Appendix 1: simulation study details

Model overview - In a first step, we generated independent species pools by simulating phylogenies and trait evolution along these phylogenies. As a result we obtained for each pool a set of 400 species. Each species, *i,* was characterised by a single trait that defined the species-specific niche optimum (opt_i, ranging from 0 to 100 in every species pool, assuming no intraspecific variability) and niche breadth (NBd) that was equal for all species (NBd=10, Table S1). In a second step, species' individuals from the species pool could enter a local community according to the chosen assembly rules. Each community, *k*, was characterised by a carrying capacity (C=100) and environmental conditions that were optimal for one specific trait value (E_k= 50). All communities were independent from each other as they all come from different species pool.

Generating the species pool - The phylogenetic relationships among species were simulated using a Yule process with a branching rate of 0.1. Species' traits evolved along the phylogenetic trees following a Brownian motion process with varying rates of trait evolution (δ) over evolutionary time. This is also called a δ–model (when $\delta = 1$ the model describes a constant rate of trait evolution, $\delta < 1$ the model describes a decreasing rate over time and inversely, $\delta > 1$ for an increasing rate overtime, Pagel 1997). The phylogenetic signal for these traits, i.e. the trend for closely related species to be more similar than distantly related species, is expected to increase when the rate of trait evolution decreases over evolutionary time (i.e. $\delta < 1$). It was measured using Blomberg's K index, where K can be superior or equal to 0: K=0 means that the trait is independent from the phylogeny, K=1 suggests a phylogenetic signal expected under the Brownian Motion Model (BM) with constant evolutionary rates over time, and K>1 suggests that the trait values between closely related species are higher than expected under BM (Blomberg et al. 2003, Münkemüller et al. 2012).

Assembling the communities - Each community was initialised with C individuals randomly drawn

from the species pool. Then, for each simulation step, C random individuals were sequentially removed from the communities and replaced by individuals from the species pool (asynchronous updating). The probability of an individual from species *i* entering the community *k*: $P_{all,i,k}$, depended on the specified assembly rules and their relative importance defined by the factors B_{env}, B_{comp} and B_{abun}.

$$P_{all,i,k} = \exp\left(B_{env} \times \log\left(P_{env,i,k}\right) + B_{comp} \times \log\left(P_{comp,i,k}\right) + B_{abun} \times \log\left(P_{abun,i,k}\right)\right) \quad \text{(Equation 1)}$$

The habitat filter ($P_{env,i,k}$) described the normalised probability of an individual *i* entering the community *k* given the local environment E_k. It was defined by a normal law centred on the niche optimum of the species (opt_i) and of standard deviation equal to the niche breadth of the species (NBd). The closer the species trait value (i.e. niche optimum of the species) was to the environmental conditions of the community *k*, the higher was its probability to enter (Figure S1).

$$P_{env,i,k} = f\left(E_k; opt_i, NBd\right) / f\left(E_k, E_k, NBd\right) \quad \text{(Equation 2)}$$

With $f(x; \mu, \sigma)$ being the density function of a Normal law $N(\mu, \sigma)$.

The competition filter ($P_{comp,i,k}$) was the normalised probability of an individual entering the community given competition. This probability was proportional to the sum of niche overlaps between the individual and all other individuals present in the community (Figure S1). The closer the species trait value was to those of the individuals already present, the lower was its probability of entering. In this way, competition between species was defined as symmetric. Note that as conspecifics individuals had the same trait value, intraspecific competition was stronger than interspecific competition.

$$P_{comp,i,k} = 1 - \frac{1}{K} \sum_j a_{ij} N_{jk} \,;\; a_{ij} = F\left(opt_i - opt_j; 0, NBd\right) \quad \text{(Equation 3)}$$

With $F(x; \mu, \sigma)$ being the cumulative distribution function of a Normal law $N(\mu, \sigma)$.

The recruitment filter ($P_{abun,i,k}$) described the probability of an individual entering the

community through the reproduction of conspecifics already present. The more abundant the species was in the community, the higher its probability of entering. This term counteracted the high intraspecific competition value generated by the competition filter.

$$P_{abun,i,k} = N_{i,k} / K \quad \text{(Equation 4)}$$

The factors B_{abun}, B_{env} and B_{comp} weighted the importance of the three filters in community assembly. In the special case of B_{env} and B_{comp} equaling one, the equation was comparable to a Lotka-Volterra equation with inter and intra-specific competition and a maximal growth rate dependent on environmental suitability.

We repeated each combination of the parameters B_{env}, B_{comp} and δ (Table S1) 100 times, leading to a total of 10,000 simulated communities with different assembly rules and different phylogenetic contexts.

References

Blomberg, S.P., Garland Jr, T., & Ives, A.R. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57: 717–745.

Münkemüller, T., Lavergne, S., Bzeznik, B., Dray, S., Jombart, T., Schiffers, K. and Thuiller, W. 2012. How to measure and test phylogenetic signal? *Methods in Ecology and Evolution*: in press.

Pagel, M. 1997. Inferring evolutionary processes from phylogenies. *Zoologica Scripta* 26: 331–348.