

## Research Article

**Cite this article:** Krasnov BR, Shenbrot GI, Khokhlova IS (2022). Regional flea and host assemblages form biogeographic, but not ecological, clusters: evidence for a dispersal-based mechanism as a driver of species composition. *Parasitology* **149**, 1450–1459. <https://doi.org/10.1017/S0031182022000907>

Received: 9 May 2022

Revised: 14 June 2022

Accepted: 17 June 2022

First published online: 5 July 2022

**Key words:**


Biogeographic realms; biome; dispersal; fleas; hosts; step-down factor analysis

**Author for correspondence:**

Boris R. Krasnov,

E-mail: [krasnov@bgu.ac.il](mailto:krasnov@bgu.ac.il)

# Regional flea and host assemblages form biogeographic, but not ecological, clusters: evidence for a dispersal-based mechanism as a driver of species composition

Boris R. Krasnov<sup>1</sup> , Georgy I. Shenbrot<sup>1</sup> and Irina S. Khokhlova<sup>2</sup>

<sup>1</sup>Mitrani Department of Desert Ecology, Swiss Institute for Dryland Environmental and Energy Research, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Sede Boqer Campus, 8499000 Midreshet Ben-Gurion, Israel and <sup>2</sup>Wyler Department of Dryland Agriculture, French Associates Institute for Agriculture and Biotechnology of Drylands, Ben-Gurion University of the Negev, Sede Boqer Campus, 8499000 Midreshet Ben-Gurion, Israel

**Abstract**

We used data on the species composition of regional assemblages of fleas and their small mammalian hosts from 6 biogeographic realms and applied a novel method of step-down factor analyses (SDFA) and cluster analyses to identify biogeographic (across the entire globe) and ecological (within a realm across the main terrestrial biomes) clusters of these assemblages. We found that, at the global scale, the clusters of regional assemblage loadings on SDFA axes reflected well the assemblage distribution, according to the biogeographic realms to which they belong. At the global scale, the cluster topology, corresponding to the biogeographic realms, was similar between flea and host assemblages, but the topology of subtrees within realm-specific clusters substantially differed between fleas and hosts. At the scale of biogeographic realms, the distribution of regional flea and host assemblages did not correspond to the predominant biome types. Assemblages with similar loadings on SDFA axes were often situated in different biomes and vice versa. The across-biome, within-realm distributions of flea vs host assemblages suggested weak congruence between these distributions. Our results indicate that dispersal is a predominant mechanism of flea and host community assembly across large regions.

**Introduction**

It is commonly accepted that the species composition of a biological community is driven by multiple mechanisms that act at both evolutionary and ecological scales (e.g. Connor and Simberloff, 1979; Chesson, 2000; Kraft and Ackerly, 2014). The main mechanisms determining species assembly are associated with either (a) the environmental requirements of the species coupled with interspecific interactions (e.g. Ackerly and Cornwell, 2007) or (b) stochastic speciation/extinction processes and biogeographic histories such as dispersal (e.g. Hubbell, 2001). An example of the former mechanism is the so-called ‘environmental filtering’ (e.g. van der Valk, 1981), in which the environment dictates the community composition by permitting it to be composed only of species possessing certain morphological, physiological, behavioural and/or ecological traits that allow a species to tolerate the biotic and abiotic conditions in that environment (e.g. Kraft and Ackerly, 2014). However, between-species similarities in these traits likely lead to an increase in the intensity of interspecific competition (e.g. Schoener, 1974). As a result, the overlap in the resource use by coexisting species must be limited (MacArthur and Levins, 1967). A community’s composition is thus seen as being driven by deterministic niche-based processes (Hubbell, 2001). In contrast, the predominance of stochastic speciation/extinction and biogeographic processes results in ‘neutral’ communities (Hubbell, 2001, 2006). Species composing ‘neutral communities’ and belonging to the same trophic level are considered equivalent in their fitness and competitive abilities, so that ‘neutral communities’ are shaped by dispersal limitations and population dynamics. Efforts aimed at reconciling niche-based and dispersal-based (= neutral) perspectives of community assembly have resulted in the conclusion that communities may be assembled by both types of processes, so that niche-assembled and dispersal-assembled communities form opposite ends of a continuum (Gravel *et al.*, 2006).

Studies aimed at elucidating assembly processes in parasite communities have produced contradictory results. The important role played by niche-based processes has been demonstrated in some parasite communities (e.g. Gutiérrez and Martorelli, 1999; Mouillot, 2007; Krasnov *et al.*, 2015a), whereas the predominance of dispersal-based assembly has been indicated in other ones (Mouillot *et al.*, 2003; Krasnov *et al.*, 2015b). Furthermore, the relative importance of niche-based vs dispersal-based assembly processes may be scale-dependent and may vary in dependence on whether the infra- (in a host individual), component (in a host population) or compound (in a locality) parasite community is considered (Holmes and Price, 1986). For example, both stochastic and niche-based mechanisms of species

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

assembly have been indicated for infracommunities of fleas harboured by small mammalian hosts (Krasnov *et al.*, 2006 *vs* Surkova *et al.*, 2018). The same was true for compound communities of these parasites (Krasnov *et al.*, 2015a *vs* Krasnov *et al.*, 2015b).

The relative importance of niche-based *vs* dispersal-based processes in parasite community assembly has rarely been specifically studied (Mouillot, 2007; Gibert *et al.*, 2021). Recently, Gibert *et al.* (2021) applied a permutation-based algorithm to infer whether regional flea communities parasitic on small mammals in 4 biogeographic realms are predominantly niche- or dispersal-assembled. They found that these communities' assembly is, to a great extent, governed by the dispersal processes mediated *via* their hosts and, to a much lesser extent, by the niche-based processes. The influence of hosts on the dispersal process is not surprising because, obviously, fleas are not likely to disperse on their own and strongly rely on hosts as dispersal vehicles. The disadvantage of Gibert *et al.*'s (2021) approach is that the input data are strictly limited to species composition in the adjacent regions. This may mask the general pattern and lead to confounded results. Indeed, the indication of dispersal-based assembly held for 3 of the 4 biogeographic realms, whereas in the Nearctic, a stronger niche-based than dispersal-based assembly mechanism, was found.

Another way to distinguish between dispersal-based *vs* niche-based processes in community assembly is to ordinate biological communities using data on species composition and compare the ordination results with the spatial (e.g. across geographic regions) *vs* the ecological (e.g. across biome habitat types) distribution of these communities. Congruence between the resultant community clusters and their geographic locations, but not habitat/biome types, would indicate the predominance of dispersal-based assembly, whereas the opposite would indicate the predominance of niche-based assembly. Congruence between the ordination results and both geographic and habitat/biome-associated distribution would indicate the action of both processes. Furthermore, a comparison between the ordination results for parasite and host species composition would provide insight into the mediating role played by hosts in parasite community assembly.

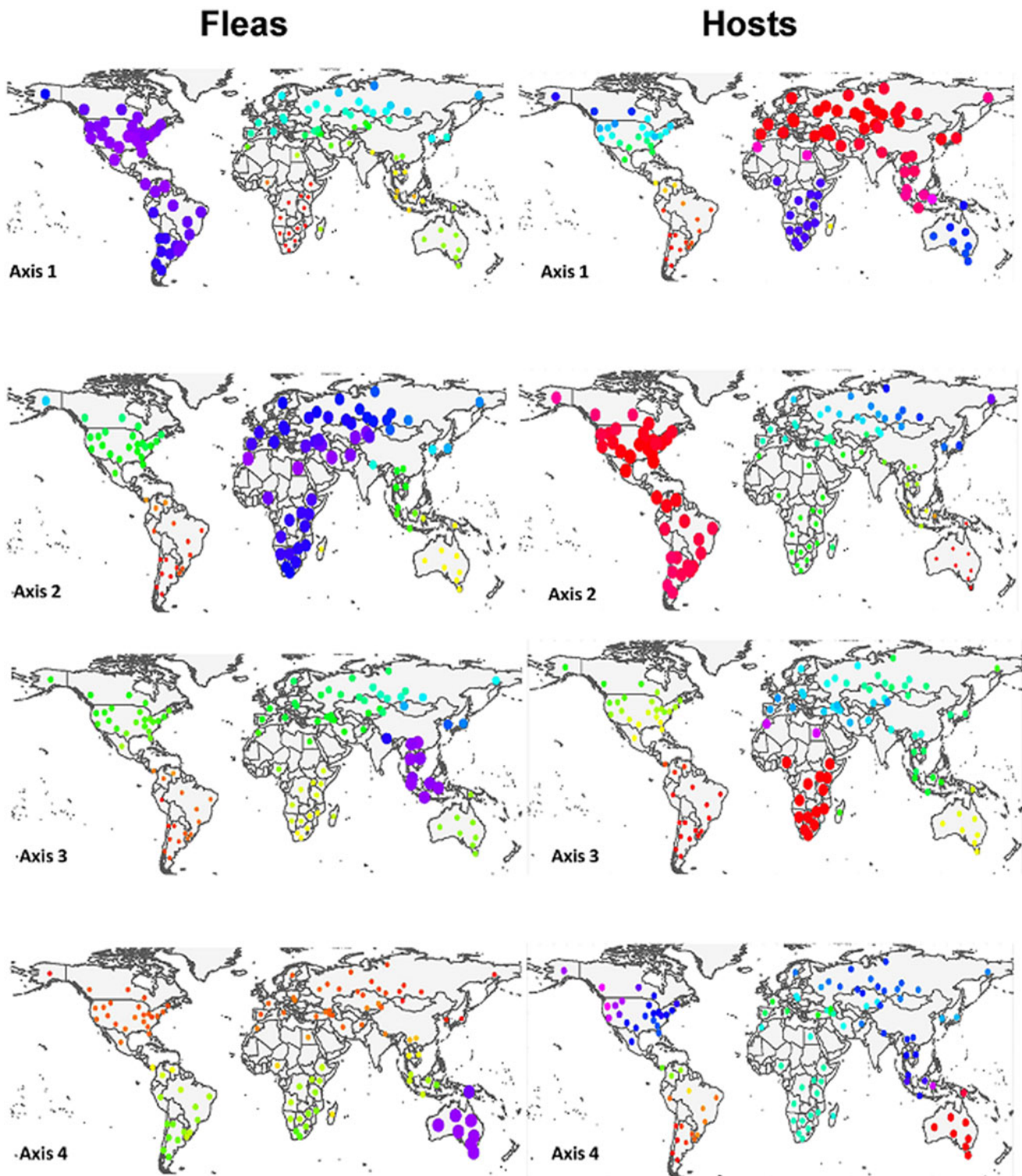
Among multiple ordination methods (see Borcard *et al.*, 2018), the principal component analysis (PCA) and the factor analysis (FA) are the most popular. The aim of both methods is to identify

the underlying variables (= factors = axes) explaining the pattern of correlations between the observed variables and, thus, to produce a small number of factors that explain most of the variance in a much larger number of original variables. The main difference between the PCA and the FA is that the former is a variance-oriented technique, whereas the latter decomposes co-variance (e.g. Shaukat and Uddin, 1989). The traditional FA has a long history of application in ecology, psychology and sociology (e.g. Dagnélie, 1960; Goff and Cottam, 1967; Cattell, 1978). One of the FA's advantages is that each produced factor (= axis) suggests groupings of sites or samples as a single entity, without an *a priori* assumption of how many groups exist or definitions of discrete clusters (Alroy, 2019). However, implementing FA for presence-absence data is problematic when the gradients are long and there are many absences so that the majority of pairs of sites/samples have no shared species. As a result, the sites/samples at gradient extremes bend inwards and appear closer than other pairs of samples (horseshoe effect; Hill, 1973; ter Braak, 1985; Borcard *et al.*, 2018). Alroy (2019) proposed a modification of the FA, called the step-down factor analysis (SDFA), which aims to resolve this issue by assigning a negative value to an absence in which the missing species never co-occurs with the species found in the relevant sample. The results of implementing the SDFA on both simulated and real data have appeared to be superior to those produced by the traditional FA, as well as by other multiple ordination methods (see details in Alroy, 2019).

Here, we applied the SDFA to data on the species composition of regional assemblages of fleas and their small mammalian hosts from 6 biogeographic realms (the Afrotropics, Australasia, the Indo-Malay, the Nearctic, the Neotropics and the Palearctic). First, we aimed to identify biogeographic (across the entire globe) and ecological (within a realm across the main terrestrial biomes) clusters of flea and host assemblages. At the global scale, we expected that the clusters of regional assemblage loadings on SDFA axes would reflect the distribution of the assemblages according to the biogeographic realms to which they belong and, thus, indicate dispersal-based assembly. At the realm scale, we expected that these clusters would reflect the distribution of regional assemblages according to the predominant biomes in the respective regions and, thus, indicate niche-based assembly. Second, we aimed to understand whether biogeographic or ecological clusters of flea and host assemblages from the same regions match each other.

**Table 1.** Proportion of variance explained by 2, 3 or 4 axes of the step-down factor analyses for flea and host assemblages across the entire world and 5 biogeographic realms (the Australasia was not analysed separately; see text for explanation)

Assemblage	Fleas or hosts	Axis 1	Axis 2	Axis 3	Axis 4	Total
World	Fleas	0.39	0.38	0.14	0.09	1.00
	Hosts	0.46	0.40	0.12	0.02	1.00
Afrotropics	Fleas	0.57	0.43	–	–	1.00
	Hosts	0.54	0.24	0.22	–	1.00
Indo-Malay	Fleas	0.45	0.38	0.17	–	1.00
	Hosts	0.49	0.43	0.09	–	1.00
Nearctic	Fleas	0.48	0.38	0.13	–	1.00
	Hosts	0.51	0.33	0.15	–	1.00
Neotropics	Fleas	0.47	0.32	0.13	0.07	1.00
	Hosts	0.36	0.3	0.26	0.08	1.00
Palearctic	Fleas	0.34	0.32	0.17	0.17	1.00
	Hosts	0.47	0.28	0.21	0.05	1.00



**Fig. 1.** Distribution of regional flea and host assemblages across the globe according to their loadings on axes 1–4 of the step-down factor analyses. Point size and colours scale to the assemblage loading on the respective axis.

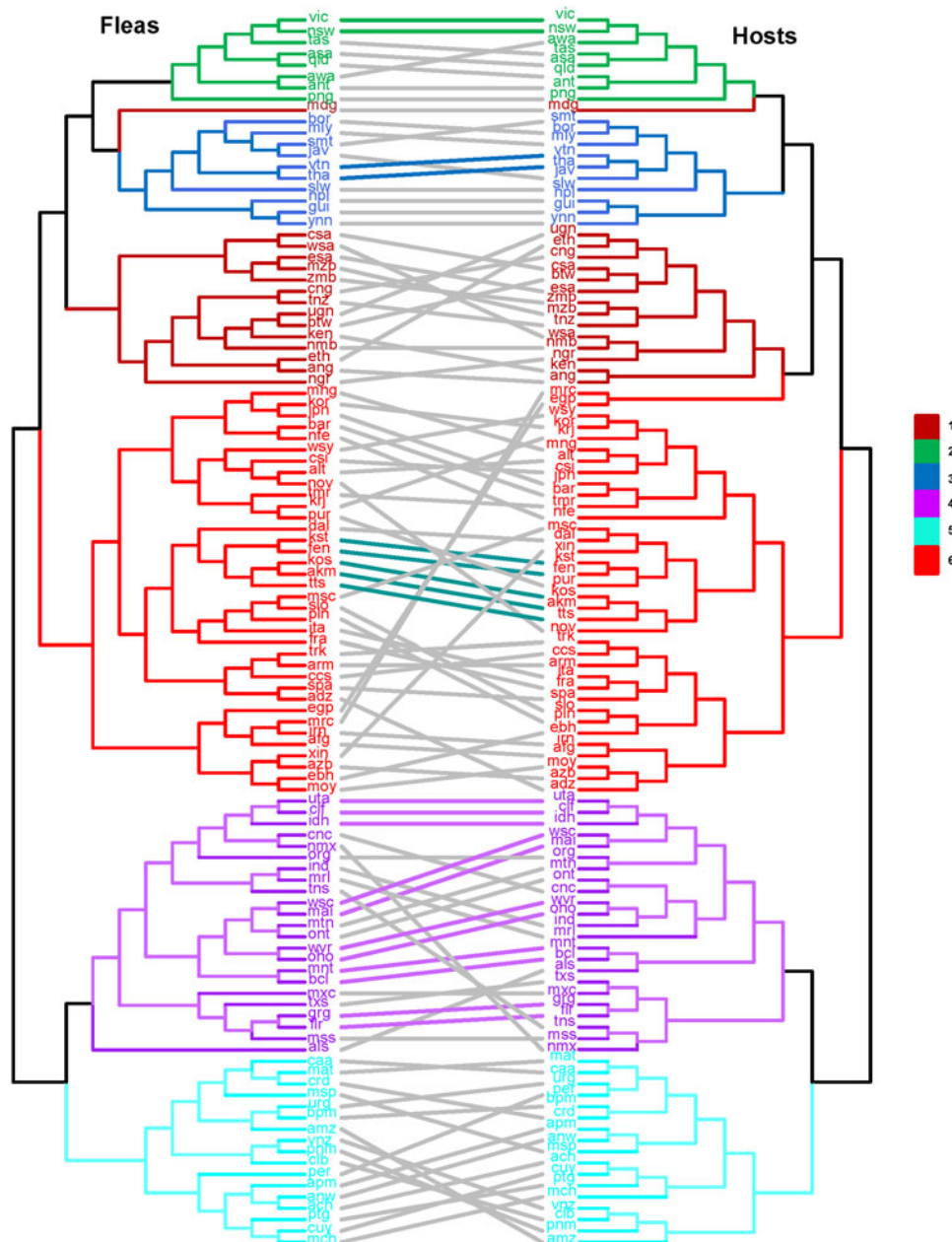
## Materials and methods

### Data on flea and host species composition

We used data obtained from various literature sources (including many ‘grey’ publications) on the species composition of fleas and their small mammalian hosts (tachyglossid Monotremata, Dasyuromorphia, Paramelomorphia, Diprotodontia, Macropodiiformes, Didelphimorphia, Paucituberculata, Macroscelidea, Eulipotyphla, Rodentia and the ochotonid Lagomorpha) from

109 regions of the world (15 different regions in the Afrotropics, 8 regions in the Australasia, 10 regions in the Indo-Malay, 23 regions in the Nearctic, 17 regions in the Neotropics and 36 regions in the Palearctic) (see lists of regions and references in Supplementary Materials, Appendices 1 and 2). We took into the analyses only host species that harboured fleas. Synanthropic ubiquitous rodents (*Rattus norvegicus*, *Rattus rattus* and *Mus musculus*) and the ubiquitous fleas associated with these rodents (*Xenopsylla cheopis*, *Xenopsylla brasiliensis*, *Nosopsyllus fasciatus*





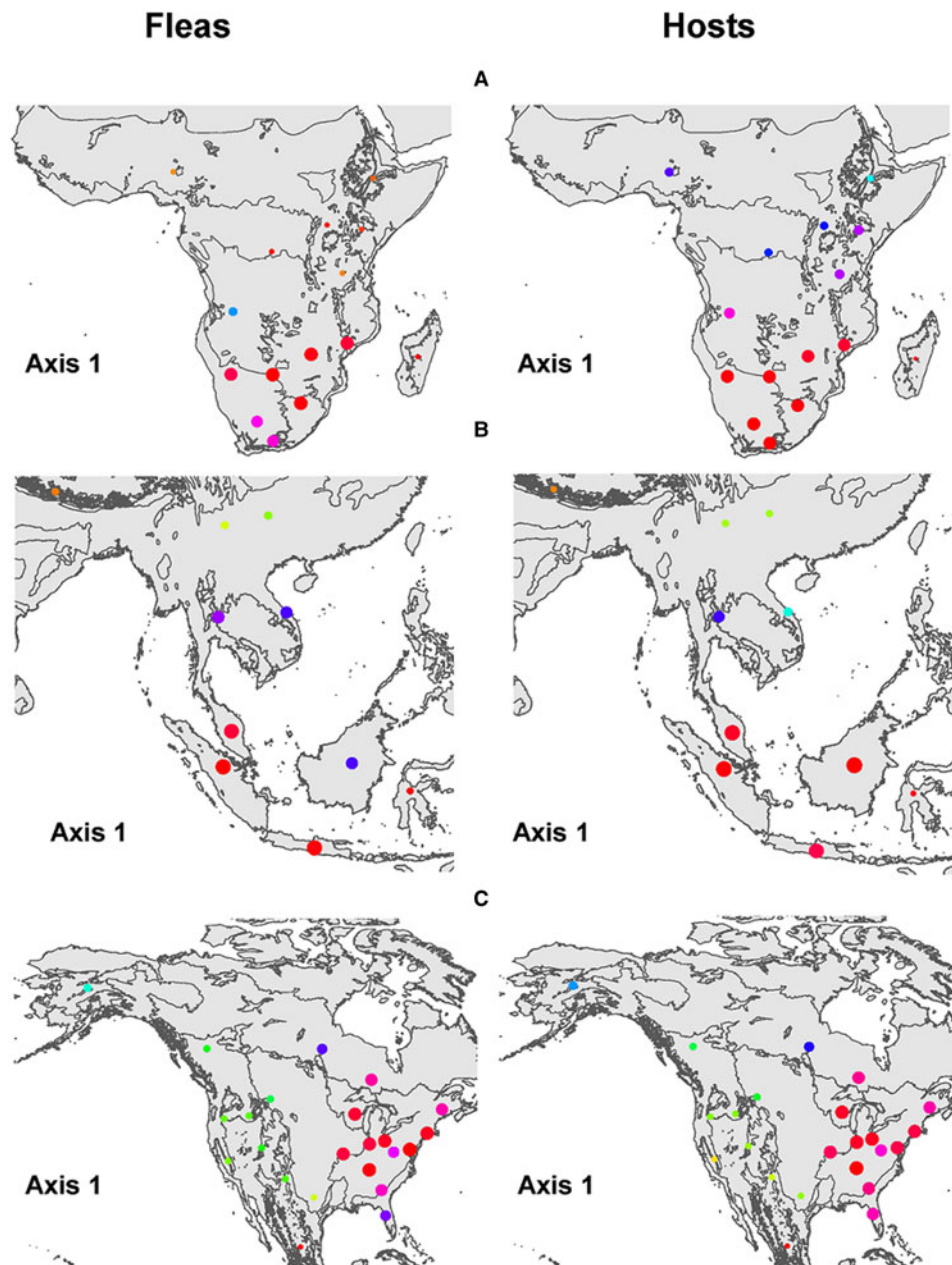
**Fig. 2.** Tanglegram of the results of the cluster analyses of regional flea and host assemblage loadings on the axes of step-down factor analyses across the globe. Colours correspond to biogeographic realms as follows: (1) the Australasia, (2) the Afrotropics, (3) the Indo-Malay, (4) the Nearctic, (5) the Neotropics, (6) the Palearctic. Coloured lines represent subtrees common to the 2 dendrograms. See Table S1, Appendix 1, Supplementary Materials for the abbreviations of region names.

and *Nosopsyllus londiniensis*) were excluded from the analyses. In total, our data included 1313 flea and 1153 host species. The numbers of flea and host species in each region are presented in Table S1, Appendix 1, Supplementary Materials.

### Data analyses

We constructed matrices of either flea species  $\times$  regions or host species  $\times$  regions for either the entire globe or separately for each of the biogeographic realms (except for Australasia because there were data for only 8 regions) and applied the SDFA for fleas and hosts separately using the function 'stepdown' proposed by Alroy (2019) and the function 'fa' of the package 'psych' (Revelle, 2022) implemented in the R statistical environment (R Core Team, 2021). We ran this function with the options (rotate = 'varimax') and (fm = 'minres'). Initially, we specified

10 factors. Then, we removed the factors that cumulatively explained less than 10% of the variation and ran the SDFA with the number of factors remaining after this winnowing. The similarity between factors that was produced by the SDFA for fleas and hosts was further assessed using Tucker's congruence coefficient (TCC; see details in Lovik *et al.*, 2020). It represents a cosine between 2 vectors defined by the matrix of loadings and based at the origin, and it ranges from  $-1$  to  $+1$ . Lorenzo-Seva and ten Berge (2006) established cut-off values in the range of 0.85–0.94 as corresponding to a fair similarity and a value higher than 0.95 as an indicator that the 2 factors are equal. We estimated the TCC using the function 'factor.congruence' of the R package 'psych'. We visualized the geographic distribution of regional flea and host assemblages according to their loadings on axes produced by the SDFA, using the R packages 'sp' (Pebesma and Bivand, 2005; Bivand *et al.*, 2013), 'sf' (Pebesma, 2018),



**Fig. 3.** Distribution of regional flea and host assemblages across the Afrotropics (A), the Indo-Malay (B) and the Nearctic (C), according to their loadings on axis 1 of the step-down factor analyses. Point size and colours scale to the assemblage loading on the respective axis. Borders of terrestrial biomes, according to Olson *et al.* (2001), are shown.

‘rnatualearch’ (South, 2017) and ‘ggplot2’ (Wickham, 2016). Mappings of regional flea and host assemblages within the biogeographic realms were done for the maps of the main terrestrial biomes according to Olson *et al.*’s (2001) classification.

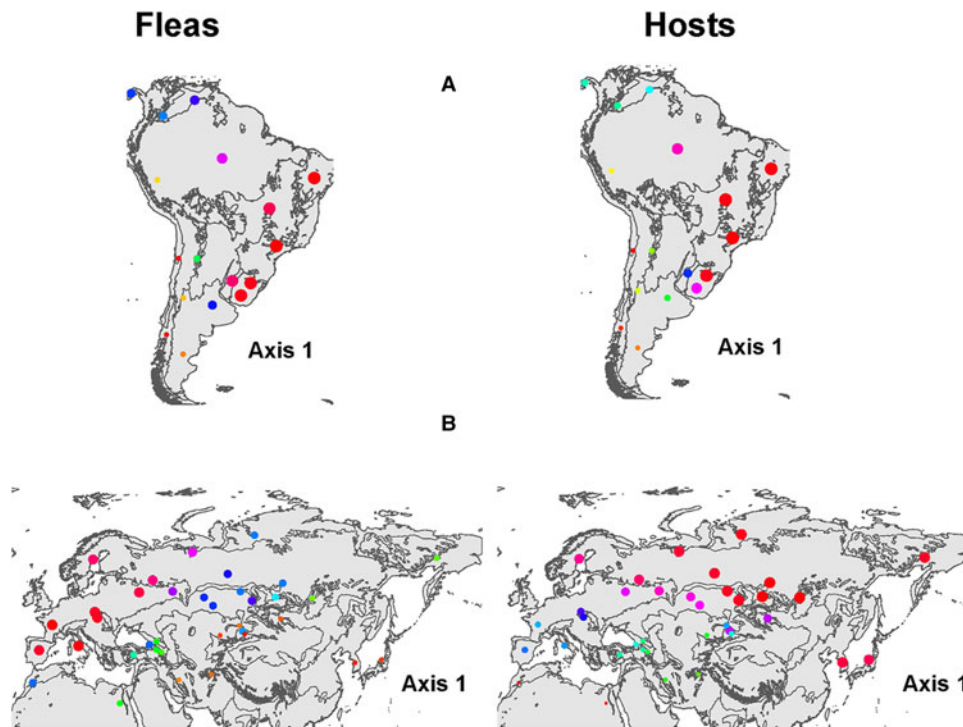
Then, we applied hierarchical cluster analyses with Ward’s linkage method to regional flea and host assemblages using the R package ‘dendextend’ (Galili, 2015). We visualized the clustering of flea *vs* host assemblages for comparison with the function ‘tanglegram’ using a stepwise rotation of the 2 trees (option method = step2side). The similarity between each pair’s (i.e. flea and host) dendrograms was estimated using Baker’s gamma correlation (Baker, 1974).

## Results

The SDFA of regional flea and host assemblages produced from 2 to 4 factors (= axes). The proportion of variance explained by

these axes, both at the global scale and at the scale of individual biogeographic realms, was high (Table 1). Two (for the Afrotropical fleas) and 3 (for the entire globe and the remaining realms) SDFA axes cumulatively explained from 83 to 100% of the variation. In general, the axes were congruent between fleas and hosts (see Tables S2–S7, Appendix 3, Supplementary Material) with the TCC ranging from 0.85 to 1.00 between the first 2 axes (except for the second axes for fleas and hosts in the Afrotropics).

At the global scale, axes 1 and 2, taken together, generally recognized flea and host clusters as belonging to the 6 biogeographic realms (Fig. 1). For fleas, however, axis 1 did not distinguish between (a) the northern part of the Nearctic and the southern and the western part of the Neotropics and (b) the Australasia and the southern part of the Palearctic, whereas axis 2 did not distinguish between (a) most of the Palearctic and the Afrotropics and (b) the Nearctic and the Indo-Malay. For hosts,



**Fig. 4.** Distribution of regional flea and host assemblages across the Neotropics (A) and the Palearctic (B), according to their loadings on axis 1 of the step-down factor analyses. Point size and colours scale to the assemblage loading on the respective axis. Borders of terrestrial biomes, according to Olson *et al.* (2001), are shown.

axis 1 combined (a) the Australasia and the Afrotropics and (b) the Indo-Malay and the eastern Palearctic. Host assemblages of the Neotropics and the northern Nearctic had similar loadings on axis 2. Interestingly, flea assemblages in North Africa and/or southern Asia were clearly distinguished from other assemblages in the Palearctic along both axes 1 and 2, whereas this was true for host assemblages along axis 1. For fleas, axes 3 and 4 combined the Holarctic flea assemblages, whereas for hosts, axes 3 and 4 recognized the difference between (a) the Afrotropics, the Neotropics and the Australasia (axis 4) and (b) the Nearctic, the Palearctic and the Indo-Malay (Fig. S1, Appendix 4, Supplementary Materials).

The results of the cluster analyses of loadings on the S DFA axes demonstrated clear clusters of flea and host assemblages respective to the 6 biogeographic realms except for fleas and hosts from Madagascar (belonging to the Afrotropics but clustered with the Indo-Malay and the Australasian assemblages, respectively) and hosts from Morocco and Egypt (both belonging to the Palearctic, but with the Afrotropical assemblages) (Fig. 2). The topology of the clusters corresponding to the biogeographic realms was similar between flea and host assemblages (Baker's gamma index = 0.98). However, the topology of subtrees within realm-specific clusters substantially differed between fleas and hosts (there were only 9 identical subtrees).

At the scale of biogeographic realms, the distribution of regional flea and host assemblages did not correspond to the predominant biome types. Assemblages with similar loadings on the S DFA axes were often situated in different biomes and vice versa. Distributions of assemblages according to their loadings on axis 1 of the S DFA, across biome types and within a realm, are presented in Figs 3–4 (see Figs S1–S4 of Appendix 4, Supplementary Material for distributions according to the loadings on axes 2–4 of the S DFA). A visual inspection of the across-biome, within-realm distributions of flea vs host assemblages (Figs 3–4 and Figs S1–S4 of Appendix 4, Supplementary Material) suggested a generally weak congruence between these distributions. This

was also supported by the results of the cluster analysis of loadings on the S DFA axes (Fig. 5). Baker's gamma indices for 4 of 5 realms ranged from 0.29 in the Indo-Malay to 0.78 in the Nearctic and reached as high as 0.88 only in the Palearctic.

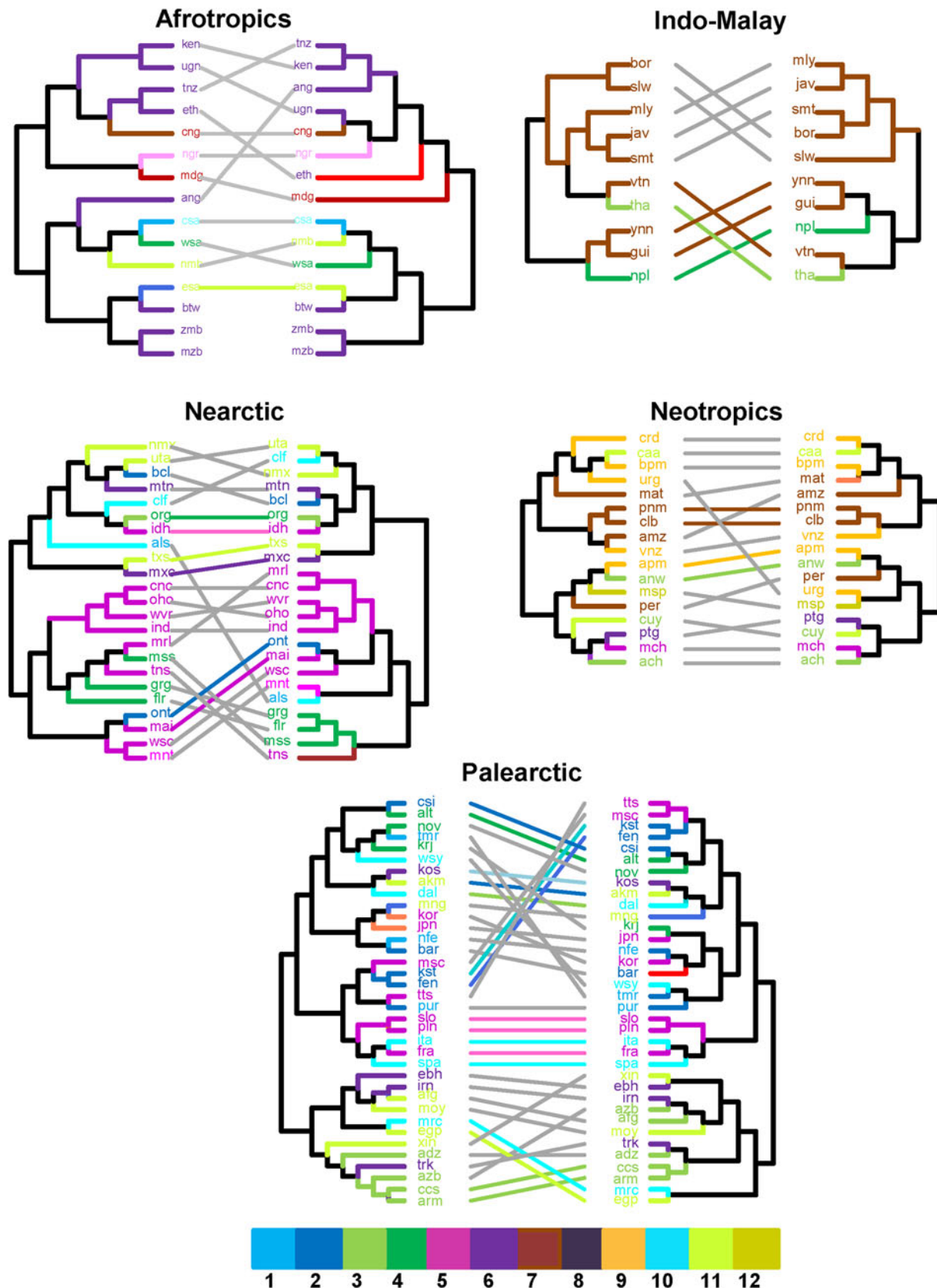
## Discussion

We found that, at the global scale, clusters of regional assemblage loadings on the S DFA axes reflected well the distribution of the assemblages according to the biogeographic realms to which they belong. This, however, was not the case for assemblage distribution across the main biome types within a realm. This suggests that dispersal is a predominant mechanism of flea and host community assembly across large regions. In other words, the species composition of regional flea and host assemblages was determined first and foremost by historical processes.

The historical biogeography of mammals has been repeatedly studied, starting from the pioneering work of Simpson (1940). Although some of Simpson's (1940) conclusions and proposed mechanisms of mammalian distributions have been criticized (Cracraft, 1974; Nelson, 1978), the application of modern molecular and analytical tools reinforced Simpson's (1940) ideas about the importance of dispersal and vicariance in mammals' biogeographic history (e.g. Springer *et al.*, 2011). As obligate parasites of mammalian (although avian as well, but to a much lesser extent) hosts, fleas have closely followed the distribution of their hosts. Both earlier narrative studies of fleas' geographic distribution (e.g. Traub, 1972, 1980; Medvedev, 1996, 2005) and modern sophisticated molecular analyses (Whiting *et al.*, 2008; Zhu *et al.*, 2015) have provided strong support for flea biogeographic patterns mirroring those of their hosts.

In this study, the input data were merely inventories of flea and host species without any association between the species composition of a regional assemblage and its geographic location. Nevertheless, the ordination results provided clear groupings of both fleas and hosts that corresponded well to Wallacean





**Fig. 5.** Tanglegram of the results of the cluster analyses of regional flea (left dendrograms) and host (right dendrograms) assemblage loadings on the axes of step-down factor analyses within each of the 5 biogeographic realms. Colours correspond to the predominant biome of a region according to the classification of Olson *et al.* (2001) as follows: (1) tundra, (2) boreal forests/taiga, (3) montane grasslands and shrublands, (4) temperate coniferous forests, (5) temperate broadleaf and mixed forests, (6) temperate grasslands, savannas and shrublands, (7) tropical and subtropical moist broadleaf forests, (8) tropical and subtropical dry broadleaf forests, (9) tropical and subtropical grasslands, savannas and shrublands, (10) Mediterranean forests, woodlands and scrub, (11) deserts and xeric shrublands, (12) flooded grasslands and savannas. Coloured lines represent subtrees common to a pair of dendrograms. See Table S1, Appendix 1, Supplementary Materials for the abbreviations of region names and predominant biome types.

biogeographic realms. Moreover, different axes of the SDFA captured somewhat different aspects of flea and host distributions. For example, axis 1 isolated North African and Middle Eastern flea communities, whereas axis 3 isolated the flea communities of Nepal from the rest of the Indo-Malay (Fig. 1). Similarly, Panamanian, Venezuelan and Colombian host communities were distinguished from the rest of the Neotropics along axis 1, and Moroccan and Egyptian communities from the rest of the Palearctic along axes 1 and 3 (Fig. 1). This emphasizes the substantial difference in species composition between these and other assemblages of the same realm.

The congruence between distributions of flea and host assemblages across biogeographic realms was high. This, however, did not hold for the across-biome, within-realm scale. In other words, similar regional assemblages of hosts do not necessarily harbour similar assemblages of fleas, suggesting that the dispersal patterns of fleas and hosts within a realm might be, to some extent, independent. The most likely mechanism for this independence is frequent host switching by fleas. The idea that cospeciation was the main event during the common history of the host and parasite lineages (Hafner and Nadler, 1988; Ronquist, 2003) has been undermined by many studies that described frequent host switching among both related and unrelated hosts in various parasite–host associations (Paterson *et al.*, 1993; Beveridge and Chilton, 2001; Roy, 2001), including fleas and their mammalian hosts (Krasnov and Shenbrot, 2002; Lu and Wu, 2005; Whiting *et al.*, 2008). In fact, host switching as a highly probable evolutionary event has been demonstrated not only in phylogenetic, biogeographic and theoretical studies (e.g. Boeger *et al.*, 2003; Meinilä *et al.*, 2004; Araujo *et al.*, 2015) but also in experimental studies (e.g. Dick *et al.*, 2009; Arbiv *et al.*, 2012). Ecological fitting (Janzen, 1985) is considered as the initial (and the main) mechanism of host switching (Agosta and Klemens, 2008; Hoberg and Brooks, 2008; Agosta *et al.*, 2010). Originally, the concept of ecological fitting is related to a preadaptation scenario in which an organism (e.g. a parasite) exploits its environment (e.g. hosts), using some traits that suggest a shared evolutionary history, whereas these traits evolved in response to a different set of conditions. Imagine that a flea's main requirement is the resource (blood) presented by a host rather than a specific host species producing the resource. Given that all terrestrial vertebrates share this resource (and blood's biochemistry and nutritional value are similar among many mammals), a flea may follow the resource rather than its original source (a host species to which the flea is adapted) and may, for example, invade new areas where the resource is present despite the absence of its original source. This, however, does not mean that any flea can switch to any vertebrate host because a host must present fleas not only with its blood as a food resource but also with its burrow or nest, which is necessary for the majority of flea species as a microhabitat where they (most species) oviposit and where their pre-imaginal development takes place (see Marshall, 1981; Krasnov, 2008). In addition, the success of host switching may depend on the fleas' ability to extract blood from a host. This ability depends both on a host's morphological (e.g. skin depth, hair structure), physiological (e.g. immunocompetence) and behavioural (patterns of anti-ectoparasitic grooming) traits and on fleas' morphological (e.g. possession of ctenidia), physiological (e.g. energetic cost of digestion) and behavioural (e.g. singular or multiple feeding bouts necessary for egg maturation) traits (see multiple examples in Krasnov, 2008).

The results of both the SDFA and cluster analyses within realms demonstrated that the distribution of both fleas and hosts was not associated with biome types. Instead, spatial clusters could be envisaged within each biome. For example, flea and host assemblages with similar species compositions appeared to be

distributed across various biomes in eastern South America (Fig. 4). The same was true for fleas in most of Europe and hosts in the eastern Palearctic (Fig. 4). This again testifies to the importance of historical processes as assembly mechanisms of flea and host communities at a large scale, whereas the role of niche-based processes, at least at this scale, is minor, supporting the idea of the scale dependence of a dispersal–niche continuum in which the role of historical processes increases with the increasing scale (Pearson and Dawson, 2003; Gibert *et al.*, 2021).

Indeed, niche-based processes are likely very important for community assembly distribution at a smaller scale, such as across habitat types within a region of a particular biome. This is especially true for ectoparasites such as fleas that are influenced not only by their hosts but, to a great extent, by off-host environmental factors, such as, for example, air temperature, humidity and soil structure. Consequently, niche-based processes might be associated with the intertwined effects of environment and hosts. Moreover, the relative roles of host species vs environment differ between flea species. Fleas in Israel's Negev Desert present 3 possible scenarios for this (Krasnov *et al.*, 1998, 1999). *Parapulex chephrenis* occurs wherever its specific hosts, rodents of genus *Acomys*, are found, independently of habitat type. *Synosternus cleopatrae* resides exclusively in sandy habitats, independently of host species composition. *Xenopsylla ramesis* parasitizes the rodent *Meriones crassus* but only in loessy habitats, whereas in dry riverbeds, it is replaced by *Xenopsylla conformis*. The latter example illustrates the combined host–environment effect that occurs due to across-habitat variation in the same host species' shelter structures (Shenbrot *et al.*, 2002), which may or may not be suitable for a given flea that requires certain levels of air temperature and humidity for successful pre-imaginal development, as well as to the efficiency of physiological processes such as digestion (see Krasnov, 2008).

From the methodological perspective, the SDFA appeared to be useful for revealing large-scale (global or continental) patterns and for recognizing spatially consistent clusters of samples or communities (see also Alroy, 2019). However, the SDFA's ability to recognize such clusters at a fine-grain scale (e.g. local) is still unknown and warrants further investigation.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/S0031182022000907>.

**Data availability.** Raw data are contained in the sources cited in Appendix 2.

**Acknowledgements.** We thank Uri Roll for help with R code. This is publication no. 1116 of the Mitrani Department of Desert Ecology.

**Author contributions.** B. R. K. conceived and designed the study. All authors collected the data and performed statistical analyses. B. R. K. wrote the first draft of the article. All authors finalized the article.

**Financial support.** This study did not receive any specific financial support.

**Conflict of interest.** None.

**Ethical standards.** This study is based on published data, and therefore, ethical standards are not applicable.

## References

- Ackerly DD and Cornwell WK (2007) A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components. *Ecology Letters* **10**, 135–145.
- Agosta SJ and Klemens JA (2008) Ecological fitting by phenotypically flexible genotypes: implications for species associations, community assembly and evolution. *Ecology Letters* **11**, 1123–1134.
- Agosta SJ, Janz N and Brooks DR (2010) How specialists can be generalists: resolving the 'parasite paradox' and implications for emerging infectious disease. *Zoologia (Curitiba)* **27**, 151–162.



- Alroy J (2019) Discovering biogeographic and ecological clusters with a graph theoretic spin on factor analysis. *Ecography* **42**, 1504–1513.
- Araujo SBL, Braga MP, Brooks DR, Agosta SJ, Hoberg EP, von Hartenthal FW and Boeger WA (2015) Understanding host-switching by ecological fitting. *PLoS ONE* **10**, e0139225.
- Arviv A, Khokhlova IS, Ovadia O, Novoplansky A and Krasnov BR (2012) Use it or lose it: reproductive implications of ecological specialization in a haematophagous ectoparasite. *Journal of Evolutionary Biology* **25**, 1140–1148.
- Baker FB (1974) Stability of two hierarchical grouping techniques. Case I: sensitivity to data errors. *Journal of the American Statistical Association* **69**, 440–445.
- Beveridge I and Chilton NB (2001) Co-evolutionary relationships between the nematode subfamily Cloacininae and its macropodid marsupial hosts. *International Journal for Parasitology* **21**, 976–996.
- Bivand RS, Pebesma EJ and Gomez-Rubio V (2013) *Applied Spatial Data Analysis with R*, 2nd Edn. NY: Springer.
- Boeger WA, Kritsky DC and Pie MR (2003) Context of diversification of the viviparous Gyrodactylidae (Platyhelminthes, Monogeneoidea). *Zoologica Scripta* **32**, 437–448.
- Borcard D, Gillet F and Legendre P (2018) *Numerical Ecology with R*, 2nd Edn. NY: Springer.
- Cattell RB (1978) *The Scientific Use of Factor Analysis in Behavioral and Life Sciences*. NY: Plenum.
- Chesson P (2000) Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics* **31**, 343–366.
- Connor EF and Simberloff D (1979) The assembly of species communities: chance or competition. *Ecology* **60**, 1132–1140.
- Cracraft J (1974) Continental drift and vertebrate distribution. *Annual Review of Ecology and Systematics* **5**, 215–261.
- Dagnélie P (1960) Contribution à l'étude des communautés végétales par l'analyse factorielle. *Bulletin du Service de la Carte Phytogéographique, Série B, Carte des Groupements Végétaux* **5**, 7–71.
- Dick CW, Esberard CEL, Gracioli G, Bergallo HG and Gettinger D (2009) Assessing host specificity of obligate ectoparasites in the absence of dispersal barriers. *Parasitology Research* **105**, 1345–1349.
- Galili T (2015) dendextend: An R package for visualizing, adjusting, and comparing trees of hierarchical clustering. *Bioinformatics* **31**, 3718–3720.
- Gibert C, Shenbrot GI, Stanko M, Khokhlova IS and Krasnov BR (2021) Dispersal-based versus niche-based processes as drivers of flea species composition on small mammalian hosts: inferences from species occurrences at large and small scales. *Oecologia* **197**, 471–484.
- Goff FG and Cottam G (1967) Gradient analysis: the use of species and synthetic indices. *Ecology* **48**, 793–806.
- Gravel D, Canham CD, Beaudet M and Messier C (2006) Reconciling niche and neutrality: the continuum hypothesis. *Ecology Letters* **9**, 399–409.
- Gutiérrez PA and Martorelli SR (1999) Niche preferences and spatial distribution of Monogenea on the gills of *Pimelodus maculatus* in Río de la Plata (Argentina). *Parasitology* **119**, 183–188.
- Hafner MS and Nadler SA (1988) Phylogenetic trees support the coevolution of parasites and their hosts. *Nature* **332**, 258–259.
- Hill MO (1973) Reciprocal averaging: an eigenvector method of ordination. *Journal of Ecology* **61**, 237–249.
- Hoberg EP and Brooks DR (2008) A macroevolutionary mosaic: episodic host-switching, geographical colonization and diversification in complex host-parasite systems. *Journal of Biogeography* **35**, 1533–1550.
- Holmes JC and Price PW (1986) Communities of parasites. In Kittawa J and Anderson DJ (eds), *Community Ecology: Pattern and Process*. Oxford: Blackwell, pp. 187–213.
- Hubbell SP (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton: Princeton University Press.
- Hubbell SP (2006) Neutral theory and the evolution of ecological equivalence. *Ecology* **87**, 1387–1398.
- Janzen DH (1985) On ecological fitting. *Oikos* **45**, 308–310.
- Kraft NJB and Ackerly DD (2014) Assembly of plant communities. In Monson RK (ed.), *Ecology and the Environment. The Plant Sciences*, vol. 8. NY: Springer, pp. 68–88.
- Krasnov BR (2008) *Functional and Evolutionary Ecology of Fleas. A Model for Ecological Parasitology*. Cambridge: Cambridge University Press.
- Krasnov BR and Shenbrot GI (2002) Coevolutionary events in history of association of jerboas (Rodentia: Dipodidae) and their flea parasites. *Israel Journal of Zoology* **48**, 331–350.
- Krasnov BR, Shenbrot GI, Medvedev SG, Khokhlova IS and Vatschenok VS (1998) Habitat-dependence of a parasite-host relationship: flea assemblages in two gerbil species of the Negev Desert. *Journal of Medical Entomology* **35**, 303–313.
- Krasnov BR, Hastriter MW, Medvedev SG, Shenbrot GI, Khokhlova IS and Vatschenok VS (1999) Additional records of fleas (Siphonaptera) on wild rodents in the southern part of Israel. *Israel Journal of Zoology* **45**, 333–340.
- Krasnov BR, Shenbrot GI, Khokhlova IS, Hawlena H and Degen AA (2006) Temporal variation in parasite infestation of a host individual: does a parasite-free host remain uninfested permanently? *Parasitology Research* **99**, 541–545.
- Krasnov BR, Shenbrot GI, Khokhlova IS, Stanko M, Morand S and Mouillot D (2015a) Assembly rules of ectoparasite communities across scales: combining patterns of abiotic factors, host composition, geographic space, phylogeny and traits. *Ecography* **38**, 184–197.
- Krasnov BR, Shenbrot GI and Khokhlova IS (2015b) Historical biogeography of fleas: the former Bering Land Bridge and phylogenetic dissimilarity between the Nearctic and Palearctic assemblages. *Parasitology Research* **114**, 1677–1686.
- Lorenzo-Seva U and ten Berge JMF (2006) Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology* **2**, 57–64.
- Lovik A, Nassiri V, Verbeke G and Molenberghs G (2020) A modified Tucker's congruence coefficient for factor matching. *Methodology* **16**, 59–74.
- Lu L and Wu H (2005) Morphological phylogeny of *Geusibia* Jordan, 1932 (Siphonaptera: Leptopsyllidae) and the host-parasite relationships with pikas. *Systematic Parasitology* **61**, 65–78.
- MacArthur RH and Levins R (1967) The limiting similarity, convergence, and divergence of coexisting species. *American Naturalist* **101**, 377–385.
- Marshall AG (1981) *The Ecology of Ectoparasitic Insects*. London: Academic Press.
- Medvedev SG (1996) Geographical distribution of families of fleas (Siphonaptera). *Entomological Review* **76**, 978–992.
- Medvedev SG (2005) *An attempted system analysis of the evolution of the order of fleas (Siphonaptera)*. *Lectures in Memoriam N. A. Kholodkovsky, No. 57*. Saint Petersburg: Russian Entomological Society and Zoological Institute of Russian Academy of Sciences, in Russian.
- Meinilä M, Kausela J, Ziętara MS and Lumme J (2004) Initial steps of speciation by geographic isolation and host switch in salmonid pathogen *Gyrodactylus salaris* (Monogenea: Gyrodactylidae). *International Journal for Parasitology* **34**, 515–526.
- Mouillot D (2007) Niche-assembly vs dispersal-assembly rules in coastal fish metacommunities: implications for management of biodiversity in brackish lagoons. *Journal of Applied Ecology* **44**, 760–767.
- Mouillot D, George-Nascimento M and Poulin R (2003) How parasites divide resources: a test of the niche apportionment hypothesis. *Journal of Animal Ecology* **72**, 757–764.
- Nelson G (1978) From Candolle to Croizat: comments on the history of biogeography. *Journal of the Historical of Biology* **11**, 269–305.
- Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'Amico JA, Itoua I, Strand HI, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P and Kassem KR (2001) Terrestrial ecoregions of the World: a new map of life on Earth: a new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience* **51**, 933–938.
- Paterson AM, Gray RD and Wallis GP (1993) Parasites, petrels and penguins: does louse presence reflect seabird phylogeny? *International Journal for Parasitology* **23**, 515–526.
- Pearson RG and Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography* **12**, 361–371.
- Pebesma EJ (2018) Simple features for R: standardized support for spatial vector data. *The R Journal* **10**, 439–446.
- Pebesma EJ and Bivand RS (2005) Classes and methods for spatial data in R. *R News* **5**, 9–13.
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at <https://www.R-project.org/>.
- Revelle W (2022) *psych: Procedures for personality and psychological research*, Northwestern University, Evanston, Illinois, USA. Available at <https://CRAN.R-project.org/package=psych> Version = 2.2.3.

- Ronquist F (2003) Parsimony analysis of coevolving species associations. In Page RDM (ed.), *Tangled Trees: Phylogeny, Cospeciation, and Coevolution*. Chicago: University of Chicago Press, pp. 22–64.
- Roy BA (2001) Patterns of association between crucifers and their flower-mimic pathogens: host jumps are more common than coevolution or cospeciation. *Evolution* **55**, 41–53.
- Schoener TW (1974) Resource partitioning in ecological communities. *Science* **185**, 27–39.
- Shaukat SS and Uddin M (1989) A comparison of principal component and factor analysis as ordination models with reference to a desert ecosystem. *Coenoses* **4**, 15–28.
- Shenbrot GI, Krasnov BR, Khokhlova IS, Demidova T and Fielden LJ (2002) Habitat-dependent differences in architecture and microclimate of the Sundevall's jird (*Meriones crassus*) burrows in the Negev Desert, Israel. *Journal of Arid Environments* **51**, 265–279.
- Simpson GG (1940) Mammals and land bridges. *Journal of Washington D.C. Academy of Sciences* **30**, 137–163.
- South A (2017). rnatuarearth: World map data from Natural Earth. R package version 0.1.0. Available at <https://CRAN.R-project.org/package=rnatuarearth>.
- Springer MS, Meredith RW, Janecka JE and Murphy WJ (2011) The historical biogeography of Mammalia. *Philosophical Transactions of the Royal Society B* **366**, 2478–2502.
- Surkova EN, Korralo-Vinarskaya NP, Vinarski MV, van der Mescht L, Warburton EM, Khokhlova IS and Krasnov BR (2018) Body size distribution in flea communities harboured by Siberian small mammals as affected by host species, host sex and scale: scale matters the most. *Evolutionary Ecology* **32**, 43–662.
- ter Braak CJF (1985) Correspondence analysis of incidence and abundance data: properties in terms of a unimodal response model. *Biometrics* **41**, 859–873.
- Traub R (1972) The zoogeography of fleas (Siphonaptera) as supporting the theory of continental drift. *Journal of Medical Entomology* **9**, 584–589.
- Traub R (1980) The zoogeography and evolution of some fleas, lice and mammals. In Traub R and Starcke H (eds), *Fleas. Proceedings of the International Conference on Fleas, Ashton Wold, England, June, 1977*. Rotterdam: A. A. Balkema, pp. 93–172.
- van der Valk AG (1981) Succession in wetlands – a Gleasonian approach. *Ecology* **62**, 688–696.
- Whiting MF, Whiting AS, Hastriter MW and Dittmar K (2008) A molecular phylogeny of fleas (Insecta: Siphonaptera): origins and host associations. *Cladistics* **24**, 677–707.
- Wickham H (2016) *ggplot2: Elegant Graphics for Data Analysis*. NY: Springer.
- Zhu Q, Hastriter MW, Whiting MF and Dittmar K (2015) Fleas (Siphonaptera) are cretaceous, and evolved with Theria. *Molecular Phylogenetic and Evolution* **90**, 129–139.