1-c)}{r}}{1-\bar{d}_{loc}+\bar{d}_{reg}\frac{(1-c)}{r}} \\ \delta_i(\vec{x}) = x_i\left(1-\dfrac{x_i(1-d_i)+f_i d_i\frac{(1-c)}{r}}{1-\bar{d}_{loc}+\bar{d}_{reg}\frac{(1-c)}{r}}\right) \end{cases} \tag{3.2}$$

with $\bar{d}_{loc}$ (resp. $\bar{d}_{reg}$) the average dispersal trait of the individuals in the community (resp. in the regional pool).

Plugging (3.2) in (3.1) yields the system:

$$\forall i, (1-x_i)\left(x_i(1-d_i)+f_i d_i\frac{(1-c)}{r}\right) = x_i\left(1-\bar{d}_{loc}+\bar{d}_{reg}\frac{(1-c)}{r}-x_i(1-d_i)-f_i d_i\frac{(1-c)}{r}\right)$$

$$\forall i, x_i = \frac{f_i d_i\frac{(1-c)}{r}}{d_i-\bar{d}_{loc}+\bar{d}_{reg}\frac{(1-c)}{r}} \tag{3.3}$$

which implies:

$$\forall i, f_i d_i\frac{(1-c)}{r} < d_i - \bar{d}_{loc} + \bar{d}_{reg}\frac{(1-c)}{r}$$

$$\forall i, \bar{d}_{loc} < d_i + \frac{(1-c)}{r}\left(\bar{d}_{reg}-f_i d_i\right)$$

$$\bar{d}_{loc} < \min_i\left[d_i + \frac{(1-c)}{r}\left(\bar{d}_{reg}-f_i d_i\right)\right]$$

In addition, $\bar{d}_{loc} > \min_i[d_i]$ so that:

$$\min_i[d_i] < \bar{d}_{loc} < \min_i\left[d_i + \frac{(1-c)}{r}\left(\bar{d}_{reg}-f_i d_i\right)\right]$$

We define the function $S_r(\delta)$ for $\delta \in \left[\min_i(d_i), \min_i\left[d_i + \frac{(1-c)}{r}\left(\bar{d}_{reg}-f_i d_i\right)\right]\right]$ as:

$$S_r(\delta) = \sum_i \frac{f_i d_i\frac{(1-c)}{r}}{d_i-\delta+\bar{d}_{reg}\frac{(1-c)}{r}}$$

To verify the system (3.3), $\bar{d}_{loc}$ must necessarily verify the equation:

$$S_r(\bar{d}_{loc}) = 1 \tag{3.4}$$

Function S_r verifies the following three properties:

12

$$S_r(\min_i(d_i)) \leq 1$$

$$S_r\left(\min_i\left[d_i + \frac{(1-c)}{r}(\bar{d}_{reg} - f_i d_i)\right]\right) \geq 1 \tag{3.5}$$

$$\forall \delta \in \left[\min_i(d_i), \min_i\left[d_i + \frac{(1-c)}{r}(\bar{d}_{reg} - f_i d_i)\right]\right], S_r'(\delta) > 0$$

The intermediate value theorem then ensures that a unique value $\bar{d}_{loc}$ verifying (3.4) exists. This value depends on the parameter r of the considered community so that we note it $\bar{d}_{loc}(r)$. The goal of this section is then to show that $\bar{d}_{loc}(r)$ is a decreasing function of r. We define the bivariate function $\varphi(r, \delta)$ as follows:

$$\varphi(r, \delta) = \sum_i \frac{f_i d_i \frac{(1-c)}{r}}{d_i - \delta + \bar{d}_{reg}\frac{(1-c)}{r}}$$

In particular, φ verifies:

$$\varphi\left(r, \bar{d}_{loc}(r)\right) = 1 \tag{3.6}$$

A differentiation of (3.6) along variable r yields:

$$\partial_r \varphi\left(r, \bar{d}_{loc}\right) + \partial_\delta \varphi\left(r, \bar{d}_{loc}\right)\bar{d}_{loc}'(r) = 0$$

$$\bar{d}_{loc}'(r) = -\frac{\partial_r \varphi(r, \bar{d}_{loc})}{\partial_\delta \varphi(r, \bar{d}_{loc})} \tag{3.7}$$

$$\partial_r \varphi\left(r, \bar{d}_{loc}\right) = -\frac{1}{r}\sum_i \frac{f_i d_i \frac{(1-c)}{r}}{d_i - \bar{d}_{loc} + \bar{d}_{reg}\frac{(1-c)}{r}}\frac{d_i - \bar{d}_{loc}}{d_i - \bar{d}_{loc} + \bar{d}_{reg}\frac{(1-c)}{r}}$$

$$\delta \to \frac{\delta - \bar{d}_{loc}}{\delta - \bar{d}_{loc} + \bar{d}_{reg}\frac{(1-c)}{r}}$$ is a concave function so that Jensen inequality implies:

$$\sum_i \frac{f_i d_i \frac{(1-c)}{r}}{d_i - \bar{d}_{loc} + \bar{d}_{reg}\frac{(1-c)}{r}}\frac{d_i - \bar{d}_{loc}}{d_i - \bar{d}_{loc} + \bar{d}_{reg}\frac{(1-c)}{r}} < 0$$

$$\partial_r \varphi\left(r, \bar{d}_{loc}\right) > 0 \tag{3.8}$$

In addition, using (3.5) one gets:

$$\partial_\delta \varphi\left(r, \bar{d}_{loc}\right) = S_r'\left(\bar{d}_{loc}\right) > 0 \tag{3.9}$$

Plugging (3.8) and (3.9) in (3.7) yields

$$\bar{d}_{loc}'(r) = -\frac{\partial_r \varphi(r, \bar{d}_{loc})}{\partial_\delta \varphi(r, \bar{d}_{loc})} < 0$$

In other words, the expected local dispersal value decreases with K.

13

Figure 2 of main text illustrates the capacity-dispersal pattern for metacommunity 1 and 2 for dispersal cost c=0.1. Similar results are obtained when using other values of dispersal cost (Figures A3.1 and A3.2). Note that the representations in these figures are different from those of main text. In main text we had 100 independent simulations and compared the distribution of observed versus randomized statistics. Here we ran a single simulation and compare the observed value with a distribution of randomized values computed from the same observation (more details in figures legends). Figure A3.1 shows that changing the dispersal cost does not affect the results about dispersal pattern for the symmetric metacommunity. The "environmental filtering" is always significant when accounting for intraspecific variance in dispersal and randomizing individuals position (Figure A3.1A) and but not significant when overlooking intraspecific variance and randomizing only species traits (Figure A3.1B). Dispersal is always significantly higher in small communities that in high communities when accounting for intraspecific variance in dispersal and randomizing individuals position (Figure A3.1C) but not (or hardly) significant when overlooking intraspecific variance and randomizing only species traits.

Figures

Figure A3.1: $T_{IC/IR}$ and the difference in average dispersal trait between small and large communities in the symmetric metacommunity for different values of the dispersal cost c. For each value of c, we performed one simulation over 20000 generation and analysed the final state. In each panel, we presented results for low cost (c=0) on the left and results for high cost (c=0.5) on the right. Dots are used to represent the observed value of the considered statistic at this final state, while bars present the average of 100 values of the statistic obtained from independent randomization of this final state. Error bars show the 3% and 97% quantiles

of the empirical distribution of the randomized statistic. (A): $T_{IC/IR}$ when accounting for intraspecific variance in dispersal and randomizing individuals position in the metacommunity. (B): $T_{IC/IR}$ when overlooking intraspecific variance in dispersal and permuting species dispersal values while preserving species abundances in communities. (C): Average dispersal in small communities minus average dispersal in large communities (Δd) when accounting for intraspecific variance in dispersal and randomizing individuals position in the metacommunity. (D): Δd when overlooking intraspecific variance in dispersal and permuting species dispersal levels while preserving species abundances in communities.
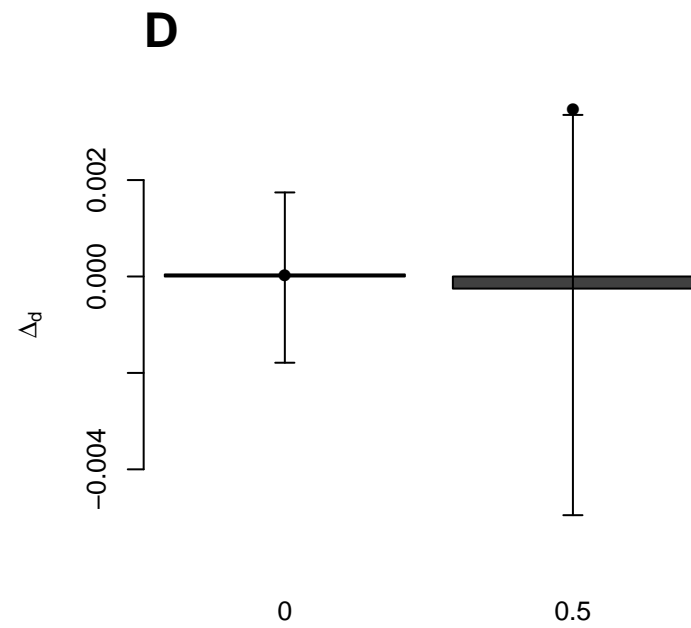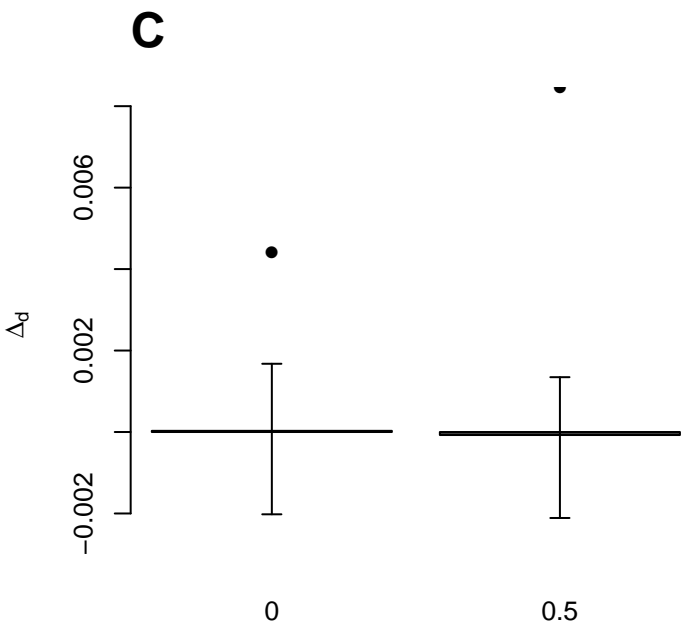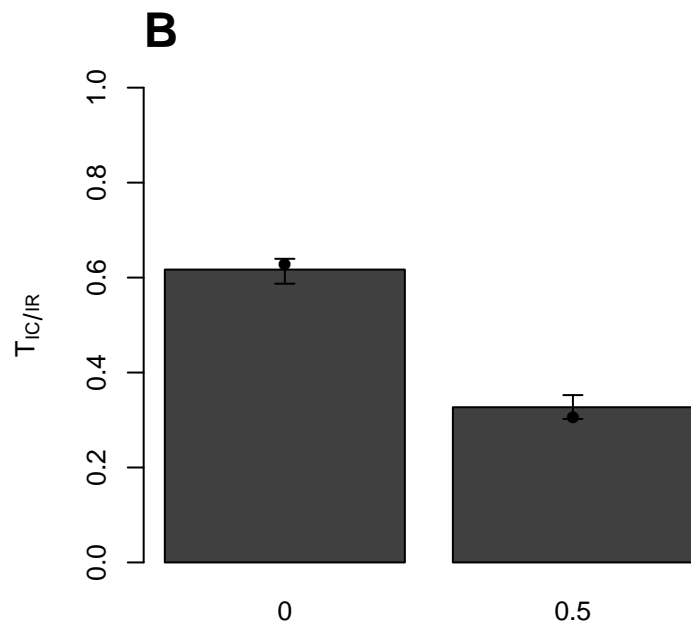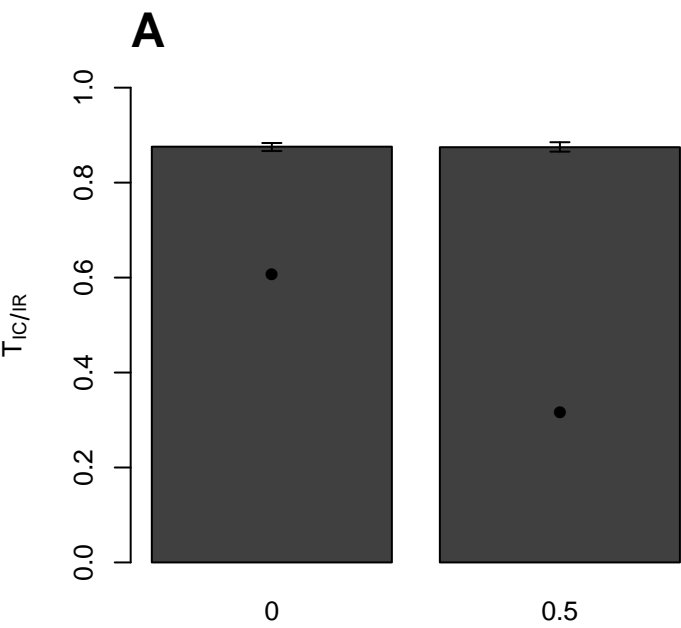
Figure A3.2: The capacity-dispersal pattern in the asymmetric metacommunity for different values of dispersal cost c. The caption is exactly the same than Figure A3.1 but for the asymmetric community.
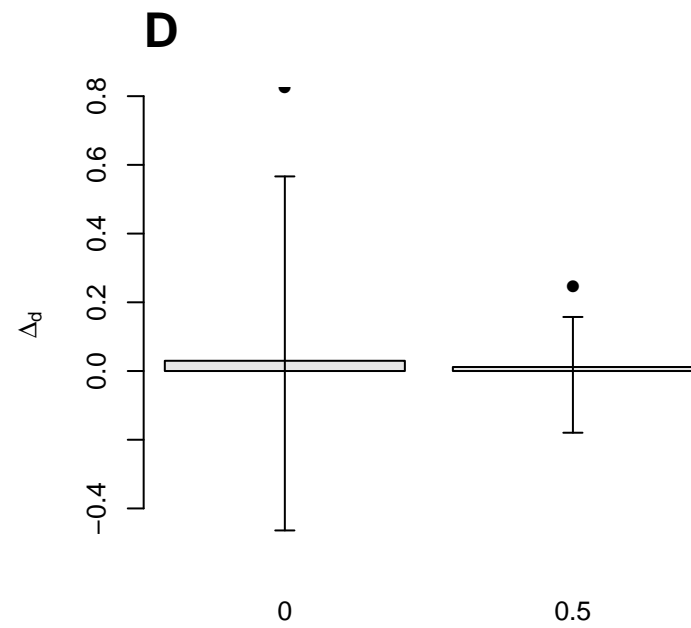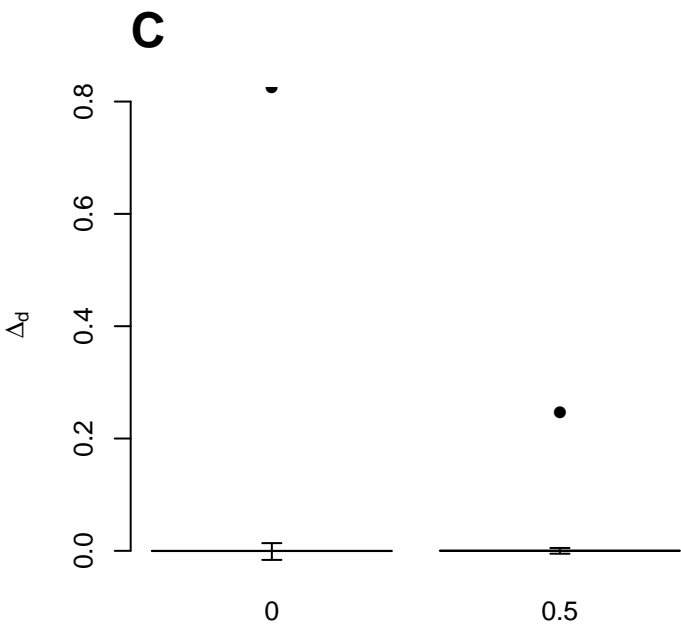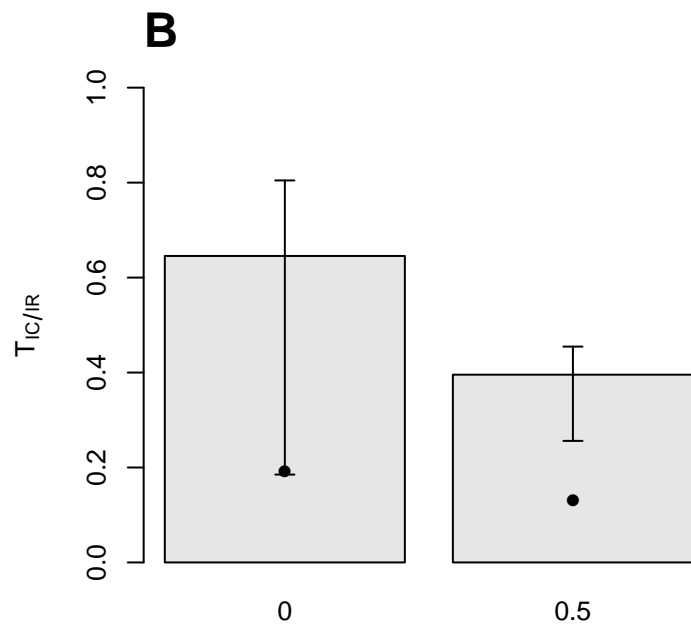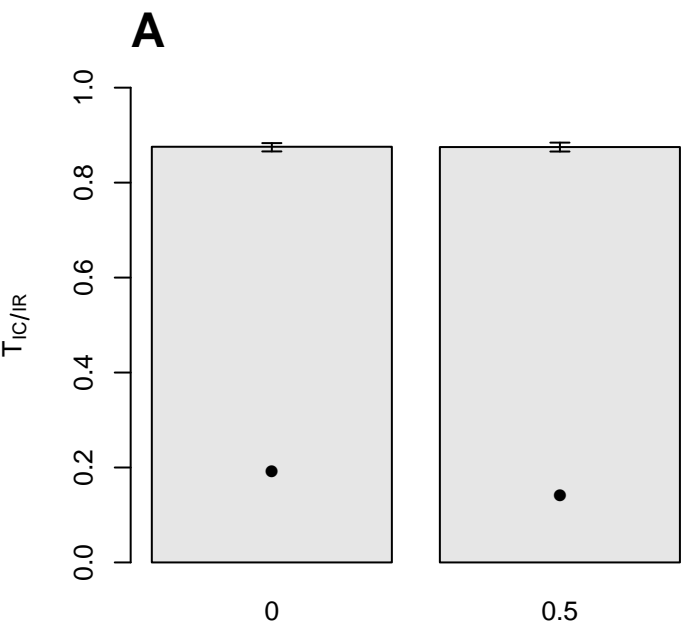
Appendix 4: Species diversity index in the neutral metacommunity – theoretical prediction

We recall that $D_\alpha(K_i)$ is the probability that two individuals sampled in the same community with carrying capacity K_i belong to different species, $D_\beta(K_i, K_j)$ is the probability that two individuals coming from distinct communities with carrying capacities K_i and K_j belong to different species and D_γ is the probability that two individuals randomly sampled in the metacommunity (any individual from any community) belong to different species.

We provide here (i) analytical expressions of $D_\alpha(K_i)$, $D_\beta(K_i, K_j)$ and D_γ in the neutral metacommunity model M_0. We first emphasize that D_γ can be deduced from the $D_\alpha(.)$s and $D_\beta(.,.)$s using the following equation:

$$D_\gamma = \left(\sum_i \frac{N_i K_i(K_i-1)}{J(J-1)} D_\alpha(K_i)\right) + \left(\sum_{i<j} 2\frac{N_i N_j K_i K_j}{J(J-1)} D_\beta(K_i, K_j)\right) + \left(\sum_i \frac{N_i(N_i-1)K_i^2}{J(J-1)} D_\beta(K_i, K_i)\right) \qquad (4.1)$$

Therefore we will focus on providing analytical expression and estimates of the $D_\alpha(.)$s and $D_\beta(.,.)$s and use equation (4.1) to infer corresponding results for D_γ.

Defining the coalescence Markov chain

The D_αs and the D_βs can be derived for any K_i, K_j values by considering the coalescence (*i.e.* the genealogy; Figure A4.1) of lineages associated to the sampled individuals. Two individuals belong to different species if and only if, going backward in time, their lineages do not merge into a single ancestral one before a speciation event occurs in either of them. Arbitrarily labelling the two lineages considered as lineage 1 and 2, the coalescence can be described as a Markov chain over the following states of the pair of lineages:

- states $C(K_i)$: the two lineages are in the same community with carrying capacity K_i;

- states $F(K_i, K_j)$: the two lineages are in different communities, lineage 1 is in a community with carrying capacity K_i while lineage 2 is in a community with carrying

capacity K_j (it is possible to have $K_i = K_j$ when the two lineages are in distinct community with the same carrying capacity);

- state M: the two lineages have merged into a single one;

- state S: one of the two lineages has undergone a speciation event.

State M and S are absorbing states of the Markov chain.

$D_\alpha(K_i)$ is the probability that the Markov chain ends in S when starting from the initial state $C(K_i)$. $D_\beta(K_i, K_j)$ is the probability that the Markov chain ends in S when starting from the initial states $F(K_i, K_j)$ or $F(K_j, K_i)$.

Computing transition probabilities of the coalescence Markov chain

For any pair of states A and B, denote $p_{A \to B}$ the probability that the Markov chain currently in state A switches to state B when a birth-death cycle occurs backward in time. Note that a lineage is affected by the next birth-death cycle to occur backward in time only if the current individual representing the lineage comes from the reproduction event of this cycle (e.g. cycles 2 and 5 in Figure A4.1). For a Markov chain in state A, its state should not change if the next cycle backward in time did not produce either of the two individuals traced in state A. In all the following, we therefore consider transition probabilities from A to B conditionally to the fact that one of the two individuals representing the lineages in A comes from the reproduction event of the next cycle backward in time. In other words, we overlook cycles that do not contribute to lineages dynamics (e.g. cycles 1, 3 and 4 in Figure A4.1).

Assume that the Markov chain is currently in state $C(K_i)$. We first assume that lineage 1 comes from the reproduction event of the next cycle backward in time (which occurs with probability $\frac{1}{2}$). We denote with exponent 1 the corresponding transition probabilities. We call C the community in which both lineages occur in current state. Various transitions can occur:

- Transition to state M occurs if the individual that reproduces at the next cycle backward in time is the individual representing lineage 2.

20

- Transition to state $C(K_i)$ (i.e. not changing state) occurs if the individual that reproduces at the next cycle backward in time is in C but is not the individual representing lineage 2.

- Transitions to state $F(K_j, K_i)$ for any K_j occur if the individual that reproduces at the next cycle backward in time is not in C and is in a community with carrying capacity K_j.

- Transition to state S occurs if a point speciation occurs at the next cycle backward in time.

The weight an individual belonging to C in the lottery for reproduction at the next cycle backward in time is the sum of its local reproductive effort $1 - d$ and the share of the propagules that fall in its own patch $\frac{d(1-c)}{N}$. $K_i - 1$ individuals belongs to C and participate in the lottery: the individual representing lineage 2 and $K_i - 2$ other individuals. In addition to those individuals in C, all the individuals occupying other communities also participate in the lottery with a per capita weight of $\frac{d(1-c)}{N}$. There are $J - K_i$ such individuals. Among them, $N_l K_l$ individuals come from communities with carrying capacity K_l for $l \neq i$ where N_l is the number of such communities. $(N_i - 1)K_i$ individuals come from communities with carrying capacity K_i (community C has been removed).

We can then identify the weight of events of interest in the lottery for reproduction:

- "the individual that reproduces is the one representing lineage 2" has weight $w_M = 1 - d + \frac{d(1-c)}{N}$.

- "the individual that reproduces is another individual from C" has weight $w_C = (K_i - 2)\left(1 - d + \frac{d(1-c)}{N}\right)$.

- "the individual that reproduces is not from C and is from a community with carrying capacity K_l, $l \neq i$" has weight $w_{F,l} = N_l K_l \frac{d(1-c)}{N}$;

- "the individual that reproduces is not from $\mathcal{C}$ and is from a community with carrying capacity K_i" has weight $w_{F,i} = (N_i - 1)K_i \frac{d(1-c)}{N}$.

If speciation occurs during the cycle, with probability v, which of this event actually occurs does not matter as the transition will be to S state in any case. The lottery is considered only if speciation does not occur, which has probability $1 - v$. Transition probability then writes:

$$p^1_{C(K_i)\to S} = v$$

$$p^1_{C(K_i)\to C(K_i)} = (1-v)\frac{w_C}{w_C+w_M+w_{F,i}+\sum_{l\neq i}w_{F,l}} = (1-v)\frac{(K_i-2)\left(1-d+\frac{d(1-c)}{N}\right)}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p^1_{\substack{C(K_i)\to F(K_j,K_i)\\ j\neq i}} = (1-v)\frac{w_{F,j}}{w_C+w_M+w_{F,i}+\sum_{l\neq i}w_{F,l}} = (1-v)\frac{N_jK_j\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p^1_{C(K_i)\to F(K_i,K_i)} = (1-v)\frac{w_{F,i}}{w_C+w_M+w_{F,i}+\sum_{l\neq i}w_{F,l}} = (1-v)\frac{(N_i-1)K_i\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p^1_{C(K_i)\to M} = (1-v)\frac{w_M}{w_C+w_M+w_{F,i}+\sum_{l\neq i}w_{F,l}} = (1-v)\frac{1-d+\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

Other transitions, provided that lineage 1 is the one coming from the reproduction at the next cycle backward in time, are not possible and have probability 0.

We denote with exponent 2 the transition probabilities when assuming that lineage 2 is the one coming from the reproduction event. Similar analysis than above yields:

$$p^2_{C(K_i)\to S} = v$$

$$p^2_{C(K_i)\to C(K_i)} = (1-v)\frac{w_C}{w_C+w_M+w_{F,i}+\sum_{l\neq i}w_{F,l}} = (1-v)\frac{(K_i-2)\left(1-d+\frac{d(1-c)}{N}\right)}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p^2_{\substack{C(K_i)\to F(K_i,K_j)\\ j\neq i}} = (1-v)\frac{w_{F,j}}{w_C+w_M+w_{F,i}+\sum_{l\neq i}w_{F,l}} = (1-v)\frac{N_jK_j\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p^2_{C(K_i)\to F(K_i,K_i)} = (1-v)\frac{w_{F,i}}{w_C+w_M+w_{F,i}+\sum_{l\neq i}w_{F,l}} = (1-v)\frac{(N_i-1)K_i\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p^2_{C(K_i)\to M} = (1-v)\frac{w_M}{w_C+w_M+w_{F,i}+\sum_{l\neq i}w_{F,l}} = (1-v)\frac{1-d+\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

Other transitions have probability 0.

The unconditional transition probability $p_{A \to B}$ can be retrieved from the transition probabilities conditionally on which lineage comes from the reproduction at next cycle backward in time as follows: $p_{A \to B} = p_{A \to B}^1 + p_{A \to B}^2$. This yields the unconditional transition probabilities:

$$p_{C(K_i) \to S} = v$$

$$p_{C(K_i) \to C(K_i)} = (1-v) \frac{(K_i-2)\left(1-d+\frac{d(1-c)}{N}\right)}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$\underset{j \neq i}{p_{C(K_i) \to F(K_i,K_j)}} = \underset{j \neq i}{p_{C(K_i) \to F(K_j,K_i)}} = \frac{1}{2}(1-v) \frac{N_j K_j \frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p_{C(K_i) \to F(K_i,K_i)} = (1-v) \frac{(N_i-1)K_i\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p_{C(K_i) \to M} = (1-v) \frac{1-d+\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

Assume that Markov chain is currently in state $F(K_i, K_j)$ with $K_i \neq K_j$. We call $\mathcal{C}_1$ (resp. $\mathcal{C}_2$) the community in which the individual representing lineage 1 (resp. lineage 2) occurs at current state.

We first assume that lineage 1 is the one involved and derive the transition probability from this situation. Various transitions can occur:

- Transition to state M occurs if the individual that reproduces at the next cycle backward in time is the individual representing lineage 2. The weight of this event in the lottery is $w_M = \frac{d(1-c)}{N}$.

- Transition to state $C(K_j)$ occurs if the individual that reproduces at the next cycle backward in time is in $\mathcal{C}_2$ but is not the individual representing lineage 2. The weight of this event in the lottery is $w_C = (K_j - 1)\frac{d(1-c)}{N}$.

- Transitions to state $F(K_l, K_j)$ for $K_l \neq K_i, K_j$ occur if the individual that reproduces at the next cycle backward in time is in a community with carrying capacity K_l (which cannot be the community of either individuals representing lineage 1 or 2 in current state). The weight of this event in the lottery is $w_{F,l} = N_l K_l \frac{d(1-c)}{N}$.

- Transition to state $F(K_i, K_j)$ occurs if the individual that reproduces at the next cycle backward in time is either in C_1 or in another a community with carrying capacity K_i. The weight of this event in the lottery is $w_{F,i} = (K_i - 1)\left(1 - d + \frac{d(1-c)}{N}\right) + (N_i - 1)K_i \frac{d(1-c)}{N}$.

- Transition to state $F(K_j, K_j)$ occurs if the individual that reproduces at the next cycle backward in time is in a community with carrying capacity K_j that is not C_2. The weight of this event in the lottery is $w_{F,j} = (N_j - 1)K_j \frac{d(1-c)}{N}$.

- Transition to state S occurs if a point speciation occurs at the next cycle backward in time.

Transition probabilities conditionally to the fact that lineage 1 is the one coming from the reproduction event at next cycle backward in time then writes:

$$p^1_{F(K_i,K_j)\to S} = v$$

$$p^1_{F(K_i,K_j)\to C(K_j)} = \frac{(1-v)w_C}{w_C+w_M+w_{F,i}+w_{F,j}+\sum_{l\neq i,j}w_{F,l}} = \frac{(1-v)(K_j-1)\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p^1_{\substack{F(K_i,K_j)\to F(K_l,K_j) \\ l\neq i,j}} = \frac{(1-v)w_{F,l}}{w_C+w_M+w_{F,i}+w_{F,j}+\sum_{l\neq i,j}w_{F,l}} = \frac{(1-v)N_l K_l\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p^1_{F(K_i,K_j)\to F(K_i,K_j)} = \frac{(1-v)w_{F,i}}{w_C+w_M+w_{F,i}+w_{F,j}+\sum_{l\neq i,j}w_{F,l}} = (1-v)\frac{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(N_i-1)K_i\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p^1_{F(K_i,K_j)\to F(K_j,K_j)} = \frac{(1-v)w_{F,j}}{w_C+w_M+w_{F,i}+w_{F,j}+\sum_{l\neq i,j}w_{F,l}} = \frac{(1-v)(N_j-1)K_j\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p^1_{F(K_i,K_j)\to M} = (1-v)\frac{w_M}{w_C+w_M+w_{F,i}+w_{F,j}+\sum_{l\neq i,j}w_{F,l}} = \frac{(1-v)\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

Similary, we derive the transition probabilities conditionally to the fact that lineage 2 comes from the reproduction event at the next cycle backward in time:

$$p^2_{F(K_i,K_j)\to S} = v$$

$$p^2_{F(K_i,K_j)\to C(K_i)} = \frac{(1-v)(K_i-1)\frac{d(1-c)}{N}}{(K_j-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_j)\frac{d(1-c)}{N}}$$

$$p^2_{\substack{F(K_i,K_j)\to F(K_i,K_l)\\l\neq i,j}} = \frac{(1-v)N_lK_l\frac{d(1-c)}{N}}{(K_j-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_j)\frac{d(1-c)}{N}}$$

$$p^2_{F(K_i,K_j)\to F(K_i,K_j)} = (1-v)\frac{(K_j-1)\left(1-d+\frac{d(1-c)}{N}\right)+(N_j-1)K_j\frac{d(1-c)}{N}}{(K_j-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_j)\frac{d(1-c)}{N}}$$

$$p^2_{F(K_i,K_j)\to F(K_i,K_i)} = \frac{(1-v)(N_i-1)K_i\frac{d(1-c)}{N}}{(K_j-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_j)\frac{d(1-c)}{N}}$$

$$p^2_{F(K_i,K_j)\to M} = \frac{(1-v)\frac{d(1-c)}{N}}{(K_j-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_j)\frac{d(1-c)}{N}}$$

We can then derive the unconditional probabilities:

$$p_{F(K_i,K_j)\to S} = v$$

$$p_{F(K_i,K_j)\to C(K_i)} = \frac{1}{2}\frac{(1-v)(K_i-1)\frac{d(1-c)}{N}}{(K_j-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_j)\frac{d(1-c)}{N}}$$

$$p_{F(K_i,K_j)\to C(K_j)} = \frac{1}{2}\frac{(1-v)(K_j-1)\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p_{\substack{F(K_i,K_j)\to F(K_i,K_l)\\l\neq i,j}} = \frac{1}{2}\frac{(1-v)N_lK_l\frac{d(1-c)}{N}}{(K_j-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_j)\frac{d(1-c)}{N}}$$

$$p_{\substack{F(K_i,K_j)\to F(K_l,K_j)\\l\neq i,j}} = \frac{1}{2}\frac{(1-v)N_lK_l\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p_{F(K_i,K_j)\to F(K_i,K_j)} = \frac{1-v}{2}\left[\frac{(K_j-1)\left(1-d+\frac{d(1-c)}{N}\right)+(N_j-1)K_j\frac{d(1-c)}{N}}{(K_j-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_j)\frac{d(1-c)}{N}}+\frac{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(N_i-1)K_j\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}\right]$$

$$p_{F(K_i,K_j)\to F(K_i,K_i)} = \frac{1}{2}\frac{(1-v)(N_i-1)K_i\frac{d(1-c)}{N}}{(K_j-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_j)\frac{d(1-c)}{N}}$$

$$p_{F(K_i,K_j)\to F(K_j,K_j)} = \frac{1}{2}\frac{(1-v)(N_j-1)K_j\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p_{F(K_i,K_j)\to M} = \frac{1}{2}\left[\frac{(1-v)\frac{d(1-c)}{N}}{(K_j-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_j)\frac{d(1-c)}{N}} + \frac{(1-v)\frac{d(1-c)}{N}}{(K_j-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_j)\frac{d(1-c)}{N}}\right]$$

Assume that Markov chain is currently in state $F(K_i, K_i)$. We call $\mathcal{C}_1$ (resp. $\mathcal{C}_2$) the community where the individual representing lineage 1 (resp. lineage 2) occurs at current state.

We first assume that lineage 1 is the one involved and derive the transition probability from this situation. Various transitions can occur:

- Transition to state M occurs if the individual that reproduces at the next cycle backward in time is the individual representing lineage 2. The weight of this event in the lottery is $w_M = \frac{d(1-c)}{N}$.

- Transition to state $C(K_i)$ occurs if the individual that reproduces at the next cycle backward in time is in $\mathcal{C}_2$ but is not the individual representing lineage 2. The weight of this event in the lottery is $w_C = (K_i - 1)\frac{d(1-c)}{N}$.

- Transitions to state $F(K_l, K_i)$ for $l \neq i$ occur if the individual that reproduces at the next cycle backward in time is in a community with carrying capacity K_l (which cannot be the community of either individuals representing lineage 1 or 2 in current state). The weight of this event in the lottery is $w_{F,l} = N_l K_l \frac{d(1-c)}{N}$.

- Transition to state $F(K_i, K_i)$ occurs if the individual that reproduces at the next cycle backward in time is either in $\mathcal{C}_1$ or in another a community with carrying capacity K_i that is not $\mathcal{C}_2$. The weight of this event in the lottery is $w_{F,i} = (K_i - 1)\left(1 - d + \frac{d(1-c)}{N}\right) + (N_i - 2)K_i\frac{d(1-c)}{N}$.

- Transition to state S occurs if a point speciation occurs at the next cycle backward in time.

Transition probabilities conditionally to the fact that lineage 1 is the one coming from the reproduction event at next cycle backward in time then writes:

$$p^1_{F(K_i,K_i)\to S} = v$$

$$p^1_{F(K_i,K_i)\to C(K_i)} = \frac{(1-v)w_C}{w_C+w_M+w_{F,i}+\sum_{l\neq i}w_{F,l}} = \frac{(1-v)(K_i-1)\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p^1_{\substack{F(K_i,K_i)\to F(K_l,K_i)\\l\neq i}} = \frac{(1-v)w_{F,l}}{w_C+w_M+w_{F,i}+\sum_{l\neq i}w_{F,l}} = \frac{(1-v)N_lK_l\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p^1_{F(K_i,K_i)\to F(K_i,K_i)} = \frac{(1-v)w_{F,i}}{w_C+w_M+w_{F,i}+\sum_{l\neq i}w_{F,l}} = (1-v)\frac{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(N_i-2)K_i\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p^1_{F(K_i,K_i)\to M} = (1-v)\frac{w_M}{w_C+w_M+w_{F,i}+\sum_{l\neq i}w_{F,l}} = \frac{(1-v)\left(1-d+\frac{d(1-c)}{N}\right)}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

Similarly, we derive the transition probabilities conditionally to the fact that lineage 2 an then the unconditional transition probabilities:

$$p_{F(K_i,K_i)\to S} = v$$

$$p_{F(K_i,K_i)\to C(K_i)} = \frac{(1-v)(K_i-1)\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p_{\substack{F(K_i,K_i)\to F(K_l,K_i)\\l\neq i}} = p_{\substack{F(K_i,K_i)\to F(K_i,K_l)\\l\neq i}} = \frac{1}{2}\frac{(1-v)N_lK_l\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p_{F(K_i,K_i)\to F(K_i,K_i)} = (1-v)\frac{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(N_i-2)K_i\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

$$p_{F(K_i,K_i)\to M} = \frac{(1-v)\frac{d(1-c)}{N}}{(K_i-1)\left(1-d+\frac{d(1-c)}{N}\right)+(J-K_i)\frac{d(1-c)}{N}}$$

Limit transition probabilities when $N \to +\infty$

We are interested in situations where the number of communities is very large (i.e. $N \to +\infty$).

We still assume that there are a distinct number of carrying capacity levels in the

metacommunity, and for level K_i, we introduce $\pi_i = \frac{N_i}{N}$. We also assume that the speciation

rate verifies: $J\nu = N\bar{K}\nu \xrightarrow[N\to+\infty]{} \theta$. The corresponding asymptotic transition probabilities are

reported in table A4.1.

Transitions from $F(.,.)$ states. Table A4.1 shows that as soon as the Markov chain reaches a

$F(.,.)$ state, it stays among $F(.,.)$ states on a large time period. We assume this period to be

long enough to ensure that that the Markov chain reaches a quasi-stationary distribution

among $F(.,.)$ states. This distribution is derived by searching the stationary distribution of the

Markov chain when neglecting rare transitions (*i.e.* with probabilities $O\left(\frac{1}{N}\right)$). We call this

stationary distribution ϕ and define ϕ_{ij} as the probability of $F(K_i, K_j)$ under ϕ. The $\phi_{ij}s$

must verify for any i, j:

$$\sum_{l\neq j} \phi_{il} \frac{p_{F(K_i,K_l)\to F(K_i,K_j)}}{1-p_{F(K_i,K_j)\to F(K_i,K_j)}} + \sum_{l\neq i} \phi_{lj} \frac{p_{F(K_l,K_j)\to F(K_i,K_j)}}{1-p_{F(K_i,K_j)\to F(K_i,K_j)}} = \phi_{ij}$$

Using analytical expressions reported in table A4.1, this rewrites:

$$\frac{\frac{\pi_j K_j}{\bar{K}-\pi_j K_j}}{\frac{\bar{K}-\pi_j K_j}{(K_j-1)(1-d)+\bar{K}d(1-c)} + \frac{\bar{K}-\pi_i K_i}{(K_i-1)(1-d)+\bar{K}d(1-c)}} \sum_{l\neq j} \frac{\phi_{il}}{(K_l-1)(1-d)+\bar{K}d(1-c)} + $$

$$\frac{\frac{\pi_i K_i}{\bar{K}-\pi_j K_j}}{\frac{\bar{K}-\pi_j K_j}{(K_j-1)(1-d)+\bar{K}d(1-c)} + \frac{\bar{K}-\pi_i K_i}{(K_i-1)(1-d)+\bar{K}d(1-c)}} \sum_{l\neq i} \frac{\phi_{lj}}{(K_l-1)(1-d)+\bar{K}d(1-c)} = \phi_{ij}$$

One can verify that the stationary probability writes

$$\phi_{ij} = C\pi_i K_i \pi_j K_j \big((K_i-1)(1-d)+\bar{K}d(1-c)\big)\big((K_j-1)(1-d)+\bar{K}d(1-c)\big)$$

where C is a normalizing constant. The expression of ϕ_{ij} implies that the two lineages have

probabilistically independent positions in the metacommunity under the stationary

distribution. Denoting κ the random variable indicating the carrying capacity of the

community occupied by a lineage in the stationary distribution ϕ_{ij} verifies:

$$\phi_{ij} = \mathbb{P}(\kappa = K_i)\mathbb{P}(\kappa = K_j)$$

with:

$$\mathbb{P}(\kappa = K_i) = \sqrt{C} \pi_i K_i \big((K_i - 1)(1 - d) + \bar{K} d(1 - c)\big)$$

We compute $\sqrt{C}$ using $\sum_i \mathbb{P}(\kappa = K_i) = 1$:

$$\sum_i \sqrt{C} \pi_i K_i \big((K_i - 1)(1 - d) + \bar{K} d(1 - c)\big) = 1$$

$$\sqrt{C} \big[(1 - d)\big(\sum_i \pi_i K_i^2\big) + \bar{K}\big(\bar{K} d(1 - c) - (1 - d)\big)\big] = 1$$

$$\sqrt{C} \big[(1 - d)(V_K + \bar{K}^2) + \bar{K}^2 d(1 - c) - \bar{K}(1 - d)\big] = 1$$

$$\sqrt{C} \big[(1 - d)\bar{K}(\gamma_2 + 1) + \bar{K} d(1 - c) - (1 - d)\big] = \frac{1}{\bar{K}}$$

$$\sqrt{C} = \frac{1}{\bar{K}[(1-d)\bar{K}(\gamma_2+1)+\bar{K}d(1-c)-(1-d)]}$$

We recall that γ_2 is the coefficient of variation in carrying capacities. We then derive a full expression of $\mathbb{P}(\kappa = K_i)$:

$$\mathbb{P}(\kappa = K_i) = \frac{\pi_i K_i\big((K_i-1)(1-d)+\bar{K}d(1-c)\big)}{\bar{K}[(1-d)\bar{K}(\gamma_2+1)+\bar{K}d(1-c)-(1-d)]}$$

We simplify the expression by introducing the effective number of migrants $I = \frac{\bar{K}d(1-c)}{1-d}$:

$$\mathbb{P}(\kappa = K_i) = \frac{\pi_i K_i}{\bar{K}} \frac{K_i - 1 + I}{\bar{K}(\gamma_2+1) - 1 + I}$$

And we obtain the expression for ϕ_{ij}:

$$\phi_{ij} = \left[\frac{\pi_i K_i}{\bar{K}} \frac{K_i - 1 + I}{\bar{K}(\gamma_2+1) - 1 + I}\right]\left[\frac{\pi_j K_j}{\bar{K}} \frac{K_j - 1 + I}{\bar{K}(\gamma_2+1) - 1 + I}\right]$$

We now consider the rare transition events that bring the Markov chain out of an F state. We derive the transition probabilities from a given $F(K_i, K_j)$ state conditionally to the fact that the Markov chain leaves an F state. This means that we only focus on the terms that were neglected in table A4.1. These probabilities are obtained by expanding these terms to order $\frac{1}{N}$ instead of order 1 (see Table A4.2). We then average the transition probabilities of Table A4.2 over the stationary distribution ϕ of possible starting states in F to obtain transition probabilities from the F state as a whole toward the other states (i.e. $C(K_i), M, S$). We note this average $\mathbb{E}_\phi$.

Transitions from $C(.)$ states. When the Markov chain is in a state $C(K_i)$ it can either leave this state toward the state M or toward a F state. When conditioning on the fact that the Markov chain leaves the state $C(K_i)$, the probability that it goes to M is:

$$\frac{p_{C(K_i)\rightarrow M}}{1-p_{C(K_i)\rightarrow C(K_i)}}=\frac{\frac{1-d}{(K_i-1)(1-d)+\bar{K}d(1-c)}}{1-\frac{(K_i-2)(1-d)}{(K_i-1)(1-d)+\bar{K}d(1-c)}}=\frac{1}{1+\frac{\bar{K}d(1-c)}{1-d}}=\frac{1}{1+I}$$

Importantly, this probability does not depend on K_i.

Simplifying the state space in the coalescence Markov chain when $N \rightarrow +\infty$

Because (i) when the coalescence Markov chain is in a $F(K_i, K_j)$ state, it reaches a stationary distribution over $F(.,.)$ states before switching to another state and (ii) when the coalescence is on a $C(K_i)$ states, the transition probabilities to other states do depend on K_i, the state space of the coalescence Markov chain can be simplified by considering only four states of interest:

- an F state in which both lineages are in distinct communities; the actual position of the lineages then follows the stationary distribution ϕ. The transition probabilities to the other states are computed by averaging transition probabilities over ϕ;

- a C state in which both lineages are in the same community; there is no need to precisely follow what is the carrying capacity of this community as it does not affect the transition probabilities out of this state;

- M and S state are defined as above.

Transition from F state to C, M and S can be computed by taking the expectations of the transition probabilities presented in Table A4.2 over a well chosen distribution of $F(K_i, K_j)$. This distribution is not directly the stationary distribution ϕ, but the stationary distribution conditionally to leaving F state, which we call Σ. Σ_{ij} is therefore the probability of being in $F(K_i, K_j)$ when leaving F state.

$$\Sigma_{ij} = \frac{\phi_{ij}\sigma_{ij}}{\mathbb{E}_{\phi}[\sigma_{kl}]}$$

where σ_{ij} is the probability of leaving F-state when in state $F(K_i, K_j)$. As we mentioned before, σ_{ij} is small, and using probabilities derived in Table A4.1, one can show that:

$$\sigma_{ij} = \frac{1}{2N\bar{K}}\left[2\theta + I\left(\frac{K_i}{K_j - 1 + I} + \frac{K_j}{K_i - 1 + I}\right)\right] + o\left(\frac{1}{N}\right)$$

Therefore:

$$\Sigma_{ij} = \left[\frac{\theta + \frac{I}{2}\left(\frac{K_i}{K_j - 1 + I} + \frac{K_j}{K_i - 1 + I}\right)}{\theta + I\mathbb{E}_\phi\left[\frac{K_k}{K_l - 1 + I}\right]}\right]\phi_{ij} + O\left(\frac{1}{N}\right)$$

Define $M_p = \sum_i \pi_i K_i^{\,p}$ the non-centered moment of carrying capacity distribution. These quantities relate to centered moments used in main text as follows: $M_1 = \bar{K}$; $M_2 = \bar{K}^2(\gamma_2 + 1)$ and $M_3 = \bar{K}^3\left[\gamma_2^{3/2}\gamma_3 + 3\gamma_2 + 1\right]$.

Going back to the calculus of Σ_{ij}:

$$\mathbb{E}_\phi\left[\frac{K_k}{K_l - 1 + I}\right] = \mathbb{E}_\phi[\boldsymbol{K_k}]\mathbb{E}_\phi\left[\frac{1}{K_l - 1 + I}\right]$$

$$\mathbb{E}_\phi[\boldsymbol{K_k}] = \sum_k \left[\frac{\pi_k K_k}{\bar{K}}\frac{K_k - 1 + I}{\bar{K}(\gamma_2 + 1) - 1 + I}\right]K_k = \frac{\sum_k\left[\pi_k K_k^2(K_k - 1 + I)\right]}{\bar{K}(\bar{K}(\gamma_2 + 1) - 1 + I)}$$

$$= \frac{\sum_k\left[\left(\pi_k K_k^3 + \pi_k K_k^2(I-1)\right)\right]}{\bar{K}(\bar{K}(\gamma_2 + 1) - 1 + I)}$$

$$= \frac{M_3 + M_2(I-1)}{M_1(M_1(\gamma_2 + 1) - 1 + I)}$$

$$= \frac{M_3 + M_2(I-1)}{M_2 + M_1(I-1)}$$

$$\mathbb{E}_\phi\left[\frac{1}{K_l - 1 + I}\right] = \sum_k \frac{\frac{\pi_k K_k}{\bar{K}}\frac{K_k - 1 + I}{\bar{K}(\gamma_2 + 1) - 1 + I}}{K_k - 1 + I} = \sum_k \frac{\pi_k K_k}{\bar{K}(\bar{K}(\gamma_2 + 1) - 1 + I)} = \frac{1}{\bar{K}(\gamma_2 + 1) - 1 + I}$$

$$= \frac{1}{\frac{M_2}{M_1} + I - 1} = \frac{M_1}{M_2 + M_1(I-1)}$$

$$\mathbb{E}_\phi\left[\frac{K_k}{K_l - 1 + I}\right] = \frac{M_1(M_3 + M_2(I-1))}{(M_2 + M_1(I-1))^2}$$

$$\Sigma_{ij} = \left[\frac{\theta + \frac{I}{2}\left(\frac{K_i}{K_j - 1 + I} + \frac{K_j}{K_i - 1 + I}\right)}{\theta + I\frac{M_1(M_3 + M_2(I-1))}{(M_2 + M_1(I-1))^2}}\right]\left[\frac{\pi_i K_i(K_i - 1 + I)}{M_2 + M_1(I-1)}\right]\left[\frac{\pi_i K_j(K_j - 1 + I)}{M_2 + M_1(I-1)}\right]$$

$$\Sigma_{ij} = \frac{[\pi_i K_i (K_i - 1 + I)][\pi_i K_j (K_j - 1 + I)]\left[\theta + \frac{I}{2}\left(\frac{K_i}{K_j - 1 + I} + \frac{K_j}{K_i - 1 + I}\right)\right]}{\theta(M_2 + M_1(I-1))^2 + IM_1(M_3 + M_2(I-1))}$$

Note that under Σ the position of the two lineages when leaving F-state are not independent anymore. The analytical expressions of the transition probabilities are provided in Table A4.3.

Computing the $D_\alpha(.)$ and $D_\beta(.,.)$

From the transition probabilities of Table A4.3, we obtain that:

1) the probability that two lineages starting in an F state undergo speciation before merging when going backward in time does not depend on the carrying capacities of the initial communities (*i.e.* at present time) and verifies:

$$\forall K_i, K_j, \qquad D_\beta(K_i, K_j) = D_\beta = \frac{\tilde{p}_{F \to S}}{1 - \tilde{p}_{C \to F}\tilde{p}_{F \to C}}$$

where $\tilde{p}$ are transition probabilities conditionally to changing state in the four state space F, C, M, S provided in table A4.3.

$$D_\beta = \frac{\mathbb{E}_\Sigma\left[\frac{2\theta}{2\theta + K_i \frac{I}{K_j - 1 + I} + K_j \frac{I}{K_i - 1 + I}}\right]}{1 - \frac{I}{1+I}\mathbb{E}_\Sigma\left[\frac{(K_j - 1)\frac{I}{K_i - 1 + I} + (K_i - 1)\frac{I}{K_j - 1 + I}}{2\theta + K_i \frac{I}{K_j - 1 + I} + K_j \frac{I}{K_i - 1 + I}}\right]}$$

$$D_\beta = \frac{\mathbb{E}_\Sigma\left[\frac{\theta}{\theta + \frac{I}{2}\left[\frac{K_i}{K_j - 1 + I} + \frac{K_j}{K_i - 1 + I}\right]}\right]}{1 - \frac{I}{1+I2}\mathbb{E}_\Sigma\left[\frac{\frac{K_j - 1}{K_i - 1 + I} + \frac{K_i - 1}{K_j - 1 + I}}{\theta + \frac{I}{2}\left[\frac{K_i}{K_j - 1 + I} + \frac{K_j}{K_i - 1 + I}\right]}\right]}$$

$$D_\beta = \frac{\Sigma_{i,j}\left[\frac{[\pi_i K_i(K_i - 1 + I)][\pi_i K_j(K_j - 1 + I)]\left[\theta + \frac{I}{2}\left(\frac{K_i}{K_j - 1 + I} + \frac{K_j}{K_i - 1 + I}\right)\right]}{\theta(M_2 + M_1(I-1))^2 + IM_1(M_3 + M_2(I-1))} \frac{\theta}{\theta + \frac{I}{2}\left[\frac{K_i}{K_j - 1 + I} + \frac{K_j}{K_i - 1 + I}\right]}\right]}{1 - \frac{I}{1+I2}\Sigma_{i,j}\left[\frac{[\pi_i K_i(K_i - 1 + I)][\pi_i K_j(K_j - 1 + I)]\left[\theta + \frac{I}{2}\left(\frac{K_i}{K_j - 1 + I} + \frac{K_j}{K_i - 1 + I}\right)\right]}{\theta(M_2 + M_1(I-1))^2 + IM_1(M_3 + M_2(I-1))} \frac{\frac{K_j - 1}{K_i - 1 + I} + \frac{K_i - 1}{K_j - 1 + I}}{\theta + \frac{I}{2}\left[\frac{K_i}{K_j - 1 + I} + \frac{K_j}{K_i - 1 + I}\right]}\right]}$$

$$D_\beta = \frac{\theta \Sigma_{i,j}\left[\frac{[\pi_i K_i(K_i - 1 + I)][\pi_i K_j(K_j - 1 + I)]}{\theta(M_2 + M_1(I-1))^2 + IM_1(M_3 + M_2(I-1))}\right]}{1 - \frac{I}{1+I2}\Sigma_{i,j}\left[\frac{[\pi_i K_i(K_i - 1 + I)][\pi_i K_j(K_j - 1 + I)]}{\theta(M_2 + M_1(I-1))^2 + IM_1(M_3 + M_2(I-1))}\left(\frac{K_j - 1}{K_i - 1 + I} + \frac{K_i - 1}{K_j - 1 + I}\right)\right]}$$

$$D_\beta = \cfrac{\cfrac{\theta(M_2+M_1(I-1))^2}{\theta(M_2+M_1(I-1))^2+IM_1(M_3+M_2(I-1))}}{1-\frac{I}{1+I}\frac{I}{2}\sum_{i,j}\left[\frac{[\pi_iK_i(K_i-1+I)][\pi_iK_j(K_j-1+I)]}{\theta(M_2+M_1(I-1))^2+IM_1(M_3+M_2(I-1))}\left(\frac{K_j-1}{K_i-1+I}+\frac{K_i-1}{K_j-1+I}\right)\right]}$$

$$D_\beta = \cfrac{\theta(M_2+M_1(I-1))^2}{\theta(M_2+M_1(I-1))^2+IM_1(M_3+M_2(I-1))-\frac{I}{1+I}\left(\frac{I}{2}\right)\sum_{i,j}\left[[\pi_iK_i(K_i-1+I)][\pi_iK_j(K_j-1+I)]\left(\frac{K_j-1}{K_i-1+I}+\frac{K_i-1}{K_j-1+I}\right)\right]}$$

Considering that:

$$\sum_{i,j}\left[[\pi_iK_i(K_i-1+I)][\pi_iK_j(K_j-1+I)]\left(\frac{K_j-1}{K_i-1+I}+\frac{K_i-1}{K_j-1+I}\right)\right]$$

$$= 2\sum_{i,j}\left[[\pi_iK_i(K_i-1+I)][\pi_iK_j(K_j-1+I)]\frac{K_j-1}{K_i-1+I}\right]$$

$$= 2\left[\sum_i\left[\frac{\pi_iK_i(K_i-1+I)}{K_i-1+I}\right]\right]\left[\sum_j[\pi_iK_j(K_j-1+I)(K_j-1)]\right]$$

$$= 2M_1[M_3+M_2(I-2)-M_1(I-1)]$$

$$= 2M_1[[M_3+M_2(I-1)]-[M_2+M_1(I-1)]]$$

One obtains a rather simplified version of D_β:

$$D_\beta = \cfrac{\theta(M_2+M_1(I-1))^2}{\theta(M_2+M_1(I-1))^2+IM_1(M_3+M_2(I-1))-\frac{I}{1+I}IM_1[[M_3+M_2(I-1)]-[M_2+M_1(I-1)]]}$$

$$D_\beta = \cfrac{1}{1+\frac{IM_1}{\theta}\left[\cfrac{\left(1-\frac{I}{1+I}\right)(M_3+M_2(I-1))+\frac{I}{1+I}(M_2+M_1(I-1))}{(M_2+M_1(I-1))^2}\right]} \qquad (4.2)$$

2) the probability that two lineages starting in C state undergo speciation before merging when going backward in time does not depend on the carrying capacities of the initial community (*i.e.* at present time) and verifies:

$$\forall K_i, \qquad D_\alpha(K_i) = D_\alpha = \frac{\tilde{p}_{C\to F}\tilde{p}_{F\to S}}{1-\tilde{p}_{C\to F}\tilde{p}_{F\to C}}$$

$$D_\alpha = \frac{I}{1+I}D_\beta \qquad (4.3)$$

Note that D_β and D_α still depend on parameters d, c through I. The impact of carrying capacity distribution can be summarized in its three lowest moments M_1, M_2 and M_3.

We report in Figure A4.2 the values of $D_\alpha s$ and $D_\beta s$ for metacommunities with parameters identical to the examples of main text except for asymmetry that we vary.

Implications for D_γ

When $N \rightarrow +\infty$, equation (2.1) becomes:

$$D_\gamma = \left(\sum_{i<j} 2 \left(\frac{\pi_i K_i}{\overline{K}} \right) \left(\frac{\pi_j K_j}{\overline{K}} \right) D_\beta \right) + \left(\sum_i \left(\frac{\pi_i K_i}{\overline{K}} \right)^2 D_\beta \right) = D_\beta$$

Tables

Table A4.1: Summary of transition probabilities of the coalescence Markov chain when $N \to +\infty$ and $N\bar{K}\nu$ is finite. O follows Landau notation.

Starting state	Ending state	Analytical expression
$C(K_i)$	S	$O\left(\frac{1}{N}\right)$
	$C(K_i)$	$\frac{(K_i-2)(1-d)}{(K_i-1)(1-d)+\bar{K}d(1-c)} + O\left(\frac{1}{N}\right)$
	$F(K_i,K_j)$ or $F(K_j,K_i)$	$\frac{1}{2}\frac{\pi_j K_j d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)} + O\left(\frac{1}{N}\right)$
	$F(K_i,K_i)$	$\frac{\pi_i K_i d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)} + O\left(\frac{1}{N}\right)$
	M	$\frac{1-d}{(K_i-1)(1-d)+\bar{K}d(1-c)} + O\left(\frac{1}{N}\right)$
$F(K_i,K_j)$	S	$O\left(\frac{1}{N}\right)$
	$C(K_i)$	$O\left(\frac{1}{N}\right)$
	$C(K_j)$	$O\left(\frac{1}{N}\right)$
	$F(K_i,K_l)$	$\frac{1}{2}\frac{\pi_l K_l d(1-c)}{(K_j-1)(1-d)+\bar{K}d(1-c)} + O\left(\frac{1}{N}\right)$
	$F(K_l,K_j)$	$\frac{1}{2}\frac{\pi_l K_l d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)} + O\left(\frac{1}{N}\right)$
	$F(K_i,K_j)$	$\frac{1}{2}\left[\frac{(K_j-1)(1-d)+\pi_j K_j d(1-c)}{(K_j-1)(1-d)+\bar{K}d(1-c)} + \frac{(K_i-1)(1-d)+\pi_i K_i d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)}\right] + O\left(\frac{1}{N}\right)$
	$F(K_i,K_i)$	$\frac{1}{2}\frac{\pi_i K_i d(1-c)}{(K_j-1)(1-d)+\bar{K}d(1-c)} + O\left(\frac{1}{N}\right)$
	$F(K_j,K_j)$	$\frac{1}{2}\frac{\pi_j K_j d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)} + O\left(\frac{1}{N}\right)$

	M	$O\left(\frac{1}{N}\right)$
$F(K_i, K_i)$	S	$O\left(\frac{1}{N}\right)$
	$C(K_i)$	$O\left(\frac{1}{N}\right)$
	$F(K_l, K_i)$ or $F(K_i, K_l)$	$\frac{1}{2}\frac{\pi_l K_l d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)} + O\left(\frac{1}{N}\right)$
	$F(K_i, K_i)$	$\frac{(K_i-1)(1-d)+\pi_i K_i d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)} + O\left(\frac{1}{N}\right)$
	M	$O\left(\frac{1}{N}\right)$

Table A4.2: Analytical expression of transition probabilities out of a given $F(K_i, K_j)$ state when conditioning on leaving the $F(.,.)$ states in asymptotically large metacommunity. o follows Landau notation.

Starting state	Ending state	Analytical expression
$F(K_i, K_j)$	S	$\dfrac{2\theta}{2\theta + K_i\frac{\bar{K}d(1-c)}{(K_j-1)(1-d)+\bar{K}d(1-c)} + K_j\frac{\bar{K}d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)}} + o\left(\frac{1}{N}\right)$
	$C(K_i)$	$\dfrac{(K_i-1)\frac{\bar{K}d(1-c)}{(K_j-1)(1-d)+\bar{K}d(1-c)}}{2\theta + K_i\frac{\bar{K}d(1-c)}{(K_j-1)(1-d)+\bar{K}d(1-c)} + K_j\frac{\bar{K}d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)}} + o\left(\frac{1}{N}\right)$
	$C(K_j)$	$\dfrac{(K_j-1)\frac{\bar{K}d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)}}{2\theta + K_i\frac{\bar{K}d(1-c)}{(K_j-1)(1-d)+\bar{K}d(1-c)} + K_j\frac{\bar{K}d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)}} + o\left(\frac{1}{N}\right)$
	M	$\dfrac{\frac{\bar{K}d(1-c)}{(K_j-1)(1-d)+\bar{K}d(1-c)} + \frac{\bar{K}d(1-c)}{(K_j-1)(1-d)+\bar{K}d(1-c)}}{2\theta + K_i\frac{\bar{K}d(1-c)}{(K_j-1)(1-d)+\bar{K}d(1-c)} + K_j\frac{\bar{K}d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)}} + o\left(\frac{1}{N}\right)$
$F(K_i, K_i)$	S	$\dfrac{\theta}{\theta + K_i\frac{\bar{K}d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)}} + o\left(\frac{1}{N}\right)$
	$C(K_i)$	$\dfrac{(K_i-1)\frac{\bar{K}d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)}}{\theta + K_i\frac{\bar{K}d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)}} + o\left(\frac{1}{N}\right)$
	M	$\dfrac{\frac{\bar{K}d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)}}{\theta + K_i\frac{\bar{K}d(1-c)}{(K_i-1)(1-d)+\bar{K}d(1-c)}} + o\left(\frac{1}{N}\right)$

Table A4.3 Transition probabilities of the asymptotic coalescence Markov chain in the simplified state space.

| Starting state	Ending state	Analytical expression
F	S	$\mathbb{E}_\Sigma\left[\dfrac{2\theta}{2\theta+K_i\frac{I}{K_j-1+I}+K_j\frac{I}{K_i-1+I}}\right]$
	C	$\mathbb{E}_\Sigma\left[\dfrac{(K_j-1)\frac{I}{K_i-1+I}+(K_i-1)\frac{I}{K_j-1+I}}{2\theta+K_i\frac{I}{K_j-1+I}+K_j\frac{I}{K_i-1+I}}\right]$
	M	$\mathbb{E}_\Sigma\left[\dfrac{\frac{I}{K_j-1+I}+\frac{I}{K_j-1+I}}{2\theta+K_i\frac{I}{K_j-1+I}+K_j\frac{I}{K_i-1+I}}\right]$
C	S	0
	F	$\dfrac{I}{1+I}$
	M	$\dfrac{1}{1+I}$

Figures

Figure A4.1: Birth-death cycles and backward dynamics of a lineage within a single community. The same community (grey ellipse) is represented at five consecutive birth-death cycles (labelled from 1 to 5 forward in time). The bottom picture of the community represents the resulting community at present time. Individuals are represented as filled blue and red circles. Crosses represent death events. Black thin arrows represent replacement of the dead individual by a neighbour. In this simple example, only one community is considered and only neighbours can replace dead individuals (no immigration). A death-birth cycle is thus depicted as the combination of a cross and the associated thin black arrow. A lineage is traced back in time from the present. Red dots represent an arbitrarily chosen individual at present time and its ancestors. Red arrows represent the dynamics backward in time of the tracked lineage. We did not represent speciation on this cartoon, although it can happen at any time step.

1

2

3

4

5

Present

Time
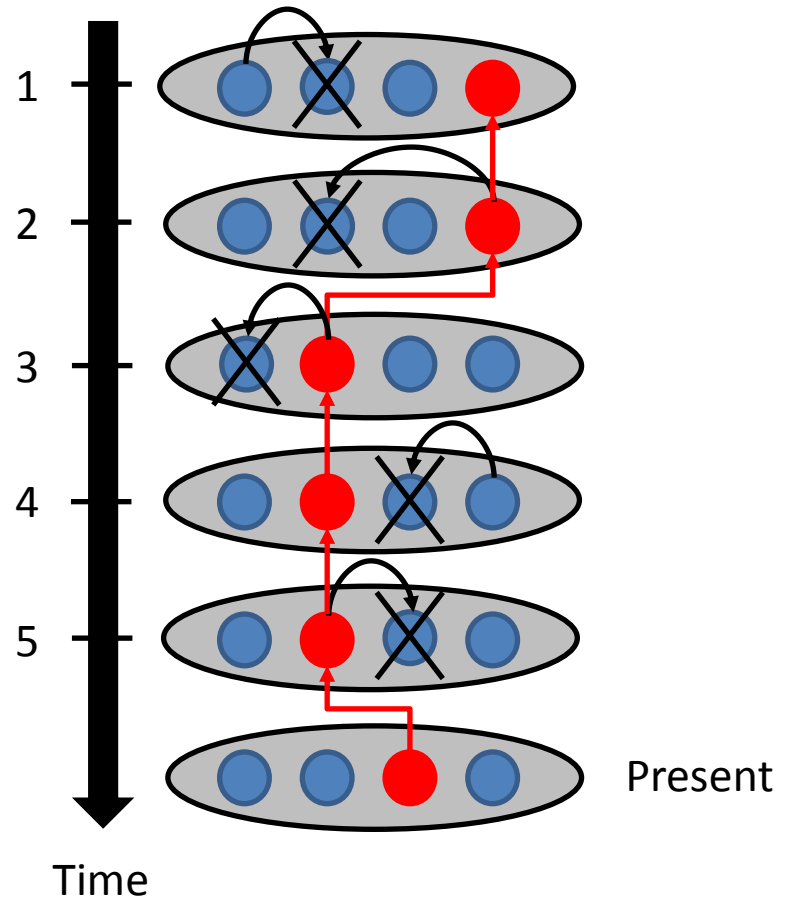
Figure A4.2: Comparing neutral predictions of D_α and D_β with simulations of model M$_0$. (A): neutral predictions obtained from equation (4.3) about D_α in small (left bars) and large (right bars) communities, when applied to metacommunity 1 (dark grey bars) and 2 (light grey bars). Black dots represent the estimated $\widehat{D}_\alpha$s (see Appendix 5) averaged over 100 simulations of model M$_0$ at stationary state. (B): neutral predictions obtained from equation (4.2) about D_β among small (left), between small and large (middle) and among large (right) communities, when applied to metacommunity 1 (dark grey bars) and 2 (light grey bars). Black dots represent the estimated $\widehat{D}_\beta$s (see Appendix 5) in small and large communities averaged over 100 simulations of model M$_0$ at stationary state. In both panels, error bars show corresponding 95% confidence intervals.
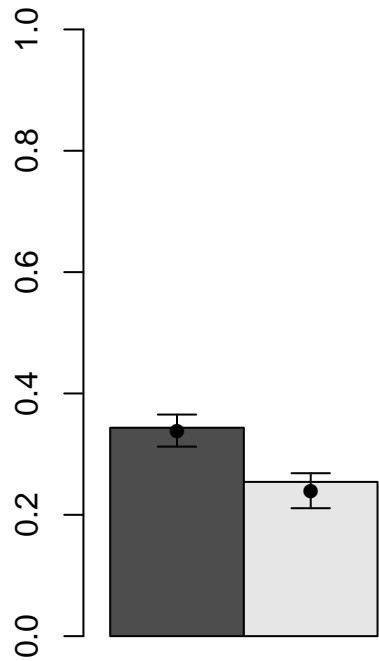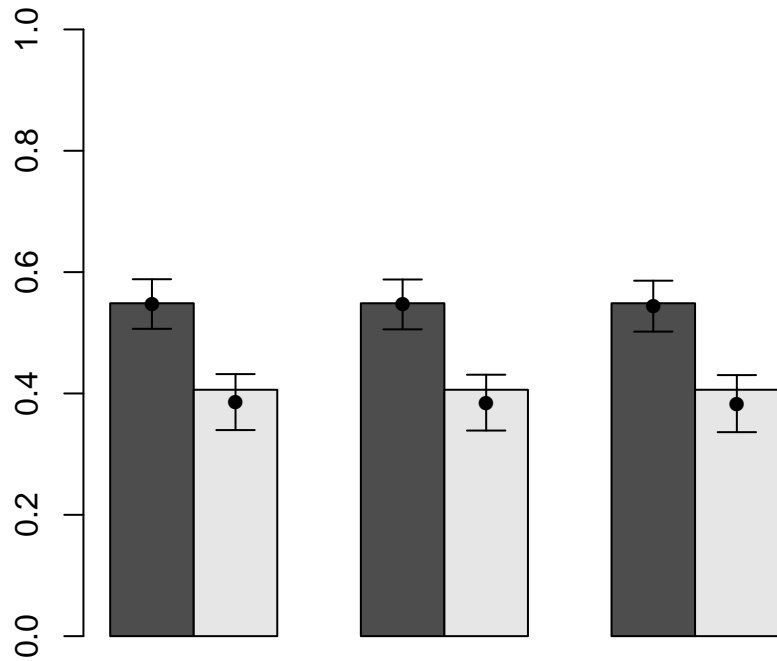
Figure A4.3 Predicted neutral diversity patterns as a function of the variance and asymmetry of carrying capacity distribution. Using $\bar{K} = 8$ and $c = 0.1$ (values of main text), we computed the neutral predictions for D_α (panel A) and D_β (panel B) for several values of the coefficient of variation γ_2 and standardized asymmetry γ_3 of the carrying capacity distributions. Values corresponding to metacommunities 1 and 2 are indicated with dots.
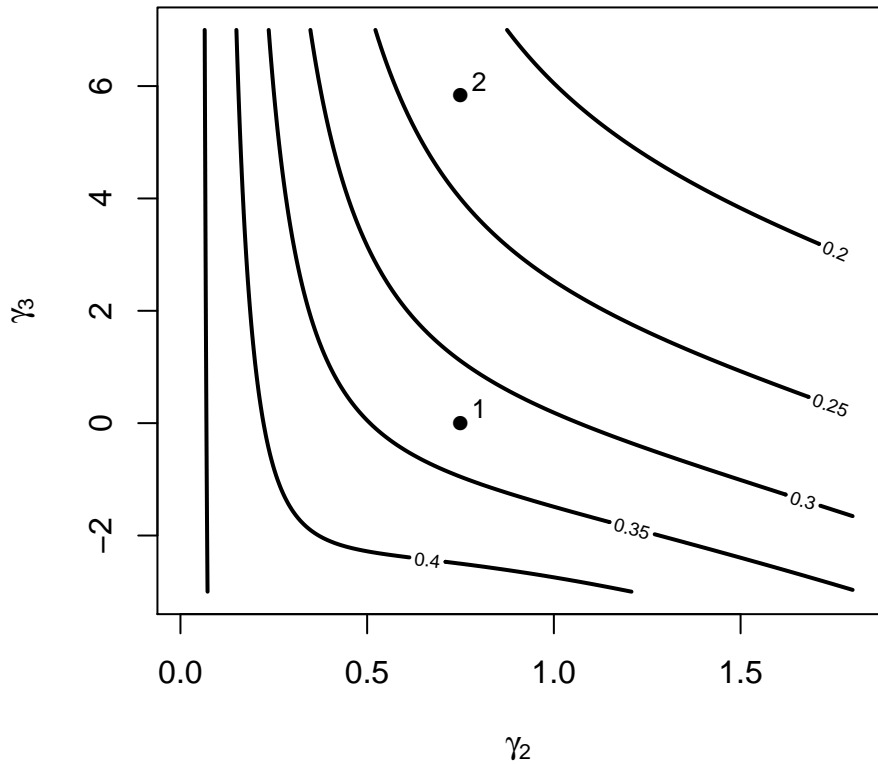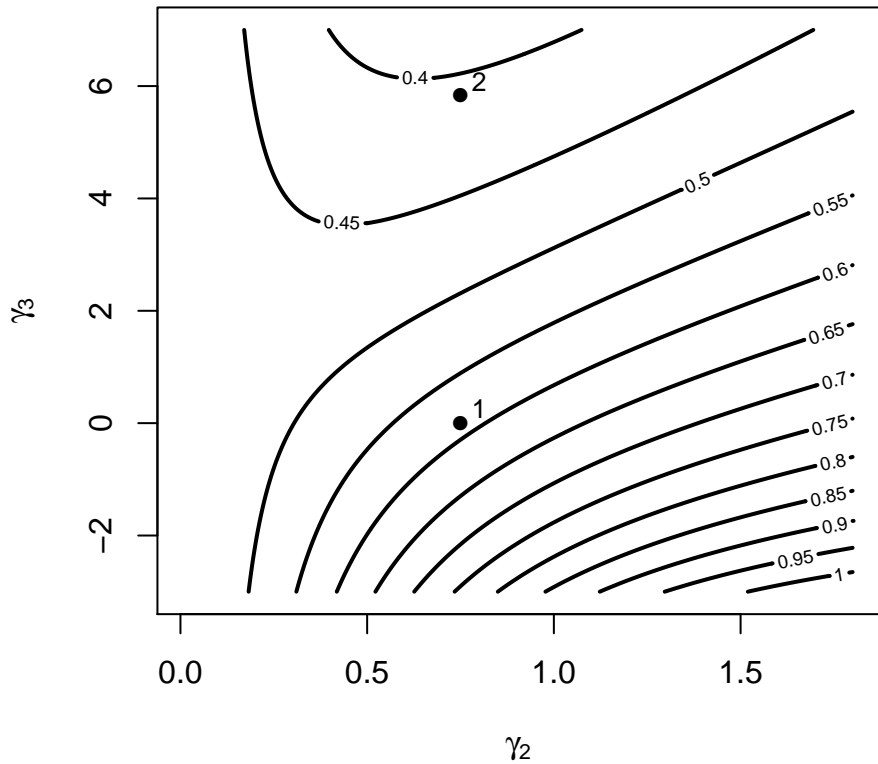
Appendix 5: $D_\alpha(.)$, $D_\beta(.,.)$ and D_γ in simulated metacommunities – estimates and robustness to dispersal cost

The aim of this appendix is to provide methods for estimating $D_\alpha(.)$, $D_\beta(.,.)$ and D_γ in a simulated metacommunity (neutral or not).

Building estimates

We consider a community A with carrying capacity K_i. We label individuals from 1 to K_i in A. $D_\alpha(K_i)$ can be estimated using:

$$\delta_\alpha(A) = 1 - \frac{\sum_{j=1}^{K_i} \sum_{\substack{k=1 \\ k>j}}^{K_i} \mathbb{I}(E_j = E_k)}{\frac{K_i(K_i-1)}{2}}$$

with

- $\mathbb{I}(X) = 1$ when X is true, 0 otherwise

- E_j is a random variable designating the species of the jth individual in community A.

$\delta_\alpha(A)$ is clearly unbiased and can be quickly calculated by noting that:

$$\delta_\alpha(A) = 1 - \frac{\sum_{l=1}^{S} N_{lA}(N_{lA} - 1)}{K_i(K_i - 1)}$$

where:

- S is the total number of species in community A

- N_{lA} is the number of individuals belonging to species l in community A.

An unbiased estimate of $D_\alpha(K_i)$ with lower variance can then be obtained by averaging the $\delta_\alpha(A)$ over all communities A with carrying capacity K_i. We use this estimate, called $\widehat{D_\alpha}(K_i)$, in each of our simulations.

We now consider two communities A and B with carrying capacities K_i and K_j. We label individuals from 1 to K_i in community A and from 1 to K_j in community B. $D_\beta(K_i, K_j)$ can be estimated using:

$$\delta_\beta(A, B) = 1 - \frac{\sum_{k=1}^{K_i} \sum_{l=1}^{K_j} \mathbb{I}(E_k = E_l)}{K_i K_j}$$

$\delta_\beta(A, B)$ is clearly unbiased and can be quickly calculated by noting that:

$$\delta_\beta(A, B) = 1 - \frac{\sum_{l=1}^{S} N_{lA} N_{lB}}{K_i K_j}$$

where:

- S is the total number of species when pooling communities A (carrying capacity K_i) and B (carrying capacity K_j) together

- N_{lA} (resp. N_{lB}) is the number of individuals belonging to species l in community A (resp. B).

An unbiased estimate of $D_\beta(K_i, K_j)$ with lower variance can then be obtained by averaging $\delta_\beta(A, B)$ over all the pairs A and B with carrying capacities K_i and K_j respectively. We use this estimate, noted $\widehat{D_\beta}(K_i, K_j)$, in each of our simulations.

For each example of metacommunity (symmetric or asymmetric) considered in main text, we ran 100 independent simulations. For each run, we computed $\widehat{D_\alpha}(K_1)$. From the 100 independent values of $\widehat{D_\alpha}(K_1)$, we computed $\widehat{\mu_\alpha(K_1)}$ and $\widehat{\sigma_\alpha(K_1)}$ the empirical mean and the empirical variance of $\widehat{D_\alpha}(K_1)$.

We use the central limit theorem and approximate the true variance of $\widehat{D_\alpha}(K_1)$ by its empirical estimate $\widehat{\sigma_\alpha(K_1)}$ to get a confidence interval of $D_\alpha(K_1)$:

$$D_\alpha(K_1) \in \left[\widehat{\mu_\alpha(K_1)} \pm 1.96 \times \frac{\widehat{\sigma_\alpha(K_1)}}{\sqrt{100}}\right]$$

and $\widehat{\mu_\alpha(K_1)}$ is an unbiased estimator of $D_\alpha(K_1)$.

We obtain confidence intervals and estimates for the other indices using the same approach.

We use equation (4.1) to obtain an estimate of D_γ:

$$\widehat{D}_\gamma = \left(\sum_i \frac{N_i K_i (K_i - 1)}{J(J-1)} \widehat{D}_\alpha(K_i) \right) + \left(\sum_{i<j} 2 \frac{N_i N_j K_i K_j}{J(J-1)} \widehat{D}_\beta(K_i, K_j) \right) + \left(\sum_i \frac{N_i (N_i - 1) K_i^2}{J(J-1)} \widehat{D}_\beta(K_i, K_i) \right)$$

Confidence intervals are obtained using the same method than for $\widehat{D}_\alpha(K_i)$ and $\widehat{D}_\beta(K_i, K_j)$.

Robustness of patterns to dispersal cost

We compared estimates of $\widehat{D}_\alpha(K_i)$ and $\widehat{D}_\beta(K_i, K_j)$ to neutral predictions in simulations presented in Figure A2.1. Contrary to Figure 4 of main text, we did not perform 100 replicates. We simply look instead at the temporal trajectory of patterns in time on a single simulation. Results are in line with those shown in main text. In the symmetric metacommunity (Figure A5.1A and A5.2A), the $\widehat{D}_\alpha(K_i)$ took similar average (and also instantaneous) values for K_1 and K_2, and so did the $\widehat{D}_\beta(K_i, K_j)$ (Figure A5.1B and A5.2B). Temporal variance of estimates precludes clear conclusion about diversity being lower than neutral expectation though. In the asymmetric metacommunity, we retrieved that $\widehat{D}_\beta(K_1, K_1)$ fluctuates around the neutral prediction and is consistently lower than $\widehat{D}_\beta(K_1, K_2)$ and $\widehat{D}_\beta(K_2, K_2)$ which are both consistently higher than neutral prediction. $\widehat{D}_\beta(K_2, K_2)$ is consistently smaller than $\widehat{D}_\beta(K_1, K_2)$

Figures

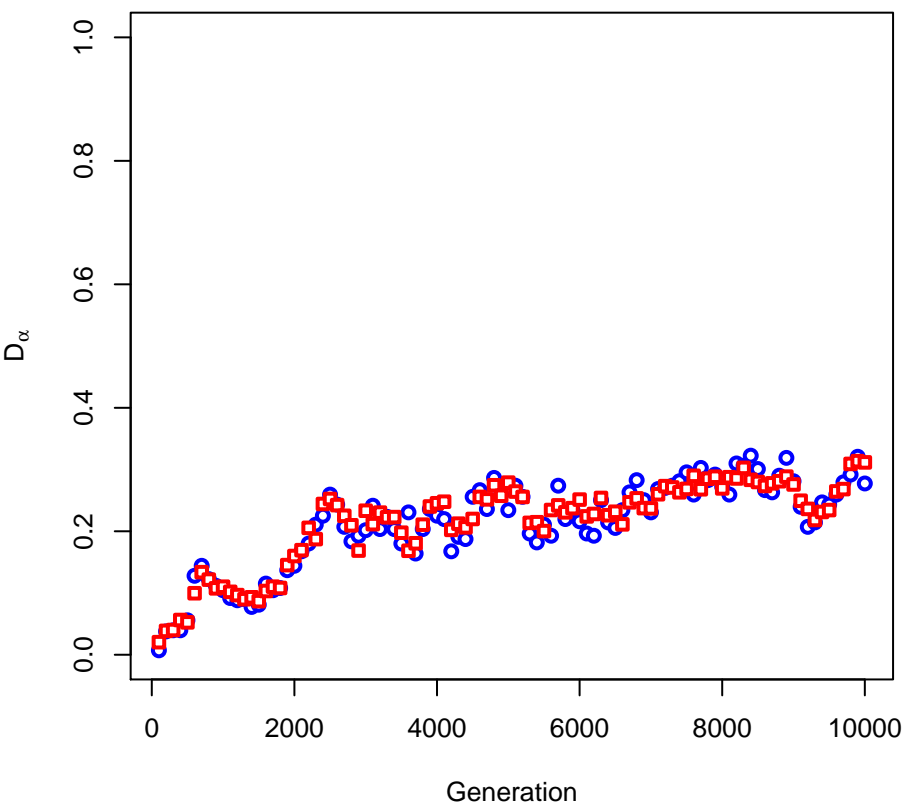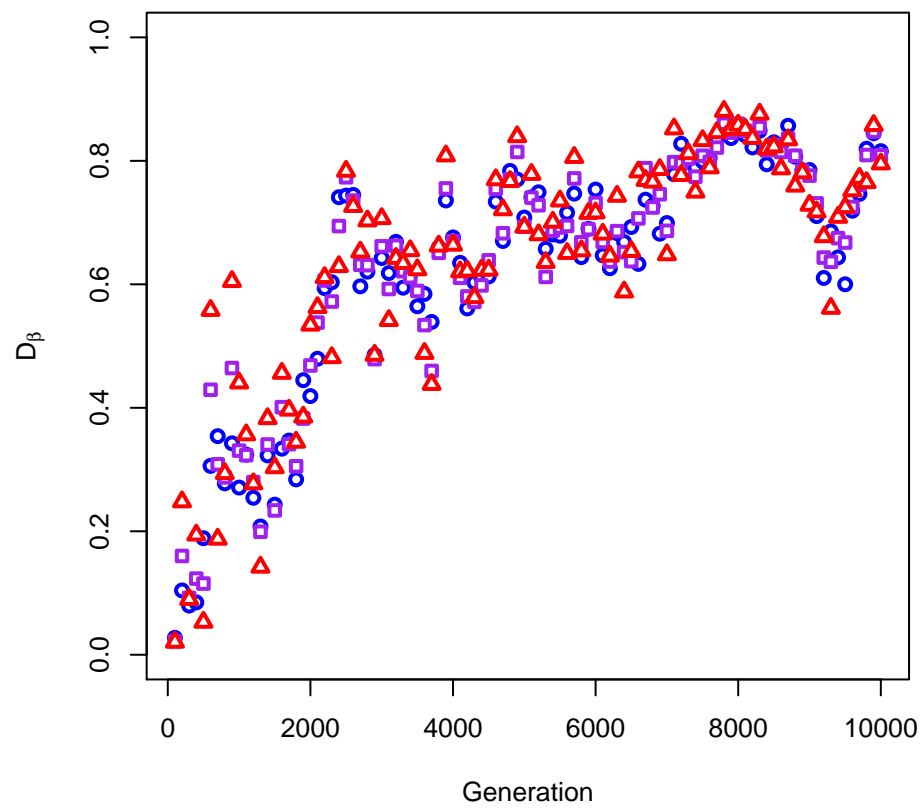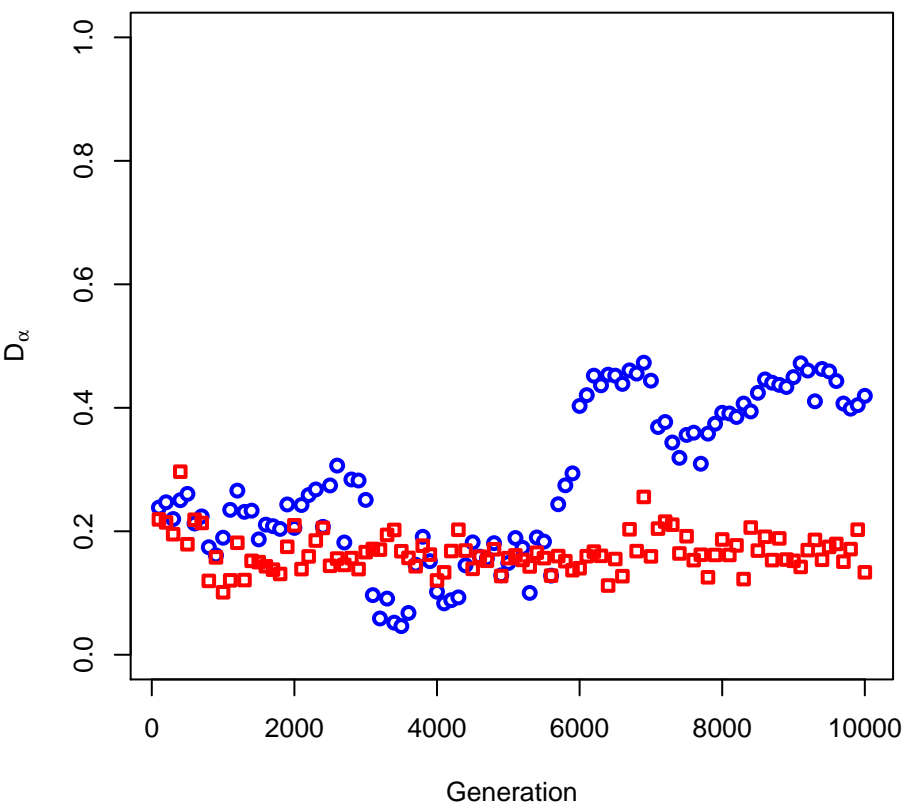**Figure A5.1: Temporal trajectory of species diversity patterns in simulations of Figure A2.1.** (A): every 200th generation, we computed $\widehat{D}_\alpha(K_1)$ (blue circles) and $\widehat{D}_\alpha(K_2)$ (red squares) in the symmetric metacommunity. (B): every 200th generation, we computed $\widehat{D}_\beta(K_1, K_1)$ (blue circles) and $\widehat{D}_\beta(K_1, K_2)$ (purple squares) and $\widehat{D}_\beta(K_2, K_2)$ (red triangles) in the symmetric metacommunity. (C): same as (A) with asymmetric metacommunity. (D): same as (B) with asymmetric metacommunity. $c = 0$ in all panels.

A

D_α vs Generation

B

D_β vs Generation

C

D_α vs Generation

D

D_β vs Generation
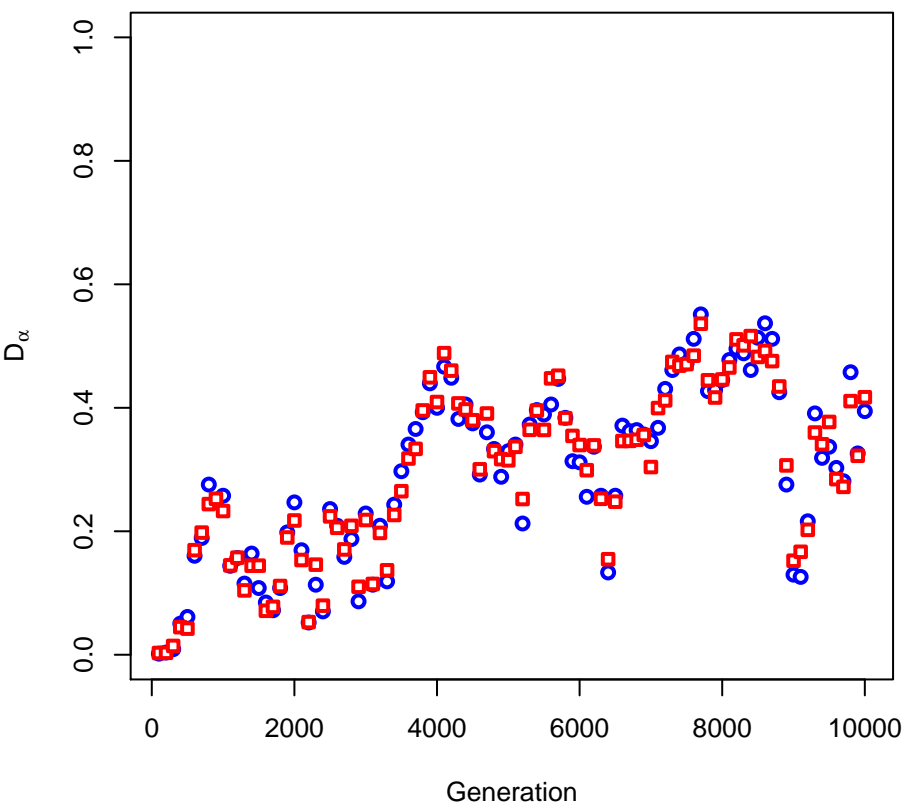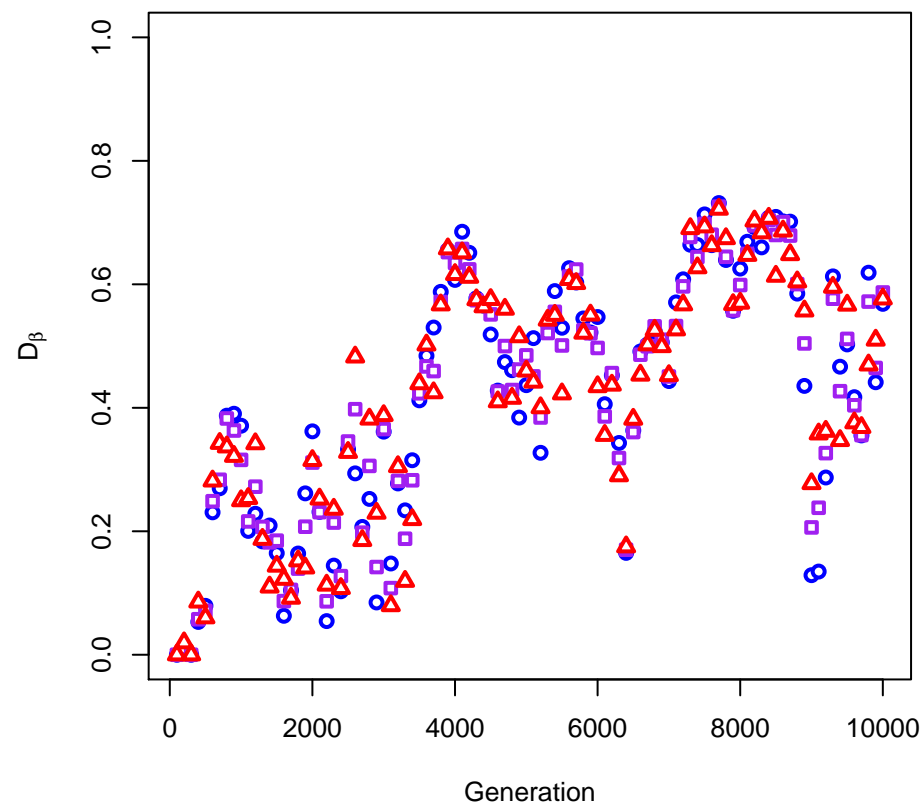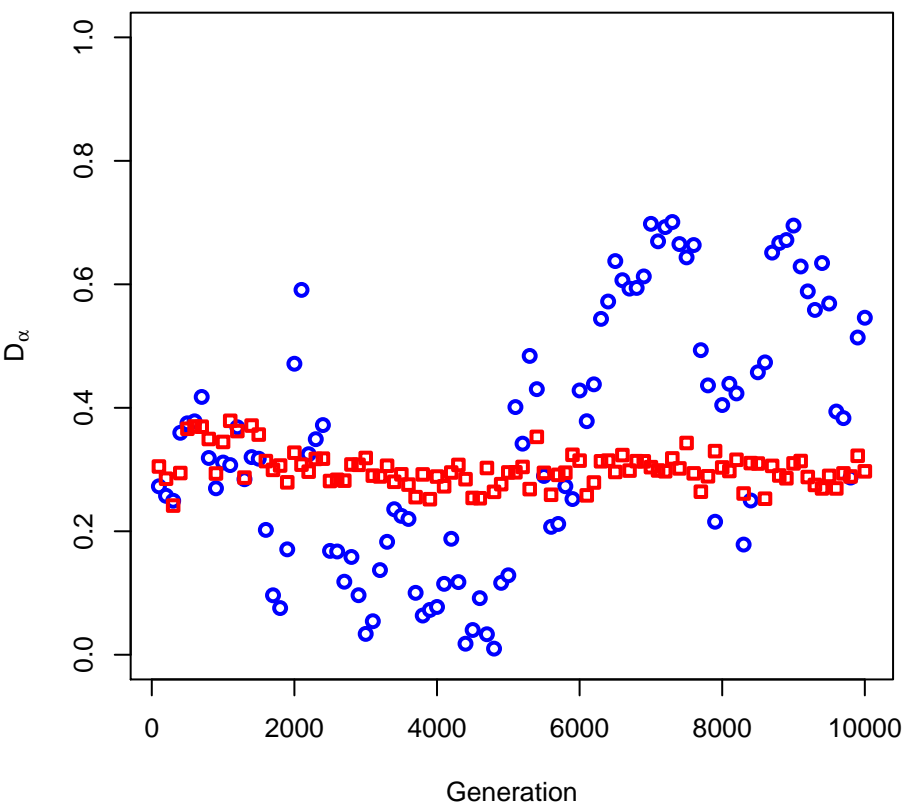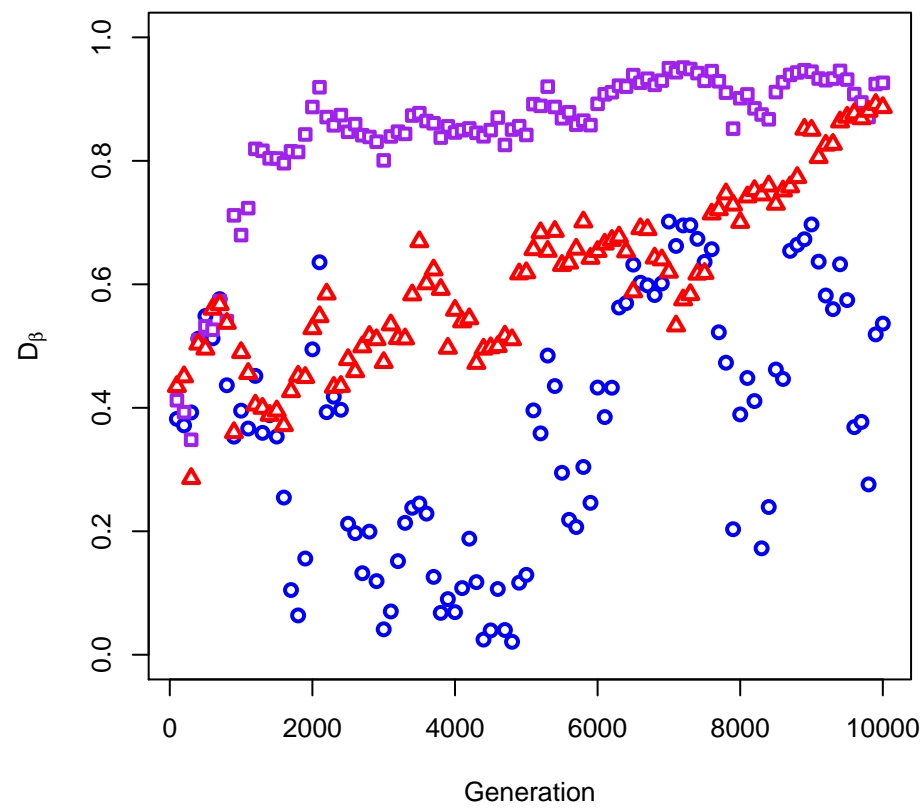
Figure A5.2: Temporal trajectory of species diversity patterns in simulations of Figure A2.1. Legend as in Figure A5.1. $c = 0.5$ in all panels.