Community size affects the signals of ecological drift and niche selection on biodiversity

Tadeu Siqueira[1], Victor S. Saito[2], Luis M. Bini[3], Adriano S. Melo[3], Danielle K. Petsch[3], Victor L. Landeiro[4], Kimmo T. Tolonen[5], Jenny Jyrkänkallio-Mikkola[6], Janne Soininen[6], Jani Heino[7]

[1]Institute of Biosciences, São Paulo State University (UNESP), Av. 24 A 1515, Rio Claro, SP, 13506-900, Brazil

[2]Universidade Federal de São Carlos (UFSCar), Departamento de Ciências Ambientais, São Carlos, Brazil

[3]Departamento de Ecologia, ICB, Universidade Federal de Goiás, Goiânia, GO, Brazil

[4]Departamento de Botânica e Ecologia, IB, Universidade Federal de Mato Grosso, Cuiabá, MT, Brazil

[5]University of Jyväskylä, Department of Biological and Environmental Sciences, Jyväskylä, Finland

[6]University of Helsinki, Department of Geosciences and Geography, Helsinki, Finland

[7]Finnish Environment Institute, Freshwater Centre, Oulu, Finland

Keywords: aquatic insects; beta diversity deviation; community assembly; demographic stochasticity; dispersal; metacommunities; null models

Abstract. Ecological drift can override the effects of deterministic niche selection on small populations and drive the assembly of small communities. We tested the hypothesis that smaller local communities are more dissimilar among each other because of ecological drift than larger communities, which are mainly structured by niche selection. We used a unique, comprehensive dataset on insect communities sampled identically in a total of 200 streams in climatically different regions (Brazil and Finland) that differ in community size by fivefold. Null models allowed us to estimate the magnitude to which beta diversity deviates from the expectation under a random assembly process while taking differences in species richness and relative abundance into account, i.e., beta deviation. Beta diversity of small tropical communities was consistently higher but closer to null expectations than β-diversity of large communities. However, although β-deviation and community size were strongly related in both regions, the direction of the relationship varied according to dissimilarity metrics. While incidence-based β-diversity was lower than expected (communities were less dissimilar than null expectations) and negatively related to community size in Brazil, abundance-based β-diversity was higher than expected (communities were more dissimilar than null expectations) and positively related to community size in both regions. We suggest that ecological drift drives variation in small communities by increasing the chances of species with low abundance and narrow distribution to occur within the metacommunity. Also, while weak niche selection and high dispersal rates likely reduced variation in community structure among large tropical streams, niche selection was likely sufficient to cause non-random variations in the relative abundances of genera among large communities in both regions. Habitat destruction, overexploitation, pollution, and reductions in connectivity have been reducing the size of biological communities. These environmental pressures will make smaller communities more vulnerable to novel conditions and render community dynamics more unpredictable, as random demographic processes should prevail under these conditions. Incorporation of community size into ecological models should provide conceptual, empirical and applied insights into a better understanding of the processes driving changes in biodiversity.

53 **Introduction**

54 Recent syntheses in community ecology propose that four main processes drive the
55 dynamics of metacommunities – deterministic niche selection, ecological drift, dispersal and
56 speciation (Vellend 2010, 2016, Leibold and Chase 2018). At broad spatial scales, dispersal,
57 colonization history, and spatial heterogeneity are thought to produce major biodiversity patterns
58 (Leibold and Chase 2018), while speciation also plays a role by altering the composition of regional
59 species pools at longer time frames (Vellend 2010). Within localities, niche selection determines
60 community structure mainly through species interactions, differential use of resources and species
61 responses to environmental gradients (Leibold and Chase 2018). However, stochasticity may also
62 play a major role in driving local community dynamics, for example, when demographic events
63 occur at random with respect to species identities (Vellend et al. 2014). Indeed, mechanistic models
64 (Orrock and Fletcher 2005, Orrock and Watling 2010) and recent empirical evidence (Gilbert and
65 Levine 2017) suggest that ecological drift can even override the effects of niche selection under
66 certain circumstances, such as when species populations in local communities are small and isolated
67 from other populations.

68 Small communities have few individuals per unit area, and thus random birth and death
69 events are likely to have a high impact on their structure (species composition and relative
70 abundances). For example, the species composition of a local community would change if all
71 individuals of a species die before reproducing. This is likely to happen in nature especially among
72 small populations or in those large populations in which only a reduced fraction of adults
73 successfully reproduce (Bunn and Hughes 1997). Some models indicate that ecological drift can
74 even reduce competition asymmetries in small communities to a level that strong and weak
75 competitors become effectively neutral (Orrock and Watling 2010) – i.e., the negative effect of a
76 superior competitor on other species is relatively small compared to the effects of demographic
77 stochasticity. Indeed, in an experiment with annual plants, Gilbert and Levine (2017) found that
78 larger communities converged to a state in which a strong competitor dominated after three years,

79 whereas in smaller communities the strong competitor co-occurred with other species at different
80 densities. Thus, both theory and experimental evidence suggest that ecological drift can change the
81 structure of local communities by altering species relative abundances as well as species
82 occurrences. Investigating the relationship between variation in community structure and
83 community size would help us better understanding the relative importance of deterministic and
84 stochastic processes on beta diversity. This should also be relevant from an applied perspective, as
85 many types of environmental disturbances, such as deforestation and floods, tend to reduce the size
86 of local communities (Barnes et al. 2014, Petsch et al. 2015).

87 Most methods commonly used to estimate β-diversity are affected by differences in species
88 richness or species abundance distributions (Chase and Myers 2011). This is undesirable for
89 understanding the beta diversity-community size relationship for two main reasons. First, estimates
90 of β-diversity can be influenced by random sampling effects that are neutral with respect to species
91 identity (Chase and Myers 2011, Myers et al. 2013, 2015). For example, let us assume that the
92 species composition of 10 local communities embedded in a larger metacommunity represents each
93 a random subset (say 5 species) of the regional species pool with, for instance, 100 species. This
94 would likely result in high β-diversity as, by chance, many pairs of local communities would not
95 share species. Second, studies employing incidence-based estimates of β-diversity (e.g., Jaccard or
96 Sørensen dissimilarity indices) do not capture variation in species relative abundances (Anderson et
97 al. 2011). In this case, if the same species occur at two sites but with different relative abundance,
98 β-diversity should not be equal. A solution to the first issue is to use a null model to produce
99 expected values under random assembly from a large species pool, contrast observed and expected
100 values and use the difference between them as estimates of β-deviation; i.e., the extent to which β-
101 diversity differs from null expectations (Kraft et al. 2011, Myers et al. 2013, 2015, Catano et al.
102 2017). In this case, positive and negative values of β-deviation indicate that communities are more
103 dissimilar and less dissimilar than expected by chance, respectively. Beta deviation values close to
104 zero indicate communities are as dissimilar as expected by chance (Kraft et al. 2011, Chase et al.

105 2011, Catano et al. 2017). Beta deviation provides a means to determine whether β-diversity
106 deviates from the expectations of stochastic assembly and if the magnitude of the deviation is
107 related to community size. A solution to the second issue is to analyze the data with dissimilarity
108 coefficients that take into account species composition and species relative abundance (e.g., Bray-
109 Curtis; Anderson et al. 2011). Such dissimilarity coefficients provide complementary information
110 regarding the main mechanisms responsible for β-diversity patterns (e.g., Siqueira et al. 2015).
111 In this study, we tested the hypothesis that ecological drift is a major process causing
112 variation among small communities. To reach our goal, we used a unique, comprehensive dataset
113 on aquatic insect communities sampled using identical methods in a total of 200 streams in two
114 climatically highly different regions (100 in Brazil and 100 in Finland). The sampling design
115 included 5 streams (communities) per watershed (20 in each region) and provided us replicates of
116 metacommunities (watersheds; Fig. S1). This allowed us to make specific predictions considering a
117 community size gradient. Our previous study showed that local community size is, on average, five-
118 fold larger in boreal than in tropical streams, with the smallest boreal stream communities being as
119 large as the largest tropical communities (Heino et al. 2018). Thus, we expected that β-diversity
120 would be high (E1) and close (E2) to null expectations (random assembly from the regional pool) in
121 watersheds with the smallest communities (some watersheds in Brazil only; Fig. 1). This would
122 indicate that ecological drift plays a major role in structuring these small subtropical communities.
123 Second, all else being equal, we expected that (E2a) β-deviation would be positive in all
124 watersheds, but higher in the largest communities (Fig. 1). This would indicate that niche selection
125 and sufficient dispersal rates are the main processes resulting in large communities to be more
126 dissimilar than expected by chance, as species sorting occurs when dispersal is sufficient to allow
127 individuals to reach sites that match their ecological requirements (Winegardner et al. 2012, Leibold
128 and Chase 2018). Although dispersal limitation can also cause positive values of β-deviation, this
129 was unlikely in our study system as the watersheds we studied are not large enough to lead to strong
130 dispersal limitation (see details below). Taken together, these two expectations would lead to a

negative relationship between β-diversity (before controlling for sampling effects) and community size, but to a positive relationship between β-deviation and community size (Fig. 1). Finally, because environmental heterogeneity and spatial extent are also expected to increase β-diversity (Heino et al. 2015a), we included them as co-variates in our models (Fig. S1).

Material and methods

Study area and sampling

The streams sampled in Brazil and Finland were distributed across a strong gradient of land cover among watersheds. It is well established that watershed land cover influences the structure of stream communities (Hynes 1975, Allan 2004, Roque et al. 2010, Siqueira et al. 2015). In Brazil, we sampled 100 streams distributed in 20 watersheds located in the southeastern region of the country – i.e., five streams per watershed, with spatial extent of 70 km in north-south and 120 km in east-west directions (Fig. S2). These streams drain through three major Atlantic Forest protected areas ('Carlos Botelho', 'Intervales' and 'Alto Ribeira' State Parks; São Paulo State). The watersheds are mainly dominated by pastures and planted forests (*Eucalyptus* and *Pinus*). The region has a dry season from April to August (average rainfall from 45 to 80 mm per month; average temperature from 16 to 20 ^{o}C) and a wet season from September to March (average rainfall from 105 mm to 180 mm per month; average temperature from 20 to 23 ^{o}C). Sampling was done between September and November in 2015.

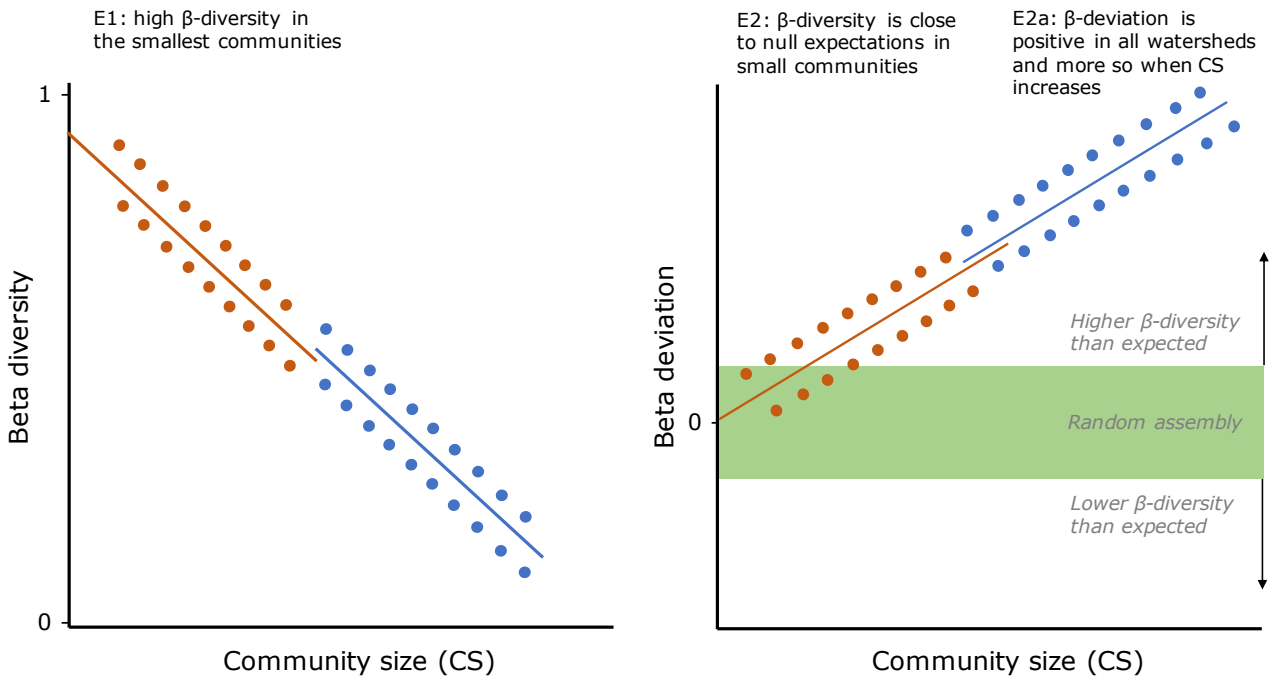


Figure 1. Graphical representation describing *a priori* expectations (E) about the relationship between beta diversity and beta deviation with community size in Brazil (vermilion) and Finland (blue).

The study sites in Finland were situated in the western part of the country. We sampled 100 streams that were distributed in 20 watersheds (as described above) with spatial extents of ca. 500 km and 300 km in north-south and in east-west directions (Fig. S2). The streams drain watersheds covered with agriculture and boreal forests. Western Finland has four typical seasons: a long winter that lasts from November to March, a short spring from April to May, a short summer from June to August, and an autumn period from September to October. Sampling was done in September 2014. Streams were generally of the same order within each watershed, but varied among watersheds, including 2nd and 3rd order streams in Brazil, and a few 4th order streams in Finland. In Brazil, maximum distances between pairs of streams within watersheds varied from 2.48 to 8.86 km, whereas in Finland it varied from 12.77 to 109.5 km. Most in-stream abiotic variables varied within a similar range and had similar mean values between regions, except conductivity, total nitrogen

168 and total phosphorus, which were much higher in Finland. A more detailed description of both
169 regions can be found in Heino et al. (2018).
170 At each of the 100 stream sites in both regions, we took a 2-minute kick-net sample (net
171 mesh size: 0.5 mm), which was composed of four 30-seconds sample units obtained in the main
172 microhabitats at a riffle site (e.g., which considered differences in current velocity, depth, benthic
173 particle size and macrophyte cover). The four sample units were pooled, preserved in alcohol in the
174 field and taken to the laboratory for further processing and identification. All insects were separated
175 from debris and the following taxonomic orders were identified to genus level: Ephemeroptera,
176 Odonata, Plecoptera, Megaloptera, Trichoptera and Coleoptera. We sampled 16,113 individuals,
177 distributed among 83 genera in Brazil, and 86,048 individuals (77 genera) in Finland. The mean
178 number of genera per stream was 17.84 (standard deviation = 7.46) in Brazil and 14.01 (sd = 5.07)
179 in Finland, while the mean number of individuals per stream was 181.50 (sd = 111.38) and 886.57
180 (sd = 700.73), respectively (Heino et al. 2018).
181 We adopted the definition of community size provided by Orrock and Watling (2010) and
182 estimated local community size as the number of individuals sampled in a stream site. However, as
183 β-diversity was estimated for each watershed (i.e., considering 5 stream sites; see below), we
184 averaged the number of individuals across five streams within each watershed. This resulted in 20
185 values of community size, one per watershed (Fig. S1).
186
187 *Beta diversity and beta deviations*
188 We first estimated β-diversity by using the Sørensen incidence-based coefficient and the
189 Bray-Curtis abundance-based coefficient. To do that, we calculated dissimilarity values between all
190 pairs of five streams within each of the 20 watersheds, separately for the tropical and boreal
191 datasets. Our measure of β-diversity was represented by the mean of these values in each watershed.
192 Thus, we estimated 20 values of Sørensen β-diversity and 20 values of Bray-Curtis β-diversity per
193 region (Fig. S1).

We estimated β-deviations by using two procedures based on null models. To estimate incidence-based β-deviations that accounted for random sampling effects, we used a modified version of the Raup-Crick coefficient following Chase et al. (2011) and the procedures described in the package vegan (Oksanen et al. 2018): (i) we defined the genus pool as all genera occurring in each region and their observed occupancy across all 100 streams; (ii) an algorithm calculated the number of genera that every pair of streams share (SG_{obs}); (iii) and assembled local stream communities by randomly sampling genus from the pool until reaching the local (observed) genus richness and by using the observed genus' occupancy frequency to determine the probability to sample a genus; (iv) step (iii) was repeated 10,000 times to generate 10,000 random matrices of genus composition and, posteriorly, the number of genera shared between each pair of streams within each watershed (SG_{exp}); (v) β-deviations were calculated as the number of random draws in which $SG_{exp} \geq SG_{obs}$ divided by the total number of random draws. This index was rescaled to range between -1 and 1, "indicating whether local communities are more dissimilar (approaching 1), as dissimilar (approaching 0), or less dissimilar (approaching -1), than expected by random chance" within each watershed (Chase et al. 2011). As with β-diversity, this procedure resulted in 20 values of incidence-based β-deviation per region, with one value per watershed (Fig. S1).

As selection and drift act on individuals, but not adding or removing entire populations, the use of abundance-based metrics should provide complementary information about assembly mechanisms. Thus, to estimate abundance-based β-deviations, we followed the procedure described by Kraft et al. (2011): (i) we defined the genus pool as all genera and their abundances across all 100 streams in each region; (ii) an algorithm then randomly assigned individuals to stream sites while preserving the overall genus-abundance distribution in the region and the total number of individuals in each stream site (local community size); (iii) step (ii) was repeated 10,000 to generate a distribution of pairwise dissimilarities within each watershed assuming random colonization of individuals from the genera pool; (iv) β-deviations were calculated as the difference between the observed Bray-Curtis dissimilarity and the mean expected (i.e. simulated) dissimilarity, divided by

220 the standard deviation of the (null) simulated distribution within each watershed. Positive values
221 and negative values indicate greater and lower dissimilarity than expected from random assembly,
222 respectively. This null model allows one to analyze how β-diversity differs from patterns generated
223 without the processes that cause clumping of species across the landscape (Kraft et al. 2011, Myers
224 et al. 2015). These procedures also resulted in 20 values of abundance-based β-deviation per region,
225 one value per watershed (Fig. S1).

226 We used linear regression models to test whether β-diversity (and β-deviation) was related
227 to community size. Environmental heterogeneity within watersheds was estimated using
228 environmental data and a Permutational analysis of multivariate dispersions (PERMDISP), as
229 described by Anderson et al. (2006). This analysis was based on a Euclidean distance calculated
230 with the standardized values (zero mean and unit variance) of the following variables: current
231 velocity (m/s), depth (cm), stream width (cm), % of sand (0.25-2 mm), gravel (2-16 mm), pebble
232 (16-64 mm), cobble (64-256 mm), and boulder (256-1024 mm), % of canopy cover by riparian
233 vegetation, pH, conductivity, total nitrogen, and total phosphorus. These were all measured at the
234 stream riffle scale (for details, see Heino et al. 2018). We also included the following watershed-
235 scale variables estimated through satellite images within a 400-m buffer along tracts of the sampled
236 streams: average slope, % of native forest cover, pasture, agriculture, planted forests, urban areas,
237 mining, water bodies, bare soil, secondary forest cover, and mixed land uses. We mapped land use
238 and land cover of Brazilian and Finnish watersheds using 5-m resolution RapidEye multispectral
239 imagery and the CORINE database (https://land.copernicus.eu/pan-european/corine-land-cover),
240 respectively. Geographical coordinates of the sampling sites were transformed to a Euclidean
241 distance matrix and then submitted to a PERMDISP procedure (Anderson et al. 2006), using
242 watershed as a grouping variable, to estimate mean spatial extent. Thus, we regressed our response
243 variables (β-diversity and β-deviation; with 20 values; one per watershed) against community size,
244 environmental heterogeneity, and spatial extent separately for each region (Fig. S1). We used the
245 *vegan* package (Oksanen et al. 2018) in R (version 3.5.0; R Core Team 2018) to estimate β-

diversity, Raup-Crick β-deviation, environmental heterogeneity, and spatial extent. In the tropical data, the strongest correlation between explanatory variables was between environmental heterogeneity and spatial extent (r = 0.47), whereas in the boreal data, it was between community size and spatial extent (r = 0.36). All other correlations were lower than 0.07. Thus, multicollinearity was not an issue in our study. Code and data that fully reproduce the analyses are archived on Zenodo (https://doi.org/10.5281/zenodo.2620550).

Results

There was a strong negative relationship between β-diversity (both with incidence- and abundance-based data) and community size in Brazil (b = -0.87, t = -7.47, $p < 0.001$, R^2 = 0.74; and b = -0.78, t = -5.32, $p < 0.001$, R^2 = 0.59, respectively; Fig. 2), indicating that small communities were more dissimilar among each other than larger communities, as we expected (E1). However, we found that neither incidence-based (Sørensen) nor abundance-based (Bray-Curtis) β-diversity were related to community size in Finland (b = -0.14, t = -0.59, p = 0.563; and b = -0.22, t = -0.94, p = 0.357, respectively).

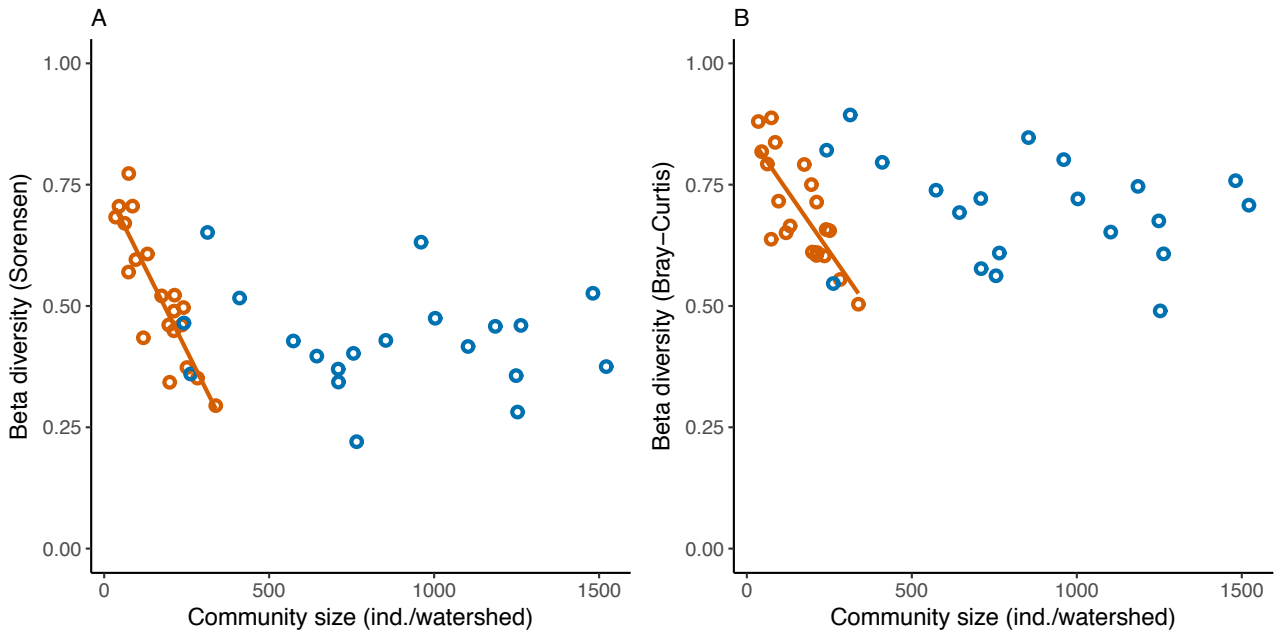


Figure 2. (A) Incidence-based (Sørensen) and (B) abundance-based (Bray-Curtis) beta diversity-community size (average number of individuals per watershed) relationships within tropical (vermilion) and boreal (blue) stream watersheds (n = 20 for each region). The average number of individuals per watershed was calculated with a sample size of 5 streams.

We found strong relationships between β-deviation (i.e., β-diversity apart from null expectations) and community size in both regions, but, contrary to our expectations, they varied according to the type of dissimilarity coefficient (incidence- and abundance-based) and the direction of relationship (negative and positive; Table 1, Fig. 3). As we expected (E2), beta diversity of tropical smaller communities was closer to null expectations than those of larger communities (Fig. 3). However, differently from what we expected, mean Raup-Crick β-deviations were negative in Brazil, varying from -0.98 to -0.02, indicating that tropical communities were less dissimilar in genus composition than expected by random sampling from the species pool. Neither environmental heterogeneity nor spatial extent were related to Raup-Crick β-deviation in Brazil. Raup-Crick β-deviation in boreal watersheds (ranging from -0.99 to 0.30) was not related to community size (Fig. 3A) but it was to environmental heterogeneity (Table 1).

Table 1. Relationship between beta deviation (incidence-based [Raup-Crick] and abundance-based [Bray-Curtis]) and community size, environmental heterogeneity and spatial extent (n = 20 watersheds in each region). CS = Community size; EH = Environmental heterogeneity; SE = Spatial extent. R^2 and adj. R^2 = coefficient of determination and adjusted coefficient of determination of the full model, respectively. b = standardized regression coefficient.

			b	*se*	*t*	*P*	R^2	adj. R^2
Brazil	Raup-Crick	CS	**-0.737**	**0.164**	**-4.505**	**<0.001**	0.510	0.419
		EH	-0.114	0.200	-0.572	0.575		
		SE	0.057	0.200	0.289	0.776		
	Bray-Curtis	CS	**0.777**	**0.147**	**5.279**	**<0.001**	0.658	0.594
		EH	0.211	0.167	1.267	0.223		
		SE	-0.087	0.167	-0.522	0.609		

Finland	Raup-Crick	CS	-0.252	0.221	-1.137	0.271	0.319	0.192
		EH	**0.461**	**0.207**	**2.229**	**0.040**		
		SE	0.338	0.222	1.527	0.146		
	Bray-Curtis	CS	**0.728**	**0.168**	**4.328**	**<0.001**	0.607	0.534
		EH	-0.115	0.157	-0.731	0.475		
		SE	0.087	0.168	0.518	0.611		

Bray-Curtis (abundance-based) β-deviation was positively related to community size in both regions (Table 1; Fig. 3B), supporting our expectations (E2a). In tropical watersheds, mean Bray-Curtis β-deviations ranged from 3.08 to 16.66, indicating that communities were more dissimilar than expected by random changes in species abundances and genus composition. Again, β-deviations of smaller communities were closer to zero than those of larger tropical communities, supporting our expectation E1. In boreal watersheds, mean Bray-Curtis β-deviation varied from 17.40 to 102.51, also indicating that communities were more dissimilar than expected by random changes in genus abundances and composition. Neither environmental heterogeneity nor spatial extent were related to Bray-Curtis β-deviation in any region.

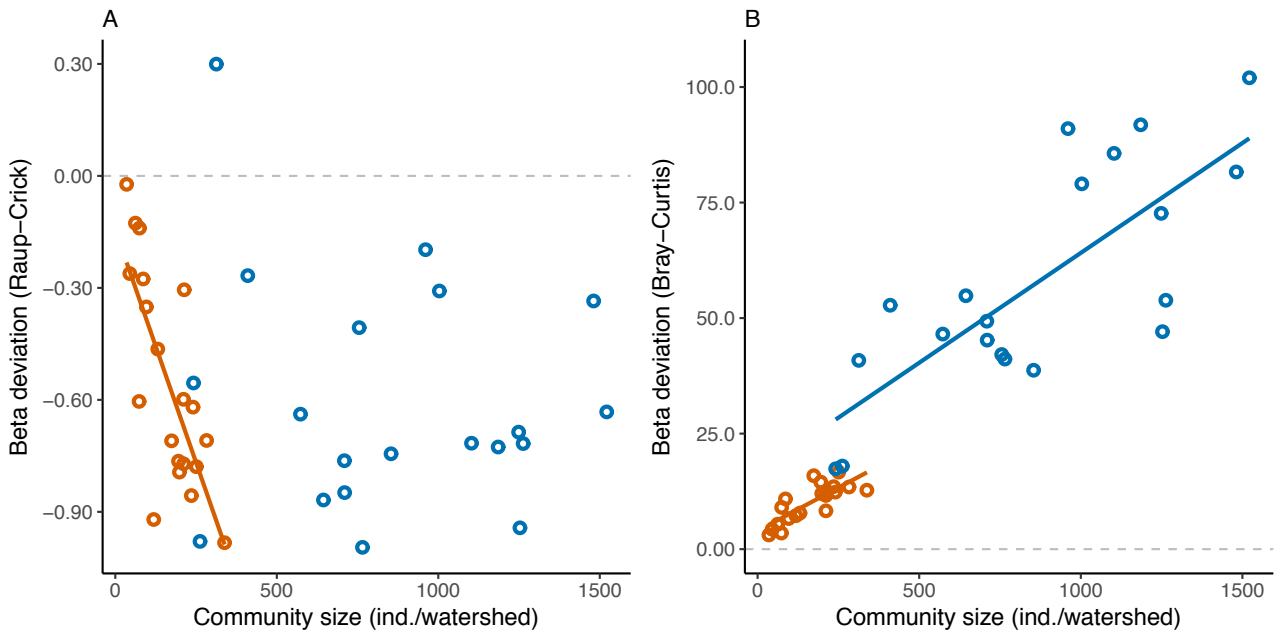


Figure 3. (A) Incidence-based (Raup-Crick) and (B) abundance-based (Bray-Curtis) beta deviation-community size (average number of individuals per watershed) relationships within tropical

301 (vermilion) and boreal (blue) stream watersheds (n = 20 for each region). The average number of
302 individuals per watershed was calculated with a sample size of 5 streams. The dashed grey line
303 indicates expected beta diversity under null assembly.
304
305 Although there was a negative relationship between Raup-Crick β-deviation and community
306 size in tropical watersheds only, mean incidence-based β-deviation in Brazil was not different from
307 mean incidence-based β-deviation in Finland (Fig. 3A). This suggests that the departure from null
308 expectations related to compositional change among streams is similar in tropical and boreal
309 regions. On the other hand, mean abundance-based β-deviation was five times higher in Finland
310 than in Brazil, suggesting that departures from null expectations are much higher in boreal than in
311 tropical streams (Fig. 3B).
312
313 **Discussion**
314 The role of stochastic community assembly processes in arranging species within and
315 among communities has gained support from models (Mouquet and Loreau 2003, Orrock and
316 Fletcher 2005, Durães et al. 2016) and data (Cottenie 2005, Lancaster and Downes 2017, Germain
317 et al. 2017, Swan and Brown 2017, Valente-Neto et al. 2017). Here, we provide empirical evidence
318 that community size, a simple characteristic of communities, may mediate the interplay between
319 niche selection and ecological drift as drivers of β-diversity in tropical and boreal
320 metacommunities. Based on null models, our estimates of variation in community structure
321 accounted for both differences in species richness and species relative abundance that deviate from
322 random assembly (i.e., β-deviations). We showed that β-diversity of smaller tropical communities
323 deviates less from null expectations than larger communities. This means that the high β-diversity
324 we observed among smaller communities was, to some extent, indistinguishable from patterns
325 generated by random changes in local genus richness and community size. As findings generated by
326 null models are our best approximation of patterns generated by stochastic processes (Kraft et al.

327 2011, Chase et al. 2011), we suggest that demographic stochasticity plays a major role in small
328 ecological communities (Orrock and Fletcher 2005, Orrock and Watling 2010, Gilbert and Levine
329 2017).

330 Explanations for the major role of demographic stochasticity in small communities involve
331 the alteration of the occupancy frequency and relative abundance of species with different fitness
332 (Orrock and Watling 2010). When local communities are small, even species with high fitness are
333 at a high risk of extinction due to demographic stochasticity in comparison to a situation when
334 communities have large populations. Consequently, species with low relative abundances have a
335 chance to increase their populations in small communities (Orrock and Watling 2010, Gilbert and
336 Levine 2017). As this would be the result of a local demographic random process, and because
337 dispersal ability is highly variable among species in riverine systems (Heino et al. 2015b, Tonkin et
338 al. 2018a), the outcome of the assembly would likely differ among local communities, increasing β-
339 diversity within metacommunities. Indeed, in Brazil, the three most abundant and widespread
340 genera (Heino et al. 2018) were not dominant in smaller communities: the mayfly *Farrodes* was not
341 among the most abundant in any of the smallest communities; the beetle *Heterelmis* was among the
342 most abundant in two communities only; and the caddisfly *Smicridea* was among the most abundant
343 in three of the five smallest communities. These three genera were dominant in four of the five
344 largest communities in Brazil. The smallest communities in Brazil were dominated by genera with
345 intermediate regional abundance and occupancy. For example, *Gripopteryx* (stonefly), *Cloeodes*
346 (mayfly) and *Callibaetis* (mayfly) were the sixth, eleventh and sixteenth most abundant genera in
347 Brazil, respectively. In Finland, these differences were less evident, as the three most abundant and
348 widespread genera dominated small and large communities: the beetle *Elmis* (two of the smallest
349 vs. three of the largest communities); the mayfly *Baetis* (three vs. three); and stonefly *Nemoura*
350 (one vs. one). This is likely because the smallest boreal stream communities were as large as the
351 largest tropical communities (Heino et al. 2018).

352 The positive relationship between abundance-based β-deviation and community size in both
353 regions is in line with our expectations. The positive slope and values of β-deviation indicate that
354 the variation in community structure among streams increased with community size more than what
355 was expected under random assembly. Thus, in terms of which genera were more abundant or less
356 abundant, and more aggregated or less aggregated, communities within the same watershed differed
357 from each other more than expected by chance, especially in mid to large communities. As dispersal
358 within watersheds was likely not limited, this positive relationship indicates that niche selection was
359 sufficient to cause non-random variations in genera relative abundance and aggregation patterns
360 among large communities. We suggest that as community size increases, demographic stochasticity
361 becomes less important and niche selection determines which species are more abundant locally and
362 widely distributed within the metacommunity. In this case, small random variations in the number
363 of individuals of relatively abundant genera occurring in larger communities, which can only be
364 detected with abundance-based β-diversity metrics, do not result in major changes in genus
365 occurrence.

366 Although small communities had incidence-based β-deviations values close to random
367 assembly expectations, as we expected, the relationship between β-deviation and community size
368 was negative in Brazil. This result contradicts our expectations and indicates that the genus
369 composition of streams harboring mid to large communities in tropical watersheds is less dissimilar
370 (negative values of incidence-based β-deviation) than patterns predicted by random assembly – i.e.,
371 these communities share more genera than expected. In general, communities can be less dissimilar
372 than random expectations in at least two cases. First, dissimilarity should be low when niche
373 selection is spatially constant, as the environment maximizes the fitness of a few species (Vellend
374 2016). Within watershed environmental heterogeneity was weakly and positively related with
375 incidence-based β-deviation in Finland only, where community sizes are on average five-fold larger
376 than in tropical streams (Heino et al. 2018). This suggests a tendency towards environmental
377 determinism; the higher the environmental variation within watersheds, the higher their β-deviation.

Thus, it is likely that even the large size of boreal communities was not sufficient to allow niche selection to be the main driver of spatial variation in genus composition among those communities. Alternatively, we might have failed to measure key abiotic and biotic variables (e.g., biological interactions) underlying community variation, or despite including key correlates of beta diversity, our snapshot analysis might not have been enough to represent the complexity of all mechanisms involved in niche selection in watersheds.

A second reason for communities to be less dissimilar than random expectations is when dispersal rates are high within the metacommunity (Mouquet and Loreau 2003, Leibold and Chase 2018). By distributing organisms among communities, dispersal can reduce β-diversity if it is excessive and is combined with a source-sink system, where populations with high growth rates supply individuals to other localities where they would otherwise be excluded by niche selection (Mouquet and Loreau 2003, Siqueira et al. 2014). Although dispersal across watersheds tends to be limited for many aquatic species, those with an adult flight stage may have higher dispersal rates especially along the stream channel (Hughes 2007, Lancaster and Downes 2017). A meta-analysis by Muehlbauer et al. (2014) showed that most adult aquatic insects tend to fly ca. 1.5 m around their natal stream, but that a few individuals can fly 550 m away from the stream, with some caddisflies being able to reach sites distant more than 650 m. Also, while Macneale et al. (2005) found that the stonefly *Leuctra ferruginea* was able to fly across headwater forested catchments, Flenner and Sahlin (2008) estimated annual range expansions of up to 88 km in non-migratory dragonflies in a boreal region of Sweden. The northern faunas of Europe should also be composed of species with good dispersal ability, as they have reached these areas since the Last Glacial Period. On the other hand, many tropical aquatic insects have multiple reproduction events per year, i.e., many are multivoltine species (Wallace and Anderson 1996, Vásquez et al. 2009). Unless local population growth rate is low, multivoltine species should have many opportunities per year to disperse from their natal streams. If a few of these multivoltine species can fly 250 m (a conservative estimate considering the results by Muehlbauer et al. 2014), after some generations,

404 some individuals of these species could reach distant sites, at a velocity of one kilometer per year.
405 We thus suggest that long-term availability for colonization and multivoltinism allowed some
406 genera to reach widespread distribution within tropical watersheds (Saito et al. 2016). Similarly,
407 good dispersal abilities of northern species may have masked the role of within-watershed
408 environmental heterogeneity in the boreal region, making these communities less dissimilar than
409 expected. Thus, as communities become larger, demographic stochasticity plays a less important
410 role, while dispersal surplus plays a major role in homogenizing the genus composition of streams
411 within tropical watersheds. However, this, as almost everything in ecology, should be scale
412 dependent (Chase et al. 2018). If our watersheds were larger (in extent), dispersal limitation would
413 likely have played a role and niche selection could have been the major driver of community
414 structure. Our results reinforce previous suggestions that dispersal should be viewed as a process
415 occurring over time and not only considering the distance individuals can move in one dispersal
416 event during one generation (Saito et al. 2015).

417 Predicting community assembly dynamics has been a major challenge for ecology in a
418 changing world (Mouquet et al. 2015). Many drivers of environmental change, such as habitat
419 destruction, overexploitation, pollution, and reductions in landscape connectivity, are likely to cause
420 reductions in community size. For example, Hallmann et al. (2017) estimated a decline of more than
421 70% in insect biomass between 1989 and 2016 in Germany and Lister and Garcia (2018) found that
422 arthropod biomass has fallen 10 to 60 times since 1970 in Puerto Rico. It is possible, then, that
423 results from previous studies associating high β-diversity with environmental changes were related
424 to small communities being more variable. For example, Hawkins et al. (2015) showed that
425 disturbed sites had high β-diversity due to a decreased prevalence of more common taxa. Although
426 they did not analyze community size per se, the mechanism they evoked to explain this pattern –
427 i.e., the relative abundance of species with lower fitness become progressively higher – is what has
428 been suggested as the reason for a major role of ecological drift in small communities (Orrock and
429 Watling 2010, Gilbert and Levine 2017). Additionally, the effects of environmental changes on

430 naturally small communities might be combined with the effects of ecological drift in non-obvious
431 ways, making community dynamics even harder to predict (e.g., Bini et al. 2014) and smaller
432 communities more vulnerable to novel environmental conditions, such as altered flow regimes
433 (Tonkin et al. 2018b, Ruhi et al. 2018). This should also be relevant for the conservation and
434 management of ecosystems. For example, stream restoration efforts tend to be directed towards
435 sites with low habitat quality and with reduced number of species and low-density populations.
436 Even if restoration efforts overcome dispersal constraints and barriers to recolonization, which are a
437 major cause for unsuccessful stream restoration (Bond and Lake 2003, Sundermann et al. 2011,
438 Tonkin et al. 2014), restored communities will likely be small at the earlier stages of community
439 assembly and, thus, more prone to ecological drift.

440 Our study provides empirical evidence of the role that community size can play in mediating
441 stochastic and deterministic process as drivers of metacommunity dynamics in tropical and boreal
442 streams, providing one possible solution to a long-standing debate about why some communities
443 are apparently more influenced by stochastic processes than others. The magnitude of β-diversity
444 deviation from the null models (negative and positive values) and how it related with community
445 size (negative and positive slopes) indicate that stochastic and deterministic processes affect species
446 occurrences and their abundances differently. While incidence-based β-deviation was negative and
447 decreased with community size in Brazil, abundance-based β-deviation was positive and increased
448 with community size in both regions. These results indicate that: (1) as communities get larger,
449 demographic stochasticity plays a less important role and excessive dispersal combined with niche
450 selection tend to homogenize the genus composition of larger communities; and (2) niche selection
451 was likely strong enough to produce variations in genera relative abundances, but not to cause
452 compositional changes among communities within the same watershed. However, we acknowledge
453 that these are, at least partially, *a posteriori* explanations as we expected that the results would be
454 similar (i.e., a positive relationship between beta deviation and community size) independently of
455 the type of coefficient. On the other hand, these contrasting results appear interesting avenues for

future research. Our findings complement previous empirical and conceptual efforts that indicated that ecological drift can drive biodiversity patterns even in communities where species have clear differences in life-history traits, resource use and competitive abilities (i.e., those communities assumed to be non-neutral). Incorporation of community size into ecological models should provide conceptual, empirical and applied insights towards better understanding of the processes driving changes in biodiversity.

Acknowledgements. This research originated during a joint research project funded by grant #2013/50424-1 (São Paulo Research Foundation, FAPESP) to TS, and by grants #273557 and #273560 to JH and JS (Academy of Finland). ASM and LMB are supported by research grants from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq no. 307587/2017-7 and 304314/2014-5, respectively). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 (scholarship to DKP). This work was also developed in the context of the National Institutes for Science and Technology (INCT) in Ecology, Evolution and Biodiversity Conservation, supported by MCTIC/CNPq (proc. 465610/2014-5) and FAPEG. We thank Christopher Catano and Jonathan Myers for gently providing R code to generate abundance-based beta deviations, and Mauricio Vancine for preparing a map of the regions. We thank Eric Harvey, Kevin Cazelles, and Romain Bertrand for their comments and suggestions on earlier versions of this manuscript, which was submitted to PCI Ecology.

Author contributions. TS conceived the idea, ran the analyses, and wrote the manuscript with substantial support from VSS, LMB, ASM, and JH. TS, LMB, VLL, ASM, JS, and JH planned the sampling design. DKP and JJ-M sampled the data. DKP, VSS and KTT identified the insects. DKP, JJ-M, KT, VLL and JS commented on the manuscript.

References

Allan, J. D. 2004. Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems. Annual Review of Ecology, Evolution, and Systematics 35:257–284.

Anderson, M. J., T. O. Crist, J. M. Chase, M. Vellend, B. D. Inouye, and A. L. Freestone. 2011. Navigating the multiple meanings of b diversity: a roadmap for the practicing ecologist. Ecology Letters 14:19–28.

Anderson, M. J., K. E. Ellingsen, and B. H. McArdle. 2006. Multivariate dispersion as a measure of beta diversity. Ecology Letters 9:683–693.

Barnes, A. D., M. Jochum, S. Mumme, N. F. Haneda, A. Farajallah, T. H. Widarto, and U. Brose. 2014. Consequences of tropical land use for multitrophic biodiversity and ecosystem functioning. Nature Communications 5:5351.

Bini, L. M., V. L. Landeiro, A. A. Padial, T. Siqueira, and J. Heino. 2014. Nutrient enrichment is related to two facets of beta diversity for stream invertebrates across the United States. Ecology 95:1569–1578.

Bond, N. R., and P. S. Lake. 2003. Local habitat restoration in streams: Constraints on the effectiveness of restoration for stream biota. Ecological Management & Restoration 4:193–198.

Bunn, S. E., and J. M. Hughes. 1997. Dispersal and Recruitment in Streams: Evidence from Genetic Studies. Journal of the North American Benthological Society 16:338–346.

Catano, C. P., T. L. Dickson, and J. A. Myers. 2017. Dispersal and neutral sampling mediate contingent effects of disturbance on plant beta-diversity: a meta-analysis. Ecology Letters 20:347–356.

Chase, J. M., N. J. B. Kraft, K. G. Smith, M. Vellend, and B. D. Inouye. 2011. Using null models to disentangle variation in community dissimilarity from variation in α-diversity. Ecosphere 2:art24.

Chase, J. M., B. J. McGill, D. J. McGlinn, F. May, S. A. Blowes, X. Xiao, T. M. Knight, O. Purschke, and N. J. Gotelli. 2018. Embracing scale-dependence to achieve a deeper understanding of biodiversity and its change across communities. Ecology Letters 21:1737–1751.

Chase, J. M., and J. A. Myers. 2011. Disentangling the importance of ecological niches from stochastic processes across scales. Philosophical Transactions of the Royal Society B: Biological Sciences 366:2351–2363.

Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community dynamics. Ecology Letters 8:1175–1182.

Durães, L., F. O. Roque, T. Siqueira, A. M. Santos, M. A. Borges, and R. S. Rezende. 2016. Simulating the role of connectivity in shaping stream insect metacommunities under colonization cycle dynamics. Ecological Modelling 334:19–26.

Flenner, I., and G. Sahln. 2008. Dragonfly community re-organisation in boreal forest lakes: rapid species turnover driven by climate change? Insect Conservation and Diversity 1:169–179.

Germain, R. M., S. Y. Strauss, and B. Gilbert. 2017. Experimental dispersal reveals characteristic scales of biodiversity in a natural landscape. Proceedings of the National Academy of Sciences 114:4447–4452.

Gilbert, B., and J. M. Levine. 2017. Ecological drift and the distribution of species diversity. Proceedings of the Royal Society B: Biological Sciences 284:20170507.

Hallmann, C. A., M. Sorg, E. Jongejans, H. Siepel, N. Hofland, H. Schwan, W. Stenmans, A. Müller, H. Sumser, T. Hörren, D. Goulson, and H. de Kroon. 2017. More than 75 percent decline over 27 years in total flying insect biomass in protected areas. PLOS ONE 12:e0185809.

Hawkins, C. P., H. Mykrä, J. Oksanen, and V. L. J. J. 2015. Environmental disturbance can increase beta diversity of stream macroinvertebrate assemblages. Global Ecology and Biogeography 24:483–494.

533 Heino, J., A. S. Melo, and L. M. Bini. 2015a. Reconceptualising the beta diversity-environmental
534 heterogeneity relationship in running water systems. Freshwater Biology 60:223–235.
535 Heino, J., A. S. Melo, J. Jyrkänkallio-Mikkola, D. K. Petsch, V. S. Saito, K. T. Tolonen, L. M. Bini,
536 V. L. Landeiro, T. S. F. Silva, V. Pajunen, J. Soininen, and T. Siqueira. 2018. Subtropical
537 streams harbour higher genus richness and lower abundance of insects compared to boreal
538 streams, but scale matters. Journal of Biogeography 45:1983–1993.
539 Heino, J., A. S. Melo, T. Siqueira, J. Soininen, S. Valanko, and L. M. Bini. 2015b. Metacommunity
540 organisation, spatial extent and dispersal in aquatic systems: patterns, processes and
541 prospects. Freshwater Biology 60:845–869.
542 Hughes, J. M. 2007. Constraints on recovery: using molecular methods to study connectivity of
543 aquatic biota in rivers and streams. Freshwater Biology 52:616–631.
544 Hynes, H. B. N. 1975. The stream and its valley. Internationale Vereinigung Fur Theoretische und
545 Angewandte Limnologie, Verhandlungen 19:1–15.
546 Kraft, N. J. B., L. S. Comita, J. M. Chase, N. J. Sanders, N. G. Swenson, T. O. Crist, J. C. Stegen,
547 M. Vellend, B. Boyle, M. J. Anderson, H. V. Cornell, K. F. Davies, A. L. Freestone, B. D.
548 Inouye, S. P. Harrison, and J. A. Myers. 2011. Disentangling the Drivers of β Diversity
549 Along Latitudinal and Elevational Gradients. Science 333:1755–1758.
550 Lancaster, J., and B. J. Downes. 2017. A landscape-scale field experiment reveals the importance of
551 dispersal in a resource-limited metacommunity. Ecology 98:565–575.
552 Leibold, M. A., and J. M. Chase. 2018. Metacommunity Ecology. Princeton University Press,
553 Princeton, NJ.
554 Lister, B. C., and A. Garcia. 2018. Climate-driven declines in arthropod abundance restructure a
555 rainforest food web. Proceedings of the National Academy of Sciences 115:E10397–
556 E10406.

Macneale, K. H., B. L. Peckarsky, and G. E. Likens. 2005. Stable isotopes identify dispersal patterns of stonefly populations living along stream corridors. Freshwater Biology 50:1117–1130.

Mouquet, N., Y. Lagadeuc, V. Devictor, L. Doyen, A. Duputié, D. Eveillard, D. Faure, E. Garnier, O. Gimenez, P. Huneman, F. Jabot, P. Jarne, D. Joly, R. Julliard, S. Kéfi, G. J. Kergoat, S. Lavorel, L. Le Gall, L. Meslin, S. Morand, X. Morin, H. Morlon, G. Pinay, R. Pradel, F. M. Schurr, W. Thuiller, and M. Loreau. 2015. REVIEW: Predictive ecology in a changing world. Journal of Applied Ecology 52:1293–1310.

Mouquet, N., and M. Loreau. 2003. Community Patterns in Source-Sink Metacommunities. The American Naturalist 162:544–557.

Muehlbauer, J. D., S. F. Collins, M. W. Doyle, and K. Tockner. 2014. How wide is a stream? Spatial extent of the potential "stream signature" in terrestrial food webs using meta-analysis. Ecology 95:44–55.

Myers, J. A., J. M. Chase, R. M. Crandall, and I. Jiménez. 2015. Disturbance alters beta-diversity but not the relative importance of community assembly mechanisms. Journal of Ecology 103:1291–1299.

Myers, J. A., J. M. Chase, I. Jiménez, P. M. Jørgensen, A. Araujo-Murakami, N. Paniagua-Zambrana, and R. Seidel. 2013. Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. Ecology Letters 16:151–157.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, and H. Wagner. 2018. vegan: Community Ecology Package.

Orrock, J. L., and R. J. Fletcher Jr. 2005. Changes in Community Size Affect the Outcome of Competition. The American Naturalist 166:107–111.

Orrock, J. L., and J. I. Watling. 2010. Local community size mediates ecological drift and competition in metacommunities. Proceedings of the Royal Society B: Biological Sciences 277:2185–2191.

Petsch, D. K., G. D. Pinha, J. D. Dias, and A. M. Takeda. 2015. Temporal nestedness in Chironomidae and the importance of environmental and spatial factors in species rarity. Hydrobiologia 745:181–193.

Roque, F. O., T. Siqueira, L. M. Bini, M. C. Ribeiro, L. R. Tambosi, G. Ciocheti, and S. Trivinho-Strixino. 2010. Untangling associations between chironomid taxa in Neotropical streams using local and landscape filters. Freshwater Biology 55:847–865.

Ruhi, A., X. Dong, C. H. McDaniel, D. P. Batzer, and J. L. Sabo. 2018. Detrimental effects of a novel flow regime on the functional trajectory of an aquatic invertebrate metacommunity. Global Change Biology 24:3749–3765.

Saito, V. S., M. V. Cianciaruso, T. Siqueira, A. A. Fonseca-Gessner, and S. Pavoine. 2016. Phylogenies and traits provide distinct insights about the historical and contemporary assembly of aquatic insect communities. Ecology and Evolution 6:2925–2937.

Saito, V. S., J. Soininen, A. A. Fonseca-Gessner, and T. Siqueira. 2015. Dispersal traits drive the phylogenetic distance decay of similarity in Neotropical stream metacommunities. Journal of Biogeography 42:2101–2111.

Siqueira, T., L. D. Durães, and F. de O. Roque. 2014. Predictive Modelling of Insect Metacommunities in Biomonitoring of Aquatic Networks. Pages 109–126 *in* C. P. Ferreira and W. A. C. Godoy, editors. Ecological Modelling Applied to Entomology. Springer International Publishing.

Siqueira, T., C. G.-L. T. Lacerda, and V. S. Saito. 2015. How Does Landscape Modification Induce Biological Homogenization in Tropical Stream Metacommunities? Biotropica 47:509–516.

Sundermann, A., S. Stoll, and P. Haase. 2011. River restoration success depends on the species pool of the immediate surroundings. Ecological Applications 21:1962–1971.

Swan, C. M., and B. L. Brown. 2017. Metacommunity theory meets restoration: isolation may mediate how ecological communities respond to stream restoration. Ecological Applications 27:2209–2219.

Tonkin, J. D., F. Altermatt, D. S. Finn, J. Heino, J. D. Olden, S. U. Pauls, and David. A. Lytle. 2018a. The role of dispersal in river network metacommunities: Patterns, processes, and pathways. Freshwater Biology 63:141–163.

Tonkin, J. D., D. M. Merritt, J. D. Olden, L. V. Reynolds, and D. A. Lytle. 2018b. Flow regime alteration degrades ecological networks in riparian ecosystems. Nature Ecology & Evolution 2:86.

Tonkin, J. D., S. Stoll, A. Sundermann, and P. Haase. 2014. Dispersal distance and the pool of taxa, but not barriers, determine the colonisation of restored river reaches by benthic invertebrates. Freshwater Biology 59:1843–1855.

Valente-Neto, F., L. Durães, T. Siqueira, and F. O. Roque. 2017. Metacommunity detectives: Confronting models based on niche and stochastic assembly scenarios with empirical data from a tropical stream network. Freshwater Biology 63:86–99.

Vásquez, D., R. W. Flowers, and M. Springer. 2009. Life history of five small minnow mayflies (Ephemeroptera: Baetidae) in a small tropical stream on the Caribbean slope of Costa Rica. Aquatic Insects 31:319–332.

Vellend, M. 2010. Conceptual synthesis in community ecology. The Quarterly review of biology 85:183–206.

Vellend, M. 2016. The Theory of Ecological Communities. Mpb Series: 57 ed. edition. Princeton University Press, Princeton.

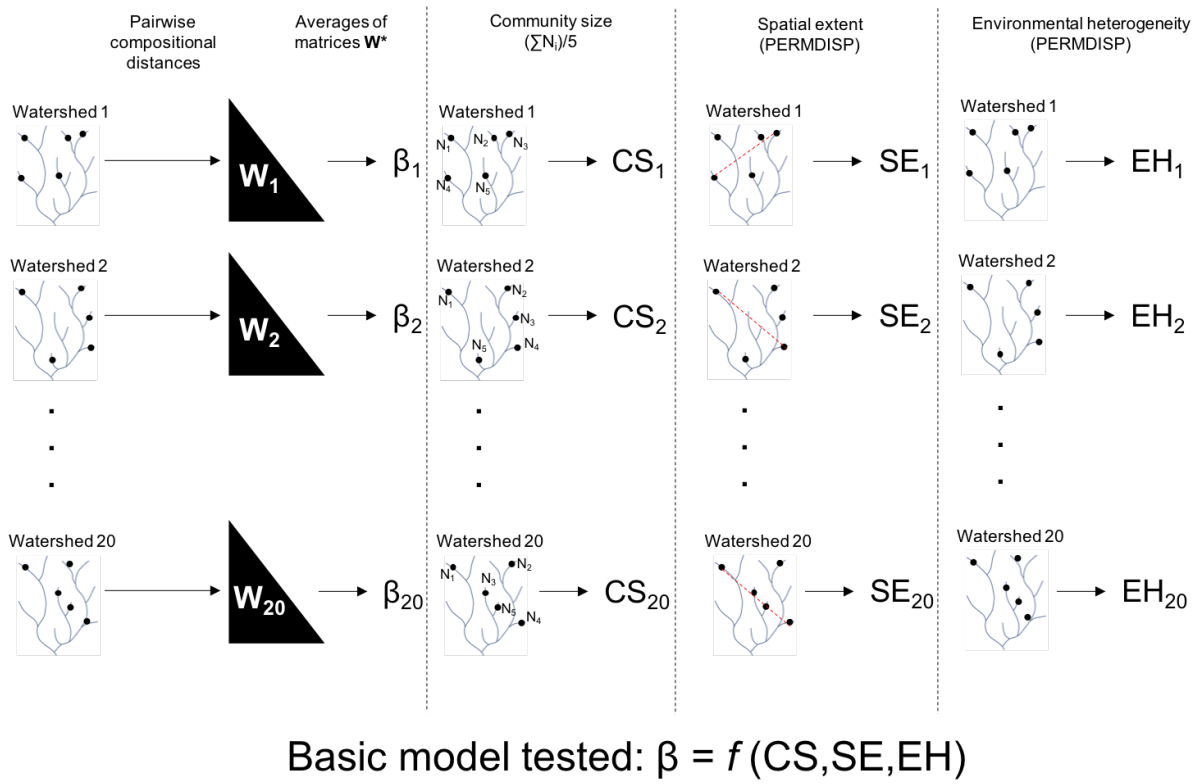
Vellend, M., D. S. Srivastava, K. M. Anderson, C. D. Brown, J. E. Jankowski, E. J. Kleynhans, N. J. B. Kraft, A. D. Letaw, A. A. M. Macdonald, J. E. Maclean, I. H. Myers-Smith, A. R. Norris, and X. Xue. 2014. Assessing the relative importance of neutral stochasticity in ecological communities. Oikos 123:1420–1430.

Wallace, J. B., and N. H. Anderson. 1996. Habitat, life history, and behavioral adaptations of aquatic insects. Pages 41–73 An Introduction to the Aquatic Insects of North America. Third edition. Kendall/Hunt Publishing Company, Dubuque, Iowa.

Winegardner, A. K., B. K. Jones, I. S. Y. Ng, T. Siqueira, and K. Cottenie. 2012. The terminology of metacommunity ecology. Trends in Ecology & Evolution 27:253–254.

Supplementary material

Figure S1. Graphical representation describing the general procedures for estimating beta diversity (β), community size (CS), spatial extend (SE) and environmental heterogeneity (EH) within watersheds (n = 20) in Brazil and Finland.



* This was repeated for different dissimilaity coefficients, and for null model analyses

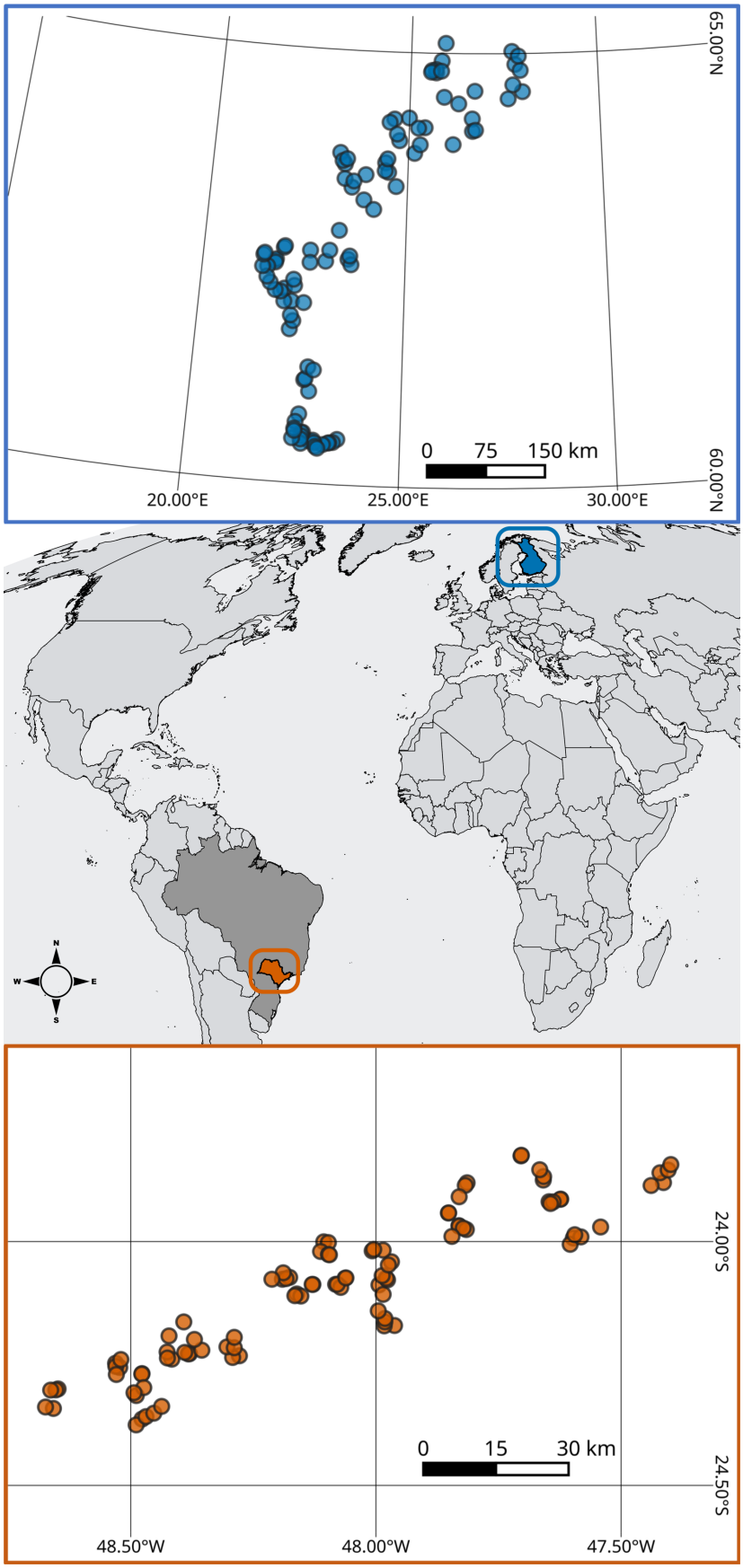



Figure S2. Geographical location and spatial distribution of streams in São Paulo State, Brazil (circles in vermilion) and Finland (circles in blue).