Oecologia
https://doi.org/10.1007/s00442-021-04954-3

COMMUNITY ECOLOGY – ORIGINAL RESEARCH

Modelling selection, drift, dispersal and their interactions in the community assembly of Amazonian soil mites

Pedro A. C. L. Pequeno[1] · **Elizabeth Franklin**[2] · **Roy A. Norton**[3]

Received: 12 September 2020 / Accepted: 26 May 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

Three processes can explain contemporary community assembly: natural selection, ecological drift and dispersal. However, quantifying their effects has been complicated by confusion between different processes and neglect of expected interactions among them. One possible solution is to simultaneously model the expected effects of each process within species, across communities and across species, thus providing more integrative tests of ecological theory. Here, we used generalized linear mixed models to assess the effects of selection, drift and dispersal on the occurrence probability of 135 soil oribatid mite species across 55 sites over an Amazonian rainforest landscape (64 km^2). We tested for interactions between process-related factors and partitioned the explained variation among them. We found that occurrence probability (1) responded to soil P content and litter mass depending on body size and reproductive mode (sexual or parthenogenetic), respectively (selection); (2) increased with community size (drift); and (3) decreased with distance to the nearest source population, and more so in rare species (dispersal limitation). Processes did not interact significantly, and our best model explained 67% of the overall variation in species occurrence probability. However, most of the variation was attributable to dispersal limitation (55%). Our results challenge the seldom-tested theoretical prediction that ecological processes should interact. Rather, they suggest that dispersal limitation overrides the signatures of drift and selection at the landscape level, thus rendering soil microarthropod species ecologically equivalent and possibly contributing to the maintenance of metacommunity diversity.

Keywords Community structure · Ecological niche · Environmental filtering · Functional traits · Dispersal limitation

Introduction

Community assembly can be understood in terms of four higher-level processes: natural selection, ecological drift, dispersal, and speciation (Hanson et al. 2012; Vellend 2016). Over observable periods, speciation is generally negligible, and meta-analyses have suggested a predominant role for environmental selection over dispersal limitation (Cottenie 2005; Hanson et al. 2012; Soininen 2016). However, the standard variance partitioning used in metacommunity analyses assumes additive effects for environmental and spatial predictors despite the theory predicting ecological processes should interact (Hanson et al. 2012; Vellend 2016; Siqueira et al. 2020). Further, drift and dispersal have often been lumped as "stochastic" or "neutral" processes that contrast with "deterministic" niche processes (i.e. selection), which underestimates that drift and dispersal should have different effects (Hanson et al. 2012; Vellend et al. 2014). More generally, stochasticity (random variation) is not the same as neutrality (similar mean fitness across species), and the same process can have stochastic and deterministic components (e.g. selection driven by environmental stochasticity). Hence, there remain important knowledge gaps about the mode and strength with which ecological processes shape natural metacommunities.

Communicated by Stefan Scheu.

✉ Pedro A. C. L. Pequeno
pacolipe@gmail.com

Elizabeth Franklin
beth@inpa.gov.br

Roy A. Norton
ranorton@esf.edu

[1] Natural Resources Program, Federal University of Roraima, Boa Vista, Brazil

[2] Biodiversity Coordination, National Institute for Amazonia Research, Manaus, Brazil

[3] College of Environmental Science and Forestry, State University of New York, Syracuse, USA

Published online: 03 June 2021

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

One way of approaching this issue is to use hierarchical models to estimate species-specific relationships while accounting for their dependence on species traits and community features so that the predicted independent and interactive effects of ecological processes can be more easily quantified (Ovaskainen and Soininen 2011; Sydenham et al. 2019). In general, species are adapted to certain environmental conditions, and this is reflected in their traits. Accordingly, the environmental response of species occurrence probability should depend on their traits, reflecting natural selection (Keddy 1992). However, species may also be ecologically equivalent, i.e. perform similarly along environmental gradients on average, so that their local occurrence probability depends more on randomness in births and deaths than on fitness differences (i.e. ecological drift) (Hubbell 2001). If so, occurrence probability should decrease as community size (i.e. total number of individuals) decreases, due to more frequent random extinction (Hubbell 2001; Vellend 2016). When communities are linked by dispersal, the immigration rate is expected to decrease with distance to the nearest source population, so that species occurrence probability should decline similarly (i.e. dispersal limitation) (MacArthur and Wilson 1967). Further, for a given dispersal rate, dispersal limitation should increase as species commonness (i.e. number of individuals in the metacommunity) decreases, as rarer species have fewer dispersers and, thus, lower chance of reaching a given site over a given period (Hubbell 2001). Likewise, dispersal limitation also could depend on species traits (Soininen 2016).

Processes themselves can interact variously depending on model details. Generally, though, stronger drift in smaller communities should weaken selection effects by introducing noise in species responses to trait × environment interactions (Siqueira et al. 2020). Likewise, high dispersal rates closer to dispersal sources may weaken selection effects by increasing species occurrence in suboptimal habitats ("mass effect"; Shmida and Wilson 1985), or actually strengthen them by mixing species and facilitating species sorting (Ron et al. 2018). Lastly, dispersal and drift can also interact: local occurrence probability depends on the product between species commonness in the metacommunity and local community size, so that common species benefit more from increases in the latter (Hubbell 2001; Waller et al. 2018).

Smaller organisms are expected to be more affected by selection than by drift or dispersal limitation due to their higher abundance and passive dispersal (Kaspari et al. 2010; Soininen 2016). By contrast, studies on terrestrial microarthropods such as mites and springtails (<2 mm) have often found much variation in species composition that is unrelated to the environment (Borcard and Legendre 1994; Caruso et al. 2012; Gao et al. 2014; Widenfalk et al. 2016; Magilton et al. 2019; Zinger et al. 2019). Yet, such studies have typically focused on very small spatial grains (soil cores <100 cm^2) and extents (<1 km, often much less), where smaller community size may increase the relative importance of drift, even if absolute community size is large (Fodelianakis et al. 2021). While there is some evidence for stronger effects of selection and dispersal limitation over larger spatial extents (Mumladze et al. 2013; Caruso et al. 2019), few studies with reasonable sampling design have addressed this issue at the landscape level (1–10 km) (e.g. Ingimarsdóttir et al. 2012), and it is unclear the extent to which dispersal effectively connects site-level communities and modifies the effects of selection and drift. While there is strong spatial autocorrelation in soil fauna from tenths to a few hundred meters (Ettema and Wardle 2002; Minor 2011), passive dispersal through wind, water currents or transport by larger animals (Costa et al. 2013; Schuppenhauer et al. 2019) may connect communities over much larger ranges.

Here, we investigated the relative importance of selection, drift, and dispersal in the site-level community assembly of oribatid mites, a numerically dominant group of soil microarthropods. We used data from a unique, large sampling effort over an Amazonian rainforest landscape (1100 soils cores across 55 sites over 64 km^2). Two traits have been highlighted as mediators of oribatid environmental responses: (1) body size, with larger species possibly being favoured by soils with larger interstices, or more resources given higher metabolic demand (Brückner et al. 2018); and (2) reproductive mode (sexual or parthenogenetic), as the prevalence of parthenogens varies widely across local oribatid communities, with parthenogens possibly better coping with disturbance through faster reproduction while suffering stronger resource limitation due to clonal competition (Maraun et al. 2019). Body size may also affect dispersal capacity, although the relationship is seemingly complex and hard to predict (Soininen 2016; Schuppenhauer et al. 2019).

To test for the predicted effects of ecological processes, we modelled species occurrence probability as a function of: (1) the interactions body size × environmental variables and reproductive mode × environmental variables, representing selection; (2) community size, representing ecological drift; and (3) the interactions commonness × distance-from-source and body size × distance-from-source, representing dispersal. We also tested for the expected effects of interactions between selection and drift (trait × environment × community size), selection and dispersal (trait × environment × distance) and drift and dispersal (community size × commonness). Then, we partitioned the variation in species occurrence explained by significant predictors. We expected stronger effects of dispersal limitation and selection over the landscape, with weaker effects of drift and selection between nearby sites due to higher dispersal rates.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Materials and methods

Study site

The data used in this study were obtained as part of the Brazilian Program for Biodiversity Research (PPBio) in Reserva Ducke (Pequeno et al. 2021), a large reserve of primary tropical rainforest (10 × 10 km) in Manaus, northern Brazil (2°57′ S, 59°56′ W) (Costa and Magnusson 2010). The reserve is managed by the National Institute for Amazonia Research (INPA). Local terrain is traversed by a dense drainage network, where bottomlands have almost pure sands and uplands have clayish soils (soil clay content > 70%). Bottomlands are dominated by ferns, herbs and arborescent palms, whereas in uplands, trees establish a closed canopy with some individuals reaching 40–45 m, and the undergrowth is dominated by stemless palms. Mean daily temperature and mean annual rainfall during 1992–2002 were 26.7 °C and 2479 mm, respectively, with a relatively dry season (less than 100 mm) from July to September (Coordination for Research on Climate and Hydric Resources, INPA, unpublished data).

Mite data

Mite sampling was carried out from March 18 to May 13, 2002. Mites were collected from 55 sites distributed over a grid in the reserve, with at least 1 km between them (Fig. 1). On each site, one 250-m transect was established along a topographic contour lines, to minimize environmental variation within it (Costa and Magnusson 2010). Then, one soil core (3.5 cm × 3.5 cm × 5 cm) was sampled each 12.5 m along the transect, with 20 cores per transect and 55 sites × 20 cores = 1100 cores overall. Within-site soil cores thus provided a representative sample of the local, site-level community (Ettema and Wardle 2002; Minor 2011), which was the sampling unit of the study. To reduce the large sample processing load, each four consecutive soil cores within transects were combined as a compound soil sample. Compound samples were kept in plastic containers and transported to the Laboratory of Systematics and Ecology of Terrestrial Arthropods at INPA's campus in Manaus, where animals were extracted using a modified Berlese-Tullgren apparatus (Franklin and de Morais 2006). Samples were gradually heated from 28 to 45 °C until they were completely dry, which took from six to seven days. Extracted animals were preserved in glass vials containing 4% formaldehyde solution.

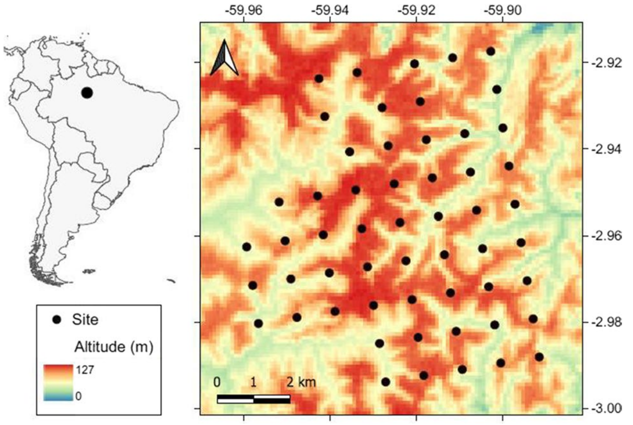

Fig. 1 Geographic disposition of sampling sites across an Amazonian rainforest landscape in Manaus, Northern Brazil

All adult oribatid mites were sorted into morphospecies and identified whenever possible using taxonomic keys (Online Resource 1). Identification proceeded by clarifying specimens with lactic acid, followed by temporary slide-mounting and examination under a compound microscope. Immatures were not considered but represented only 8% of extracted individuals. Voucher specimens were deposited in the Entomological Collection of INPA.

We estimated the mean body mass of each sampled species by measuring 1–15 individuals of each sampled species, depending on their abundance. For each individual, body length and width (μm) were measured under a microscope, and body mass (μg) was predicted using a well-established allometric equation ($R^2 = 0.98$; Caruso and Migliorini 2009):

$$\text{Mass} = -17.17 + 3.0\log(\text{length} + width) \tag{1}$$

Then, the mean body mass was calculated for each species.

Species reproductive mode (sexual or parthenogenetic) was inferred using published records (Maraun et al. 2019). In oribatid mites, parthenogenesis is estimated to occur in nearly 10% of species (Norton and Palmer 1991). When present, it is obligatory and thelytokous, i.e. all individuals are diploid female clones produced from unfertilized eggs (Heethoff et al. 2009). When a species' reproductive mode was unknown (i.e. morphospecies with a single individual), it was inferred from closely related species, if part of a taxonomic group that is not known to vary in reproductive mode. Otherwise, the species was assumed to be sexual.

Overall, 2046 adult individuals in 135 (morpho)species were found, of which 898 individuals in 22 species were parthenogenetic, representing only 16% of the species but 44% of the individuals (Online Resource 1, Table S1, Fig. S1). There was no relationship between body mass and reproductive mode (Linear Mixed Model with taxonomic ranks as nested random factors, $n = 135$, $t = 0.96$, $P = 0.34$; Online Resource 1, Fig. S2).

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Environmental data

Data were obtained on soil texture (clay content, in %), soil contents of water (%) and nutrients (C, N and P, in g kg^{-1}) and litter dry mass (g). Litter dry mass was measured during mite surveys by marking 50×50 cm squares along transects, one each 50 m, within which all litter was harvested. Litter samples were transported to the laboratory in plastic containers and dried to constant weight. Further, in each transect, six soil cores (one each 50 m) were collected to a depth of 5 cm. Soil cores were pooled in a plastic container and transported to INPA for granulometric analysis, and to the Brazilian Agricultural Research Corporation (EMBRAPA), also in Manaus, for nutrient analyses. Soil samples were oven-dried, cleaned of stones and roots, and passed through a 2 mm sieve. Soil granulometry was determined using the hydrometer method (clay: <0.002 mm; silt: 0.002–0.05 mm; sand: 0.05–2 mm). As clay and sand contents were highly correlated ($r = -0.99$) and silt content was negligible, we used clay to describe soil texture. Total organic carbon was measured by wet oxidation, using acid dichromate solution followed by titration with 0.5 N $FeSO_4$ and o-phenalphthroline. Total nitrogen was estimated using the wet oxidation (Kjeldahl method), by converting organic N to ammonium (NH4+) for measurement. Available phosphorus was estimated using the ammonium molybdate–ascorbic acid method, by reading the blue complex formed at 712 nm under a spectrophotometer. Soil water content was obtained by comparison between the wet and dry weights of soil samples. Soil and litter measurements were averaged by site.

Statistical analyses

We built a generalized linear mixed model that combines species-specific models in a common model for all species by assuming that species responses to a common predictor depend on species traits and features of local communities. By doing so, we can simultaneously test for relationships within species, across species, and across local communities. This approach also overcomes the difficulty in separately modelling rare species, by gaining information from the other species (Ovaskainen and Soininen 2011), and has been shown to reliability identify prevailing assembly processes in simulated metacommunities (Ovaskainen et al. 2019; Sydenham et al. 2019). Here, we extended this approach by including model terms representing predicted interactions among all three ecological processes, and partitioning the variation explained by them (see below). Because of the large number of parameters of interest, we divided the analysis in two parts: we (1) built a global model testing for the separate effects of drift, selection, and dispersal, and simplified it by excluding parameters that were not statistically supported; (2) created three alternative versions of the simplified model representing the interactions between drift and selection, selection and dispersal, and drift and dispersal, and determined the most supported model among these four.

We first concatenated species occurrences (presence or absence) into a single occurrence variable with 135 species $\times$ 55 sites = 7425 observations. Then, this variable was modelled in relation to the predictors of interest, assuming binomial errors and complementary log–log link function to account for asymmetric, sigmoid response curves (Zuur et al. 2009), and “species” and “site” as random factors to account for autocorrelation within species and sites, respectively. To test for selection, we included interactions between species traits (body mass and reproductive mode) and environmental variables. Because we had six environmental variables, there were 12 possible trait × environment interactions. To reduce the number of parameters, we summarized environmental variables using Principal Component Analysis (PCA), with soil P log-transformed to account for its skewed distribution. PCA revealed three major environmental gradients: one representing soil texture and organic matter (C and N) (PC1); one representing soil P (PC2); and another representing litter mass (PC3) (Table 1). Thus, we used these three principal components in body mass × PC and reproductive mode × PC interactions (six interactions). To test for drift, we included community size (total number of individuals in a site) as a predictor. To test for dispersal limitation, we calculated species commonness (total number of individuals in the metacommunity) and the Euclidian distance between each site and the nearest site occupied by the species. Then, we included commonness × distance and body mass × distance interactions as predictors. Community size and species commonness were log-transformed to account for their highly skewed distributions. All numeric predictors were scaled to zero mean and unit variance, and parameters were estimated by maximum likelihood.

When a species occurred at a single site, either the nearest occupied site was outside the sampled area (and its distance was unknown) or it was the single site itself (and the

Table 1 Principal Component Analysis (PCA) of environmental variables used in this study

Variable	Range	r_{PC1} (52%)	r_{PC2} (20%)	r_{PC3} (16%)
Soil clay (%)	2–87	0.49	− 0.27	0.25
Soil water (%)	12–65	0.46	0.22	− 0.09
Soil C (g kg^{-1})	0.73–7.30	0.48	0.28	− 0.15
Soil N (g kg^{-1})	0–13	0.52	0.13	− 0.16
log Soil P (g kg^{-1})	15–79	− 0.22	0.77	− 0.29
Litter dry mass (g)	131–482	0.02	0.42	0.89

All variables are site averages ($n = 55$). r_{PC} is the correlation between the variable and ordination scores. Numbers in brackets indicate the variance proportion explained by each axis

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

distance should be zero). We ran the global model including and excluding such species and found similar results (Table 2; Online Resource 1, Table S2), so we report the analysis including all species. The global model was simplified by excluding non-supported interactions ($P > 0.05$) and refitting the model to test for independent effects, which were similarly excluded if not supported.

Having the simplified model, we tested for an interaction between drift and selection by adding community size × trait × PC interactions to the simplified model. Likewise, to test for an interaction between selection and dispersal, we added distance × trait × PC interactions. To test for an interaction between drift and dispersal, we included a community size × commonness interaction in the simplified model. These three models, along with the simplified model assuming independent effects, were compared using the Bayesian Information Criterion (BIC), and the model with the lowest BIC was judged the most supported.

We assessed the amount of variation explained by the final model using the marginal $\left(R^2_{\mathrm{m}}\right)$ and conditional coefficients of determination $\left(R^2_{\mathrm{c}}\right)$, which measure the variation explained by predictors vs. predictors plus random factors in GLMMs, respectively (Nakagawa and Schielzeth 2013). Then, we decomposed the marginal R^2 into fractions explained by separate predictor terms using hierarchical partitioning (Chevan and Sutherland 1991). If there were interactions among drift, selection and dispersal, we computed the fraction explained by the interaction (i.e. sum of fractions explained by model terms related to both processes); otherwise, we computed fractions for each process separately. Predictor effects were visualized using conditional plots (Breheny and Burchett 2017).

Table 2 Generalized linear mixed model of oribatid mite occurrence probability

Process	Predictor	Coefficient	z	P
	Intercept	– 2.89	–	–
Selection	PC1	0.063	0.905	0.366
	PC2	– 0.081	1.094	0.274
	PC3	– 0.197	2.545	**0.011**
	log Mass	– 0.05	– 0.708	0.479
	Mode	– 0.023	– 0.238	0.812
	log Mass × PC1	– 0.03	– 0.801	0.423
	log Mass × PC2	0.126	– 3.246	**0.001**
	log Mass × PC3	– 0.011	0.279	0.78
	Mode × PC1	– 0.051	– 0.638	0.523
	Mode × PC2	0.099	– 1.177	0.239
	Mode × PC3	0.234	– 2.744	**0.006**
Drift	log Community size	0.598	14.339	**< 0.001**
Dispersal	Distance	– 0.743	– 8.289	**< 0.001**
	log Commonness	1.164	14.213	**< 0.001**
	log Commonness × distance	0.362	3.824	**< 0.001**
	log Mass × distance	– 0.049	– 0.563	0.573

Models assumed binomial errors, complementary log–log link, and "site" and "species" as random intercepts ($n = 55$ sites × 135 species = 7425 observations). All numerical predictors were scaled to zero mean and unit variance; reproductive mode was coded as 0 (parthenogenetic) or 1 (sexual). Bold numbers indicate statistically significant effects ($P < 0.05$)

Our models assumed no residual correlation among sites and among species. This is reasonable because spatial proximity among sites was modelled directly, and because species interactions that could drive microarthropod species correlations independent of their measured environmental responses occur at much finer spatial resolutions (Ettema and Wardle 2002), which should erase their statistical signal (Thuiller et al. 2015). We validated both assumptions by extracting Pearson residuals (on link scale) and (1) creating a spatial correlogram of such residuals using Moran's I, and (2) applying PCA to species residuals to sequentially test PCs (from PC1 onwards) for significant information content using permutation tests, until the first non-significant PC was found (Dray 2008). If species residuals were generally correlated, their shared variation should allow reducing their dimensionality. As expected, there was no evidence for either spatial (Online Resource 1, Fig. S3) or species correlation in residuals ($P = 0.30$ for PC1). All computations were performed in R 3.6.3 (Team RC 2020), with aid of packages "glmmTMB" (Brooks et al. 2017), "MuMIn" (Barton 2019), "ncf" (Bjornstad 2020), "ade4" (Bougeard and Dray 2018) and "visreg" (Breheny and Burchett 2017).

Results

The global model revealed that species occurrence probability was related to predictors representing selection, drift and dispersal limitation (Table 2). Model simplification revealed that PC1 had no effect whatsoever ($z = 0.550$, $P = 0.582$), so this variable was excluded from the simplified model. Comparison among the simplified model and alternative models including pairwise interactions between selection, drift and dispersal revealed stronger support for the simplified model (Table 3; Online Resource 1, Table S3).

According to the best model, species occurrence probability depended on trait × environment interactions: P-rich soils favoured larger over smaller species (Fig. 2a), whereas sites with more litter selected against parthenogenetic species (Fig. 2b). Second, occurrence probability increased with community size (Fig. 2c). Third, occurrence probability decreased with distance to the nearest source, with rare species reaching relatively low occurrence probability farther from the source faster (Fig. 2d). This model explained 67% of the variation in species occurrence probability, all of

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Table 3 Comparison of generalized linear mixed models on oribatid mite occurrence probability using the Bayesian Information Criterion (BIC)

Model	Model notation	BIC
Simplified	log Mass × PC2 + Mode × PC3 + Community size + Commonness × Distance	4097.427
Dispersal × drift	log Mass × PC2 + Mode × PC3 + **Commonness × Community size** + Commonness × Distance	4102.883
Selection × drift	**log Mass × PC2 × Community size** + **Mode × PC3 × Community size** + Commonness × Distance	4144.686
Selection × dispersal	**log Mass × PC2 × Distance** + **Mode × PC3 × Distance** + Community size + Commonness × Distance	4147.470

Models assumed binomial errors, complementary log–log link, and “site” and “species” as random intercepts ($n = 55$ sites × 135 species = 7425 observations). The simplified model included only the significant interactions in Table 1 plus their lower-order terms; remaining models included additional terms representing pairwise interactions between ecological processes (highlighted in bold). Variables were coded as in Table 1. Model notation omits intercepts and lower-order interaction terms for clarity

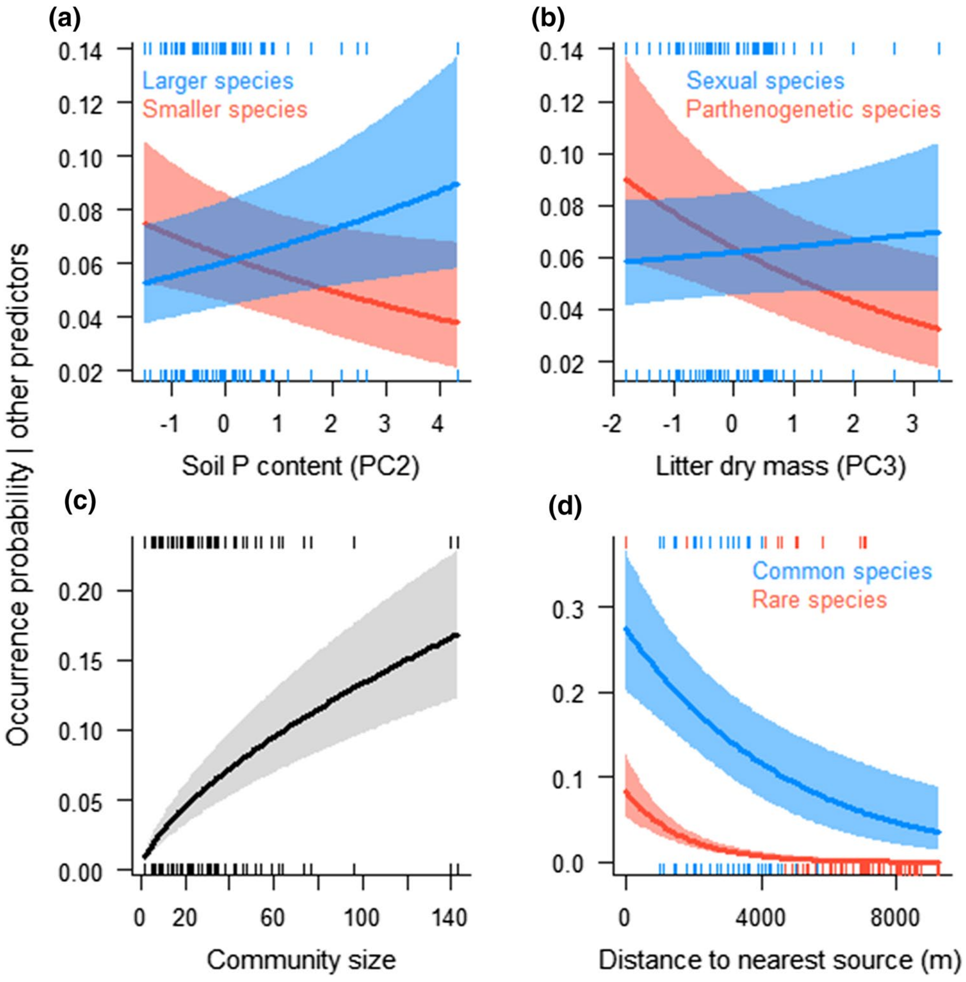

Fig. 2 Effects of selection (**a**, **b**), drift (**c**) and dispersal limitation (**d**) on oribatid mite occurrence probability. Lines represent means predicted by the best generalized linear mixed model, which assumed binomial errors, complementary log–log link, and “site” and “species” as random intercepts ($n = 55$ sites × 135 species = 7425 observations). Coloured bands represented 95% confidence intervals. Upper and bottom tick marks represent presence and absence, respectively. Conditional plots show the effect of each predictor or interaction while holding other predictors at their medians, and sexual species for reproductive mode. Species body mass and commonness were divided as above or below their medians to facilitate visualization

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

which was accounted for by predictors only ($R_{\mathrm{m}}^2 = R_{\mathrm{c}}^2$). Hierarchical partitioning revealed that model terms representing dispersal limitation explained most of this variation, with drift- and particularly selection-related terms contributing much less (Fig. 3).

Discussion

Our analysis revealed clear patterns in the occurrence probability of oribatid mite species that were consistent with the expected effects of selection, drift and dispersal on community assembly. However, most of the variation in occurrence probability was attributable to dispersal limitation, suggesting this is the dominant process shaping community composition across sites over the landscape. Further, there was no evidence for the predicted interactions between the effects of selection, drift and dispersal, supporting the common assumption of additive effects.

The decline in occurrence probability with increasing distance to nearest source reached a “floor” only after a few kilometres (Fig. 2d), suggesting that communities are relatively well connected by dispersal up to this distance. This is consistent with the ease of passive dispersal in soil microarthropods (Costa et al. 2013; Schuppenhauer et al. 2019) and with similar findings from a contrasting, glacial landscape (Ingimarsdóttir et al. 2012). However, the decay in occurrence probability with increasing distance to source was stronger in rarer species (Fig. 2d), as predicted by neutral models: rarer species are less likely to reach farther sites over a given period due to fewer migrants (Hubbell 2001). As the distance to nearest source explained much of the variation in occurrence probability independently of environmental predictors, species may often occur in suboptimal habitats (i.e. mass effect; Shmida and Wilson 1985).

Lower species occurrence probability in smaller communities is consistent with ecological drift: given that births and deaths are not completely predictable from traits (i.e. demographic stochasticity), and that communities are finite, species will be more likely to go extinct in communities with fewer individuals, thereby decreasing their occurrence probability (MacArthur and Wilson 1967; Hubbell 2001; Vellend 2016). For instance, this is the reason for the well-known species-area relationship in Island Biogeography Theory. Although soil arthropods can reach high densities and our site-level communities were aggregates of 20 soil cores, their size ranged from a few to around 140 individuals only (Fig. 2c), which is within the range over which drift has been found to have significant effects in simulated communities (Hubbell 2001; Vellend 2016). Yet, drift has been demonstrated even in bacterial communities of millions of cells per litre (Fodelianakis et al. 2021). Thus, the relevant question is not whether drift occurs, but how much it affects community structure compared to other processes, namely selection and dispersal.

Most metacommunity analyses of soil macroarthropods focus on rather small spatial grains (e.g. 10 cm) and extents (< 1 km), and often find that most variation in community structure is unrelated to environment or space (Caruso et al. 2012; Gao et al. 2014; Magilton et al. 2019; Zinger et al. 2019). While part of this unexplained variation is probably sampling error, it might also reflect the relatively stronger effect of drift in fine-grained communities. By contrast, our results point to a much weaker effect of drift relative to dispersal limitation, which is consistent with the larger spatial grain and extent analyzed here.

There were detectable trait × environment interactions underlying species occurrence probability, consistent with selection. Occurrence probability increased with soil P content for larger species but decreased for smaller species. P strongly limits detrital food webs in lowland tropical forests (Camenzind et al. 2018). Further, P total amount increases while P content decreases with body size in arthropods (Woods et al. 2004). On the one hand, this suggests that P limits larger species simply due to their higher whole-body nutritional reqruiements. On the other hand, species may compete differently for P as function of their size, e.g. smaller species may better forage for P-rich foods (Woods et al. 2004) and thus perform better under low P (Fig. 2a). Likewise, occurrence probability changed little with litter mass in sexual species and decreased in parthenogenetic species. Assuming that most soil oribatid mites are fungivores (Heethoff et al. 2009; Potapov et al. 2019), litter should mainly provide a habitat for mites. Hence, less litter may provide less shelter, e.g. against predators (Wilson 2005). Parthenogens are thought to reproduce faster by dispensing with males and mating, and may be favoured when mortality is high (Maraun et al. 2019), e.g. in more exposed soil.

Regardless of the precise mechanisms driving trait × environment interactions, the low variation explained by them suggests that species were largely distributed as if they were ecologically equivalent, i.e. had similar performance under local environmental conditions (Shmida and Wilson 1985). Ecological equivalence is the key assumption of neutral models of community assembly, as compared to niche models (Hubbell 2001; Vellend 2016). Ecological equivalence is probably an unrealistic assumption in general, given widespread evidence for niche partitioning in plants (Johnson et al. 2017) and animals, including soil fauna (Potapov et al. 2019). However, dispersal limitation can reduce co-occurrence among species that differ in performance when together, thereby delaying competitive exclusion to levels compatible with ecological equivalence (Munoz and Huneman 2016). This might explain why trait × environment

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

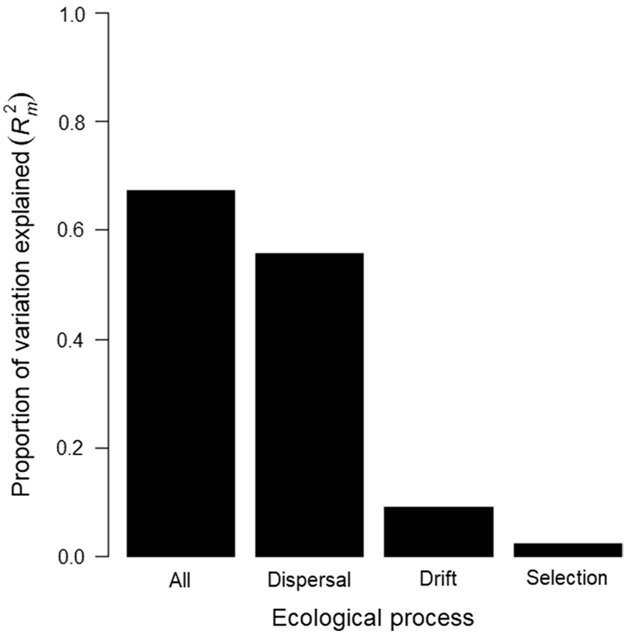

Fig. 3 Hierarchical partitioning of variation explained by ecological processes under the best generalized linear mixed model for oribatid mite occurrence probability, which assumed binomial errors, complementary log–log link, and "site" and "species" as random intercepts ($n = 55$ sites × 135 species = 7425 observations). Partitioning was based on the marginal coefficient of determination (R^2_m)

interactions, although supported, only weakly explained occurrence probability.

Our results challenge theoretical predictions that ecological processes should interact in community assembly (Hanson et al. 2012; Vellend et al. 2014; Vellend 2016; Ron et al. 2018). Indeed, the few experimental studies that explicitly tested for effects of community size (drift) and distance to source habitat (dispersal) on soil microarthropod community structure found evidence for both independent and interacting effects (Hoyle 2004; Åström and Bengtsson 2011). Yet, most metacommunity analyses have not explicitly quantified drift and have assumed independent effects for selection and dispersal limitation (Cottenie 2005; Hanson et al. 2012; Soininen 2016). Our results support this assumption for the conditions we studied. Perhaps interactions are only relevant at wider environmental or geographic extents. For instance, species environmental responses of aquatic invertebrates were weaker in tropical than in boreal streams, with community size being five times larger in the latter (Siqueira et al. 2020). Likewise, drift may be relatively stronger in tropical soil animal communities given their lower density (Takeda and Abe 2001), which might account for the relatively weak selective effects detected here.

Several caveats should be considered. First, it is difficult to infer causality from observational data. Accordingly, we focused on non-overlapping predictions from higher-level ecological processes (MacArthur and Wilson 1967; Hubbell 2001; Vellend 2016; Waller et al. 2018) and applied a modelling framework that has been shown by simulation to reliably infer such processes under known conditions (Ovaskainen et al. 2019; Sydenham et al. 2019). Yet, experiments will be ultimately required to validate our interpretations. Second, important predictors can always go unmeasured, especially regarding environmental variables and traits. However, as predictors related to dispersal and drift accounted for most of the variation in species occurrence, adding predictors is unlikely to change this result. Third, we focused on species occurrence rather than abundance, but the latter is harder to predict. For instance, neutral theory predicts that both equilibrium species abundance and its variance should increase with community size, making average trends less discernible (Hubbell 2001). It is possible that factors other than the ones analyzed here affect local abundance once species have established. Lastly, rare species probably looked even rarer due to imperfect detection. This could bias the estimated distance to the nearest site occupied by a species. Yet, it has been shown that rare species contribute little to observed multivariate patterns in species occurrence (Franklin et al. 2013), and simulations suggest that ecological relationships can be reasonably recovered despite imperfect detection (Barker et al. 2018; Yamaura et al. 2019). The fact that analyses excluding singletons provided similar results corroborates that overall patterns are robust to variability in rare species.

Our study suggests that dispersal limitation is the major driver of oribatid species occurrence probability over a landscape in central Amazonia, although dispersal could still connect local communities up to a few kilometres, possibly by passive dispersal. By contrast, drift- and especially selection-related predictors accounted for little variation in occurrence probability. Our results also challenge the expectation that higher-level ecological processes should interact, supporting instead the assumption of additive effects typical of variance partitioning methods. Overall, the high predictive power of distance to nearest source despite variation in species environmental responses suggests that dispersal limitation renders species ecologically equivalent at the landscape level, possibly contributing to the maintenance of metacommunity diversity.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00442-021-04954-3.

Acknowledgements We thank José Wellington de Morais, Rozilete Guimarães and Edilson Fagundes for their assistance in field work; Jamile de Moraes for her support with species identification and counting; Breno Azevedo for mite measurements; Flávio Luizão (deceased) for the soil data; and Mariane Bosholn for the study area map. The first author received a PNPD fellowship from the Brazilian Coordination

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

for Training of Higher Education Personnel (CAPES) during this study (Grant No. 88887.477573/2020-00).

Author contributions PACLP and EF conceived the ideas and designed methodology; EF and RAN collected the data; PACLP analysed the data; PACLP led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Data availability statement All data used in this manuscript are available from the figshare repository (Pequeno et al. 2021).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All applicable institutional and/or national guidelines for the care and use of animals were followed.

References

Åström J, Bengtsson J (2011) Patch size matters more than dispersal distance in a mainland-island metacommunity. Oecologia 167:747–757. https://doi.org/10.1007/s00442-011-2024-y

Barker RJ, Schofield MR, Link WA, Sauer JR (2018) On the reliability of N-mixture models for count data. Biometrics 74:369–377. https://doi.org/10.1111/biom.12734

Barton K (2019) MuMIn: multi-model inference. R package version 1.43.15. https://CRAN.R-project.org/package=MuMIn

Bjornstad ON (2020) ncf: spatial covariance functions. R package version 1.2-9. https://CRAN.R-project.org/package=ncf

Borcard D, Legendre P (1994) Environmental control and spatial structure in ecological communities: an example using oribatid mites (Acari, Oribatei). Environ Ecol Stat 1:37–61

Bougeard S, Dray S (2018) Supervised multiblock analysis in R with the ade4 package. J Stat Softw 86:1–17. https://doi.org/10.18637/jss.v086.i01

Breheny P, Burchett W (2017) Visualization of regression models using visreg. R J 9:56–71. https://doi.org/10.32614/RJ-2017-046

Brooks ME, Kristensen K, van Benthem KJ et al (2017) glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J 9:378–400. https://doi.org/10.32614/rj-2017-066

Brückner A, Heethoff M, Norton RA, Wehner K (2018) Body size structure of oribatid mite communities in different microhabitats. Int J Acarol 44:367–373. https://doi.org/10.1080/01647954.2018.1508248

Camenzind T, Hättenscwiler S, Treseder K et al (2018) Nutrient limitation of soil microbial processes in tropical forests. Ecol Monogr 88:4–21. https://doi.org/10.1111/ijlh.12426

Caruso T, Migliorini M (2009) Euclidean geometry explains why lengths allow precise body mass estimates in terrestrial invertebrates: the case of oribatid mites. J Theor Biol 256:436–440. https://doi.org/10.1016/j.jtbi.2008.09.033

Caruso T, Taormina M, Migliorini M (2012) Relative role of deterministic and stochastic determinants of soil animal community: a spatially explicit analysis of oribatid mites. J Anim Ecol 81:214–221. https://doi.org/10.1111/j.1365-2656.2011.01886.x

Caruso T, Schaefer I, Monson F, Keith AM (2019) Oribatid mites show how climate and latitudinal gradients in organic matter can drive large-scale biodiversity patterns of soil communities. J Biogeogr 46:611–620. https://doi.org/10.1111/jbi.13501

Chevan A, Sutherland M (1991) Hierarchical partitioning. Am Stat 45:90–96. https://doi.org/10.1080/00031305.1991.10475776

Costa FRC, Magnusson WE (2010) The need for large-scale, integrated studies of biodiversity—the experience of the Program for Biodiversity Research in Brazilian Amazonia. Braz J Nat Conserv 8:3–12

Costa D, Timmermans MJTN, Sousa JP et al (2013) Genetic structure of soil invertebrate populations: collembolans, earthworms and isopods. Appl Soil Ecol 68:61–66. https://doi.org/10.1016/j.apsoil.2013.03.003

Cottenie K (2005) Integrating environmental and spatial processes in ecological community dynamics. Ecol Lett 8:1175–1182. https://doi.org/10.1111/j.1461-0248.2005.00820.x

Dray S (2008) On the number of principal components: a test of dimensionality based on measurements of similarity between matrices. Comput Stat Data Anal 52:2228–2237. https://doi.org/10.1016/j.csda.2007.07.015

Ettema CH, Wardle DA (2002) Spatial soil ecology. Trends Ecol Evol 17:177–183

Fodelianakis S, Valenzuela-Cuevas A, Barozzi A, Daffonchio D (2021) Direct quantification of ecological drift at the population level in synthetic bacterial communities. ISME J 15:55–66. https://doi.org/10.1038/s41396-020-00754-4

Franklin E, de Morais JW (2006) Soil mesofauna in central Amazon. In: Moreira FMS, Siqueira JO, Brussaard L (eds) Soil biodiversity in Amazonian and other Brazilian ecosystems. CABI Publishing, Oxfordshire, pp 142–162

Franklin E, De Moraes J, Landeiro VL et al (2013) Geographic position of sample grid and removal of uncommon species affect multivariate analyses of diverse assemblages: The case of oribatid mites (Acari: Oribatida). Ecol Indic. https://doi.org/10.1016/j.ecolind.2013.04.024

Gao M, He P, Zhang X et al (2014) Relative roles of spatial factors, environmental filtering and biotic interactions in fine-scale structuring of a soil mite community. Soil Biol Biochem 79:68–77. https://doi.org/10.1016/j.soilbio.2014.09.003

Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. Nat Rev Microbiol 10:497–506. https://doi.org/10.1038/nrmicro2795

Heethoff M, Norton RA, Scheu S, Maraun M (2009) Parthenogenesis in oribatid mites (Acari: Oribatida): evolution without sex. In: Schön I, Martens K, van Dijk P (eds) Lost sex: the evolutionary biology of parthenogenesis. Springer, Dordrecht, pp 241–258

Hoyle M (2004) Causes of the species-area relationship by trophic level in a field-based microecosystem. Proc R Soc B Biol Sci 271:1159–1164. https://doi.org/10.1098/rspb.2004.2704

Hubbell SP (2001) The unified theory of biodiversity and biogeography. Princeton University Press, Princeton

Ingimarsdóttir M, Caruso T, Ripa J et al (2012) Primary assembly of soil communities: disentangling the effect of dispersal and local environment. Oecologia 170:745–754. https://doi.org/10.1007/s00442-012-2334-8

Johnson DJ, Condit R, Hubbell SP, Comita LS (2017) Abiotic niche partitioning and negative density dependence drive tree seedling survival in a tropical forest. Proc R Soc B Biol Sci. https://doi.org/10.1098/rspb.2017.2210

Kaspari M, Stevenson BS, Shik J, Kerekes JF (2010) Scaling community structure: how bacteria, fungi, and ant taxocenes differentiate along a tropical forest floor. Ecology 91:2221–2226

Keddy PA (1992) Assembly and response rules: two goals for predictive community ecology. J Veg Sci 3:157–164. https://doi.org/10.2307/3235676

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

MacArthur RH, Wilson EO (1967) The theory of island biogeography. Princeton University Press, Princeton

Magilton M, Maraun M, Emmerson M, Caruso T (2019) Oribatid mites reveal that competition for resources and trophic structure combine to regulate the assembly of diverse soil animal communities. Ecol Evol 9:8320–8330. https://doi.org/10.1002/ece3.5409

Maraun M, Caruso T, Hense J et al (2019) Parthenogenetic vs. sexual reproduction in oribatid mite communities. Ecol Evol 2019:7324–7332. https://doi.org/10.1002/ece3.5303

Minor MA (2011) Spatial patterns and local diversity in soil oribatid mites (Acari: Oribatida) in three pine plantation forests. Eur J Soil Biol 47:122–128. https://doi.org/10.1016/j.ejsobi.2011.01.003

Mumladze L, Murvanidze M, Behan-Pelletier V (2013) Compositional patterns in Holarctic peat bog inhabiting oribatid mite (Acari: Oribatida) communities. Pedobiologia 56:41–48. https://doi.org/10.1016/j.pedobi.2012.10.001

Munoz F, Huneman P (2016) From the neutral theory to a comprehensive and multiscale theory of ecological equivalence. Q Rev Biol 91:321–342. https://doi.org/10.1086/688098

Nakagawa S, Schielzeth H (2013) A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol Evol 4:133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Norton RA, Palmer SC (1991) The distribution, mechanisms and evolutionary significance of parthenogenesis in oribatid mites. In: Schuster R, Murphy PW (eds) The Acari: reproduction, development and life history strategies. Springer, Dordrecht, pp 107–136

Ovaskainen O, Soininen J (2011) Making more out of sparse data: hierarchical modeling of species communities. Ecology 92:289–295. https://doi.org/10.1890/10-1251.1

Ovaskainen O, Rybicki J, Abrego N (2019) What can observational data reveal about metacommunity processes? Ecogr (cop) 42:1877–1886. https://doi.org/10.1111/ecog.04444

Pequeno PACL, Franklin E, Norton, RA (2021) Data on: modelling selection, drift, dispersal and their interactions in the community assembly of Amazonian soil mites. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14605581.v1

Potapov AM, Tiunov AV, Scheu S (2019) Uncovering trophic positions and food resources of soil animals using bulk natural stable isotope composition. Biol Rev 94:37–59. https://doi.org/10.1111/brv.12434

Ron R, Fragman-Sapir O, Kadmon R (2018) Dispersal increases ecological selection by increasing effective community size. Proc Natl Acad Sci USA 115:11280–11285. https://doi.org/10.1073/pnas.1812511115

Schuppenhauer MM, Lehmitz R, Xylander WER (2019) Slow-moving soil organisms on a water highway: aquatic dispersal and survival potential of Oribatida and Collembola in running water. Mov Ecol 7:1–14. https://doi.org/10.1186/s40462-019-0165-5

Shmida A, Wilson MV (1985) Biological determinants of species diversity. J Biogeogr 12:1–20

Siqueira T, Saito VS, Bini LM et al (2020) Community size can affect the signals of ecological drift and niche selection on biodiversity. Ecology 101:3. https://doi.org/10.1002/ecy.3014

Soininen J (2016) Spatial structure in ecological communities—a quantitative analysis. Oikos 125:160–166. https://doi.org/10.1111/oik.02241

Sydenham MAK, Moe SR, Steinert M, Eldegard K (2019) Univariate Community Assembly Analysis (UniCAA): combining hierarchical models with null models to test the influence of spatially restricted dispersal, environmental filtering, and stochasticity on community assembly. Ecol Evol 9:1473–1488. https://doi.org/10.1002/ece3.4868

Takeda H, Abe T (2001) Templates of food-habitat resources for the organization of soil animals in temperate and tropical forests. Ecol Res 16:961–973. https://doi.org/10.1046/j.1440-1703.2001.00450.x

R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Thuiller W, Pollock LJ, Gueguen M, Münkemüller T (2015) From species distributions to meta-communities. Ecol Lett 18:1321–1328. https://doi.org/10.1111/ele.12526

Vellend M (2016) The theory of ecological communities. Princeton University Press, Princeton

Vellend M, Srivastava DS, Anderson KM et al (2014) Assessing the relative importance of neutral stochasticity in ecological communities. Oikos 123:1420–1430. https://doi.org/10.1111/oik.01493

Waller DM, Mudrak EL, Rogers DA (2018) Do metacommunity mass effects predict changes in species incidence and abundance? Ecogr (cop) 41:11–23. https://doi.org/10.1111/ecog.02539

Widenfalk LA, Malmström A, Berg MP, Bengtsson J (2016) Small-scale Collembola community composition in a pine forest soil—overdispersion in functional traits indicates the importance of species interactions. Soil Biol Biochem 103:52–62. https://doi.org/10.1016/j.soilbio.2016.08.006

Wilson EO (2005) Oribatid mite predation by small ants of the genus Pheidole. Insectes Soc 52:263–265. https://doi.org/10.1007/s00040-005-0802-4

Woods HA, Fagan WF, Elser JJ, Harrison JF (2004) Allometric and phylogenetic variation in insect phosphorus. Funct Ecol 18:103–109

Yamaura Y, Blanchet FG, Higa M (2019) Analyzing community structure subject to incomplete sampling: hierarchical community model vs. canonical ordinations. Ecology 100:1–14. https://doi.org/10.1002/ecy.2759

Zinger L, Taberlet P, Schimann H et al (2019) Body size determines soil community assembly in a tropical forest. Mol Ecol 28:528–543. https://doi.org/10.1111/mec.14919

Zuur A, Ieno E, Walker N et al (2009) Mixed effects models and extensions in ecology with R. Springer, New York

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com