Competition and immigration lead to clusters of similar species, not trait separation

Rafael D'Andrea,[1,*] Maria Riolo[2] & Annette Ostling[1]
[1]Dept. Ecology & Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA
[2]Center for the Study of Complex Systems, University of Michigan, Ann Arbor, MI, USA
[*]rdandrea@illinois.edu, ORCID 0000-0001-9687-6178

Abstract

Patterns of trait distribution among coexisting species can potentially reveal the processes by which species assemble into communities. The dominant paradigm, that competition causes species to differ more than expected by chance, has limited empirical support. Here we show that when competition acts in concert with stochasticity and dispersal, communities spontaneously organize into clusters of similar species. Further, we show clusters feature generally across different niche mechanisms, and even under the confounding influence of environmental filters. While clusters have been previously reported as transient or else persisting under restricted circumstances, our results show they persist broadly under immigration. Previous stochastic niche studies missed this effect because they only explored extremely low immigration. We provide parameter-free metrics for detecting clusters in field data, which we validate using simulations. We conclude that clusters are a more general pattern than overdispersion, and trait-based searches for niche differentiation may be more successful once they account for this fact.

Keywords: Trait-based patterns | Species clustering | Self-organized similarity | Stochastic niche assembly | Competitive coexistence | Community assembly

Statement of authorship: RD and AO conceptualized the study, RD performed modeling work and analyzed output data. RD wrote the first draft of the manuscript, and RD and AO contributed substantially to revisions. MR developed the Ripley's K-based metric.

Introduction

Patterns in the characteristics of co-occurring species can reflect the processes by which species assemble into communities[1–5]. In particular, species traits are generalizable indicators of ecological strategy[6,7], and hence trait patterns may reveal coexistence mechanisms. A key coexistence mechanism is niche differentiation, whereby species minimize competition through differences in ecological strategies. Ecologists traditionally expect niche-differentiated species to exhibit trait overdispersion—coexisting species should differ more from one another than expected by chance[8–16]. Although intuitive, this hypothesis has limited empirical support[2,3,5]. One possible reason for this is that the overdispersion hypothesis ignores the role of chance and dispersal, which are increasingly recognized as important players in shaping communities[17–20]. So-called stochastic niche studies have attempted to resolve the action of niche differentiation in concert with these forces, and yet found little difference in trait patterns relative to classical niche studies[21–23]. However, these studies explored only very low immigration regimes.

Here we show that stochastic niche communities broadly exhibit a pattern that is in many ways opposite to classical expectations. Rather than maximizing trait differences, species self-organize into multiple clusters of similar species. The clusters occur because the species most similar to the ultimate survivors are paradoxically the last to go extinct[24,25]. Immigration maintains many species that would otherwise be excluded—thus prolonging the clusters indefinitely. Previous studies have shown that clusters arise in transients, and persist under special circumstances such as an array of specialized enemies[26]. Here we show clusters should be expected broadly, because immigration alone causes them to persist, and because they occur across niche mechanisms, and remain distinctive even under the confounding influence of environmental filters.

We also provide two metrics to detect clustering in field data. Existing reports of clustering in nature rely on tools that require arbitrary binning of traits or additional data on species persistence[26–30]. Our metrics have no free parameters and require only trait and abundance data. We show they successfully distinguish simulated communities under niche differentiation from other communities.

Methods

Models and simulation design

In Appendix A we provide a detailed description of our simulation design and niche models, which we now summarize. Schematic illustrations of our models are shown in Fig. 1.

We first consider a heuristic stochastic niche model where trait similarity determines the degree of competition. This has been observed in nature[31–33] and features in influential niche models[34]. Using this model we consider pattern under a range of scenarios of immigration, regional diversity, environmental filtering, the speed of competitive sorting, and departures from the exclusive association between competition and trait similarity. Next, we consider stochastic models with more specific mechanisms of niche differentiation, namely differences in preferred resource, differences in preferred environmental conditions, and differences in life history strategies. To facilitate comparison across scenarios, we tune parameters in all models so that between 12 and 15 species would stably coexist under deterministic niche dynamics.

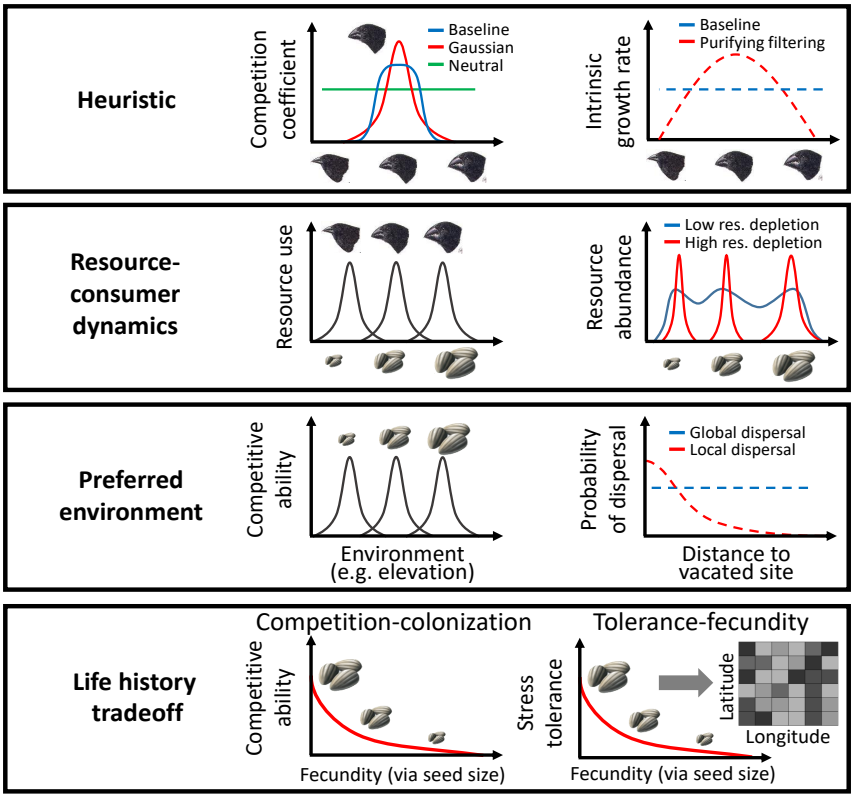


Figure 1: Schematic representation of niche models and scenarios. Heuristic: strength of competition depends upon trait similarity, for example beak size (left). Curves show competitive impact on the focal species on top caused by each species in the community. If the environment selects for a particular trait value (environmental filtering), some species may have a higher intrinsic growth rate than others (right). In this example, the environment selects for birds of intermediate beak size. In our baseline case, all species have identical intrinsic growth rates. We also looked at a case where filtering is unrelated to the trait at hand (not shown in the diagram); **Resource-consumer**: We assume a continuous resource axis, and species trait value (beak size) determines their use of the different resources (seeds of different sizes, left). Our two scenarios differ by whether or not consumption depletes some resources to the point of extirpation (right); **Preferred environment**: species niche-differentiate based on suitability to local environments (left). In this example, plants with seeds of different sizes specialize to different elevations. In this spatially explicit model, our two scenarios differ by whether an individual's distance to and a newly vacated site affects its chances to colonize it with its offspring (right); **Life history tradeoffs**: Species niche-differentiate by trading off fecundity with ability to displace other species (competition-colonization, left) or recruitment in high-stress sites (tolerance-fecundity, right). Here we assume species with smaller seeds have higher fecundity. The patchwork of squares on the right represent a local community, where each square is a local site that can be colonized by a single individual, and the shade of grey represents local stress level, assumed fixed.

We use a lottery model framework to implement stochastic dynamics in a local community subject to immigration from a regional pool. The local community is saturated with J individuals. A proportion m of deaths are replaced by immigrants from the pool, and the remainder by local offspring. This is analogous to Hubbell's[35] neutral community model, except here the niche mechanism sets the probabilities of birth and death across different species. We use the 50-hectare plot of tropical forest on Barro Colorado Island, Panama, as a reference point for our community size ($J = 21,000$ individuals,[36]), default immigration rate ($m = 0.08$ individuals per recruitment event,[37]), and regional abundance distribution (neutral metacommunity with biodiversity parameter $\theta = 50$,[35]. We use 150,000 individuals, leading to approximately 400 species. Higher numbers

had little impact on pattern.) We start the local community with a random draw of offspring from the regional pool, and then simulate local dynamics for 100 million death and replacement events, at which point species abundance distributions are stationary. All simulations and statistical manipulations are performed in R[38].

Heuristic model

The probability that species i is selected for a death event is $\sum_j A_{ij}N_iN_j / \sum_i \sum_j A_{ij}N_iN_j$, where N_i is the abundance of species i and the competition coefficient A_{ij} quantifies the competition caused on species i by species j. The sums are over all species present in the local community, and the denominator normalizes the probabilities to 1. The probability that species i is chosen at the subsequent recruitment stage is $m\, q_i + (1-m)\, r_iN_i / \sum_j r_jN_j$, where q_i is the relative abundance of species i in the regional pool, and r_i is the rate at which it grows when not being limited by competition. (We are placing competitive interactions on mortality, but results are unchanged when we place them on births instead.)

Theoretically, one expects competition to decline with increasing trait differences, as traits determine species' strategies with regards to consumption of resources, or interaction with other regulating factors[39]. We tie competition to traits by setting $A_{ij} = \exp[-(|x_i - x_j|/w)^p]$, where x_i and x_j are the trait values of species i and j. A_{ij} is high when species are similar and low when they are different. The scaling parameter w defines "similar" and "different" by controlling how large $|x_i - x_j|$ must be before competition drops substantially. The shape parameter p sets the smoothness of the transition between the high- and low-competition regimes (a small p gives a tent-like function, while a high p gives a box-like function).

Our model can be compared with the Lotka-Volterra model used by MacArthur and Levins[34] to derive the principle of limiting similarity. The MacArthur-Levins model predicts that species coexisting in equilibrium will be evenly spaced on the trait axis, and is the theoretical basis for much pattern-based research[12]. It is also the model in which transient clustering has been demonstrated[26,40]. Strictly speaking, the differential equations affiliated with our model are not Lotka-Volterra, but we verified numerically that their dynamics are virtually identical (Appendix B). Our approach is computationally expedient and allows easy comparison between different scenarios, as we can directly control community size.

We start with a baseline scenario where numerical exploration revealed strong pattern. We then consider variants of the baseline in order to verify sensitivity to our parameter choices and delineate conditions under which pattern is stronger. For a summary of the heuristic scenarios and corresponding parameter choices, see Table 1. For more details, see Appendix A.

Niche differentiation by preferred resource

Our heuristic model can in principle represent competition based on resource overlap, with trait similarity indicating similar resource preferences,[34,41]. However, it ignores resource dynamics. Here we consider explicit stochastic resource dynamics based on[42], and in particular allow for resource extinction. Abrams showed that resource depletion affects competition-similarity relationships[43] and coexistence outcomes[44]. Because these relationships are drivers of pattern, here we consider two scenarios of this model, differing by whether resource depletion is low or high.

Niche differentiation by preferred abiotic environment

Our next niche model was introduced by Schwilk et al.[22] It is similar to Tilman's stochastic niche model[21], and was also studied by Gravel et al.[23] It assumes a linear array of sites forming a gradient

Table 1: Heuristic model scenarios. Pairwise competition coefficients form the matrix $A_{ij} = \exp(-(d_{ij}/w)^p)$, where $d_{ij} = |x_i - x_j|$ is the absolute trait difference between species i and j, with $x_i \in [0, 1]$. Parameter r_i is the intrinsic growth rate of species i when free from competition. Shape parameter p controls the transition between high- and low-competition regimes as trait differences increase. Scaling parameter w controls the reach of competitive interactions along the trait axis. Immigration parameter m controls the proportion of recruitment events from the regional pool. Biodiversity parameter θ controls regional species diversity S. All variants have the same parameter values as the baseline, except for those specified in the Parameters column. For more details on the variants, see Appendix A.

Name	**Description**	**Parameters**
Baseline	No filtering, Linear trait axis, Quartic competition function, Uniform specialization, Trait similarity $\leftrightarrow$ competition	$r = 1$ $p = 4$ $w = 0.063$ $m = 0.08$ $\theta = 50$ ($S \approx 400$)
Immigration rate	Different immigration rates	$m = 0.005, 0.01, 0.05, 0.15$
Regional diversity	Different regional species diversity	$\theta = 5, 10, 20, 30$
Purifying filtering	Environmental filtering favors species with intermediate trait value	$r_i = x_i(1 - x_i)$
Random filtering	Favored species bear no relation to trait value	$r_i \sim U(0, 1)$
Circular axis	Species with extreme trait values have as many competitors as intermediate species (i.e. no edge effects)	$d_{ij} =$ $\min(\lvert x_i - x_j\rvert, 1 - \lvert x_i - x_j\rvert)$
Gaussian competition	Competition declines more gradually with trait differences than Baseline	$p = 2$
Specialization gradient	Species with low trait value are more specialized and face less competition from others	$w_i = 0.015\frac{S-i}{S-1} + 0.15\frac{i-1}{S-1}$

of fixed environmental conditions. A species' trait determines its optimal environment, and its ability to thrive in other local environments declines with the difference between those environments and its optimum. We explore two dispersal scenarios in this model: local dispersal, where the probability of arrival of propagules to a site quickly drops with the distance to the parent, and global dispersal, where the probability of arrival is independent of distance to the parent.

Niche differentiation by life history strategy

We examine two niche models where species differ by life-history strategy. In the competition-colonization tradeoff[45,46], fecundity trades off with competitive ability, defined as the ability of propagules to displace heterospecific individuals. In the tolerance-fecundity tradeoff[47], fecundity trades off with propagule stress tolerance in a heterogeneous environment. In both models the trait is fecundity, i.e. per-capita propagule production. We note that treating demography as a trait is not standard in trait-based ecology[7]. Alternatively, one can take seed size as the trait, as it correlates (negatively) with fecundity[48]. Unlike the previous mechanisms, here there is a competitive hierarchy: some species are better competitors than others. We use revised formulations of[47,49] without discontinuities in competitive ability, as those inflate coexistence in the original models ([50–52]; see also Appendix A).

Clustering metrics: the gap statistic method

Clusters have only recently been associated with niche differentiation, and the ecological literature lacks good tools for quantifying them. Existing metrics are mostly qualitative or ignore species abundances. Vergnon et al.[27] introduced a test based on variation of species richness along the trait axis. However, the test ignores species abundances, and requires information on permanent and temporary species, which is typically unavailable except for systems well sampled over time. Segura et al.[28] compare Shannon diversity along bins (octaves) on the trait axis against a null expectation of uniform diversity. However, it is not clear that this method is robust to the particular choice of bins, and it has not been tested on model communities generated by niche differentiation.

Here we introduce a variation of the gap statistic method[53], which we apply using two different clustering measures: k-means dispersion[54] and Ripley's K function[55]. The first provides a direct estimate of the number of clusters, whereas the second focuses on distance scales on the trait axis. In general terms, we feed the method the list of traits and abundances of each species in our community, and it returns the estimated number of clusters (k-means) or distance scale of spacing between clusters (Ripley's K), the magnitude of the clustering, and the statistical significance of the result. We explain the technical details of both versions in Appendix C and will provide code for applying our clustering metrics to field data upon manuscript acceptance. In Box 1 we summarize the k-means version and give a step-by-step recipe for its implementation, and illustrate it at work on two example communities.

We assess statistical significance and degree of clustering by comparing our communities against null models. Those we generate by randomly shuffling abundances across our species. This allows us to test specifically for a nonrandom association between traits and abundances, as opposed to a nonrandom distribution of traits or abundances per se. We are thus assuming no relation between a species trait and its local abundance, while making no presumptions regarding the null distribution of abundances and traits. One alternative null model is a neutral community, where abundances would follow a characteristic dispersal-limited multinomial distribution[56].

Because our niche models are stochastic, we run multiple replicates of each scenario to account for variation within the same niche model and scenario. For each niche scenario we report the

average z-score across replicates, as well the percentage of replicates that were significantly clustered (i.e. had a p-value ≤ 0.05).

We check for false positives by testing our metric on neutral communities and communities where differences in species performance are due strictly to environmental filtering. From the former we expect no clustering, and from the latter we expect that species will cluster around the favored trait. Thus we distinguish niche differentiation from pure environmental filtering and neutrality by the presence of multiple clusters as opposed to a single one or none.

Relation between clustering and niche dynamics

In addition to potentially driving community structure, niche differentiation plays a critical role in stabilizing community dynamics. In this study we check whether those two effects are correlated using persistence times—the amount of time species remain in the local community between colonization and extinction events. We compare species' persistence times against their traits in assemblages with and without niche differentiation. Based on the idea that strong niche dynamics will dampen mass effects, we also compare persistence against regional abundances, and check whether the link is weaker when clustering is stronger. In particular, we verify whether species at the center of their clusters have higher persistence than would be predicted based on their regional abundance.

We note that persistence is a coarse measure of dynamics relative to an abundance-weighted measure, as it makes no distinction between rare and common species. Because mass effects can permanently maintain local populations regardless of niche relations, in this part of the study we consider scenarios with lower immigration rate.

Results

Clustering in heuristic model

The baseline scenario of our heuristic niche model (Table 1, Fig. 1) shows clear abundance-by-trait structure (Fig. 2A). There are clusters in lieu of what would be single species in the equilibrium of the Lotka-Volterra model under similar parameter values (Fig. S7). Those clusters occur in the transient state of that model, whereas in ours they are sustained by immigration. The clusters spread evenly along the trait axis, and are each dominated by one or few species. The k-means gap between the community and null communities is insignificant below $k = 12$ clusters, then shoots up above null expectations and peaks at $k = 12$ (Fig. 2B). The gap statistic is the gap value at the peak, $G \simeq 0.7$. Comparison with the null communities yields a z-score $Z > 5$ and a p-value $P < 0.01$, indicating strong and significant clustering (this was a typical run. All 10 runs were significant, with an average z-score of 4.9 and estimated number of clusters between 12 and 13, see Fig. 3C and Fig. S8). By contrast, the neutral community lacks trait structure (Fig. 2C) and yields no clustering (Fig. 2D; average z-score across 100 runs was 0.05). The community under environmental filtering but no niche differentiation (Fig. 2E) shows a single cluster (Fig. 2F), in contrast to the multiple clusters in the niche-differentiated community.

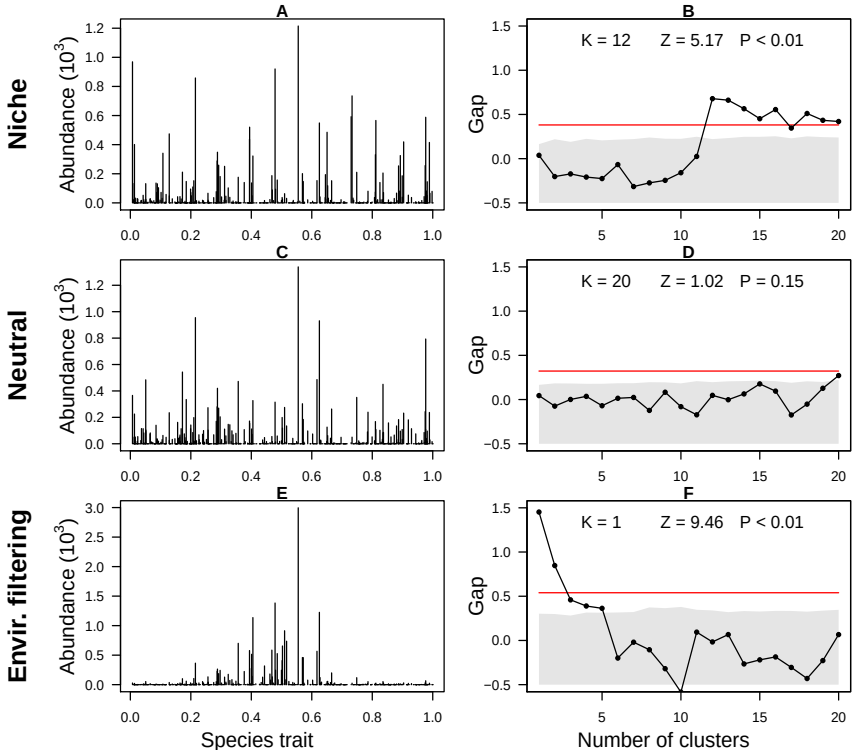


Figure 2: Simulation outcomes and clustering: niche vs neutral. **A**: Visible abundance-by-trait structure in the baseline scenario of the heuristic niche model. **B**: Dispersion gap between the community in A and randomized null communities, plotted against candidate number of clusters k. The gray area marks points below the 95th percentile of the dispersion gap across the null communities. The red line shows the 95th percentile of the gap statistic across the null communities. Starting from $k = 1$, the dispersion gap is low and within null expectations. The curve then shoots up and peaks ($G \simeq 0.7$) at $k = 12$, indicating 12 clusters in A. The high z-score $Z = 5.17$ indicates a strong difference from the null expectation. None of 1,000 null communities had a higher gap statistic than G (p-value $P < 0.01$), signifying high statistical significance. **C, D**: Under neutral dynamics, no significant clustering is observed. **E, F**: Under neutral competition coupled with environmental filtering for intermediate trait values, species cluster around the favored trait. Gap curve shows a very strong single cluster, in contrast with the multiple clusters in the niche-differentiated community.

Our variants of the heuristic model (stem plots in Figs. S9, S10) reveal that clustering is stronger under the following conditions:

Intermediate immigration. As we consider scenarios with increasing immigration, the z-score of the gap statistic initially rises and then falls (Fig. 3A). This indicates that while immigration sustains the clusters, it also adds noise as it is unrelated to local niche dynamics.

Higher regional diversity. When we increase species richness in the regional pool while keeping local parameters unchanged, clustering becomes stronger (Fig. 3B). This occurs because there are more species to occupy the same number of clusters, and therefore more species per cluster, which helps the metric pick up the pattern. As seen above, the number of clusters corresponds to the number of species that can survive without immigration—i.e. the number of locally available niches.

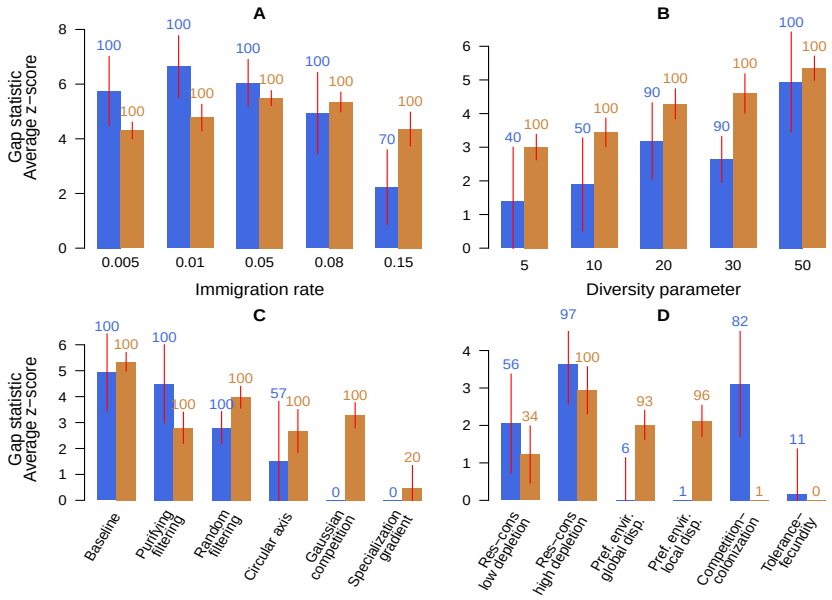


Figure 3: Clustering results across our niche scenarios (k-means in blue, Ripley's K in orange). Results shown are averaged across 10 runs in the heuristic scenarios and 100 runs in the explicit-mechanism scenarios. Bars show the z-scores averaged across runs. Error bars show standard deviation. Numbers over each bar show percentage of runs where clustering was significant (p-value < 0.05). **A**: Baseline heuristic model with increasing immigration rates. Labels show immigration rate used in simulation. **B**: Baseline heuristic model with increasing regional diversity. Labels show neutral diversity parameter θ used to generate regional pool. Higher values correspond to higher regional diversity. **C**: Variants with environmental filtering, circular axis, Gaussian competition, and specialization gradient (see Table 1 and Appendix A for full description of each scenario). By comparison, the average z-score of the k-means gap statistic across 100 neutral runs was 0.05, and the standard deviation was 0.85. **D**: Models with explicit niche mechanisms ($m = 0.01$ and $\theta = 50$).

Low environmental filtering. When the environment selected for species with intermediate trait values, those ended up with higher abundance (Fig. S10). This abundance envelope interfered with niche-driven clustering, lowering the z-score (Fig. 3C). An extreme envelope will eclipse niche differentiation and force a single cluster, as in the case of pure filtering (see Fig. S1 and Fig. S8). Severe filtering can also impair the metric's ability to detect clustering (Appendix C). Filtering unrelated to the niche-trait also weakened clustering, as expected, but did not wipe it out completely (Fig. 3C).

Edge effects. Removing edge effects by making the axis circular lowered the average z-score and led to a lower percentage of significantly clustered runs (Fig. 3C). Edge effects break the otherwise perfect symmetry of the axis and set the positions of the niches: there is a niche at each edge, and as many niches in between as allowed by the width of the competition curve. On a circular axis, the niche positions shift through time with stochastic fluctuations, leading to inconsistency in which species are favored and which are disfavored, and lessening pattern.

Steep transition between high and low competition. Changing the shape parameter p from 4 to 2 (Gaussian competition) makes competition decline more gradually with trait differences relative to the box-like baseline, $p = 4$. This leads to weaker clustering (Fig. 3C) because niche dynamics are slower and hence more easily overwhelmed by immigration (Fig. S11).

No specialization gradient. We saw no clustering in our specialization gradient scenario—where species with small trait values consume fewer resources than those with large trait values (Fig. 3C). The stem plot (Fig. S10) suggests that clustering may still be occurring, but the spacing between

them increases towards high trait values. This lack of consistency in cluster spacing may be throwing off the metrics we used, and it is possible that a different metric of clustering may capture it.

Clustering under explicit niche mechanisms

Clustering occurred across niche mechanisms, albeit with considerable variation between them (Figs. 3D, 4). It was also generally less pronounced than in the heuristic model, and stronger under lower immigration—compare Fig. 3D ($m = 0.01$) and Fig. S13 ($m = 0.08$).

Clustering was strong among consumers under niche differentiation for resources, particularly when resource depletion is high (Fig. S12 shows both consumers and resources). The extirpation of most resources forces consumers to compete for the remaining few, and species then cluster based on specialization to specific resources. Abrams and coauthors have argued that resource depletion fundamentally changes the competitive interactions between consumers[43,44], but our results indicate that with regards to clustering the outcome is qualitatively similar, if quantitatively different.

Our k-means implementation failed to find clustering under niche differentiation by preferred environment. However, our Ripley's K implementation indicated clustering in both dispersal scenarios of this model. Clustering was stronger when dispersal was local rather than global. This dispersal limitation may effectively reduce immigration. The discrepancy in results between the measures may be related to parallels between the Schwilk-Ackerly model we used and the heuristic model with Gaussian competition, where Ripley's K finds clustering but k-means does not. In the Schwilk-Ackerly model, overlap in environmental preference is a Gaussian function of species similarity. The model therefore inherits the slow niche sorting properties of the Gaussian scenario, and the niche dynamics get swamped by immigration.

The life history tradeoff models were characterized by striking asymmetries on the niche axis (stem plots in Fig. 4). The competition-colonization tradeoff led to strong clustering according to the k-means measure, although for a number of clusters substantially lower than the number of species found at equilibrium in the deterministic model with similar parameter values (Fig. S7). We note that only high-tolerance species survived in our simulations of the tolerance-fecundity tradeoff, despite immigration (Fig. 4; the same was true of the deterministic immigration-free model, see Fig. S7). We expect higher coexistence and potentially stronger trait pattern from a spatially explicit formulation of this niche mechanism, where the higher fecundity would provide better compensation for low stress tolerance.

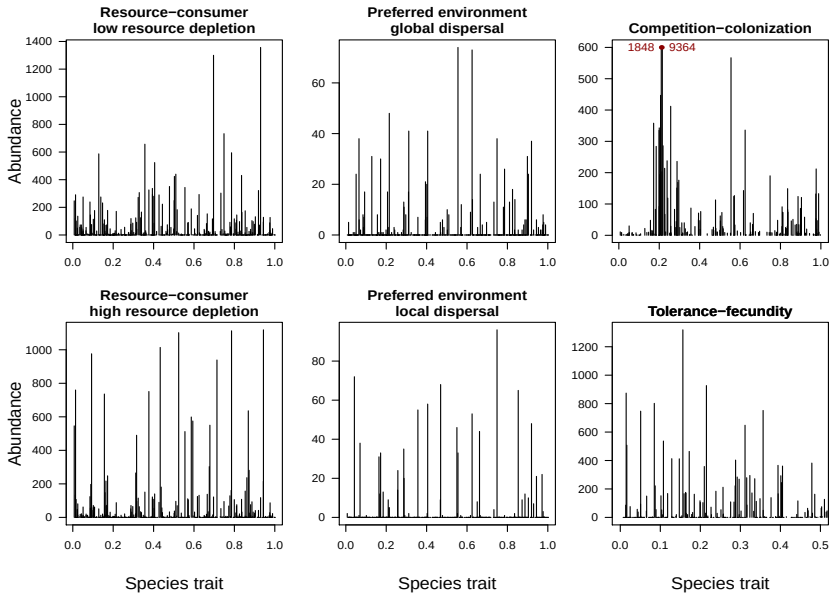


Figure 4: Simulation outcomes of models with explicit niche mechanisms. Left: Niche differentiation by preferred resource. In the scenario shown in the top panel, parameters are set so that resources are not severely depleted by consumption; in the bottom scenario, consumption causes extensive resource depletion and extinction. **Center**: Niche differentiation by preferred environmental environment. In the top scenario, species are not limited by dispersal: the offspring of every individual compete for every vacancy. In the bottom scenario, dispersal declines with distance to the vacated site. **Right**: Niche differentiation by life history strategy. Top: Competition-colonization tradeoff. Numbers in red show abundances of species exceeding 1,000 individuals (marked by red dots). Bottom: Tolerance-fecundity tradeoff. Notice truncation of trait axis. Species with higher fecundity had only background level populations kept by mass effects, and were not included in the clustering analysis.

Relation between clustering and niche dynamics

In addition to being a recognizable snapshot pattern, clustering also reflects the impact of competition on community dynamics. Fig. 5 shows the distribution of persistence times plotted against trait values and overlaid on species abundances. In a neutral assemblage, persistence times have no relation to traits (Fig. 5A). In an assemblage under environmental filtering but otherwise neutral competition, species favored by the local environment tend to persist longer (Fig. 5B). Under niche differentiation, however, even species disfavored by the environment can have high persistence, in accordance with their niche standing. Persistence times follow the same periodic distribution against traits as abundances (Fig. 5C), indicating that species near the center of their clusters are well adapted to their niches and persist longer than those located in the gaps between the clusters/niches. In fact, persistence times in assemblage C are also clustered, with 13 clusters, a k-means gap z-score of 2.5 and a p-value of 0.001. By comparison, the abundances form 13 clusters with a z-score of 4.29 and p-value < 0.001.

Because of propagule pressure, regional abundances influence local persistence times. The relationship is tight in the neutral model (Fig. 5D), but looser under filtering and niche differentiation (Fig. 5E-F). Note that species sitting at the center of their clusters, marked by red dots, have very long persistence times even when they are not regionally abundant. Furthermore, niche scenarios where clustering is stronger tend to also be those where local persistence is less related to regional abundances, relative to scenarios with comparable niche dynamics but weaker clustering (Fig. S14).

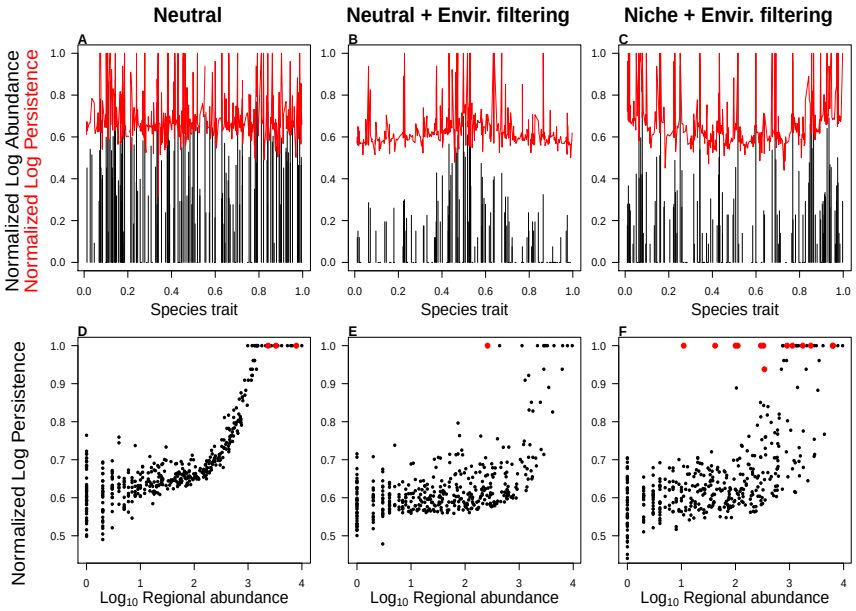


Figure 5: Persistence times. **A-C**: Abundances and persistence times plotted against species traits. We log quantities to facilitate visualization, and normalize in order to overlay the plots. The black stems represent $y = \log(N)/\log(N_{\max})$, where N are species abundances and $N_{\max}$ is the maximum abundance observed in the respective community. The red lines represent $y = \log(T)/\log(T_{\max})$, where T refers to the persistence time in simulation steps, and $T_{\max} = 5 \times 10^7$ is the maximum possible persistence time, corresponding to the number of simulation steps during which we tracked persistence. Environmental filtering used here is of the purifying type, i.e. species with intermediate trait values have higher intrinsic growth rates. Immigration rate in all three assemblages is $m = 0.005$. **D-F**: Persistence times plotted against regional abundances (log-log scale). Because more regionally abundant species immigrate more frequently, they drive local persistence times in all three assemblages. However, the connection is much stronger in the neutral assemblage (D) than the other two (E, F). Red dots represent species in the center of each cluster, meaning those whose individuals have the lowest average trait-distance to all other individuals in their cluster (technically the medoids, see[57]). In the niche community, those species have the highest possible persistence even when they are regionally rare. This did not occur in the neutral community.

Overdispersion and even spacing

There was little to no overdispersion or even spacing between species in our niche scenarios (Fig. 6). We used two common metrics of dispersion and evenness, detailed in Appendix D: the average trait separation between a randomly chosen pair of individuals (Rao's quadratic index[9]), and the coefficient of variation (CV) of trait separation between adjacent species. Results are shown in Fig. 6. We discuss them in detail in Appendix D, and summarize them below.

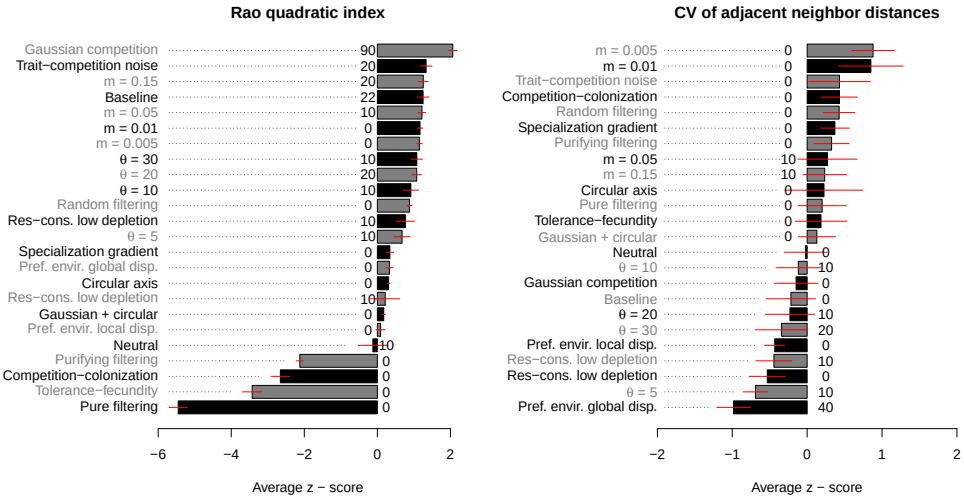


Figure 6: Average z-scores of Rao's quadratic index and the coefficient of variation (CV) in trait separation between adjacent species. Numbers near the bars show percentage of runs that were significant in the expected direction (high Rao, low CV). Only niche scenarios with at least one significant run are shown. With one exception (Gaussian competition), significance levels and effect size are small in all niche scenarios, with mean z-scores under magnitude 1.6 and significance percentages under 50%.

Species in our niche scenarios do not show a general tendency to overdisperse in trait space in the sense of Rao's quadratic index. In fact, in some scenarios they were underdispersed (Fig. 6). This was the case of our tradeoff scenarios and the heuristic model with purifying filtering, where intermediate-trait species have a fitness advantage. These niche scenarios had a lower Rao index than the fully neutral communities. Only one case showed substantial overdispersion, namely that of Gaussian competition with a finite trait axis, in which edge effects were strong due to slow niche sorting. The Rao index has been used to show minimization of niche overlap in marine phytoplankton[27], but we note that in that study the number of niches in the community seemed low. Our results suggest that Rao may not be suitable for a higher number of niches.

Species are also not evenly spaced along the trait axis. The coefficient of variation (CV) is often indistinguishable from random or neutral communities (Fig. 6). None of our niche scenarios had high z-scores nor a high percentage of significance under this metric. Metrics of even spacing that do not account for abundances are severely limited in the context of open communities, as they cannot distinguish between dominant species and rare species maintained by mass effects.

We conclude that functional dispersion and even spacing are often poor indicators of niche differentiation because they are prone to false negatives. Functional dispersion is sensitive to differences in intrinsic growth rates and other types of filtering that may counter the effects of niche differentiation, and even spacing is more suitable for closed communities.

Discussion

Competition is a driving force in nature, and understanding its impact on species coexistence is a longstanding goal in community ecology. Our study indicates that communities shaped by competition, chance, and dispersal, are generally characterized by clusters of species similar in traits. When the number of species immigrating into the community exceeds the number of niches locally available, it pays to be similar rather than to be different. Clusters then represent successful niche strategies which species do well to adopt. Some authors in fact see clusters as a bridge between

niche and neutral theory, and term the formation of clusters emergent neutrality[18,28,58]. Our finding that clusters appear under various niche mechanisms and can be easily maintained by immigration, even when confounding forces are at play, suggests that clusters are a likely feature of nature. Incorporating clusters into trait-based searches of niche differentiation in communities will represent a substantial departure from the dominant hypothesis of maximal trait separation.

Our tests indicate that clustering originates from a strong link between competitive interactions and trait similarity. This suggests that any niche mechanism with that property will produce clustering. On the other hand, clusters are less likely to appear in low-diversity systems or when competitive sorting is so slow and weak as to be overwhelmed by stochastic noise and washed out by immigration, or so fast that immigration cannot maintain the clusters. Finally, we found that clustering relates to the impact of niches on species persistence, indicating a link between the impact of niches on static and dynamic aspects of community structure.

Our results apply in principle not only to traits but any measure of species similarity reflecting degree of niche overlap. In systems with phylogenetic niche conservatism[59], competition may be stronger between closer relatives[60]. Our results suggest that in these cases competition may cause phylogenetic clusters among coexisting species, challenging Darwin's expectation of greater coexistence between distant relatives[61–63]. One could look for these clusters by applying our metrics to measures of phylogenetic distance. Analogously, in microbial communities where genotype similarity leads to competition, clustered patterns of abundance in genotype space might provide evidence of niche differentiation[64,65].

Our analysis can be extended to competition in higher dimensions. Competitive interactions may often be mediated by variation in multitrait rather than single-trait phenotypes[66]. Theoretical work on simple models indicates multidimensional niche space leads to multidimensional clustering[67]. Our metric can quantify clusters in any dimension, using generalized measures of niche separation. The main challenge is to connect multitrait phenotypes to competitive relations, in order to define the correct measure of distance in multitrait space. This requires natural history expertise.

The theoretical underpinnings for clustering go back to MacArthur and Levins[34]. They showed that an intermediate-trait species can invade two residents only if the latter are sufficiently separated on the trait axis. However when that condition is not met, the invader has higher fitness if it is similar to one of the residents than if it falls right in between them. The argument extends to multi-species communities: competition is overall lower near the niche-differentiated residents than in the midpoints. Species in the midpoints are therefore excluded faster, and the remaining ones form clusters. More recently, the formation of transient clusters was connected with the structure of competitive interactions in the MacArthur-Levins model[40,67]. Our study shows that this effect also holds under more complex niche mechanisms, and under more general circumstances.

Other authors have suggested that transient clusters can persist through highly specialized enemies, or species differences on implicit additional niche axes[26,58,68]. Our work highlights that those are not necessary, as immigration alone can cause clusters to persist. Tilman[21] found little impact of immigration on resulting trait pattern, unless the adult seedling biomass is unrealistically small, or disturbance events are large or frequent. However, successful immigration was highly infrequent in his model. Resources made available through mortality were assumed to be redistributed broadly, so that available resource levels remain low in any one location, making recruitment of new individuals highly unlikely. In contrast, in our model, mortality events create local opportunities for recruitment. Schwilk and Ackerly[22] also considered very low immigration. Though in their model mortality did open up chances for recruitment, when immigration occurred it took the form of a single immigrant seed being added to a large pool of local seeds competing for the site. Not surprisingly, they also found little influence of immigration on pattern. Gravel et al.[23] tested both very high and very low immigration, and also found even spacing at the low end and random-looking patterns at the

high end. Their competition kernel is a Gaussian function of trait separation, which leads to weak niche sorting dynamics, easily overwhelmed under high immigration. This might explain the lack of clustering in their study.

Alternative sources of clustering include one-sided competition, environmental filters, speciation, and mutation. One-sided competitive dynamics, where a particular trait outcompetes all others without a balancing tradeoff[69], will produce a single cluster as opposed to multiple clusters. Local environmental filters will also produce a single cluster through selection for the best trait. At larger spatial scales a heterogeneous environment will filter for different trait values at different sites, potentially causing multiple clusters. We note that competitive exclusion and environmental filtering are not necessarily different biological processes, but rather a matter of scale[70]. Indeed, heterogeneous filtering may comprise a niche mechanism, as in the Schwilk-Ackerly model. Sympatric speciation leading to similar parent-daughter species may also generate clusters without niche differentiation[65], though only if community size and speciation rates are sufficiently small[71]. Similarly, neutral mutation is another possible source of clustering in genotype space[65,71]. Ruling out alternative sources of clustering could require further testing at a spatial scale larger than expected from the niche mechanism[5,72], or corroborating evidence from system-specific methods such as measurement of competitive effects or verification of frequency dependence[4].

We assumed sufficiently small spatial scales such that species can easily disperse within the local community, with a substantial proportion of diversity being maintained through dispersal from outside. It could in principle be argued that at larger spatial scales diversity can only be explained by niche differences. Species locally maladapted might survive because they are niche-differentiated elsewhere. Thus once larger scales are considered, every species could be found to occupy their own niche. However, this does not mean that even spacing (or any pattern in general) is expected at these scales. The set of species stabilized by trait differences will differ from one location to the next, and these differences may have more to do with spatial isolation and historical contingency than niche differences. Ultimately, it is possible that the sum total of heterogeneous local communities, each with distinctive local trait pattern, results in no discernible pattern at a regional scale.

Clusters bring a paradigm shift to our understanding of how competition shapes communities. We showed that they appear across different niche mechanisms and are sustained by immigration, despite demographic stochasticity and environmental filters. We provided parameter-free tools to find and quantify them in species assemblages, which can be adapted for use with genetic data. We showed that clusters are more likely to occur when competition arises from trait similarity, niche sorting is fast relative to immigration rates, and regional diversity is high. Future work will clarify how evolutionary forces contribute to cluster emergence, and how the relative spatial scales of dispersal and niche differentiation determine the spatial scale at which clustering should be most visible.

Acknowledgments

We thank David Ackerly for helpful discussions, and Hugo Fort, Brian Sedio, James O'Dwyer, Stephen Smith, and Lacey Knowles for comments on previous versions of the manuscript. This work was supported by the National Science Foundation under grant 1038678, "Niche versus neutral structure in populations and communities" funded by the Advancing Theory in Biology program, as well as by the Danish National Research Foundation through support of the Center of Macroecology, Evolution and Climate (grant DNRF96), and by the Miller Institute for Basic Research in Science at the University of California, Berkeley.

References

[1] D. Simberloff and W. Boecklen, "Santa Rosalia Reconsidered: Size Ratios and Competition," *Evolution*, vol. 35, no. 6, pp. 1206–1228, 1981.

[2] J. M. Chase and M. Leibold, *Ecological Niches: Linking Classical and Contemporary Approaches*. The University of Chicago Press, 2003.

[3] L. Götzenberger, F. de Bello, K. A. Bråthen, J. Davison, A. Dubuis, A. Guisan, J. Lepš, R. Lindborg, M. Moora, M. Pärtel, L. Pellissier, J. Pottier, P. Vittoz, K. Zobel, and M. Zobel, "Ecological assembly rules in plant communities-approaches, patterns and prospects," *Biological Reviews*, vol. 87, no. 1, pp. 111–127, 2012.

[4] J. HilleRisLambers, P. B. Adler, W. Harpole, J. M. Levine, and M. M. Mayfield, "Rethinking Community Assembly through the Lens of Coexistence Theory," *Annual Review of Ecology, Evolution, and Systematics*, vol. 43, pp. 227–248, dec 2012.

[5] R. D'Andrea and A. M. Ostling, "Challenges in linking trait patterns to niche differentiation," *Oikos*, vol. 125, no. 10, pp. 1369–1385, 2016.

[6] B. J. McGill, B. J. Enquist, E. Weiher, and M. Westoby, "Rebuilding community ecology from functional traits.," *Trends in ecology & evolution*, vol. 21, pp. 178–85, apr 2006.

[7] C. Violle, M.-L. Navas, D. Vile, E. Kazakou, C. Fortunel, I. Hummel, and E. Garnier, "Let the concept of trait be functional!," *Oikos*, vol. 116, pp. 882–892, may 2007.

[8] P. R. Grant, "Colonization of islands by ecologically dissimilar species of birds bill length feeding," *Canadian Journal of Zoology*, vol. 47, no. 1, pp. 41–43, 1969.

[9] Z. Botta-Dukát, "Rao's quadratic entropy as a measure of functional diversity based on multiple traits," *Journal of Vegetation Science*, vol. 16, no. 5, pp. 533–540, 2005.

[10] E. Laliberté and P. Legendre, "A distance-based framework for measuring functional diversity from multiple traits.," *Ecology*, vol. 91, pp. 299–305, jan 2010.

[11] F. de Bello, "The quest for trait convergence and divergence in community assembly: are null-models the magic wand?," *Global Ecology and Biogeography*, pp. no–no, jun 2011.

[12] W. J. Stubbs and J. B. Wilson, "Evidence for limiting similarity in a sand dune community," *Journal of Ecology*, vol. 92, pp. 557–567, aug 2004.

[13] W. K. Cornwell, D. W. Schwilk, and D. D. Ackerly, "A trait-based test for habitat filtering: Convex hull volume," *Ecology*, vol. 87, no. 6, pp. 1465–1471, 2006.

[14] T. Ingram and J. B. Shurin, "Trait-based assembly and phylogenetic structure in northeast Pacific rockfish assemblages.," *Ecology*, vol. 90, pp. 2444–53, sep 2009.

[15] T. Astor, J. Strengbom, M. P. Berg, L. Lenoir, B. Marteinsd??ttir, and J. Bengtsson, "Under-dispersion and overdispersion of traits in terrestrial snail communities on islands," *Ecology and Evolution*, vol. 4, no. 11, pp. 2090–2102, 2014.

[16] M. Pontarp and O. L. Petchey, "Community trait overdispersion due to trophic interactions : concerns for assembly process inference," *Proc Roy Soc Lond. B*, vol. 283, p. 20161729, 2016.

[17] K. J. Gaston and S. L. Chown, "Neutrality and the niche," *Functional Ecology*, vol. 19, no. 1, pp. 1–6, 2005.

[18] R. D. Holt, "Emergent neutrality.," *Trends in Ecology & Evolution*, vol. 21, pp. 531–3, oct 2006.

[19] M. Vellend, "Conceptual synthesis in community ecology.," *The Quarterly review of biology*, vol. 85, pp. 183–206, jun 2010.

[20] J. Rosindell, M. O. Hunter, and R. S. Etienne, "The unified neutral theory of biodiversity and biogeography at age ten.," *Trends in ecology & evolution*, vol. 26, pp. 340–8, jul 2011.

[21] D. Tilman, "Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, pp. 10854–61, jul 2004.

[22] D. W. Schwilk and D. D. Ackerly, "Limiting similarity and functional diversity along environmental gradients," *Ecology Letters*, vol. 8, pp. 272–281, jan 2005.

[23] D. Gravel, C. D. Canham, M. Beaudet, and C. Messier, "Reconciling niche and neutrality: the continuum hypothesis.," *Ecology letters*, vol. 9, pp. 399–409, apr 2006.

[24] R. D'Andrea and A. M. Ostling, "Biodiversity maintenance may be lower under partial niche differentiation than under neutrality. doi: 10.1002/ecy.2020," *Ecology*, 2017.

[25] K. Scranton and D. A. Vasseur, "Coexistence and emergent neutrality generate synchrony among competitors in fluctuating environments," *Theoretical Ecology*, vol. 9, no. 3, pp. 353–363, 2016.

[26] M. Scheffer and E. H. van Nes, "Self-organized similarity, the evolutionary emergence of groups of similar species.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 16, pp. 6230–6235, 2006.

[27] R. Vergnon, N. K. Dulvy, and R. P. Freckleton, "Niches versus neutrality: uncovering the drivers of diversity in a species-rich community.," *Ecology letters*, vol. 12, pp. 1079–90, oct 2009.

[28] A. M. Segura, D. Calliari, C. Kruk, D. Conde, S. Bonilla, and H. Fort, "Emergent neutrality drives phytoplankton species coexistence.," *Proceedings. Biological sciences / The Royal Society*, vol. 278, pp. 2355–61, aug 2011.

[29] Z. Yang, J. Powell, C. Zhang, and G. Du, "The effect of environmental and phylogenetic drivers on community assembly in an alpine meadow community," *Ecology*, vol. 93, no. 11, pp. 2321–2328, 2012.

[30] A. M. Segura, C. Kruk, D. Calliari, F. García-Rodriguez, D. Conde, C. E. Widdicombe, and H. Fort, "Competition Drives Clumpy Species Coexistence in Estuarine Phytoplankton," *Scientific Reports*, vol. 3, no. 1, p. 1037, 2013.

[31] M. E. Johansson and P. A. Keddy, "Intensity and asymmetry of competition between plant pairs of different degrees of similarity: an experimental study on two guilds of wetland plants," *Oikos*, vol. 60, no. 1, pp. 27–34, 1991.

[32] L. Jiang, J. Tan, and Z. Pu, "An experimental test of Darwin's naturalization hypothesis.," *The American Naturalist*, vol. 175, no. 4, pp. 415–23, 2010.

[33] J. Burns and S. Strauss, "More closely related species are more ecologically similar in an experimental test," *Proceedings of the National Academy of Sciences*, vol. 108, no. 13, pp. 5302–5307, 2011.

[34] R. H. MacArthur and R. Levins, "The Limiting Similarity, Convergence, and Divergence of Coexisting Species," *The American Naturalist*, vol. 101, no. 921, pp. 377–385, 1967.

[35] S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography.* Princeton, NJ: Princeton University Press, 2001.

[36] I. Volkov, J. R. Banavar, M. O. Hunter, and A. Maritan, "Neutral theory and relative species abundance in ecology.," *Nature*, vol. 424, pp. 1035–7, aug 2003.

[37] R. A. Chisholm and J. W. Lichstein, "Linking dispersal, immigration and scale in the neutral theory of biodiversity.," *Ecology Letters*, vol. 12, pp. 1385–93, dec 2009.

[38] R Core Team, "R: A Language and Environment for Statistical Computing," 2017.

[39] S. Levin, "Community equilibria and stability, and an extension of the competitive exclusion principle," *The American Naturalist*, vol. 104, no. 939, pp. 413–423, 1970.

[40] H. Fort, M. Scheffer, and E. H. van Nes, "The paradox of the clumps mathematically explained," *Theoretical Ecology*, vol. 2, no. 3, pp. 171–176, 2009.

[41] E. Hernández-García, C. López, S. Pigolotti, and K. H. Andersen, "Species competition: coexistence, exclusion and clustering.," *Philosophical Transactions of the Royal Society A*, vol. 367, pp. 3183–95, aug 2009.

[42] M. L. Rosenzweig and R. H. MacArthur, "Graphical Representation and Stability Conditions of Predator-Prey Interactions," *American Naturalist*, vol. 97, no. 895, pp. 209–223, 1963.

[43] P. Abrams, C. Rueffler, and R. Dinnage, "Competition-similarity relationships and the nonlinearity of competitive effects in consumer-resource systems.," *The American Naturalist*, vol. 172, pp. 463–74, oct 2008.

[44] P. Abrams and C. Rueffler, "Coexistence and limiting similarity of consumer species competing for a linear array of resources.," *Ecology*, vol. 90, pp. 812–22, mar 2009.

[45] R. Levins and D. Culver, "Regional Coexistence of Species and Competition between Rare Species," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 68, no. 6, pp. 1246–1248, 1971.

[46] D. Tilman, *Resource competition and community structure.* Princeton University Press, monographs ed., 1982.

[47] H. C. Muller-Landau, "Tolerance-Fecundity Tradeoff Supporting Information," *PNAS*, pp. 1–5, 2010.

[48] M. L. Henery and M. Westoby, "Seed Mass and Seed Nutrient Content as Predictors of Seed Output Variation between Species," *Oikos*, vol. 92, no. 3, pp. 479–490, 2001.

[49] a. P. Kinzig, S. Levin, J. Dushoff, and S. Pacala, "Limiting Similarity, Species Packing, and System Stability for Hierarchical Competition-Colonization Models," *The American Naturalist*, vol. 153, pp. 371–383, apr 1999.

[50] G. Barabás, R. D'Andrea, and A. M. Ostling, "Species packing in nonsmooth competition models," *Theoretical Ecology*, vol. 6, pp. 1–19, jan 2013.

[51] R. D'Andrea, G. Barabás, and A. M. Ostling, "Revising the tolerance-fecundity trade-off; or, on the consequences of discontinuous resource use for limiting similarity, species diversity, and trait dispersion," *The American Naturalist*, vol. 181, no. 4, pp. E91–101, 2013.

[52] B. Haegeman, T. Sari, and R. S. Etienne, "Predicting coexistence of plants subject to a tolerance-competition trade-off.," *Journal of Mathematical Biology*, vol. 68, no. 7, pp. 1815–47, 2014.

[53] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, pp. 411–423, 2001.

[54] J. B. MacQueen, "Some Methods for classification and Analysis of Multivariate Observations," *5th Berkeley Symposium on Mathematical Statistics and Probability 1967*, vol. 1, no. 233, pp. 281–297, 1967.

[55] P. M. Dixon, "Ripley's K function," in *Encyclopedia of Environmetrics* (A. H. El-shaarawi and W. W. Piegorsch, eds.), vol. 3, pp. 1796–1803, Chichester: John Wiley & Sons, 2002.

[56] R. S. Etienne, D. Alonso, and A. J. McKane, "The zero-sum assumption in neutral biodiversity theory.," *Journal of theoretical biology*, vol. 248, pp. 522–36, oct 2007.

[57] P. J. Rousseeuw and L. Kaufman, *Finding Groups in Data.* Wiley Online Library, 1990.

[58] R. Vergnon, E. H. van Nes, and M. Scheffer, "Emergent neutrality leads to multimodal species abundance distributions.," *Nature communications*, vol. 3, p. 663, jan 2012.

[59] J. Lord, M. Westoby, and M. Leishman, "Seed Size and Phylogeny in Six Temperate Floras: Constraints, Niche Conservatism, and Adaptation," *The American Naturalist*, vol. 146, no. 3, pp. 349–364, 1995.

[60] C. Violle, D. R. Nemergut, Z. Pu, and L. Jiang, "Phylogenetic limiting similarity and competitive exclusion," *Ecology Letters*, vol. 14, no. 8, pp. 782–787, 2011.

[61] C. Daehler, "Darwin's Naturalization Hypothesis Revisited," *The American Naturalist*, vol. 158, no. 3, pp. 324–330, 2001.

[62] C. O. Webb, D. D. Ackerly, M. A. McPeek, and M. J. Donoghue, "Phylogenies and Community Ecology," *Annual Review of Ecology and Systematics*, vol. 33, no. 1, pp. 475–505, 2002.

[63] O. Godoy, N. J. Kraft, and J. M. Levine, "Phylogenetic relatedness and the determinants of competitive outcomes," *Ecology Letters*, vol. 17, no. 7, pp. 836–844, 2014.

[64] P. Jeraldo, M. Sipos, N. Chia, J. M. Brulc, A. S. Dhillon, M. E. Konkel, C. L. Larson, K. E. Nelson, A. Qu, L. B. Schook, F. Yang, B. A. White, and N. Goldenfeld, "Quantification of the relative roles of niche and neutral processes in structuring gastrointestinal microbiomes," *PNAS*, vol. 109, no. 25, pp. 9692–9698, 2012.

[65] R. D'Andrea and A. M. Ostling, "Can Clustering in Genotype Space Reveal "Niches"?," *The American Naturalist*, vol. 187, no. 1, pp. 130–135, 2016.

[66] J. S. Clark, M. Dietze, S. Chakraborty, P. K. Agarwal, I. Ibanez, S. LaDeau, and M. Wolosin, "Resolving the biodiversity paradox.," *Ecology letters*, vol. 10, pp. 647–59; discussion 659–62, aug 2007.

[67] H. Fort, M. Scheffer, and E. H. van Nes, "The clumping transition in niche competition: a robust critical phenomenon," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2010, p. P05005, may 2010.

[68] G. Barabás, R. D'Andrea, R. Rael, G. Meszéna, and A. M. Ostling, "Emergent Neutrality or Hidden Niches?," *Oikos*, vol. 122, pp. 1565–1572, 2013.

[69] M. M. Mayfield and J. M. Levine, "Opposing effects of competitive exclusion on the phylogenetic structure of communities.," *Ecology letters*, vol. 13, pp. 1085–93, sep 2010.

[70] P. B. Adler, A. Fajardo, A. R. Kleinhesselink, and N. J. B. Kraft, "Trait-based tests of coexistence mechanisms.," *Ecology letters*, aug 2013.

[71] B. Houchmandzadeh, "Neutral aggregation in finite-length genotype space," *Physical Review E*, vol. 95, no. 1, pp. 1–12, 2017.

[72] N. J. B. Kraft and D. D. Ackerly, "Functional trait and phylogenetic tests of community assembly across spatial scales in an Amazonian forest," *Ecological Monographs*, vol. 80, pp. 401–422, aug 2010.

[73] P. I. Prado, M. D. Miranda, and A. Chalom, *sads: Maximum Likelihood Models for Species Abundance Distributions*, 2016.

[74] R. K. S. Hankin, "Introducing untb, an R package for simulating ecological drift under the unified nuetral theory of biodiversity," *Journal of Statistical Software*, vol. 22, sep 2007.

[75] M. Vallade and B. Houchmandzadeh, "Analytical solution of a neutral model of biodiversity.," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 68, no. 6 Pt 1, p. 061902, 2003.

[76] P. A. Keddy, "Assembly and response rules: two goals for predictive community ecology," *Journal of Vegetation Science*, vol. 3, no. 2, pp. 157–164, 1992.

[77] S. Pigolotti, C. Lopez, E. Hernández-García, and K. A. H., "How Gaussian competition leads to lumpy or uniform species distributions," *Theoretical Ecology*, vol. 3, no. 2, pp. 89–96, 2010.

[78] O. Leimar, A. Sasaki, M. Doebeli, and U. Dieckmann, "Limiting similarity, species packing, and the shape of competition kernels," *Journal of Theoretical Biology*, vol. 339, pp. 3–13, 2013.

[79] R. H. MacArthur, *Geographical ecology; patterns in the distribution of species*. Harper & Row, 1972.

[80] M. Yan and K. Ye, "Determining the number of clusters using the weighted gap statistic," *Biometrics*, vol. 63, no. 4, pp. 1031–1037, 2007.

[81] B. Yan, J. Zhang, Y. Liu, Z. Li, X. Huang, W. Yang, and A. Prinzing, "Trait assembly of woody plants in communities across sub-alpine gradients: Identifying the role of limiting similarity," *Journal of Vegetation Science*, vol. 23, pp. 698–708, aug 2012.

Box 1: Detecting clustering using the gap statistic with k-means

The first step in the k-means clustering analysis is to assign individuals of each species to one of a specified number of clusters based on their trait values in a way that maximizes the "tightness" of the clusters—i.e. the average similarity of individuals in the same cluster. This index gives us a measure of how well those clusters fit the data. We then repeat this procedure on each of a set of null communities, which we generate by randomly shuffling abundances across species. The difference in the clustering index (i.e. the gap) between the observed data and the null data gives a sense of whether the clusters fit our community better than expected.

If we knew how many clusters the community is supposed to have—i.e. how many niches our species are sorting into—we could stop here and take this gap as our summary statistic. In general we do not know the number of clusters, so we perform the comparison for a range of possibilities, and find the number of clusters that maximizes the gap between the observed data and the null data. This is then the estimated number of clusters, and the maximum gap is our summary statistic—the gap statistic[53].

Finally, we asses the magnitude of the clustering and its statistical significance by repeating the entire procedure on each of the null communities, thus obtaining a set of null values of this statistic. The standardized effect size (z-score)—defined as the difference in gap statistic from the mean null value, divided by the standard deviation of null values—gives us the magnitude of the clustering; the proportion of null values with a higher gap statistic gives us the p-value.

More specifically, the k-means clustering analysis involves the following steps:

1. Decompose the community into a list of individuals $j = 1, 2, \ldots, J$, where J is the community size. Let x_j be the trait value of the j-th individual in the community. Use the k-means algorithm[54] to partition the elements of vector x_j into a single group ($k = 1$). Calculate the clustering measure $F_k = log(1/W_k)$ for $k = 1$. (This clustering measure is inversely related to W_k, a "k-means dispersion" metric, which is specifically the total across the k clusters of the within-cluster sum of squared distances between individuals–see Appendix C for the mathematical definition.)

2. Repeat step 1 for every k between 1 and the maximum number of clusters being tested for (in our study, $k_{\max} = 20$). Obtain vector F_k, with $k = 1, 2, \ldots, k_{\max}$.

3. Repeat step 2 on each of N null communities, each generated by redistributing abundances across species. Obtain the vector of average clustering in the null communities $\widetilde{F}_k = \frac{1}{N}\sum_i \widetilde{F}_{ik}$, where $i = 1, 2, \ldots, N$ labels the null communities.

4. You now have a vector of gap values $G_k = F_k - \widetilde{F}_k$ quantifying how much better k clusters fits our data than it fits the null communities.

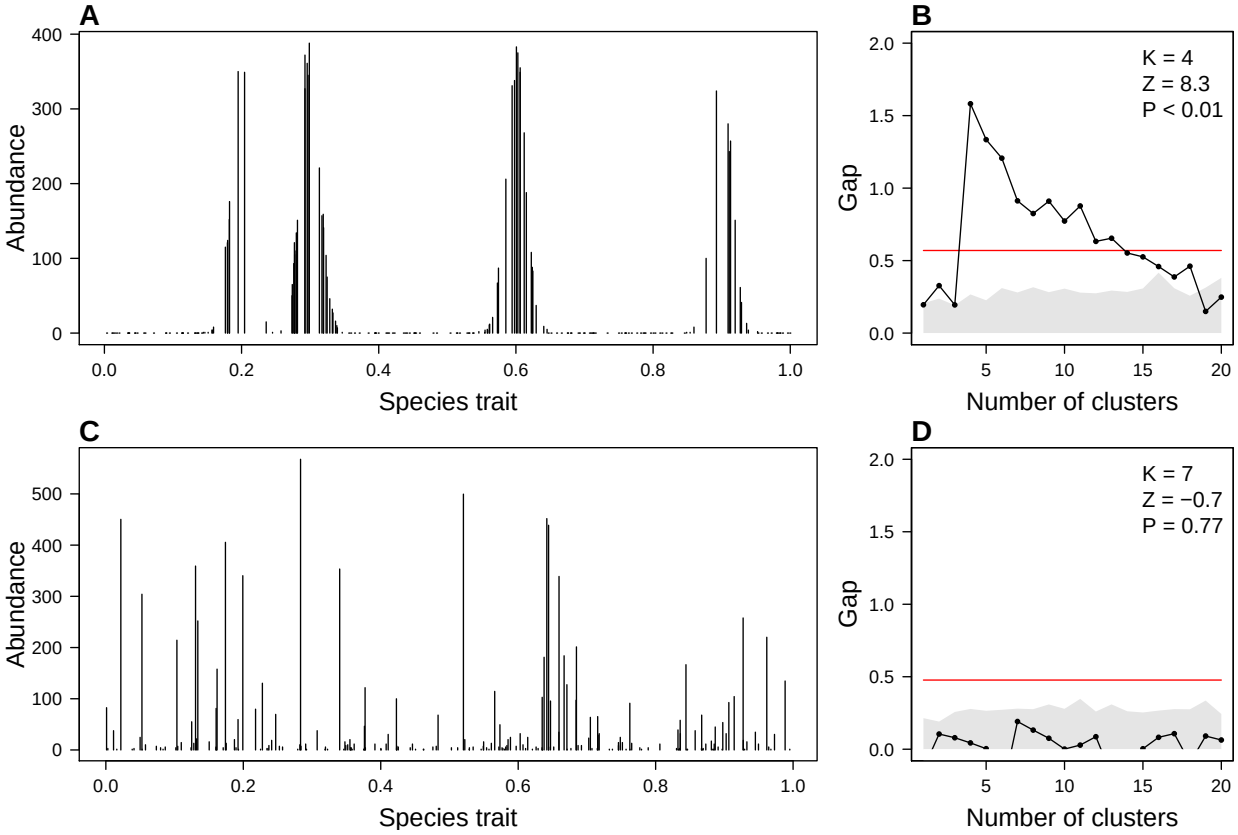
5. The best fit number of clusters is K such that $G_K = \max(G_k)$. The associated gap statistic is $G = G_K$.

6. Repeat steps 2-5 for each of the null communities. Obtain vector $\widetilde{G}_i$, with $i = 1, 2, \ldots, N$.

7. Obtain standardized effect size (z-score): $Z = (G - \mu_{\widetilde{G}})/\sigma_{\widetilde{G}}$, where $\mu_{\widetilde{G}}$ and $\sigma_{\widetilde{G}}$ are the mean and standard deviation of the $\widetilde{G}_i$. Obtain p-value: $P = \frac{1}{N}\sum_i I(\widetilde{G}_i \geq G)$, where index function I is 1 if its argument is true and 0 otherwise.

Fig. B1 shows two hypothetical test communities, and the respective results from our metric. The first community (Fig. B1A) is constructed to contain four clusters, whereas the second community (Fig. B1B) has no abundance-by-trait structure. The metric correctly diagnoses both.

Figure B1: **A**: Hypothetical species assemblage with visible clustering. There are 10,000 individuals among approximately 200 species, arranged here by trait value ranging from 0 to 1. **B**: The corresponding gap curve shows a distinct peak in the dispersion gap at $K = 4$ clusters, soaring above the 95th percentile of the gap curves of the null communities (gray band). The peak, $G \simeq 1.5$, is the gap statistic, and is well above the 95th percentile of the gap statistic across the null communities (red line). The z-score (Z) is very large, and the p-value (P) is very low. **C**: Species assemblage with same richness and size as A, but where abundances were drawn from a log-series distribution independently from species traits. No clustering is visible. **D**: Corresponding gap curve indicates that no number of clusters gives a gap beyond null expectations. The low z-score and high p-value confirm the lack of clustering in the community. [*Numerical details*: Community A was generated using a trait-dependent probability distribution given by the sum of four Gaussian curves centered at 0.2, 0.3, 0.6, 0.9 ($\sigma = 0.02$). Specifically, 200 possible trait values were randomly chosen between 0 and 1 and then 1000 individuals were drawn from each of these trait values with the probability of drawing trait value t_i being $f(x_i)/\sum_{j=1}^{200} f(x_j)$ where $f(x)$ is the sum of the aforementioned Gaussian curves. All individuals with the same trait values were taken to be of the same species. In community C we drew log-series-distributed abundances using R package "sads"[73] with parameters N=1000 and $\alpha = 50$ and assigned them to randomly chosen trait values.]



Appendices

A Detailed description of models

We simulate stochastic niche assembly using a Moran process of alternating death and birth events, with immigration from a regional pool. The fixed community size facilitates comparison of pattern across different niche scenarios, as community size will otherwise affect the visibility of trait pattern. For the same reason, we fix the size and diversity of the regional pool, and the immigration rates, unless noted otherwise. Although our study is not designed to describe any particular real-world system, we keep the 50-hectare tropical plot on Barro Colorado Island in Panama (BCI) as a reference for our parameter choices, including local and regional community size, immigration rates, and regional diversity.

Communities start as a random draw from the pool, and in each subsequent step one individual is randomly chosen to die and is replaced by a new recruit. A proportion of deaths are replaced by local offspring, and the remainder by immigrants. The particular model used dictates the probabilities that each species is selected for each death and each recruitment event. We simulate dynamics for 50×10^6 steps to allow for a stationary state to be reached, i.e. to allow for species abundance distributions averaged across a set of runs to settle to relatively stationary values. We then run for another 50×10^6 steps while keeping track of all introduction and extirpation events, for a total of 100×10^6 steps, corresponding to roughly 5,000 community turnovers. At the end of a run, trait pattern is calculated using the trait values and abundances of extant species, and each species' average persistence time is computed by calculating the average number of steps that species persisted between colonizing and going locally extinct. Unless otherwise noted for specific niche scenarios, our parameter choices were as follows: the local community has $J = 21,000$ individuals, approximately the number of trees larger than 10 cm diameter at breast height in the 50-hectare plot on BCI[36]. The regional pool, generated with R package untb[74], is a neutral metacommunity with diversity parameter $\theta = 50$ matching fits to BCI's species abundance distribution[35]. The size of the pool was set to $J_M = 150,000$ individuals, which leads to approximately $S = 410$ species. We checked that bigger sizes do not change neutral predictions for the SAD. Species trait values are randomly drawn from a uniform distribution between 0 and 1 (standard uniform distribution). Following empirical estimates for Barro Colorado Island, we set the default immigration rate m to 0.08[37].

Heuristic niche model

In our baseline scenario of the heuristic model, described in the main text and Table 1, we ensure that competition caused by species similarity is the main driver of species dynamics other than demographic stochasticity and immigration. We then consider variations of the baseline in order to verify sensitivity to our parameter choices and delineate conditions under which clustering is stronger. Each variant emphasizes a different aspect:

Immigration and regional diversity. Immigration and regional diversity are extrinsic to local niche dynamics, and hence may affect its ability to drive pattern. We test this using scenarios with increasing immigration rates and number of species in the regional pool. We use $m = 0.005$, 0.01, 0.05, 0.08, 0.15 for immigration rates, and $\theta = 5$, 10, 20, 30, 50 for regional biodiversity parameter, leading to $\langle S \rangle \approx 52$, 97, 179, 256, 401 expected species in the regional pool, respectively[75]. In all cases the number of local clusters/niches, determined by parameter w, is fixed at about 12 or 13 (the variation is due to demographic stochasticity).

Environmental filtering. Our niche mechanism does not single out a species as intrinsically more fit than any other; rather, their competitive standing is only defined in relation to their similarity to others in the community. As such, in the baseline we assume all species are equally capable of thriving when there are no competitors (r_i = constant). However, environmental filtering[76] or physiological constraints may cause some species to grow faster than others in the absence of competitors (i.e. differences in intrinsic growth rates), and this might interfere with the emergence of niche-driven clusters. For example, species with intermediate traits may grow faster because extreme trait values may have metabolic costs. We test this possibility in the Middle-trait Filtering scenario, in which we set $r_i = x_i(1 - x_i)$, where $x_i \in [0, 1]$ is the trait value of species i. Alternatively, the environment may favor species with characters unrelated to the trait responsible for niche overlap. We implement this scenario by randomly distributing intrinsic growth rates across our species, $r_i \sim U(0, 1)$.

We also look at environmental filtering on its own with no niche differentiation. In that scenario we used middle-trait filtering.

Edge effects. If competition is caused by overlap in niche strategies, then species with extreme trait values have an inherent advantage over those with intermediate trait values, as they overlap with fewer species. Some authors avoid these edge effects by imposing a circular axis. We reckon those effects are likely present in nature, so we default to a linear niche axis in our baseline. However, we test the impact of removing edge effects by comparing results against a variant scenario with a circular axis. In that scenario, we define $d_{ij} = \min(|x_i - x_j|, 1 - |x_i - x_j|)$. In words, the maximum distance on a circle is half its circumference. Because traits are defined to range between 0 and 1, the circumference is 1, and therefore the maximum distance in this scenario is 0.5. Species with extreme trait values fall close together on the circle, in contrast to the linear-axis case.

Steepness of transition between high and low competition. The competition function in our baseline scenario has a box-like shape, meaning that competition is high between species below a threshold of trait difference, and low between those above it. This relates to the idea that competition is intense between species occupying the same niche but low between species in different niches. Alternatively, the competition function can be more gradual, making for a less clear boundary between same-niche and other-niche. Here we test this scenario by changing the shape parameter from $p = 4$ (quartic) to $p = 2$ (Gaussian). We note that the exponential competition function we use is well studied in the literature[41,77,78]. The choice of $p = 2$ is particularly common, as it can be derived from arguments of resource overlap under simplified conditions[79]. However, it has been shown to have idiosyncratic properties[41,77] that do not generalize to other competition functions, which is why we set $p = 4$ in all other scenarios.

Specialization gradient. Suppose the trait in the model is beak size of seed-eating birds. Birds of similar beak size compete because they eat similar seeds. However, it is possible that small-beak birds eat a lower variety of seeds than those big-beak birds that can crack more types of seeds. This specialization gradient may affect the shape and distribution of clusters. We model this scenario via a variant where the scaling parameter w increases among species on the spectrum from lower to higher trait values. We define $w_1 = 0.01$ and $w = 0.15$ for the species with the lowest and highest trait value respectively, and every species in between have intermediate values in linear sequence. By comparison, in the Baseline scenario w is fixed at 0.063.

Noise between traits and competition. Finally, we test a scenario where the link between trait similarity and competitive interactions is imperfect or noisy because unknown factors also contribute to competition coefficients. This is arguably more realistic than the baseline case where trait similarity fully determines competitive interactions. We implement this in a variant where the scaling parameter w is a random number, w_{ij}, different for every pair of species. For each pair we draw this number from a uniform distribution from 0.015 to 0.15.

Niche differentiation by preferred resource

The Rosenzweig-MacArthur model[42] is defined by the dynamic equations

$$\begin{aligned} \frac{\mathrm{d}X_a}{\mathrm{d}t} &= r\left(1-\frac{X_a}{K}\right)X_a - \sum_i C_{ai}X_aN_i \\ \frac{\mathrm{d}N_i}{\mathrm{d}t} &= \epsilon\sum_a C_{ai}X_aN_i - \mu N_i \ , \end{aligned} \tag{1}$$

where X_a is the abundance of resource a, N_i is the abundance of consumer species i, μ is consumer mortality, identical for all species, r and K are the resources intrinsic growth rate and carrying capacity, also identical for all resources, C_{ai} quantifies the amount of resource a consumed by consumer species i, and ϵ is a conversion (efficiency) factor.

In our stochastic model we choose birth and death rates that reflect a similar process of resource-consumer interactions, and competition between consumers based on overlap in resource use. We use the same zero-sum Moran process described for our heuristic model. We assume density-independent mortality, and resource consumption affects consumer birth rates. As death rate is identical across species, the probability that species i with abundance N_i is chosen for a death event is simply $N_i/\sum_j N_j$. The probability that species i is selected for a recruitment event is $m\,q_i + (1-m)\sum_a C_{ai}X_aN_i/\sum_j\sum_a C_{aj}X_aN_j$. Resources also follow a stochastic birth-death process. Logistic growth under neutral competition between resources determines birth rates, and consumption determines death rates. Resource a is selected for a death event with probability $\sum_i C_{ai}X_aN_i/\sum_b\sum_i C_{bi}X_bN_i$, and for a birth event with probability $(1-X_a/K)X_a/\sum_b(1-X_b/K)X_b$. If a resource population is depleted to zero, the resource is extirpated and cannot be replenished.

Notice that the only parameters that affect the stochastic dynamics are the resource carrying capacity and the consumption matrix. We define $C_{ai} = C_0\,\exp[-(d_{ai}/\sigma)^2]$, where d_{ai} is the difference between resource a and the preferred resource of consumer i, σ controls how fast the consumption declines as d_{ai} increases, and C_0 is the consumption level of the preferred resource. We assume C_0 to be identical across all consumers, and therefore it cancels out of the probabilities defined above. We set $\sigma = 0.03$.

The Rosenzweig-MacArthur model simplifies to the MacArthur-Levins model if the resource dynamic equations are solved for equilibrium and the resource equilibrium abundances substituted in the consumer dynamic equations[79]. This assumes faster resource dynamics relative to consumer dynamics, and ignores the possibility of resource extinction. Abrams showed that resource depletion or exclusion has important effects on the competitive interactions between consumers[43], as well as on coexistence outcomes[44]. We therefore consider two scenarios of this model, one with $K = 100$, and the other with $K = 400$. In the first, most resources are never severely depleted by consumption, whereas in the second resource depletion is severe and many resources are extirpated. Both scenarios start with 500 resources with trait value drawn from a standard uniform distribution.

Niche differentiation by preferred abiotic environment

The third type of niche model we examined was introduced by[22], and is based on Tilman's stochastic niche model[21] and was also studied by Gravel et al.[23]. It assumes that different species are optimally adapted to different abiotic environmental conditions, and their ability to colonize different local environments declines with the difference to their optimal environment (their niche optimum). As with the stochastic models described above, Schwilk and Ackerly assumed a saturated community where every site is occupied, and each death, which occurs randomly in one of the sites, is followed by a lottery between all propagules arriving at the vacated site, some of which can be from the

regional pool. Which propagule wins the lottery is probabilistically determined by their suitability to the local environment.

A species' "trait" in this model is its niche optimum, which is drawn from a standard uniform distribution. The suitability of site a to species i is $F_{ai} = \exp[-0.5(\Delta_{ai}/0.05)^2]$, where Δ_{ai} is the difference between the local environment a and the niche optimum of species i. There are 1,000 sites arranged on a line, forming a gradient of fixed abiotic conditions between 0 and 1. An example would be an elevation gradient. Notice that here we are using a smaller community size than in the other niche models (1,000 rather than 21,000), because this model is more computationally demanding.

Dispersal limitation is implemented via a Gaussian dispersal kernel, $D_{ab} = \exp[-0.5((a-b)/\sigma)^2]$, which sets the probability that a propagule from an individual located at site b arrives at the vacated site a. We explore two dispersal scenarios of this model: local dispersal ($\sigma = 50$), where the probability of arrival of propagules to a site quickly declines with the distance to the parent, and global dispersal ($\sigma = 10^5$), where the probability of arrival is independent of distance to the parent. This corresponds to a case where the physical distances are small relative to the dispersal ability of the individuals.

The probability that vacated site a is filled by an individual of species i in each time step is then $(1-m)\sum_{b\neq a}\phi_{i,a,b}/\sum_{j=1}^{S}\sum_{c\neq a}\phi_{j,a,c} + mp_i$, where $\phi_{i,a,b} = F_{ai}\ D_{ab}$ and p_i is the regional abundance of species i.

Our model of immigration differs from Schwilk and Ackerly's. In their model, when there is an immigration event, which occurs with probability m, they add a single immigrant seed to the total seed pool, and then all seeds in the pool compete for the site. The final probability of immigrant establishment is therefore only a fraction of m, whereas in our model it is exactly m. Our default value for m is 0.08, while in their paper m ranges between 0.05 and 0.2. This means immigrant recruitment in our simulations is a much more typical event than in theirs.

Niche differentiation by life history strategy

In the competition-colonization tradeoff model[45,46], species trade off fecundity with competitive ability: the number of propagules produced trades off with the ability of propagules to displace individuals of other species. Species are ordered by fecundity value, with species 1 being the least fecund and most competitive, and species S being the most fecund and least competitive. This model as originally formulated is described by the dynamic equations[49]

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = f_i p_i (1 - \sum_{j=1}^{i} p_j) - p_i \sum_{j=1}^{i-1} f_j p_j - \mu p_i \ , \tag{2}$$

where p_i is the relative abundance of species i in the local community, f_i is its fecundity (i.e. the number of propagules produced by each individual during time interval $\mathrm{d}t$), and μ is the mortality rate, here assumed identical across all species. After some algebra, Equation (2) can be rewritten (see[51])

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = p_i \left(f_i - \mu - \sum_{j=1}^{S} \Theta_{ij}(f_i + f_j) p_j \right), \tag{3}$$

where $\Theta_{ij} = 1$ if $i > j$, $1/2$ if $i = j$, and 0 if $i < j$.[51] argued that the strict hierarchy is not realistic, because it requires that species with arbitrarily similar traits must still have large differences in competitive ability. This was shown to inflate coexistence, which can impact trait pattern. To account for this we use a probabilistic hierarchy, whereby higher-ranking species are more likely to

displace lower-ranking ones. We thus replace the abrupt function Θ_{ij} with the continuous function $G_{ij}(f_i, f_j) = 0.5\,(1 - \tanh[\nu(f_j - f_i)])$, which is equal to 0.5 when $f_i = f_j$ and asymptotes to 1 and 0 when $f_i \gg f_j$ and $f_i \ll f_j$, respectively. In the limit of $\nu \to \infty$ one recovers the step function.

Notice that Equation (3) has the same format as the Lotka-Volterra model, Equation (8). We use that correspondence in our stochastic formulation of this model. Again, we implement a saturated community under a stochastic birth-death process with immigration from a regional pool, and place the density-dependent effects on the probability of being selected for death. The probability that death befalls species i is thus $\sum_j A_{ij} N_i N_j / \sum_i \sum_j A_{ij} N_i N_j$, where $A_{ij} = G_{ij}\,(f_i + f_j)$ is the competition coefficient, and the probability that a recruit is from species i is $m\, f_i q_i / \sum_j f_j q_j + (1 - m)\, r_i N_i / \sum_j r_j N_j$, where $r_i = f_i - \mu$. Notice that immigrants are more likely to be from species with high fecundity. The logarithm of the fecundities were drawn from a uniform distribution between 0 and $\log 500$. We set mortality rate $\mu = 1$, and steepness parameter $\nu = 0.15$ in function G.

The other life-history tradeoff model we examine is Muller-Landau's tolerance-fecundity tradeoff model[47]. This model assumes that species trade off fecundity with stress tolerance, and the environment varies in stress levels. The model can be written[51]

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = \mu \sum_a h_a \frac{T_{ai} f_i N_i}{\sum_j T_{aj} f_j N_j} - \mu N_i \tag{4}$$

where f_i is species i's fecundity, T_{ai} is its tolerance to stress level a, h_a is the number of sites with stress level a, and μ is the mortality rate, common to all species. In this model, mortality is density-independent, and competition occurs at the recruitment stage.

Recruitment occurs as follows: for each recently vacated site of stress level a, all species j contribute propagules according to their fecundity f_j and abundance N_j, which are then weighted by their tolerance to stress a. A proportion $T_{ai} f_i N_i$ of those viable propagules belong to species i. The recruitment at all sites with stress a is equal to this proportion multiplied by the number of sites with that stress level, h_a. Total recruitment is then calculated by summing over all stress levels.

In our stochastic implementation of this model, which follows the same zero-sum recipe used in the other models, species are selected to die in proportion to their abundances, and the probability that species i is selected for a recruitment event is $m\, f_i q_i / \sum_j f_j q_j + (1 - m)\, I_i / \sum_j I_j$, where $I_i = \sum_a h_a (T_{ai} f_i N_i / \sum_j T_{aj} f_j N_j)$. There are 21,000 sites of 400 different stress levels in the local community. The number of sites of each stress level is uniformly distributed. The stress levels are drawn from a standard uniform distribution, as are the fecundity values (note that multiplicative factors in the fecundity values cancel out of the stochastic dynamics). The tolerance function we used was $T_{ai}(s_a, f_i) = 0.5\,(1 - \tanh[100(s_a + f_i - 1)])$, where s_a is the stress level with index a. It has a sigmoidal shape, ranges from 0 to 1, and monotonically decreases with stress level and fecundity value (In the original formulation by Muller-Landau, T_{ai} was a step function, but we follow our own advice[51] to avoid assuming that differences in tolerance strategy can remain large between arbitrarily similar species—an unrealistic assumption which has strong effects on coexistence levels and presumably on pattern).

Neutral and Pure filtering scenarios

In addition to our stochastic niche models we also tested two models without any kind of niche mechanism to make sure our clustering metric was not giving us false positives. In both, competition was uniform between all individuals irrespective of species identities or traits—neutral competition. This was done by assigning every element in the competition matrix to 1. In the Neutral scenario, all species have identical intrinsic growth rates. In the Pure filtering scenario, we model environmental

filtering on top of neutral competition by implementing differences in intrinsic growth rates, in the same form as in the Middle-trait filtering niche scenario, $r_i = x_i(1 - x_i)$.

Deterministic immigration-free simulations

We also simulated the deterministic immigration-free models defined by Equations (8, 1, 3, 4), in order to compare and contrast the outcomes with our simulations of stochastic niche assembly under immigration.[1]

We start with all of the approximately 400 species at their regional abundances, totaling 150,000 individuals. Using the Euler method, we then update abundances until equilibrium is reached. We set the time interval so that on average each update corresponds to one death and/or one birth at the beginning of the simulation. We observed that 2 to 10 million updates sufficed to reach equilibrium, depending on the model. We checked for equilibrium by verifying that all abundances changed by less than one individual across the most recent 100,000 updates. Simulation outcomes are shown in Fig. S7.

[1]Except for the Schwilk-Ackerly model of competition for suitable habitat, which is mathematically similar to the MacArthur-Levins model[34] with Gaussian competition function. The rationale here is the same behind the demonstration that the Rosenzweig-MacArthur[42] model, in the limit of fast resource dynamics, reduces to the MacArthur-Levins model: the overlap between two Gaussian resource use curves yields a Gaussian competition curve[79].

B Moran versus Gillespie implementation of niche models

Our stochastic lottery model, essentially a Moran-style process, was defined by the probabilities of species i being chosen for a death and a recruitment event, $\sum_j A_{ij}N_iN_j/\sum_i\sum_j A_{ij}N_iN_j$ and $r_iN_i/\sum_j r_jN_j$, respectively, where here we are considering the closed community case with no immigration, $m = 0$. Writing these generally as $d_i(\{N\})N_i/\sum_j d_j(\{N\})N_j$ and $b_i(\{N\})N_i/\sum_j b_j(\{N\})N_j$, the expected change in abundance of species i given its current abundance $N_i(t) = N_i$ is written

$$\begin{aligned} & \mathbb{E}[N_i(t+\Delta t) - N_i(t)|N(t) = N_i] \\ = & \left[\left(1 - \frac{d_iN_i}{D}\right)\frac{b_iN_i}{B} - \frac{d_iN_i}{D}\left(1 - \frac{b_iN_i}{B}\right)\right]\Delta t \\ = & \left(\frac{b_iN_i}{B} - \frac{d_iN_i}{D}\right)\Delta t \end{aligned} \tag{5}$$

where $B = \sum_j b_jN_j$ and $D = \sum_j d_jN_j$ are the total birth and death rates, and the interval between events Δt is Poisson-distributed. This is a probabilistic model where the expected change per event is the weighted sum of possible gains and losses, weighted by the probability of their occurrence. The affiliated differential equation is

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} \approx \frac{N_i(t+\Delta t) - N_i(t)}{\Delta t} = \frac{b_iN_i}{B} - \frac{d_iN_i}{D} \tag{6}$$

Substituting our model gives

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = N_i\left(\frac{r_i}{\sum_j r_jN_j} - \frac{\sum_j A_{ij}N_j}{\sum_k\sum_l A_{kl}N_kN_l}\right) \tag{7}$$

Therefore it appears that our zero-sum process does not correspond to the Lotka-Volterra model studied by MacArthur and Levins[34],

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = N_i\left(r_i - \sum_j A_{ij}N_j\right) \tag{8}$$

whose stochastic version could be implemented using a Gillespie algorithm with $d_i = \sum_j A_{ij}N_j$ and $b_i = r_i$.

However, we verified with simulations that the trait patterns—and indeed the dynamics—resulting from Eqns. (7) and (8) are essentially identical within the parameter space we tested. This suggests that these models actually represent the same dynamics in general, although we have not been able to prove it analytically.

C Detailed description of clustering metric

Here we describe our metric for detecting and quantifying clustering in species assemblages.

Schematically, the test works as follows: for each parameter value θ within a provided range $[\theta_{\min}, \theta_{\max}]$, we calculate the "clustering measure" $F(\theta)$ (see concrete examples below) in our species assemblage, as well as in each of the null assemblages $\widetilde{F}_i(\theta)$, $i = 1, 2, \ldots, N$, where N is the total number of null assemblages. We then take the difference (the gap) between the former and the average of the latter: $G(\theta) = F(\theta) - \frac{1}{N}\sum_i \widetilde{F}_i(\theta)$. The gap statistic is the extremum of the gap function, $G = \underset{\theta}{\text{extr}}\, G(\theta)$. If the parameter θ is the number of clusters, as is the case when using k-means dispersion, the value $\hat{\theta}$ at which the gap is maximal is the estimated number of clusters in the data.

If we knew the expected distribution of this statistic under the null hypothesis where species traits and abundances are unrelated to each other, we could stop here, as the value of G would tell us both the degree of clustering and the statistical significance. Because we do not know that distribution, we must compare it with a null set of values. We therefore calculate the metric on each of our null assemblages, and from that we obtain a standardized effect size (z-score) and a p-value. Those are defined as $Z = (G - \mu_{\widetilde{G}}) \,/\, \sigma_{\widetilde{G}}$ and $P = \frac{1}{N}\sum_i I(\widetilde{G}_i \geq G)$ where $\widetilde{G}_i$ is the gap statistic obtained for the i-th null assemblage, $\mu_{\widetilde{G}}$ and $\sigma_{\widetilde{G}}$ are the mean and standard deviation of those values, and index function I is 1 if its argument is true and 0 otherwise.

We note that 1) We can use any number of traits to describe our species, as long as we can define a "distance" between species (e.g. Euclidean distance in a high-dimensional trait space, or simple trait differences on a single trait axis); 2) The gap method can work with different dispersion measures. Here we use two such measures: k-means dispersion[54] and Ripley's K function[55]. The first provides a direct estimate of the number of clusters, whereas the second focuses on distance scales on the trait axis.

Gap statistic via k-means

Here the parameter θ is the candidate number of clusters k, and we use the k-means clustering algorithm[54]. For a given k, the algorithm finds the partition of individuals into k groups that minimizes the within-group dispersion. Let D_r be the total pairwise squared distance between members of group r, that is, $D_r = \sum_{i,j}(x_i - x_j)^2$, where x_i is the trait value of individual i. The dispersion W_k is then the sum of those numbers across all k groups, $W_k = \sum_{r=1}^{k} D_r$. The k-means algorithm assigns each individual to one of the k groups so that W_k is as low as possible[2,3]. We set $F(\theta) = F_k = \log(1/W_k)$ as the clustering measure[4]. See Box 1 for a step-by-step recipe for the metric.

[2]Dispersion can be defined in different ways, depending on how one wishes to weigh the number of members per group. While[53] choose to scale the squared distances by the number of individuals in the group,[80] argue that scaling it by the number of pairs in the group is a a better choice in some cases. In our study we chose not to scale the square distances, as that choice proved more robust to environmental filtering—see subsection below.

[3]In our model, all conspecific individuals have the same trait value, and therefore necessarily belong in the same cluster. Species are therefore trivial clusters of individuals. Ecologists are of course interested in clustering beyond the obvious grouping of individuals into species.

[4]Note that it does not make sense to compare F_k directly across different k's because F_k necessarily increases with k, as the average within-cluster distance is always lower for higher numbers of clusters. By comparing against null assemblages, the gap method finds the biggest increase in goodness of fit *beyond* what is expected from the increase in k. The reason we use $\log(1/W_k)$ rather than simply $1/W_k$ is that the expected increments in $1/W_k$ with increasing k are multiplicative rather than additive (see[53] for more details).

Weights and the impact of environmental filtering on the gap statistic

Here we defined k-means dispersion as the total within-cluster sum of squared distances between individuals, $W_k = \sum_{r=1}^{k} D_r$, where $D_r = \sum_{i,j \in C_r} d_{ij}^2$, and the sum is over all pairs of individuals i, j in the r-th cluster C_r. In the literature, dispersion has been defined with additional weights.[53] define $W_k = \sum_{r=1}^{k} \frac{D_r}{2n_r}$, and[80] use $W_k = \sum_{r=1}^{k} \frac{D_r}{2n_r(n_r-1)}$, where n_r is the number of individuals belonging to cluster C_r. Neither study offers an a priori justification for their respective choice of weights, and we do not know of any theoretical reason to favor either choice over the other, or against no weighting, as done here. However, we have observed that these weight choices affect metric performance when there is an overarching abundance trend along the trait axis.

In nature such an abundance trend could be caused by environmental filtering. Local conditions could, for example, favor species with intermediate trait values. This selection would then affect species growth in addition to and separate from competitive interactions with other species, leading to overall higher abundance to habitat-selected species. In the main text we considered two niche scenarios with environmental filtering—purifying filtering and random filtering—and found multiple clusters in both. However, as filtering is intensified (Fig. S1), the ability of the metric to detect multiple clusters is eventually compromised, and the weight used to define the dispersion index affects this metric sensitivity (Fig. S2).

The top left panel of Fig. S1 shows the distribution of intrinsic growth rate r across species in each of the niche scenarios examined here. We include five scenarios with progressively stronger filtering, from a flat distribution of r to a narrowly peaked distribution where species near the center of the trait axis have a strong advantage over species near the edge. The other panels of Fig. S1 show example outcomes of each scenario, with matching colors. The increasing differences in r are reflected in gradually stronger overarching abundance trends.

The left panel of Fig. S2 shows the percentage of runs, out of 10 in each niche scenario, where more than one cluster was detected, according to each weighting system. The right panel shows the percentage of significant runs where a single cluster was found. Under all three weighting choices, the percentage of multiple-cluster detection decreases with the degree of environmental filtering, accompanied by an increase in the percentage of single-cluster detection.

These trends reflect the fact that as filtering gets stronger, species gradually start forming one big cluster centered on the trait value selected by the environment. When this environmental filtering is the dominant ecological force, the community has a single niche. The metric is not failing as much as it is starting to identify that single cluster as the most prominent feature in the community. However, even in cases with obvious environmental filtering, substructure within the big cluster may reveal that niche differentiation is also structuring the community. From this perspective, the weight choice in[53] outperforms the weight choice in[80], and not weighting outperforms both. Given these results, we chose the latter option in the main text.

Gap statistic via Ripley's K

In addition to k-means, we also introduce another method to measure clustering, this one based on Ripley's K. The approach here is to count the number of pairs of individuals within each distance in trait space and determine whether or not it deviates significantly from what we would expect if the same set of abundances were randomly distributed among the present species. In particular, we look for significant open spaces between clusters, i.e. distances at which Ripley's K is surprisingly low. The parameter θ here is therefore the trait distance d.

Ripley's K function at distance d is defined as $K(d) = \frac{\sum_{i \neq j} I(d_{ij}<d) N_i N_j}{\sum_{i \neq j} N_i N_j}$, where N_i is the abundance of species i and d_{ij} is the distance in trait space between species i and j. We then

calculate Ripley's K in each of our null assemblages, and define the Ripley's K clustering measure as the standardized K function: $F(d) = K(d)/\sigma_{\widetilde{K}(d)}$, where $\sigma_{\widetilde{K}(d)}$ is the standard deviation of the null values. The gap function is then $G(d) = K(d)/\sigma_{\widetilde{K}(d)} - \mu_{\widetilde{K}(d)}/\sigma_{\widetilde{K}(d)}$. In Fig. S3, we plot the cumulative count of pairs versus distance for our data and 1,000 reshuffled runs, and in Fig. S4 we plot $G(d)$.

Because we are interested in significantly low counts, we define our Ripley gap statistic as the minimum value of $G(d)$ across candidate distances, $G = \min_d G(d)$. This is analogous to our implementation with the k-means algorithm (see Table S1 for a direct comparison between the two implementations).

When the observed assemblage has significant open spaces (Fig. S5A), i.e. a lower G than the null communities, it suggests multiple clusters in trait space of high abundance species, with a dearth of species in between them. This is what we would expect in niche structure. When the observed assemblage is significantly lacking in gaps (Fig. S5B), it indicates a single clump of abundant species, as could be expected when the environment filters for a single trait value.

Table S1: Comparison of our two implementations of the gap method. Both apply the gap method using different measures of dispersion. $W(k)$ is the k-means dispersion for k clusters; $K(d)$ is Ripley's K at trait distance d; $F(k)$ ($F(d)$) is the dispersion for k clusters (distance d); I is the indicator function, equal to 1 if its argument is true and zero otherwise; N_i is the abundance of species i; $G(k)$ ($G(d)$) is the gap function for k clusters (distance d); $\mu_{\widetilde{X}}$ and $\sigma_{\widetilde{X}}$ are the mean and standard deviation of quantity X taken across the null communities; $\hat{d}$ is the distance at which $G(d) = G$. N is the number of null communities.

	k-means	**Ripley's K**
Clustering measure	$F_k = \log(1/W(k))$	$F(d) = K(d)/\widetilde{\sigma}_K(d)$
Gap function	$G_k = F_k - \mu_{\widetilde{F}_k}$	$G(d) = F(d) - \mu_{\widetilde{F}(d)}$
Gap statistic	$G = \max_k G_k$	$G = \min_d G(d)$
Estimated number of clusters	k such that $G_k = G$	Number of abundance-peaks within distance $\hat{d}$ of each other
Z-score	$Z = (G - \mu_{\widetilde{G}}) \,/\, \sigma_{\widetilde{G}}$	$Z = (G - \mu_{\widetilde{G}}) \,/\, \sigma_{\widetilde{G}}$
P-value	$P = \frac{1}{N}\sum_i I(\widetilde{G}_i \geq G)$	$P = \frac{1}{N}\sum_i I(\widetilde{G}_i \geq G)$

D Functional dispersion and even spacing

We tested our niche scenarios using three metrics of niche differentiation found in the literature:

Rao's quadratic index[9]: The expected trait distance between two individuals randomly sampled with replacement from the community. A high Rao index signifies functional overdispersion, i.e. randomly sampled individuals tend to be more different than expected by chance. Mathematically defined as $Q = \sum d_{ij} p_i p_j$, where p_i and p_j are the relative abundances of species i and j, and d_{ij} is the trait distance (absolute value of trait difference) between them, and the sum is over all species pairs.

Functional dispersion[10]: the mean trait difference between each individual in the community and the average individual. Analogous to Rao, with a high value indicating trait overdispersion. $\text{FDis} = \sum_i |x_i - \bar{x}| p_i$, where x_i is the trait value of species i, and $\bar{x}$ is the trait value of the average individual.

Coefficient of variation of distances between adjacent neighbors on the trait axis: tests for even spacing[65]. Given a set of species sorted by trait value, it is defined as $\text{CV} = \sqrt{\text{Var}[d_i]}/E[d_i]$, where d_i is the trait distance between species i and species $i+1$. The CV has also been used to look for clustering, where it is expected to be high[81]. A related metric uses the trait range in the denominator instead of the mean[12,72].

Results for the Rao index and CV are shown in Fig. 6. The functional dispersion index had similar but weaker results than Rao's index, to which it is related, and is not shown. Species in our niche scenarios do not show a general tendency to overdisperse in trait space in the sense of Rao's quadratic index. Within our heuristic niche model scenarios, the only ones with at least one significant run were the Trait-competition noise, Gaussian competition, and variants of the Baseline differing by regional diversity and immigration rate. Across our models with explicit niche mechanisms, only the resource-consumer model had significant runs. The average z-score was low across the board, and the only case where significance occurred in more than half the runs was Gaussian competition in the heuristic model. Edge effects were strongest in that niche scenario due to the slow niche sorting (compare with Gaussian competition on a circular axis, Fig. 6), which probably raised the Rao index above the other scenarios.

Not only did our niche scenarios largely fail to show significantly high Rao and functional dispersal indices, some were in fact significantly low (Fig. 6). This was the case of our tradeoff scenarios and the heuristic model with purifying filtering, where intermediate-trait species have a fitness advantage. These niche scenarios had a higher Rao index than the pure filtering case with fitness asymmetries and neutral species interactions, but lower Rao index than the fully neutral communities. The Rao index has been used to show minimization of niche overlap in marine phytoplankton[27], but we note that in that study the number of niches in the community seemed low. Our results suggest that Rao may not be suitable for a higher number of niches.

Species are also not evenly dispersed along the trait axis, contrary to the common expectation that has motivated metrics of even spacing such as the variance in distances between nearest neighbors or adjacent neighbors on the trait axis[12,72]. We tested the coefficient of variation (CV) in distances between adjacent neighbors on the trait axis, and found that it is often indistinguishable from random or neutral communities (Fig. 6). None of our niche scenarios had high z-scores nor a high percentage of significance under this metric. Metrics of even spacing that do not account for abundances are severely limited in the context of open communities, as they cannot distinguish between dominant species and rare species maintained by mass effects.

We conclude that functional dispersion and even spacing are often poor indicators of niche differentiation because they are prone to false negatives. Functional dispersion is sensitive to

differences in intrinsic growth rates and other types of filtering that may counter the effects of niche differentiation, and even spacing is more suitable for closed communities.

Supplementary figures

Figure S1: Heuristic model scenarios with differences in intrinsic growth rates. Top left panel: intrinsic growth rate as a function of species trait used in scenarios A-E. Remaining panels: example simulation outcomes. Progressively steeper abundance trends reflect their respective intrinsic growth distributions. In the flat r case (A), all species have the same r. In B, they are defined as $r_i = x_i(1 - x_i)$, where $x_i \in [0, 1]$ is the trait value of species i. In C through E, we define $r_i = \exp(-(x_i - 0.5)^2/\sigma^2)$, with $\sigma = 0.5$, 0.2, 0.1 for cases C, D, and E, respectively. Note that the the r_i are defined only up to a constant factor, as only the relative values affect the dynamics.

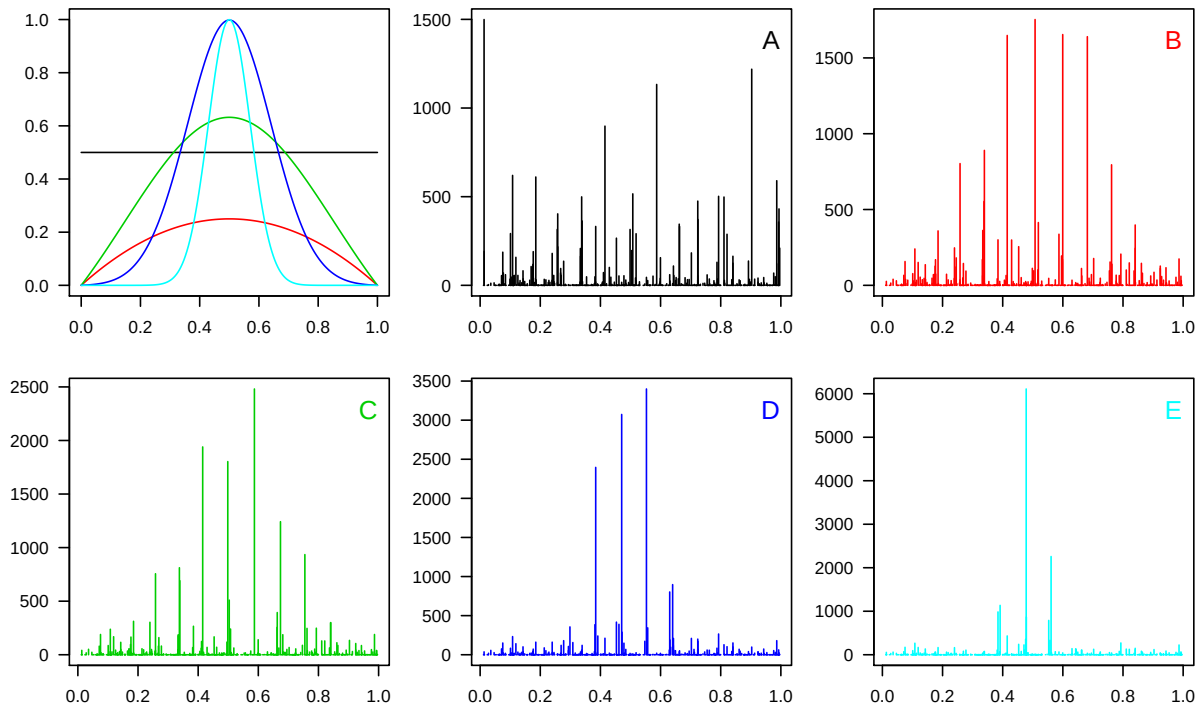


Figure S2: Percentage of runs where more than one cluster was detected (left) and percentage where a single cluster was detected (right) follow opposite trends across scenarios with gradually stronger habitat filtering. The definitions of the dispersion index used by[53] and[81] are more sensitive to the overarching abundance trend caused by environmental filtering, compared with the definition used in this paper.

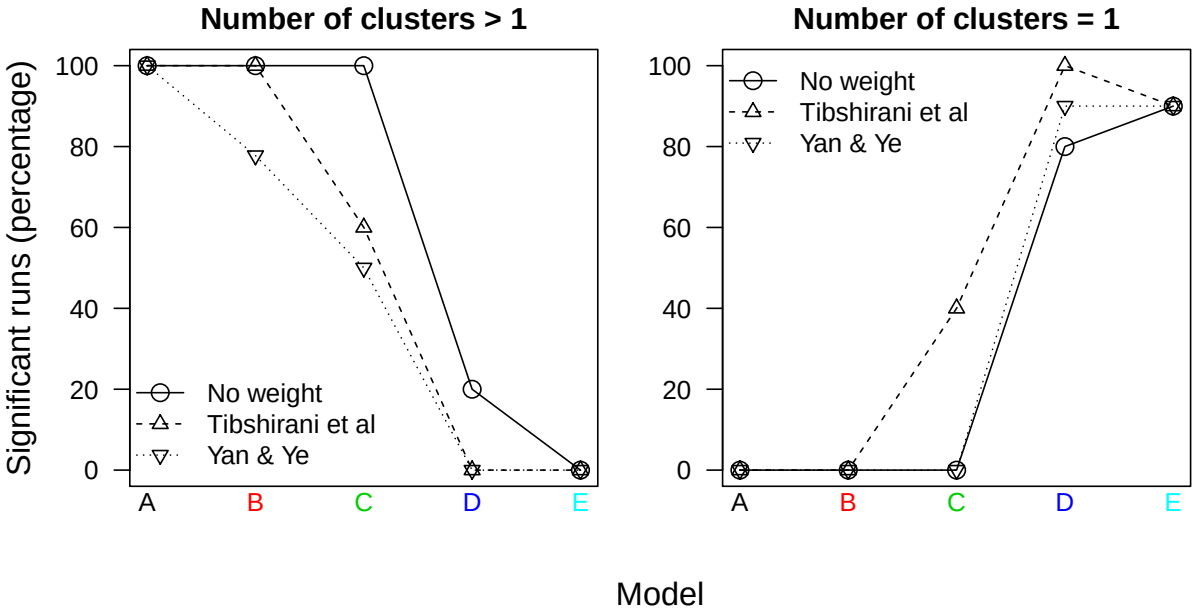


Figure S3: Top: Abundances plotted against trait values in three sample runs using (A) the baseline heuristic niche model, (B) environmental filtering, (C) a neutral model. **Bottom**: Ripley's K vs. distance for the observed and shuffled abundances in the same runs: (D) the heuristic niche model, (E) environmental filtering, (F) a neutral model. Observed abundances are plotted in black. The %95 confidence interval of $K(d)$ among reshuffled runs is shaded in gray.

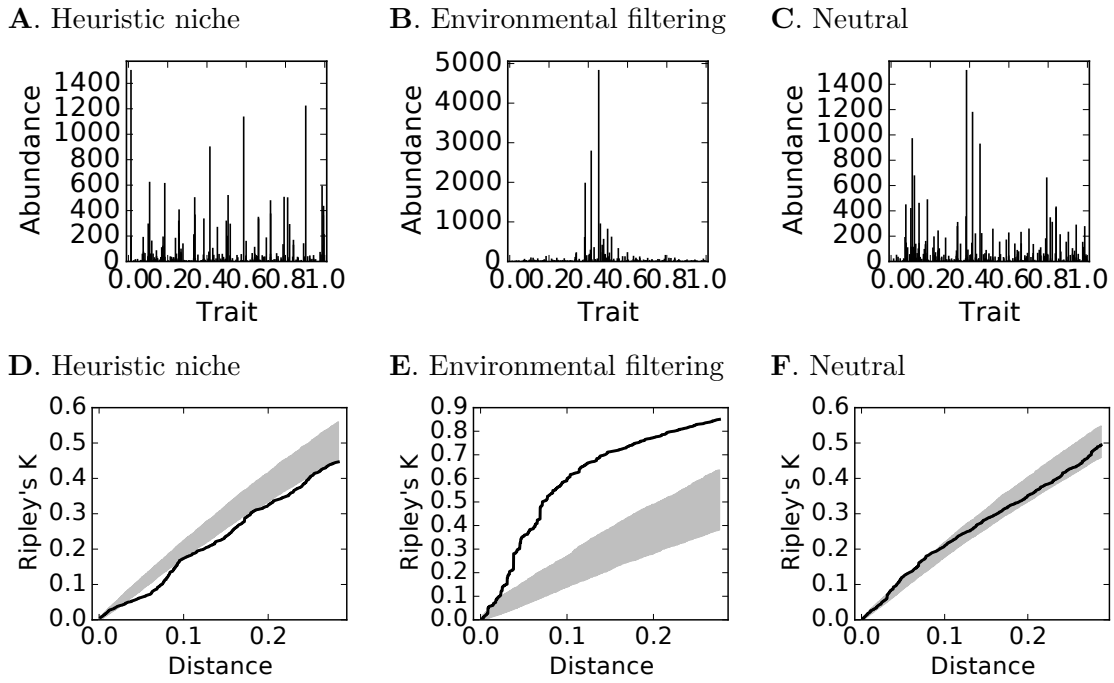


Figure S4: Cumulative pair counts vs. trait distance for the observed and shuffled assemblages. (A1-C1) with mean null value at each distance subtracted out $K(d) - \widetilde{\mu}_K(d)$; (A2-C2) rescaled by the variance at each distance $G(d) = K(d)/\widetilde{\sigma}_K(d) - \widetilde{\mu}_K(d)/\widetilde{\sigma}_K(d)$. Observed values are plotted in black, the %95 CI of reshuffled runs is shaded grey. The red line is the threshold for a significantly low gap statistic ($p < 0.05$), and the triangle indicates the distance with the most surprisingly low density of pairs, i.e. the distance d where the Ripley gap statistic G is achieved, $G(d) = G$.

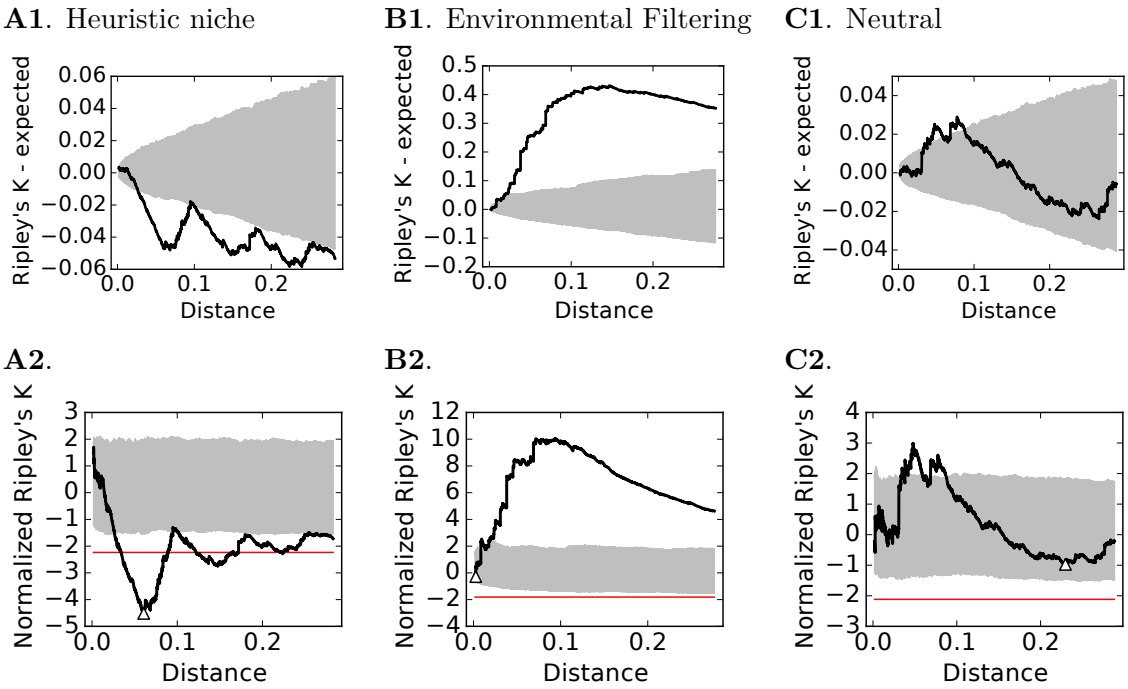


Figure S5: The distribution of G in 1,000 reshuffles of the sample runs of (A) the heuristic niche model, (B) an environmental filtering model, (C) a neutral model. The black vertical line indicates the value of G in the observed data. The red lines indicate the bounds of the %95 confidence interval.

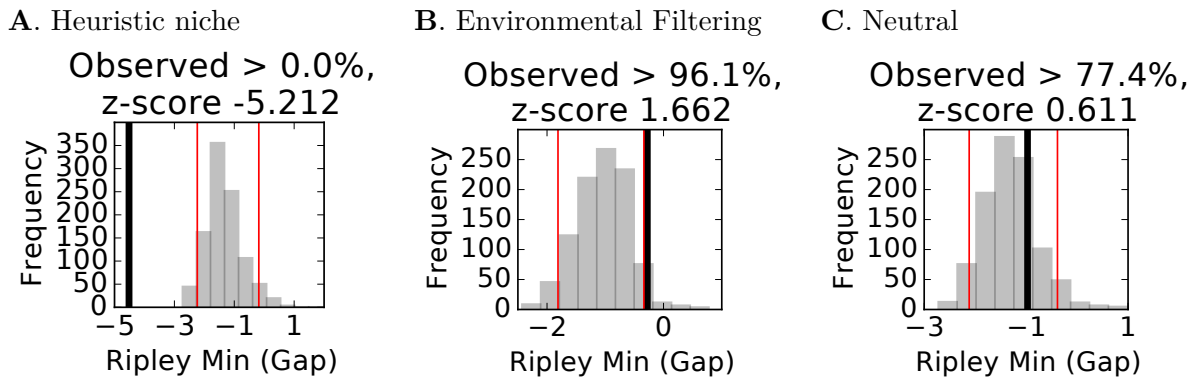


Figure S6: Abundances plotted against traits, with colors indicating the clusters detected using the Ripley gap statistic in the sample run of the heuristic niche model shown in Fig. S3.

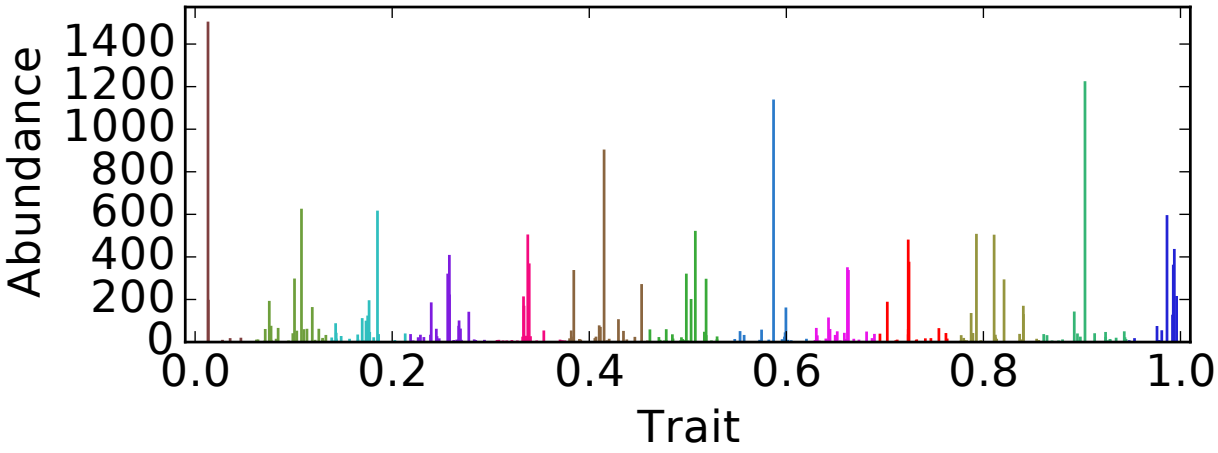


Figure S7: Simulation outcomes of deterministic immigration-free niche models (Eq. 8, 1, 3, 4), with parameter choices matching scenarios of our stochastic niche scenarios.

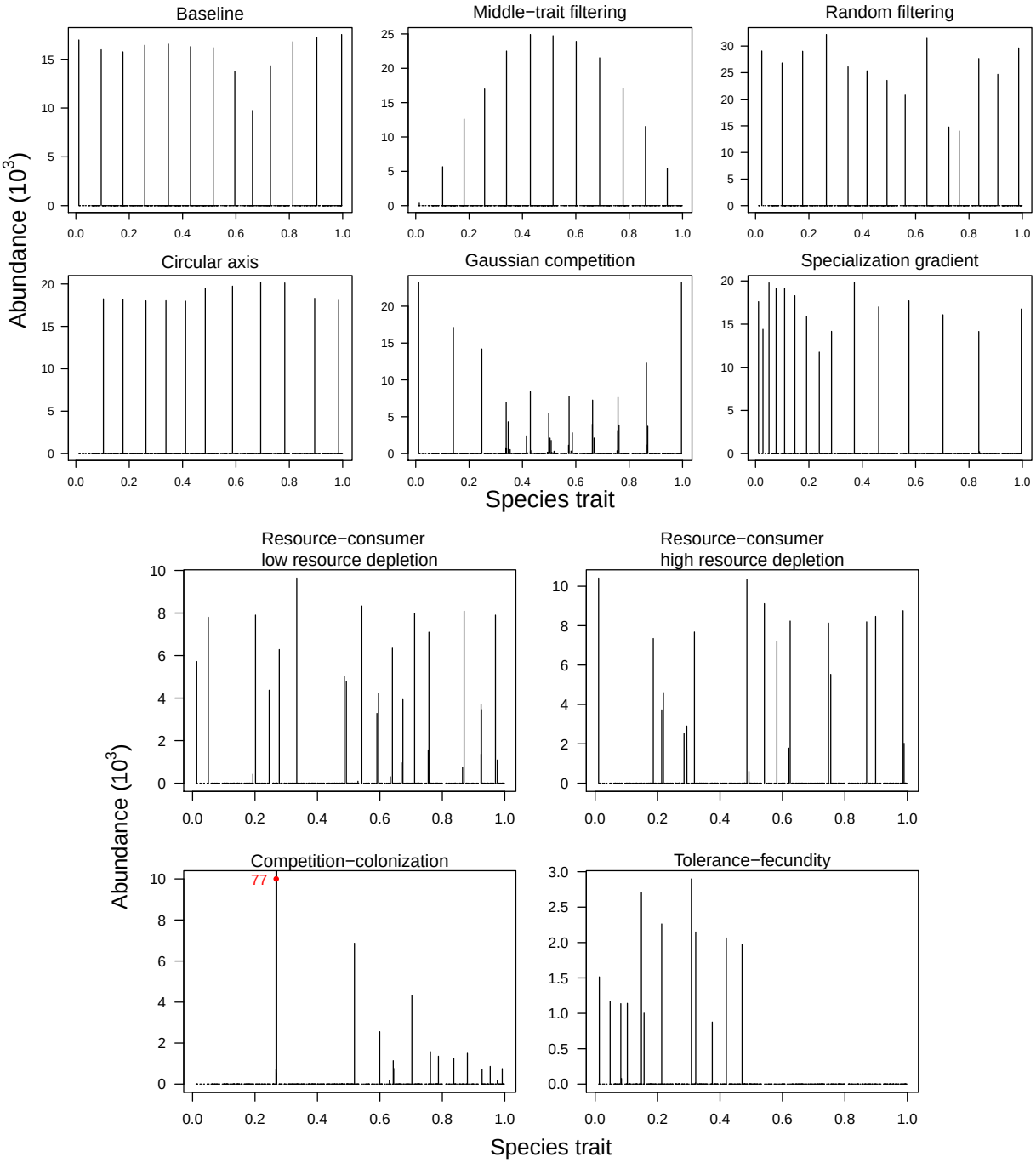


Figure S8: Number of clusters across models and scenarios, as measured by the k-means gap statistic. Squares and error bars show mean and standard deviation across runs of each scenario. We searched between 1 and 20 clusters, knowing that the deterministic immigration-free niche models under same parameters typically led to between 12 and 15 coexisting species (Fig. S7). Neutral scenario had a mean of 10 clusters and large variance, while the pure filtering scenario consistently showed a single cluster. By contrast, most niche scenarios had about the target number of clusters, with much less variation across runs. Scenarios of the heuristic model had 10 runs each, and models with explicit niche mechanism had 100 runs each.

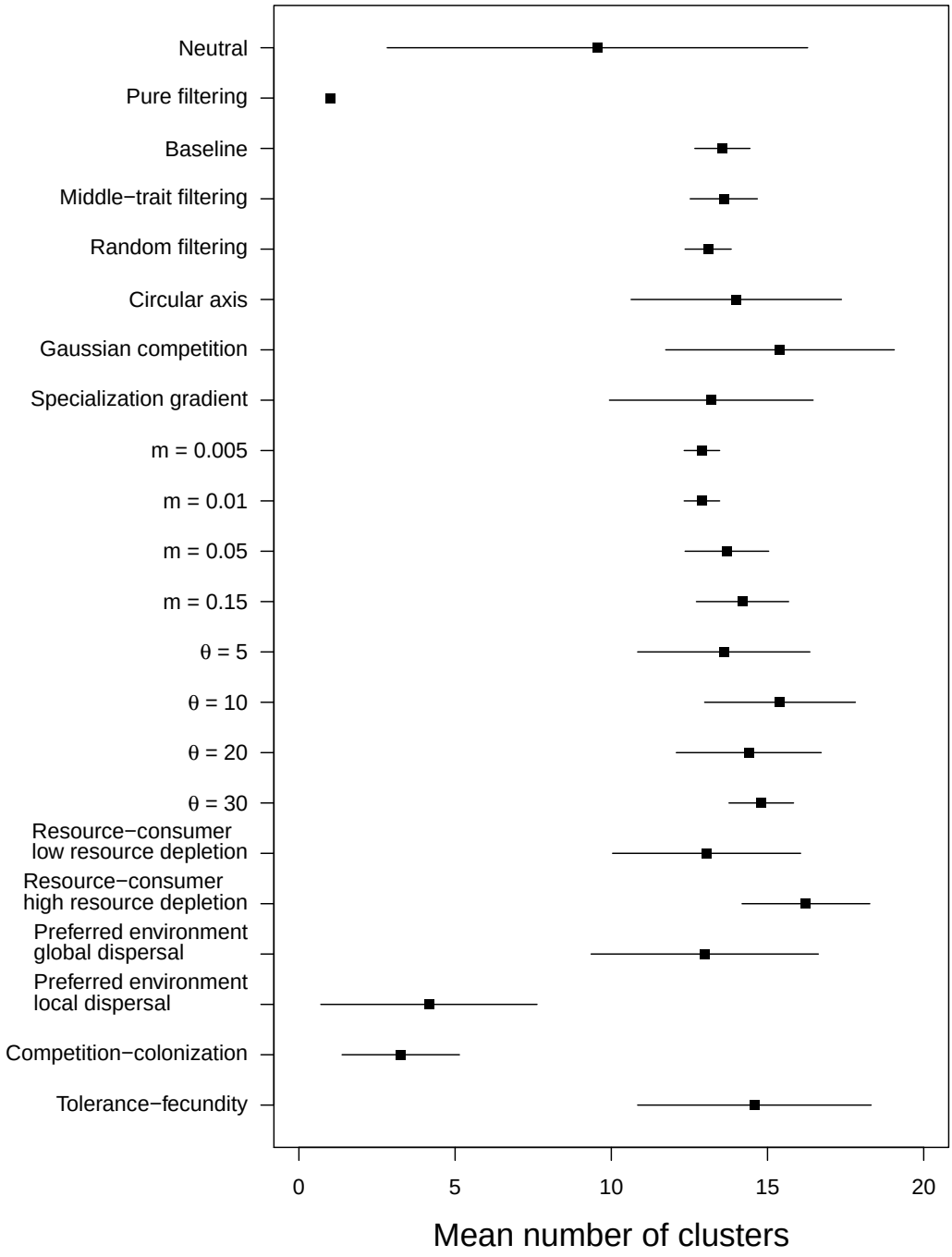


Figure S9: Example simulation outcomes of the baseline scenario of the heuristic niche model with increasing immigration rate m (left) and regional diversity parameter θ (right).

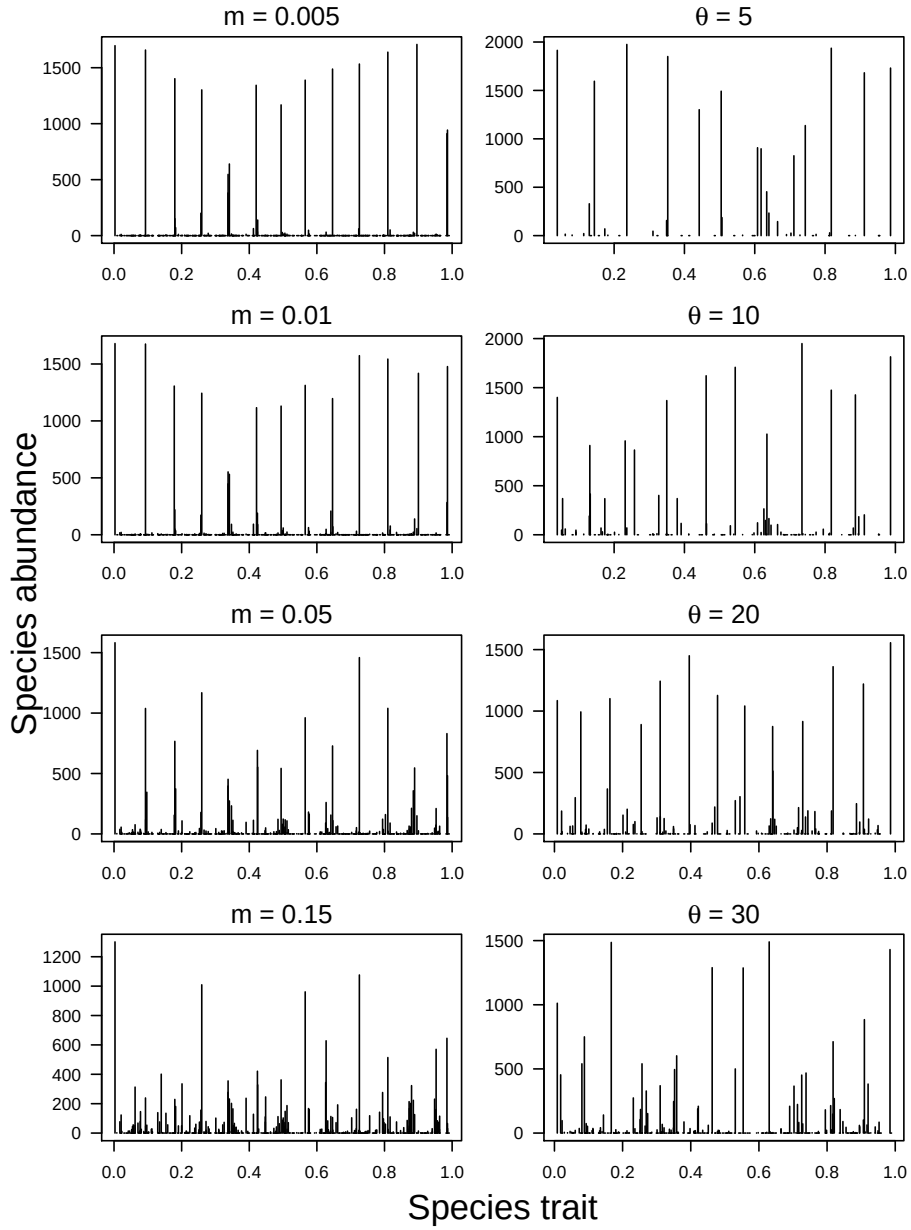


Figure S10: Example simulation outcomes of variants of our heuristic stochastic niche model.

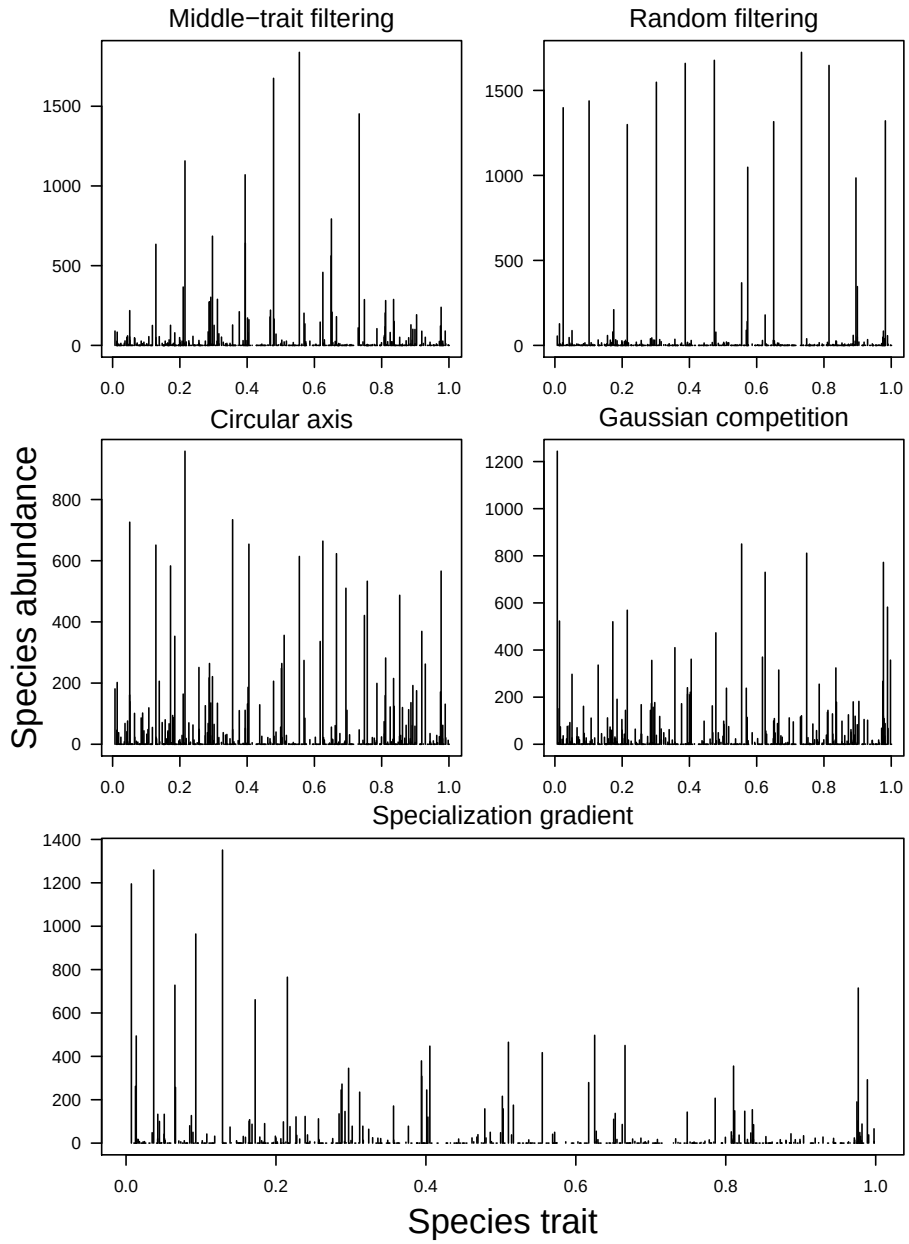


Figure S11: Comparison of rate of exclusion between heuristic model with $p = 4$ (our baseline scenario) and $p = 2$ (Gaussian competition). Both simulations start and end with the same number of species, but the sorting is much faster in the $p = 4$ case.

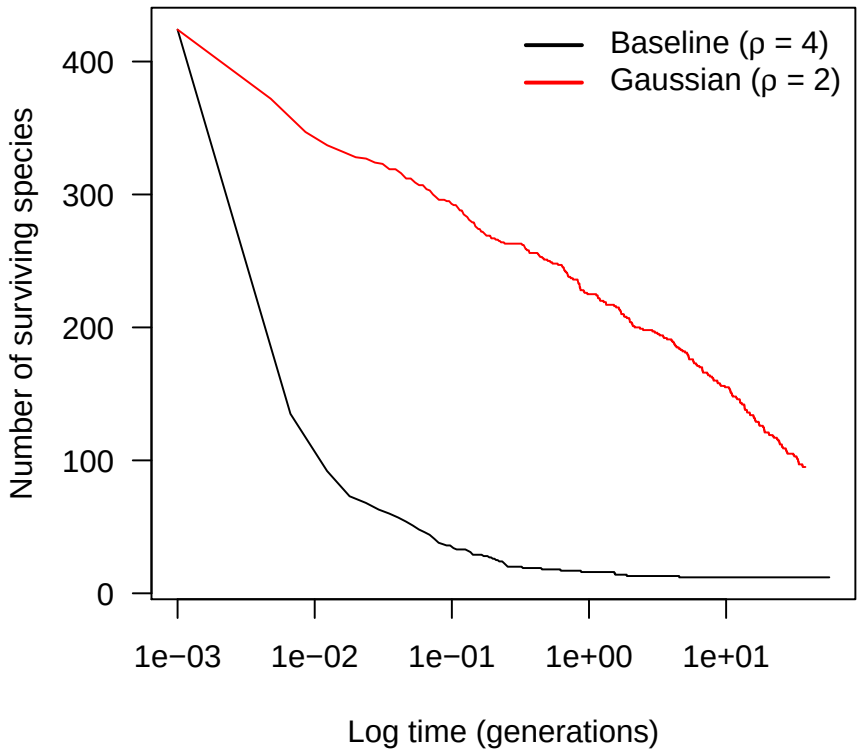


Figure S12: Example simulation outcomes of the resource-consumer model for low and high resource depletion (left and right columns, respectively). Resources are shown on top, consumers at the bottom. In the high resource depletion scenario, gaps left by resource extirpation cause corresponding gaps among consumer species. These gaps potentially explain why clustering was stronger in this scenario than under low depletion.

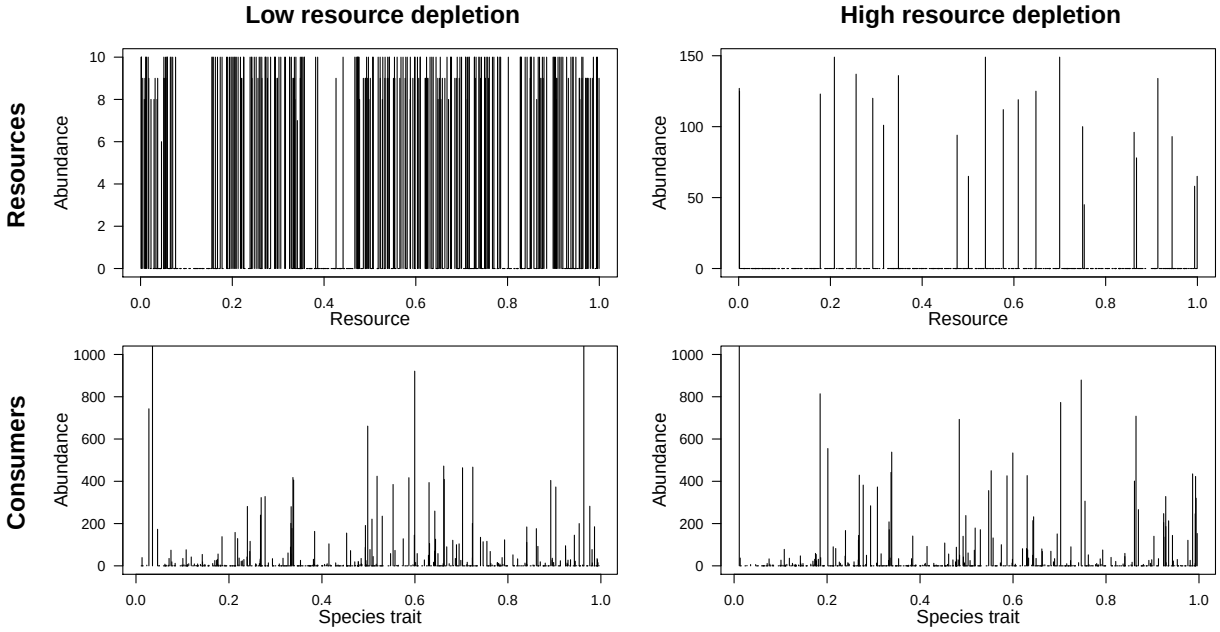


Figure S13: Clustering results for models with explicit niche mechanisms under immigration rate $m = 0.08$. Results are similar as with a lower immigration rate, although z-scores and significance are generally lower (compare against results with $m = 0.01$ in Fig. 3D in main text). Relative z-scores indicate stronger clustering when resource depletion is high in the model with differentiation for resources, and when dispersal is limited in the model with differentiation for environmental conditions. Ripley's K measure (orange) is better at picking up the pattern than the k-means measure (blue) except for the competition-colonization tradeoff model. It should be noted that in this model the estimated number of clusters is very low, at only 2 clusters.

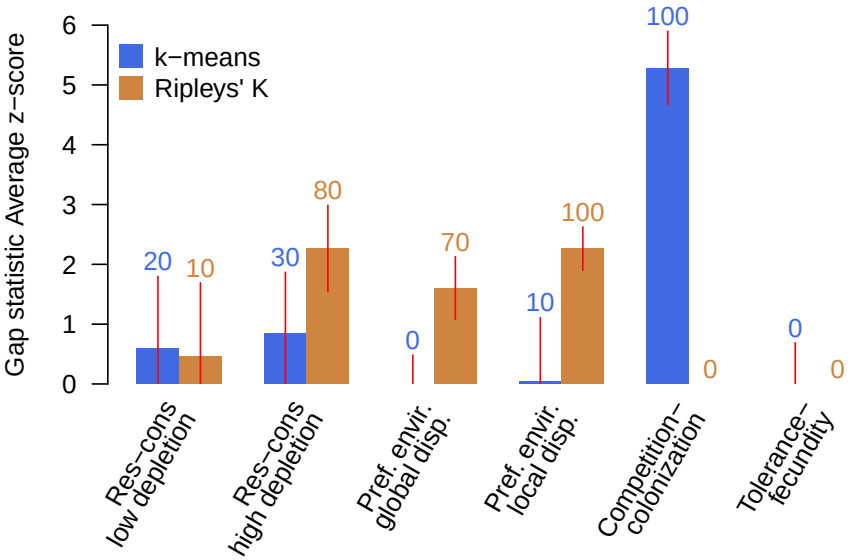
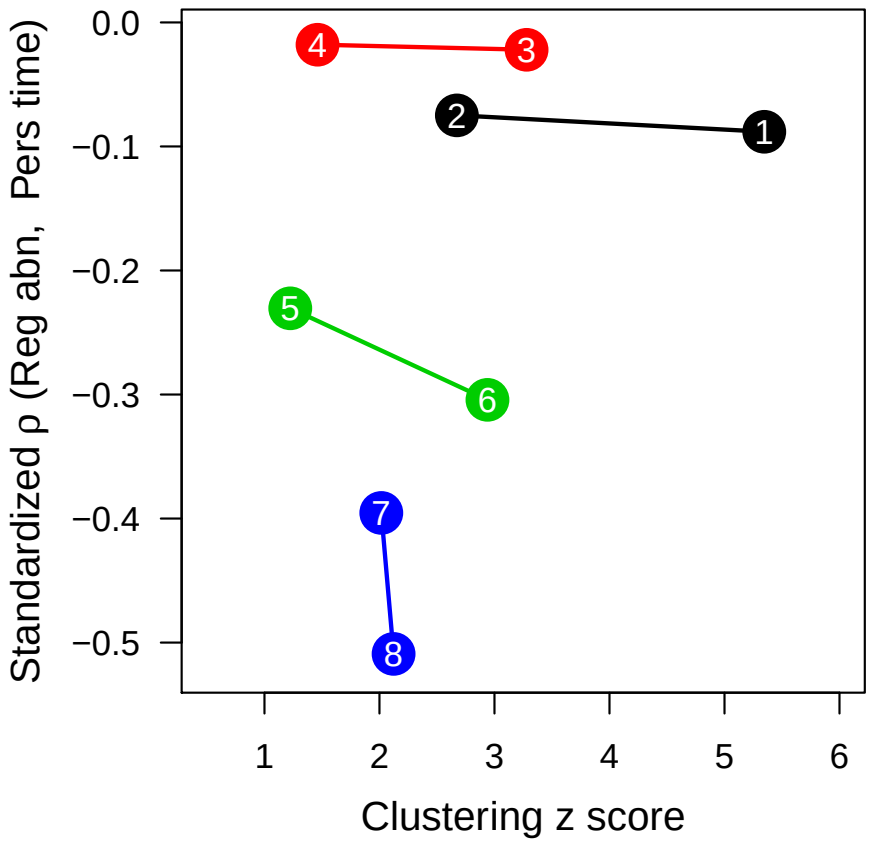



Figure S14: Comparison of the effect of niches on dynamics between niche scenarios with differing degree of clustering. x-axis shows z-score of the Ripley's K gap statistic. y-axis shows the rank correlation (Spearman's ρ) between regional abundances and local persistence across species. The coefficient is standardized by its value under pure neutrality: $\rho' = (\rho - \overline{\rho}_{\text{neutral}})/\sqrt{Var(\rho_{\text{neutral}})}$, where ρ_{neutral} was obtained from the corresponding neutral runs. Values on both axes are averaged across runs of each scenario. **Black**: Heuristic model with $p = 4$ and linear (1) or circular (2) niche axis; **Red**: Heuristic model with $p = 2$ and linear (3) or circular (4) niche axis; **Green**: Resource-consumer model with low (5) or high (6) resource depletion; **Blue**: Preferred environment model with global (7) or local (8) dispersal. Immigration rate for the heuristic scenarios (1-4) was $m = 0.08$, explaining the higher values of the standardized ρ compared with the explicit-mechanism scenarios (5-8), where we used $m = 0.01$.