Local-scale dispersal constraints promote spatial structure and arthropod diversity within a tropical sky-island

Nancy Gálvez-Reyes[1], Paula Arribas[2], Carmelo Andujar[2], Brent Emerson[2], Daniel Piñero[1], and Alicia Mastretta-Yanes[3]

[1]Universidad Nacional Autonoma de Mexico Instituto de Ecologia
[2]IPNA
[3]CONABIO

October 5, 2020

Abstract

Physical disruption of gene flow among mountains is commonly viewed as an important process for the generation of hyperdiverse tropical mountain biotas. However, the role of *in situ* diversification within mountains has been seldom explored. Here we evaluate spatially fine-scale patterns of arthropod community assembly within a single mountain to understand the role of dispersal limitation and landscape features as drivers of tropical mountain diversity. We focus on a single tropical sky-island of the Transmexican Volcanic Belt, where we sampled whole-communities of arthropods for eight orders with a comparable design at a spatial scale ranging from 50 m to 20 km, using 840 pitfall traps and whole community metabarcoding. We explored multiple hierarchical levels, from individual haplotypes to lineages at 0.5, 1.5, 3, 5, 7.5% similarity thresholds, to evaluate patterns of richness, turnover and distance decay of similarity with isolation-by-distance and isolation-by-resistance approaches. Our results showed that distance and altitude influence distance decay of similarity at all hierarchical levels. This holds for arthropod groups of contrasting dispersal abilities, but with different strength depending on the spatial scale. Our results suggest long-term persistence of lineages within sky islands, combined with local-scale differentiation, may be an important driver of high arthropod biodiversity in tropical mountains.

Introduction

Mountainous tropical regions are hyperdiverse (Fjeldså et al., 2012; Rahbek, et al., 2019a; Rahbek, et al., 2019b). Besides climate heterogeneity producing different habitats in a patchy distribution, it has been argued that two additional factors particularly increase diversity in tropical mountains. First, because tropical taxa have narrow physiological tolerances to temperature, tropical mountain passes are more effective barriers to dispersal than those in temperate regions, thus promoting speciation through physical disruption of gene flow (Janzen, 1967; Polato et al., 2018; Sheldon et al., 2018). Second, coupling altitudinal gradients with tropical latitudes allows populations within tropical mountains to persist relatively *in situ* despite global climate fluctuations (Mastretta-Yanes, et al., 2018; Rahbek, et al., 2019a; Rahbek, et al., 2019b). These processes of isolation and long-term persistence have been widely used to explain diversification among tropical mountain peaks across the world (e.g., Fjeldså et al., 2012; He et al., 2019; Knowles, 2001; Mastretta-Yanes et al., 2018; McCormack et al., 2009; Uscanga et al. *in review*). However, recent evidence suggests that neutral processes within a single mountain could also play an important role in generating endemism (Bray & Bocak, 2016), but diversification at the single mountain scale and short geographic distances has seldom been explored.

Studying evolutionary processes in mountain ranges across fine spatial scales is challenging due to their complexity with regard to topography and geological and climatic history. However, the insular nature of sky-islands reduces the complexity of mountainous regions, thus providing more simplified systems to

analyze evolutionary processes. Within true oceanic islands, habitat discontinuity has been demonstrated to have an important role in driving geographical diversification (García-Olivares et al., 2019; Goodman et al., 2012). Additionally, Salces-Castellano et al. (2020) have demonstrated that, when dispersal ability and climate tolerance are restricted, strong geographic isolation over distances of only a few kilometers can be found for multiple co-occurring arthropod species of an oceanic island. These results suggest that besides allopatric diversification between different islands, island systems could also promote high levels of intra-island geographical diversification, acting as a local-scale diversity source. If geographical diversification were occurring within individual sky-islands in this way, following the expectations of the neutral theory of biodiversity, it would be expected to find similar spatial patterns of differentiation from the haplotype to the community levels (Baselga et al., 2013).

Arthropod communities are ideal systems to test evolutionary processes at fine spatial scales within sky-islands because they are locally abundant and diverse and relatively easy to sample massively. They also harbour a broad diversity of groups with different dispersal abilities, e.g. including winged and non-winged species and varying body-sizes. However, taxonomic identification to the species level of rich communities of arthropods with traditional methods is challenging (Favreau et al,. 2006, Yu et al., 2012). In this sense, new high throughput sequencing approaches applied to the study of arthropods are revolutionizing the understanding of complex arthropod communities (Andujar et al., 2015; Arribas et al., 2016; Ji et al., 2013; Yu et al., 2012). Specially promising is the use of whole community metabarcoding (cMBC) for the bulk sequencing of the mitochondrial COI gene of mixed communities (Andujar et al., 2018). Recent improvements for denoising metabarcoding datasets (e.g., Edgar, 2016; Callahan et al., 2016) and evaluating the prevalence of sequencing errors and co-amplified pseudogenes (Andujar et al., 2020) raised the prospect of read-based, haplotype-level analyses with mitochondrial COI cMBC data, which represents a step change for the study of diversity patterns through whole-community genetic analyses (Andujar et al., 2018; Arribas et al., 2020).

Haplotype data from hyperdiverse arthropod communities can be used directly for analyses of genetic diversity, or aggregated into species-level entities for analyses of species diversity, allowing for the joint analysis of turnover (beta diversity) at multiple hierarchical levels. Local assemblages may diverge simply due to the lack of population movement which, when assessed for entire communities, results in a largely regular decay of community similarity with spatial distance for the typically neutral haplotype variation of the mitochondrial COI gene (Baselga et al., 2013). Under a scenario where dispersal constraints are a dominant driver of spatial variation in community structure, assemblage turnover at the species level should mirror these haplotype patterns, albeit at a higher level of similarity (Baselga et al., 2013; Baselga, Gomez-Rodriguez, & Vogler, 2015). This analytical approach has been exploited to determine whether the composition across multiple beetle taxa assemblages is predominantly driven by dispersal (Baselga et al., 2013; Baselga, Gomez-Rodriguez, & Vogler, 2015) and has been proposed as a useful way to compare relative dispersal constraints between lineages from different taxonomic groups (Gomez-Rodriguez et al., 2019, Murria et al., 2017). These, and other studies using a multi-hierarchical approach, have focused from regional to continental-scale distances. However recent work has also exploited this framework to analyze community assemblage at finest geographic scales (<15 km) using cMBC data, while also allowing to explore hyperdiverse communities, like soil mesofauna (Arribas et al., 2020). Therefore, applying the cMBC approach to the hyperdiverse arthropod faunas of tropical mountains offers much potential for understanding their community structure and the processes that have shaped it.

Here, we evaluate fine-spatial community patterns within the arthropod fauna of a tropical sky-island forest, to better understand the roles of dispersal limitation and landscape features as drivers of the diversity found in tropical mountains. To do this, we performed a systematic sampling consisting of 840 pitfall traps across *Abies religiosa*forests within the Nevado de Toluca, a sky-island from the Transmexican Volcanic Belt (TMVB). We generated haplotype-level metabarcoding data for 42 arthropod communities distributed in sampling blocks separated from 50 m to 19 km, and evaluate patterns of richness, turnover and distance decay in community similarity at multiple hierarchical levels (haplotype, putative species (OTU) and supra-specific levels). As we are interested in the effect of ecological and topographic features on dispersal limitation, in

addition to testing for the effect of isolation-by-distance (IBD), we also performed an isolation-by-resistance (IBR) analysis. This allows cost dispersal given by landscape features to be incorporated, yielding biologically more informative distance decay relationships than Euclidean distance alone (McRae, 2006). Our analytical framework thus allows us to assess the role of dispersal constraint within a local spatial setting on tropical mountain diversity.

Materials and methods

2.1. Study area and bulk sampling

The study area comprises *Abies religiosa* forests which grow at around 2800-3500 m.a.s.l. in the Nevado de Toluca volcano, which acts as sky-island in the TMBV (Mastretta-Yanes et al., 2015; Rzendowski, 2006). We sampled arthropods from 14 sampling points distributed in four sites: Tlacotepec (TLC; 3 sampling points), San Juan de las Huertas (SJH; 3 sampling points), San Bartolo (ASB; 5 sampling points) and Agua Bendita (AAB; 3 sampling points; Figure 1c), all within the conservation zone of Nevado de Toluca's natural protected area (Figure 1a). Sampling was performed during the rainy season of 2015 during mid August and September for all the sites. At each of the 14 sampling points, three sampling blocks of 20 x 15 m were established, resulting in a total of 42 community samples (Figure 1b). Each sample consisted in all the specimens collected by 20 pitfall traps distributed equidistantly inside each sampling, which were left in the field for 15 days, totalling 840 pitfall traps. Sampling blocks were separated by at least 50 m within each sampling point, and maximum distance among sampling points was 19 km (Figure 1b). Pitfall traps consisted in a plastic cup of 13 cm height and 10 cm diameter, with rectangular perforations of 2 x 1.5 cm length and width, respectively, 8 cm above the cup base. Lids of the pitfall traps were painted with brown color. Each perforation was 2 cm apart from one another and 10 cm above the plastic cup base. Traps were filled with a mixture of 185 ml of ethanol 70% and 15 ml of glycerin (Figure 1b; Figure S1a). Pitfall traps of each block site (20 traps) were collected after 15 days and pooled in a single bulk-sample in a bottle containing ethanol at 96%. Sampling was performed with SEMARNAT permit No. SGPA/DGVS/02641/15.

2.2 Molecular laboratory processes

We cleaned up each sample (comprised of 20 pitfalls pooled) following a Flotation-Filtration-Stereoscope protocol (FFS) that allowed us to have a 'clean' extraction (see Figure S1b-g). To prevent differences of biomass from causing biases at the DNA extraction and PCR steps (Elbrecht & Leese, 2015), we divided the specimens by size into: small (e.g., size*Drosophila melanogaster*), medium (e.g., size *Apis mellifera*), or large specimens (e.g., adult grasshopper). We then divided each sample in two subsamples, one including the complete bodies from small arthropods, and the second with the thorax (head included) from medium-sized arthropods and two legs from large arthropods. Subsamples (n=84) were processed independently for DNA extraction and library construction, but using the same library barcode identifier (thus, the final number of libraries retrieved was 42, one library per sample). For DNA extractions, we treated each bulk-sample with an electric homogenizer (Qiagen TissueLyser II: two times for 1 min at 30 Hz and 30 s at 30 Hz) using platinum beads with Qiagen DNeasy Blood & Tissue Kit and BSA buffer 1x.

For metabarcoding library construction, we used the double dual tagging method with the XT Illumina Adapter Kit as in Arribas et al. (2016). We amplified a 418-base pair region from the 5' end of mitochondrial COI gene (within the standard barcode region for metazoa) with the primers B_F 5' CCIGAYATRGCITTY-CCICG 3' (Shokralla et al., 2015) and Fol-degen-R 5' TANACYTCNGGRTGNCCRAARAAYCA 3' (Yu et al., 2012) modified to include Illumina overhang adaptors for subsequent nested PCR. For each subsample (n=84), we performed and pooled together three independent PCR replicates for each arthropod-size subsample. We included a negative control reaction with no DNA template in all experiments. All information regarding PCR reagents and conditions is available in Material S1 (Supporting Materials S1). Each pool of PCR amplicons was cleaned with Agencourt AMPure XP beads (Beckman Coulter) to purify the COI amplicon away from remaining primers and primer dimers. Then, we used primary amplicons as template for a limited-cycle PCR amplification to add dual-indices barcodes (N7 index and S5 index) and the P5 and P7 Illumina sequencing adapters using the Nextera XT Index Kit from Illumina. We sent the 42 resulting

metabarcoding libraries and a negative control to the Cornell Institute of Biotechnology, Cornell University, USA, for sequencing on a lane of Illumina MiSeq 2x300 bp (Figure S2a-f).

2.3 Bioinformatics processing to identify OTUs at different thresholds of genetic similarity

The resulting paired-end reads of the 42 samples were quality filtered following procedures described by Arribas et al. (2020). Briefly the processing included quality checking, primer removal, pair merging, quality filtering, denoising, and clustering each library independently. We checked raw reads quality with *fastqc*(*https://www.bioinformatics.babraham.ac.uk/projects/fastqc/*). We trimmed primers using fastx_trimmer option of the*fastx-toolkit* (http://hannonlab.cshl.edu/fastx_toolkit/) to trim the primer sequences (20 and 26 bases for R1 and R2, respectively). Then, we processed reads in *trimmomatic-0.36*(Bolger et al., 2014) using TRAILING:20 to cut-off bases of the end of a read, if below a threshold quality of 20. We used R1 and R2 reads to search paired sequences with *pairfq-0.17*(Staton, 2013) and we merged paired reads overlapping with *fastq_mergepairs* command,*-fastq_minovlen 50* and *-fastq_maxdiffs 15* in*usearch-9.2*(Edgar, 2013). We used quality-filtered (Maxee = 1), dereplicated (*-fastx_uniques*) and sorted (*-sortbylength*) options to keep only reads of 418 pb in *usearch-10*(Edgar & Flyvbjerg, 2015). Surviving reads were denoised to generate zero-radius OTUs (ZOTUs) with the *unoise3* and *-minsize 4* commands. ZOTUs are equivalent to Amplicon Sequence Variants (ASVs; sensu Callahan et al., 2016), and are proposed to be a set of predicted biological sequences to be used for direct analysis without the need of OTU clustering (Callahan et al., 2017; Edgar, 2016).

We then assigned high-level taxonomic categories in all of the reads using the lowest common ancestor (LCA) algorithm of *MEGAN-6*(Huson et al., 2016). Taxonomic identification of each read was done using BLAST against the*nucleotide* NCBI *nt* database (June 06 2018; blastn -outfmt 5 -evalue 0.001). We fed BLAST matches into *MEGAN*(Huson et al., 2016), and the taxonomic assignments were used to extract eight ASV datasets, each one of the following orders: Diptera, Collembola, Arachnida, Coleoptera, Hymenoptera, Hemiptera, Myriapoda, Lepidoptera. Remaining sequences were no further considered. We exported the tree (in newick format), visualised and edited it using *figtree-1.4.3* (Rambaut, 2012). Each ASV dataset was aligned in *geneious-8.0.2* with MAFFT, using the FFT-NS-1 algorithm, a scoring matrix of 200/PAM/K=2, GAP open penalty of 3, and the Translation align option. We reviewed all sequences for insertions, deletions or stop codons disrupting the reading frame, which were afterwards excluded.

Subsequently, we generated a community table with read-counts (haplotype abundance) of each retained ASV for the eight orders by matching ASVs against the complete collection of reads (i.e., reads before the dereplicating and denoising steps) using *-search_exact* command with *usearch -10*(Edgar & Flyvbjerg, 2015). Additional filtering according to AVS abundances in community tables (one ASV per taxa) was performed as described by Arribas et al., (2020). Shortly, we first removed from each library those haplotypes with abundances of four or fewer reads (same criteria than denoising). Second, we identified haplotypes contributing with less than 1% of the total reads of the library where they were present, and removed them from the analysis. The 1% cut-off value has been seen in similar datasets to remove most of the spurious ASVswhile maximizing the number of real haplotypes (Andujar et al., 2020; Arribas et al., 2020). Lastly, the community tables of filtered haplotypes were transformed into incidence (presence/absence) data and used for downstream analyses.

Finally, we defined lineages at different clustering levels for each of the orders. For this, each of the ASV filtered dataset was used to generate an UPGMA tree with corrected distances under a F84 model, and based on this tree, all haplotypes were nested into clusterin levels (CLs) following the genetic similarity at different thresholds (0.5%, 1.5%, 3%, 5% and 7.5%), plus an additional threshold corresponding to the result of a species delimitation analyses conducted with the generalized mixed Yule-coalescent (GMYC) model in *R*(Pons et al., 2006). The analyses were performed using *vegan* (Oksanen et al., 2013),*cluster* , *PMCMR* , *hier.part* , *ecodist* , and*betapart*(Baselga & Orme, 2012).

2.4 Community diversity and composition

We used the eight taxonomic subsets (arthropod orders) of ASVs to conduct analyses of community diversity

and composition among sampling sites, considering haplotypes (raw ASVs) and all CLs (0.5%, 1.5%, 3%, 5% and 7.5% lineages, plus that corresponding to the GMYC species delimitation. We first estimated total accumulation richness curves for multiple levels of genetic similarity (AVS, 3% and 5% CL) using the R package *betapart*(Baselga & Orme, 2012). Then we tested for significant differences in alpha diversity among communities by performing ANOVAs with post-hoc.kruskal (method Bonferroni).

Total beta diversity (Sorensen index, $\beta_{\sigma o \rho}$), additive turnover (Simpson index, $\beta_{\sigma \imath \mu}$; species replacement, without the effect of variation in richness) and nestedness (Sorensen–Simpson index, $\beta_{\sigma \nu \epsilon}$; pure richness effect) components were then estimated based on community compositions at different hierarchical levels. We used the *R* package*vegan* (Oksanen et al., 2019) and community composition matrices to perform non-parametric multidimensional scaling ordination (NMDS) based on Sorensen similarity. Plots were created with the *ordispider*option to visualise the compositional ordination of the communities among sites. We then compared arthropod communities composition between sites with an analysis of similarity (ANOSIM) test for each taxonomic group. ANOSIM is a non-parametric analogue for analyzing variance and testing multivariate differences between groups, based on a resemblance matrix and rank dissimilarity (Clarke, 1993). Plots were made in *R* using package *ggplot2*(Wickham et al., 2020).

2.5 Similarity distance-decay and landscape connectivity

We are interested in the effect of landscape features on the variation of arthropods community composition, but considering the multi-hierarchical approach of Baselga et al. (2013). For this, we tested for distance decay of similarity across the different clustering levels, considering as independent variable geographic distance (i.e., IBD) and effective distance, that is, resistance to dispersion considering the landscape features: slope, altitude and vegetation types (i.e., IBR). We focused on Diptera and Collembola assemblages, because they were the two sampled orders showing better completeness and it is expected they show strong differences in dispersal potential (Figure S3). As response variable we used pairwise similarity of assemblages at the haplotype, all CLs (0.5%, 1.5%, 3%, 5% and 7.5% lineages) and putative molecular species by GMYC, among all sites. The analysis was done both with the entire dataset and at finer geographical distances within the West (AAB and ASB), and East (SJH, TLC) sampling points of our study area (Figure 1a). At each CL, we calculated patterns of similarity between pair of sites using the means of the Simpson's similarity index (1- pairwise $\beta_{\sigma \imath \mu}$ = a/[a + min(b,c)] diversity) with *R* using the package *betapart*(Baselga & Orme, 2012), where *a* is the number of species present in both territories, and *b* and *c* the number of species unique to one or another, respectively.

To estimate effective distances we used the program*circuitscape-4.0*(McRae et al., 2013) to calculate "resistance distances" between pairs of nodes on a raster grid. The input files for *circuitscape* are a raster of the study area in which each cell is assigned a conductance value corresponding to the relative probability of the arthropods moving through it; and a list of focal nodes, that is the geographic coordinates of our sampling blocks. In our study, we assessed whether community similarity in different arthropod groups can be explained by landscape features, so we used conductance grids representing different levels of resistance to dispersion depending on vegetation heterogeneity, altitude and slope. To compare the effect of these landscape variables against the effect of distance alone, we also generated a 'flat' landscape; that is, a landscape in which all cells have equal conductance. This is equivalent to Euclidean distance but accounts for the finite size of the input landscape being analyzed, so therefore is more appropriate for comparison with the models using other grids (Lee-Yaw et al., 2009).

We acquired a 10 m-resolution digital land cover map of the Nevado of Toluca defining the *Abies religiosa* forest (sampled vegetation),*Pinus hartwegii* forest, alpine grassland, agriculture, urbanization, water in the crater and Nevado de Toluca's crater (González-Fernández et al., 2018; Sunny et al., 2017). After cross-validation with our sampling records, minor modifications were performed to adjust*Abies* forest distribution in areas of high topographic complexity (see Supporting Material S2 for details). Then, we built the conductance grids for the vegetation heterogeneity analyses assigning different conductance values to each vegetation type (maximum value of conductance is 1, meaning no resistance to dispersion) as detailed in Table S2.

To build the conductance grids for altitude and slope, we followed a similar approach, assigning different conductance values to altitudinal and slope ranges. In total 30 conductance grids were tested (Table S2; Figure S5). Additionally, we tested for distance decay of similarity at smaller geographic distances, for which we separated the West (AAB, ASB) and East (SJH, TLC) sites and performed the same analyses of above, but only with the "flat" raster and the one with highest explanatory level on the previous test.

We used a negative exponential function 'decay.model' included in*R* package *betapart* (Baselga & Orme, 2012), to adjust a negative exponential function to a generalized linear model (GML). We used Simpson similarity ($1 - \beta_{\sigma\imath\mu}$; Baselga, 2010) as a response variable, the pairwise effective distances of each resistance surface as predictor, log link and Gaussian error (Arribas et al., 2020; Gómez-Rodríguez & Baselga, 2018). Finally, we evaluated the fractal pattern (i.e., self-similar systems; Baselga et al., 2015) by a log–log Pearson correlation of the haplotype level and (1) number of lineages, (2) initial similarity (i.e., intercept), and (3) mean similarity independently in each group (Collembola and Diptera) for both our entire sampling and at the East and West sections of our sampling. High correlation values are indicative of self-similarity in lineage branching (i.e., number of lineages) and/or spatial geometry of lineage distributional ranges (i.e., initial and mean similarity; Baselga et al., 2015). Analyses and graphical representations of data were performed with *R* using the packages *vegan*(Oksanen et al., 2019) *betapart*(Baselga et al., 2018) and *ecodist*(Goslee & Urban, 2020).

Results

Phylogenetic groups and ASVs recovered by COI

For the 42 sample libraries, MiSeq sequencing generated a total of 11,639,999 paired reads. We obtained from 108,583 to 419,903 reads in each direction by each sample. Of these, from 39,484 to 178,720 sequences remained after quality filtering (totalling 4,990,334). After read merging and sequence filtering to a length of 418 bp, each sample comprised from 33,609 to 155,634 sequences (totalling 4,305,390) remained. Taxonomic assignments with *usearch* showed high similarity for a wide range of arthropod species (Figure 2a). The 4,305,390 sequences included 1,277 ASVs (unique variants), divided in 385 ASVs in Diptera, 270 in Collembola, 155 in Arachnida, 136 in Coleoptera, 133 in Hymenoptera, 116 in Hemiptera, 51 in Myriapods, and 31 in Lepidoptera (Figure 2b). The number of lineages decreased while increasing the hierarchical clustering (Table S1). The GMYC threshold value obtained was 0.9% in Diptera, 2.9% in Collembola, 1.3% in Arachnida, 0.7% in Coleoptera, 1% in Hymenoptera, 1.8% in Hemiptera, 1.5% in Myriapoda, and 0.3% in Lepidoptera (Table S1).

Arthropod ASVs richness at multi-hierarchical levels

The rarefaction curves indicate that the sampling effort was enough to achieve from 60 to 91% completeness at the haplotype and 3% CL in four of the eight taxa groups (Diptera, Collembola, Arachnida and Hymenoptera Figure S3). The groups that showed larger richness within the communities (alpha diversity by sample) were Diptera (mean = 87 SD +- 23.1 haplotypes and mean = 21 +- 5.15 lineages at 3%), Collembola (mean = 38 +- 9.44 haplotypes and mean = 16 +- 2.91 lineages at 3%), and Arachnida (mean = 12 +- 6.28 haplotypes and mean = 9 +- 3.64 lineages at 3%). Some groups showed significant differences between sampling sites but richness patterns across sites were not consistent for the different arthropod groups (Figure 3). The sampling site that presented the highest mean (+- SD) richness of haplotypes was TLC with 103 (+- 19.2) and 38 (+- 10.7) haplotypes for Diptera and Collembola, respectively; followed by SJH with 94 (+- 9.31) and 42 (+- 13.7) haplotypes for Diptera and Collembola. The sampling site showing the highest richness at higher clustering levels was ASB with Diptera (mean = 25 +- 2.65 and mean = 24 +- 2.45, for CLs 3 and 5%, respectively) and Collembola (mean = 17 +- 2.13 and mean = 15 +- 2.22 for CLs 3 and 5%, respectively); followed by SJH with Diptera (mean = 22 +- 4.04 and mean = 20 +- 4.24, for CLs 3 and 5% respectively), and Collembola (mean = 16 +- 2.28 and mean = 15 +- 2.12, for CLs 3 and 5% respectively; Figure 3).

Community composition at multi-hierarchical levels

The overall turnover ($\beta_{\sigma\imath\mu}$) among sampling points was high for all the arthropod groups and at the multiple

levels analysed (close to 1, $\beta_{\sigma\imath\mu}$ haplotypes from 0.904 to 0.957; $\beta_{\sigma\imath\mu}$ 3% from 0.896 to 0.945; Figure S4) and so the average pairwise $\beta_{\sigma\nu\epsilon}$ was close to zero, $\beta_{\sigma\nu\epsilon}$ haplotypes from 0.010 to 0.026;$\beta_{\sigma\nu\epsilon}$ 3% from 0.014 to 0.033. The mean value for$\beta_{\sigma\imath\mu}$ of haplotypes of Coleoptera (n=136) was 0.957, the average pairwise $\beta_{\sigma\nu\epsilon}$ was 0.014, and$\beta_{\tau o\tau\alpha\lambda}$ was 0.970 ($\beta_{\sigma o\rho}$). For Diptera (n=385), the mean value for $\beta_{\sigma\imath\mu}$ was 0.904, the average pairwise $\beta_{\sigma\nu\epsilon}$ was 0.024, and the mean for average pairwise $\beta_{\sigma o\rho}$ was 0.928.

The NMDS ordination plots showed differences in community composition among sampling sites, particularly for those located in the western (AAB and ASB) and eastern (SJH and TLC) sampling points within Nevado de Toluca (Figure 4). The ANOSIM results showed that differences were significant for the different groups and multi-hierarchical levels analysed dissimilarity, except Myriapoda (Figure 4). The orders Diptera, Collembola, Hemiptera and Coleoptera formed two groups related to the geographic origin of the samples: 1) AAB-ASB (West) and 2) SJH-TLC (East), and Arachnida had three groups: 1) AAB, 2) ASB and 3) SJH-TLC (Figure 4). The variation in Collembola, Arachnida and Diptera among sites largely contributed to the distinct spatial distribution patterns. The highest values were observed in Collembola (Haplotype$r^2 = 0.914$, $p < 0.001$; CL3$r^2 = 0.826$, $p < 0.001$; CL5$r^2 = 0.817$, $p < 0.001$), followed by Arachnida (Haplotype $r^2 = 0.569$,$p < 0.001$; CL3 $r^2 = 0.452$,$p < 0.001$; CL5 $r^2 = 0.461$,$p < 0.001$), and Diptera (Haplotype$r^2 = 0.595$, $p < 0.001$; CL3$r^2 = 0.372$, $p < 0.001$; CL5$r^2 = 0.369$, $p < 0.001$; Figure 4).

Landscape community connectivity

When analyzing the entire sampling at Nevado de Toluca (19 km), pairwise similarity within communities from the haplotype to higher CLs (0.5 to 7.5% lineages) decreased with Euclidean ("flat") distance (Figure 5ac). The fit of IBD was higher in Collembola ($r^2 = 0.704$, b = -2.07, $p< 0.001$ at the haplotype level, $r^2 = 0.599$, b = -1.13, $p < 0.001$ at GMYC, to$r^2 = 0.580$, b = -0.92, $p< 0.001$ at CL 7.5%; Table S3; Figure 5a) than Diptera ($r^2 = 0.293$, b = -0.56, $p< 0.001$ at the haplotype level, $r^2 = 0.195$, b = -0.38, $p < 0.001$ at GMYC, to$r^2 = 0.036$, b = -0.14, $p< 0.001$ at CL 7.5%; Table S4; Figure 5c). However, the explanatory power was slightly higher when considering effective distances by IBR (Figure 5bd; Figure S5). For Collembola, the resistance surface "Altitude 3,000" had the highest explanatory power among CLs (from $r^2 = 0.723$, b = -1.63, $p< 0.001$ at the haplotype level, $r^2 = 0.644$, b = -0.87, $p < 0.001$ for GMYC and$r^2 = 0.615$, b = -0.71, $p< 0.001$ at CL 7.5%; Table S3; Figure 5b). In Diptera the highest explanatory power was given by the resistance surface "Altitude B" (from $r^2 = 0.286$, b = -0.15,$p < 0.001$ at the haplotype level,$r^2 = 0.228$, b = -0.11, $p< 0.001$ for GMYC to 0.070, b = -0.05, $p< 0.001$ at CL 7.5%; Table S4, Figure 5d).

Performing DDRs analyses at even finer geographic distances in Nevado de Toluca (i.e., independently analysing communities within western or eastern sampling points), DDRs remained strong in Collembola, but were weaker or non-significant in Diptera (Figure 6). The subset of sites within the East sampling points (<5 km), also showed distance decay of similarity in Collembola (from haplotype level$r^2 = 0.485$, b = -1.22, $p< 0.01$ to 7.5% CL $r^2 = 0.141$,b = -.27, $p < 0.01$; Table S5, Figure 6a), and Diptera, but coefficients were lower in the second (from haplotype level$r^2 = 0.110$, b = -0.17, $p< 0.01$ to 7.5% CL $r^2 = 0.031$,b = -0.12, $p < 0.01$; Table S5, Figure 6c). When considering altitudinal limitations to dispersal with the IBR analysis, very similar results were found (Table S5, Figure S6ac). For the sites within the West section of our sampling (<2 km), a similar pattern was found for Collembola (from haplotype level$r^2 = 0.544$, b = -1.36, $p< 0.01$ to 7.5% CL $r^2 = 0.240$,b = -0.41, $p < 0.01$; Table S5, Figure 6b) but in Diptera the tests yielded non-significant results both for the IBD and IBR analyses (Table S5, Figure 6d, Figure S6d). The slopes of the exponential decay curves in Collembola were higher than in Diptera, scales were very similar at all threshold levels, and all similarity of assemblage increased with each multi-hierarchical level (Figure 5,6; Table S3, S4). The levels of initial genetic similarity showed a significant log–log correlation with the number of lineages, initial similarity and mean similarity of communities for both our entire sampling area and focusing on the West and East sides of it (Table S6). Thus we found that community variation across genetic similarity levels can be described by a fractal geometry in each group in all the geographic scales sampled. The log-log linear correlations suggest that the patterns of assemblage variation across hierarchical levels can be described by a fractal geometry (Baselga et al., 2013, 2015).

Discussion

We recovered 1,277 ASV, 496 lineages at 3% CL, 441 at 5% CL and 570 using GMYC for the eight arthropod orders. Overall composition of arthropod communities exhibited high turnover among sampling blocks and ordination tests also showed significant differences among sites. Within our entire sampling (maximum 19 km between sampling blocks), we found that distance (pure geographical distance and once corrected by elevation) plays a primarily role driving community structure, from the level of haplotypes through to lineages. This pattern holds at finer geographic distances (<2 km), but only in the group with considerably low dispersal ability (non-winged Collembola).

Arthropod communities recovered by metabarcoding and sampling blocks

The estimation of species richness in any ecological setting, and especially in forested environments, can be challenging due to the rarity of some species, differences in detection probabilities, and the field effort necessary to collect enough samples or species to ensure meaningful coverage (Andújar et al., 2017; Arribas et al., 2016; Creedy et al., 2019). In our study, we used pitfall traps in sampling blocks, maximizing the probability of detecting arthropod species by sampling intensively at multiple sites in one mountain, covering eight orders of arthropods from haplotypes to communities. Our sampling method and size sorting step allowed the recovery of eight major arthropod orders, congruently with other metabarcoding analyses (Elbrecht et al., 2017; Elbrecht et al., 2018; Creedy et al., 2019). According to our rarefaction curves, our sampling detected different taxa per sample (Figure S3), demonstrating the utility of cMBC and our sampling design to study a region of high biological diversity and ecological complexity. It is difficult to compare our results against morphological studies in Nevado de Toluca because there are no complete checklists of arthropods for Mexican highlands. For instance, out of 29 dung beetle species found by a recent survey of four sky-islands, more than 10% were new species (Arriaga-Jiménez et al., 2018). Comparing our results against other cMBC studies shows that more OTUs (913 at 3% CL) of different orders were found at a tropical forest canopy (Creedy et al., 2019) than what was found here for a tropical conifer forest floor (476 OTUs at 3% CL).

Additionally, the metabarcoding pipeline we followed (Arribas et al., 2020) allowed us to analyze diversity patterns for each order separately and allowed us to consider multi-hierarchical levels into community assembly, from haplotypes to 7.5% CL lineages. Our metabarcoding approach coupled with pitfall traps allowed for the automated identification of 1,277 ASV from 42 bulk samples, which contained whole organisms (or part of them). This allowed to calculate community composition and turnover without the bias introduced by traditional taxonomy (Creedy et al.,2019), including small taxa (< 0.5 mm), such as Collembola, and specimens that break easily, such as Diptera (Figure 3). Thus, combining metabarcoding and pitfall traps sampling, allows to perform large scale community composition analyses in tropical mountains.

Strong turnover of arthropods communities at multi-hierarchical levels

Our study provides haplotype-level data of entire communities, which allows surveys of community composition and species turnover of eight taxonomic orders based on presence–absence of haplotypes and higher lineages (Figure 4, Figure S4). High beta diversity was found across communities in all orders, and was dominated by lineage turnover ($\beta_{\sigma\mu}$) instead of nestedness ($\beta_{\sigma\nu\epsilon}$) from haplotypes to higher hierarchical levels (Figure S4). We found significant differentiation among sites placed in a single mountain and vegetation type (*Abies* forests) according to site haplotype composition, but also when molecular entities that conservatively represent species are considered. Diptera and Collembola presented the highest differentiation among sites, specifically dividing the East (SJH-TLC) and the West (AAB-ASB) sampling points, as shown by the NMDS and ANOSIM analyses (Figure 4). Again, this occurs at the level of haplotype, 3% CL and 5% CL (Figure 4). The West-East division in community structure within Nevado de Toluca (Figure 4) implies opposing hillsides. In mountain landscapes, East-facing slopes with morning sun may provide different conditions from cold and foggy West-facing slopes (Rahbek et al., 2019a), which could influence community assembly. While we can not discard the role of environmental heterogeneity (environmental distance) driving community differences between western-eastern sampling points, it should be emphasized that all sampling points were located within*A. religiosa* forests with no apparent differentiation within the Nevado de Toluca.

In fact, *A. religiosa* forests are expected to grow under similar conditions within a single mountain, having a quite restricted environmental niche associated with moist and cold sites (Rzendowski, 2006).

Dispersal limitations drive community structure across multi-hierarchical levels at fine geographical scales

To investigate whether dispersal limitations shape distance-decay relationships within a mountain, we performed an IBD and IBR analyses on Diptera and Collembola, which have contrasting dispersal capabilities (winged and unwinged, respectively). The unified neutral theory of biodiversity predicts that similarity in species composition decreases with distance due to dispersal limitation (Hubbell, 2001). To test this, we examined if DDRs at multi-hierarchical levels can be explained by the dispersion limitation in addition to distance. Specifically we analyzed which landscape variables (Figure S5) influenced DDRs in arthropods of contrasting dispersal abilities (Collembola and Diptera), considering our entire sampling in Nevado de Toluca (19 km max distance among sampling points) and finer geographic scales within the East (<5 km) and West (<2 km) subsets of our sampling. We found that decay of similarity of communities decreases with spatial distance at the level of haplotypes, all CLs (0.5 to 7.5% lineages) and putative molecular species by GMYC (Figure 5). This occurs both in Collembola and Diptera, but is more marked in Collembola whose dispersal abilities are more limited. Interestingly, for Collembola our results also hold both considering our entire sampling as well as finer (<2 km) geographic distances (Figure 5, 6) which is consistent with genetic studies within Collembola showing genetic differentiation over very short geographic distances (Cicconardi et al., 2013; Faria et al., 2019).

High dispersal ability is expected to enhance community similarity (Baselga et al., 2012). Our results support this, because the fit of the decay is higher in the wingless Collembola ($r^2 = 0.704$ and$r^2 = 0.599$ at the haplotype and GMYC levels, respectively; Table S3; Figure 5a) than in the winged Diptera ($r^2 = 0.293$ and $r^2 = 0.195$ at the haplotype and GMYC levels, respectively; Table S4; Figure 5c). Similar patterns of higher distance decay relationships at multi-hierarchical levels in poorly dispersing organisms than in better dispersers were found in European water beetles (Baselga et al., 2013), Iberian leaf beetles (Baselga et al., 2015) and European beetles (Gómez-Rodríguez & Baselga, 2018) at much larger (hundreds of km) geographical scales than here (but see Gómez-Rodríguez et al., 2019 where the pattern was not clear for terrestrial molluscs). Communities of good dispersers are more homogeneous not only because they can disperse larger distances, but also because they can more easily overcome geographical barriers between suitable habitat (Thompson & Townsend, 2006; Vellend, 2010). Thus, if dispersal abilities matter, then landscape features impeding dispersal may also play a role in structuring diversity, which can be explicitly tested including landscape features in analyses such as IBR.

IBR quantifies 'effective distances' between communities that may yield more biologically informative DDRs than Euclidean distance (McRae, 2006; McRae et al., 2008). Our results show positive significant correlation with different explanatory power depending on the surface used, with altitudinal differences better explaining similarity decay than distance alone ("flat" landscape), slope or vegetation type. The resistance surface "flat" (i.e., IBD) has slightly less explanatory power for Collembola ($r^2 = 0.704$ at the haplotype level,$r^2 = 0.599$ at GMYC; Table S3; Figure 5a) than "Altitude 3,000" ($r^2 = 0.723$ at the haplotype level, $r^2 = 0.644$ at GMYC; Table S3; Figure 5b), the best fitting resistance surface. This resistance surface corresponds to the elevation at which the Nevado de Toluca volcano massif begins (Figure S5; Table S2), suggesting that Collembola followed a pattern of IBD and that their limited dispersal is not impacted by landscape features. For Diptera, the highest explanatory power was provided by the resistance surface "Altitude B" ($r^2 = 0.319$ at the haplotype level,$r^2 = 0.228$ at GMYC; Table S4; Figure 5d). This resistance surface assumes maximum conductance at the mean altitude of our sampling and a gradual decrease until reaching altitudes outside of our sampling range, but still where *Abies* forest can be found (Table S2). This suggests that besides being able to disperse larger distances, Diptera moves through relatively unsuitable conditions (different altitudes) less efficiently. Therefore, for Diptera it is not distance alone that drives community structure, but also landscape features. Thus, although our sampling blocks are separated by short distances from 50 m to 19 km, connectivity among sites for Diptera depends upon the elevation model used to set the conductance values (Table S2; Figure S5). This is congruent with Janzen's prediction of "mountain

passes being higher the tropics" (Janzen, 1967), and adds to the recent empirical data (Polato et al., 2018) corroborating it. However, although our results show that landscape connectivity contributes to dispersal limitation, geographic distance seems to play a more dominant role both for both orders. This is consistent with dispersal limitation acting over evolutionary time, as has been suggested to explain the small spatial scale diversification of*Scarelus* beetles within tropical mountains (Bray & Bocak, 2016).

Distance decay patterns at the species level could reflect environmental heterogeneity spatially correlated (i.e., between western and eastern sides of the Nevado de Toluca). While some degree of environmental distance could impact on the obtained biodiversity patterns (but see above on the homogeneity of the sampling study habitat), our results on i) spatial patterns of community dissimilarity recurrently found for multiple hierarchical levels, including haplotypes which are expected to behave neutraly across environmental gradients; ii) high values of turnover and local endemicity at multiple spatial scales and iii) the consistent multihierarchical pattern of distance decay in community similarity at reduce scales (within mountain sites) for less dispersive species, while substantially diluted for the more dispersive ones, point to dispersal limitation within this single single sky-island as a major driver of community assemblage.

Multi-hierarchical approaches are useful to assess whether variation in biological assemblages driven by dispersion follow a fractal geometry where the same neutral processes underlie the distribution of haplotypes and higher clustering levels (Baselga et al., 2013, 2015). Fractal patterns have been revealed in aquatic beetles, leaf beetles, and terrestrial molluscs, highlighting an important role for neutral processes in the spatial structuring of biodiversity (Baselga et al., 2013, 2015; Gomez-Rodriguez et al., 2019). Our results also reveal the existence of a fractal pattern for DDRs, with similarity decreasing with spatial distance from the level of haplotypes to 7.5% CLs (Table S6). These patterns represent a considerably finer geographic scale than that reported in previous studies (from 820 km to 4,500 km as in Baselga et al., 2013, 2015). DDRs decreased with distance at even finer geographic scales (<5 and <2 km) in Collembola (Table S5, Figure 6 and S6), revealing that for arthropods with low dispersal ability, DDRs can occur at very fine geographic distances and at all multi-hierarchical levels, within the geographic confines of a sky island. Further to this, we reveal that DDRs can also emerge within arthropod groups that are typically considered as good dispersers (Diptera), but at comparatively larger geographic distances (Figure 5). Our results align well with analyses performed in Iberian forest and grassland mesofauna, where DDRs were found at all hierarchical levels, also in less than 15 km, for soil taxa with low dispersal abilities (Arribas et al., 2020). Given these short distances, our findings are important not only for understanding evolution, but also for biomonitoring efforts aiming to detect changes in community assembly, even in relatively short distances among sampling points.

Implications for evolution in insular systems and tropical mountains

Our results are interesting in the context of evolution within insular systems. We found that dispersal limitations drive community structure across multi-hierarchical levels within a single habitat in a geographically limited sky-island. This is congruent with analyses in oceanic islands, showing that when dispersal ability and climate tolerance are restricted, strong geographic isolation within an island can occur even in a few kms but extending back even millions of years (Salces-Castellano et al., 2019). Therefore, our results represent an additional source of evidence of how topography and neutral processes can promote biodiversity diversification even at short geographic distances within insular systems. This pattern of dispersal constraints is expected to be more pronounced in the tropics than temperate areas, as tropical species have typically narrower thermal tolerances and lower dispersal than temperate species, leading to higher isolation-by-distance and isolation-by-elevation (e.g., Polato et al., 2018).

Nevado de Toluca is part of a geographically extensive sky-islands complex, and considering this broader spatial context has interesting implications for why tropical mountains are biodiversity hotspots. It has been hypothesized that the global pattern of hyperdiverse tropical mountains likely reflects the differentiation of small, spatially isolated populations combined with the long-term maintenance of these populations, leading to speciation (Rahbek et al., 2019a). In this context, spatial isolation normally refers to habitat fragments distributed across different mountain peaks (Fjeldsa et al., 2012; Rahbek et al., 2019a, 2019b), however

our results support that the processes of differentiation and long-term persistence of small populations may also hold at local scales. Firstly, our data shows that a single sky-island harbour arthropod communities that are spatially structured, at the haplotype and lineage levels, even within a single type of forest with presumably similar environmental conditions at short geographic distances. Secondly, a previous study on sky-islands of the TMVB (including Nevado de Toluca) showed that montane ecosystems are able to persist within the same mountain during climate fluctuations, but shifting up and down slope (Mastretta-Yanes et al., 2018). Coupling these results together, it is supported that a single sky-island can act as a cradle for population differentiation and that this differentiation can persist, and accumulate, relatively *in situ* over evolutionary time scales. Previous case studies on beetle species have reached similar conclusions (Bray & Bocak, 2016), but here we show that rather than a particular case restricted to extremely poor dispersing taxa, the phenomenon could be widespread among tropical arthropods.

Conclusions

Our results are consistent with the expectations of the neutral theory of biodiversity (Hubbell, 2001) and with the idea that global patterns of hyperdiverse tropical mountains reflect the differentiation of small isolated populations combined with their long-term persistence (Bray & Bocak, 2016; Rahbek et al., 2019a). Complementing studies that focus on population isolation among different mountains, here we found that the composition of arthropods communities shows strong turnover within a sky-island and a limited geographic scale (<20 km). We also found that distance and elevation drive biodiversity structure from haplotypes to lineages levels. In groups of low dispersal ability (Collembola), this pattern holds even at a very fine spatial scale (<2 km). Therefore, our results support a general model where neutral dynamics and dispersal limitations act as a source of local-scale diversity within tropical mountains.

Acknowledgements

We are thankful to A. Sunny for providing the raster of vegetation types, to D. Ortiz and W. Tobon for help with spatial data transformations and to E. Campos for IT support. Nancy Galvez-Reyes is a doctoral student from Programa de Doctorado en Ciencias Biomedicas, Universidad Nacional Autonoma de Mexico (UNAM) and was supported by CONACYT (scholarship no. 401781). CONACYT also provided financial support with the project 178245 granted to D. Pinero. B. Emerson was supported by the Spanish Agencia Estatal de Investigacion (CGL2017-85718-P), co-financed by FEDER. P. Arribas was supported by a postdoctoral grant by the Spanish Ministry of Economy and Competitiveness (MINECO, Spain) within the Juan de la Cierva Formacion Program. C. Andujar was supported by the Spanish Ministry of Economy and Competitiveness (MINECO, Spain) (CGL2015-74178-JIN MINECO/FEDER, UE). Analyses were carried out using CONABIO's computing cluster with the assistance of E. Campos. Thanks to Santiago Tlacotepec community for allowing us to sample in their lands, and to C.C.M.S.S., A. Silva, M. Campos, V. Reyes, A. Guerra, I. Rojas, T. Garrido, G. H. Giles, A. Villarruel, M. Garduno, J. Jimenez and J. C. Garcia for sampling support, and to F. S. Galvez for assistance with pitfall traps.

DATA ACCESSIBILITY

Raw metabarcoding data, OTU tables, sampling metadata of this project have been deposited at Dryad Repository under the accession*https://doi.org/XXXXXX* (available upon acceptance). Scripts used for analyses are available at*https://github.com/XXXXX*(available upon acceptance). Laboratory protocol, primers and PCR conditions are in Supporting Materials.

AUTHOR CONTRIBUTIONS

A.M.Y., N.G.R., B.C.E., and D.P. conceived and designed the study. N.G.R, A.M.Y and D.P performed fieldwork, N.G.R. performed laboratory work, processed metabarcoding data and made all analyses. P.A., C.A., and B.C.E. supervised analyses and contributed to the discussion. Manuscript writing was led by N.G.R and A.M.Y., with contributions from all authors. All authors approved the final version of the manuscript.

References

Andujar, C., Creedy, T. J., Arribas, P., Lopez, H., Salces-Castellano, A., Perez-Delgado, A., Vogler, A. P., & Emerson, B. C. (2020). NUMT dumping: Validated removal of nuclear pseudogenes from mitochondrial metabarcode data.*BioRxiv*, 2020.06.17.157347. https://doi.org/10.1101/2020.06.17.157347

Andujar, C., Arribas, P., Gray, C., Bruce, C., Woodward, G., Yu, D. W., & Vogler, A. P. (2018). Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill. *Molecular Ecology*, *27*(1), 146–166. https://doi.org/10.1111/mec.14410

Andujar, C., Arribas, P., Yu, D. W., Vogler, A. P., & Emerson, B. C. (2018). Why the COI barcode should be the community DNA metabarcode for the metazoa. *Molecular Ecology* , *27* (20), 3968-3975. https:doi:10.1111/mec.14844

Andujar, C., Perez-Gonzalez, S., Arribas, P., Zaballos, J. P., Vogler, A. P., & Ribera, I. (2017). Speciation below ground: Tempo and mode of diversification in a radiation of endogean ground beetles.*Molecular Ecology*, *26*(21), 6053–6070. https://doi.org/10.1111/mec.14358

Andujar, C., Arribas, P., Ruzicka, F., Crampton-Platt, A., Timmermans, M. J., & Vogler, A. P. (2015). Phylogenetic community ecology of soil biodiversity using mitochondrial metagenomics. *Molecular Ecology*,*24*(14), 3603-3617. https://doi:10.1111/mec.13195 Arriaga-Jimenez, A., Ros, M., & Halffter, G. (2018). High variability of dung beetle diversity patterns at four mountains of the Trans-Mexican Volcanic Belt. *PeerJ*, *6*, e4468. https://doi.org/10.7717/peerj.4468

Arribas, P., Andujar, C., Salces-Castellano, A., Emerson, B. C., & Vogler, A. P. (2020). The limited spatial scale of dispersal in soil arthropods revealed with whole-community haplotype-level metabarcoding.*Molecular Ecology* , 2020;00:1–14. https://doi.org/10.1111/mec.15591

Arribas, P., Andujar, C., Hopkins, K., Shepherd, M., & Vogler, A. P. (2016). Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil. *Methods in Ecology and Evolution*,*7*(9), 1071–1081. https://doi.org/10.1111/2041-210X.12557

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography* , 19, 134–143. https://doi.org/10.1111/j.1466-8238.2009.00490.x

Baselga, A., Fujisawa, T., Crampton-Platt, A., Bergsten, J., Foster, P. G., Monaghan, M. T., & Vogler, A. P. (2013). Whole-community DNA barcoding reveals a spatio-temporal continuum of biodiversity at species and genetic levels.*Nature Communications*, *4*, 1892. https://doi.org/10.1038/ncomms2881

Baselga, A., Gomez-Rodriguez, C., & Vogler, A. P. (2015). Multi-hierarchical macroecology at species and genetic levels to discern neutral and non-neutral processes. *Global Ecology and Biogeography*,*24*(8), 873–882. https://doi.org/10.1111/geb.12322

Baselga, A., Lobo, J. M., Svenning, J. C., Aragon, P., & Araujo, M. B. (2012). Dispersal ability modulates the strength of the latitudinal richness gradient in European beetles. *Global Ecology and Biogeography*, *21*, 1106–1113. https://doi.org/10.1111/j.1466-8238.2011.00753.x

Baselga, A., & Orme, C. D. L. (2012). betapart: An R package for the study of beta diversity: Betapart package. *Methods in Ecology and Evolution*, *3*(5), 808–812. https://doi.org/10.1111/j.2041-210X.2012.00224.x

Baselga, A., Orme, D., Villeger, S., Bortoli, J. D., Leprieur, F., Logez, M., & Henriques-Silva, R. (2018). *betapart: Partitioning Beta Diversity into Turnover and Nestedness Components* (1.5.1) [Computer software]. https://CRAN.R-project.org/package=betapart

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bray, T. C., & Bocak, L. (2016). Slowly dispersing neotenic beetles can speciate on a penny coin and generate space-limited diversity in the tropical mountains.*Scientific Reports*, *6*(1), 33579. https://doi.org/10.1038/srep33579

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal*, *11*(12), 2639–2643. https://doi.org/10.1038/ismej.2017.119

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581–583. https://doi:10.1038/nmeth.3869

Cicconardi, F., Fanciulli, P. P., & Emerson, B. C. (2013). Collembola, the biological species concept and the underestimation of global species richness. *Molecular Ecology*, *22*(21), 5382–5396. https://doi.org/10.1111/mec.12472

Clarke, K. R. (1993). Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology*, *18*(1), 117–143. https://doi.org/10.1111/j.1442-9993.1993.tb00438.x

Creedy, T. J., Ng, W. S., & Vogler, A. P. (2019). Toward accurate species-level metabarcoding of arthropod communities from the tropical forest canopy. *Ecology and Evolution*, *9*(6), 3105–3116. https://doi.org/10.1002/ece3.4839

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature methods* , *10* (10), 996-998. https://doi. org/10.1038/nmeth.2604

Edgar, R. C. (2016). UNOISE2: Improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv*, 81257. https://doi.org/10.1101/081257

Edgar, R. C., & Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics (Oxford, England)*, *31*(21), 3476–3482. https://doi.org/10.1093/bioinformatics/btv401

Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass—sequence relationships with an innovative metabarcoding protocol. *PloS one* , *10* (7), e0130324. https://doi.org/10.1371/journal.pone.0130324

Elbrecht, V., Peinert, B., & Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecology and evolution*, *7*(17), 6918-6926. https://doi.org/10.1002/ece3.3192

Elbrecht, V., Vamos, E. E., Steinke, D., & Leese, F. (2018). Estimating intraspecific genetic diversity from community DNA metabarcoding data. *PeerJ*,*6*, e4644. https://doi.org/10.7717/peerj.4644

Faria, C. M. A., Shaw, P., & Emerson, B. C. (2019). Evidence for the Pleistocene persistence of Collembola in Great Britain. *Journal of Biogeography*, jbi.13610. https://doi.org/10.1111/jbi.13610

Favreau, J. M., Drew, C. A., Hess, G. R., Rubino, M. J., Koch, F. H., & Eschelbach, K. A. (2006). Recommendations for assessing the effectiveness of surrogate species approaches. *Biodiversity & Conservation*, *15*(12), 3949-3969. https://10.1007/s10531-005-2631-1

Fjeldsa, J., Bowie, R. C. K., & Rahbek, C. (2012). The role of mountain ranges in the diversification of birds. *Annual Review of Ecology, Evolution, and Systematics*, *43*(1), 249–265. https://doi.org/10.1146/annurev-ecolsys-102710-145113

Garcia-Olivares, V., Patino, J., Overcast, I., Salces-Castellano, A., Lopez de Heredia, U., Mora-Marquez, F., Machado, A., Hickerson, M. J., & Emerson, B. C. (2019). A topoclimate model for Quaternary insular speciation. *Journal of Biogeography*, *46*(12), 2769–2786. https://doi.org/10.1111/jbi.13689

Gomez-Rodriguez, C., & Baselga, A. (2018). Variation among European beetle taxa in patterns of distance decay of similarity suggests a major role of dispersal processes. *Ecography*, 1825–1834. https://doi.org/10.1111/ecog.03693

Gomez-Rodriguez, C., Miller, K. E., Castillejo, J., Iglesias-Pineiro, J., & Baselga, A. (2019). Understanding dispersal limitation through the assessment of diversity patterns across phylogenetic scales below the species level. *Global Ecology and Biogeography* , *28* (3), 353-364. https://doi.org/10.1111/geb.12857

Gonzalez-Fernandez, A., Manjarrez, J., Garcia-Vazquez, U., D'Addario, M., & Sunny, A. (2018). Present and future ecological niche modeling of garter snake species from the Trans-Mexican Volcanic Belt. *PeerJ*, *6*, e4618. https://doi.org/10.7717/peerj.4618

Goodman, K. R., Welter, S. C., & Roderick, G. K. (2012). Genetic divergence is decoupled from ecological diversification in the Hawaiian Nesoydne planthoppers. *Evolution*, *66*(9), 2798–2814.

Goslee, S., & Urban, D. (2020). *ecodist: Dissimilarity-Based Functions for Ecological Analysis* (2.0.5) [Computer software]. https://CRAN.R-project.org/package=ecodist

He, K., Gutierrez, E. E., Heming, N. M., Koepfli, K.-P., Wan, T., He, S., Jin, W., Liu, S.-Y., & Jiang, X.-L. (2019). Cryptic phylogeographic history sheds light on the generation of species diversity in sky-island mountains.*Journal of Biogeography*, *46*(10), 2232–2247. https://doi.org/10.1111/jbi.13664

Hubbell, S. (2001).*The Unified Neutral Theory of Biodiversity and Biogeography.* Princeton University Press; JSTOR. https://doi.org/10.2307/j.ctt7rj8w

Huson, D. H., Beier, S., Flade, I., Gorska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., & Tappu, R. (2016). MEGAN Community Edition—Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLOS Computational Biology*, *12*(6), e1004957. https://doi.org/10.1371/journal.pcbi.1004957

Janzen, D. H. (1967). Why Mountain Passes are Higher in the Tropics Author (s): Daniel H. Janzen Source: The American Naturalist , Vol. 101 , No. 919 (May—Jun ., 1967), pp. 233-249 Published by: The University of Chicago Press for The American Society of Naturalist. *The American Naturalist,101*(919), 233–249.

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P. M., Woodcock, P., Edwards, F. A., Larsen, T. H., Hsu, W. W., Benedick, S., Hamer, K. C., Wilcove, D. S., Bruce, C., Wang, X., Levi, T., Lott, M., ... Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*(10), 1245–1257. https://doi.org/10.1111/ele.12162

Knowles, L. L. (2001). Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshopprers. *Molecular Ecology,10*(3), 691–701. https://doi.org/10.1046/j.1365-294x.2001.01206.x

Lee-Yaw, J. A., Davidson, A., McRae, B. H., & Green, D. M. (2009). Do landscape processes predict phylogeographic patterns in the wood frog?*Molecular Ecology*, *18*(9), 1863–1874. https://doi.org/10.1111/j.1365-294X.2009.04152.x

Mastretta-Yanes, A., Moreno-Letelier, A., Pinero, D., Jorgensen, T. H., & Emerson, B. C. (2015). Biodiversity in the Mexican highlands and the interaction of geology, geography and climate within the Trans-Mexican Volcanic Belt.*Journal of Biogeography*, *42*(9), 1586–1600. https://doi.org/10.1111/jbi.12546

Mastretta-Yanes, A., Xue, A. T., Moreno-Letelier, A., Jorgensen, T. H., Alvarez, N., Pinero, D., & Emerson, B. C. (2018). Long-term insitu persistence of biodiversity in tropical sky islands revealed by landscape genomics.*Molecular Ecology*, *27*(2), 432–448. https://doi.org/10.1111/mec.14461

McCormack, J. E., Huang, H., Knowles, L. L., Gillespie, R., & Clague, D. (2009). Sky islands. In *Encyclopedia of islands* (pp. 841–843).

McRae, B. H. (2006). Isolation By Resistance. *Evolution*, *60*(8), 1551–1561. https://doi.org/10.1554/05-321.1

McRae, B. H., Dickson, B. G., Keitt, T. H., & Shah, V. B. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation.*Ecology*, *89*(10), 2712–2724. https://doi.org/10.1890/07-1861.1

McRae, B. H., Shah, V. B., & Mohapatra, T. K. (2013). *Circuitscape 4 User Guides* (4.0) [Computer software]. The Nature Conservancy. http://www.circuitscape.org

Murria, C., Bonada, N., Vellend, M., Zamora-Munoz, C., Alba-Tercedor, J., Sainz-Cantero, C. E., ... & Derka, T. (2017). Local environment rather than past climate determines community composition of mountain stream macroinvertebrates across Europe. *Molecular ecology ,26* (21), 6085-6099. https://doi.org/10.1111/mec.14346

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2019). *vegan: Community Ecology Package* (2.5-6) [Computer software]. https://CRAN.R-project.org/package=vegan

Polato, N. R., Gill, B. A., Shah, A. A., Gray, M. M., Casner, K. L., Barthelet, A., Messer, P. W., Simmons, M. P., Guayasamin, J. M., Encalada, A. C., Kondratieff, B. C., Flecker, A. S., Thomas, S. A., Ghalambor, C. K., Poff, N. L., Funk, W. C., & Zamudio, K. R. (2018). Narrow thermal tolerance and low dispersal drive higher speciation in tropical mountains.*Proceedings of the National Academy of Sciences*, *115*(49), 12471–12476. https://doi.org/10.1073/pnas.1809326115

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., Kamoun, S., Sumlin, W. D., & Vogler, A. P. (2006). Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects.*Systematic Biology*, *55*(4), 595–609. https://doi.org/10.1080/10635150600852011

Rahbek, C., Borregaard, M. K., Colwell, R. K., Dalsgaard, B., Holt, B. G., Morueta-Holme, N., Nogues-Bravo, D., Whittaker, R. J., & Fjeldsa, J. (2019a). Humboldt's enigma: What causes global patterns of mountain biodiversity? *Science*, *365*(6458), 1108–1113. https://doi.org/10.1126/science.aax0149

Rahbek, C., Borregaard, M. K., Antonelli, A., Colwell, R. K., Holt, B. G., Nogues-Bravo, D., Rasmussen, C. M. O., Richardson, K., Rosing, M. T., Whittaker, R. J., & Fjeldsa, J. (2019b). Building mountain biodiversity: Geological and evolutionary processes. *Science*,*365*(6458), 1114–1119. https://doi.org/10.1126/science.aax0151 Rambaut, A. (2012). FigTree v.1.4.2 Available http://tree.bio.ed.ac.uk/software/figtree/ Rzendowski, J. (2006). Bosque de coniferas. *Vegetacion de Mexico*, 295–327. Salces-Castellano, A., Patino, J., Alvarez, N., Andujar, C., Arribas, P., Braojos-Ruiz, J. J., ... & Manolopoulou, I. (2020). Climate drives community-wide divergence within species over a limited spatial scale: evidence from an oceanic island. *Ecology Letters*, *23*(2), 305-315. https://doi.org/10.1111/ele.13433 Sheldon, K. S., Huey, R. B., Kaspari, M., & Sanders, N. J. (2018). Fifty Years of Mountain Passes: A Perspective on Dan Janzen's Classic Article. *The American Naturalist*, *191*(5), 553–565. https://doi.org/10.1086/697046

Shokralla, S., Porter, T. M., Gibson, J. F., Dobosz, R., Janzen, D. H., Hallwachs, W., Golding, G. B., & Hajibabaei, M. (2015). Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform.*Scientific Reports*, *5*(1), 9687. https://doi.org/10.1038/srep09687

Staton, E. (2019).*Sestaton/Pairfq* [Perl]. https://github.com/sestaton/Pairfq (Original work published 2013)

Sunny, A., Gonzalez-Fernandez, A., & D'Addario, M. (2017). Potential distribution of the endemic imbricate alligator lizard (Barisia imbricata imbricata) in highlands of central Mexico. *Amphibia-Reptilia*, *38*(2), 225–231. https://doi.org/10.1163/15685381-00003092

Thompson, R., & Townsend, C. (2006). A truce with neutral theory: Local deterministic factors, species traits and dispersal limitation together determine patterns of diversity in stream invertebrates. *Journal of Animal Ecology*, *75*(2), 476–484. https://doi.org/10.1111/j.1365-2656.2006.01068.x

Vellend, M. (2010).*2010 Community Ecology. 85*(2), 183–206. https://doi.org/10.1086/652373

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., & RStudio. (2020). *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics* (3.3.2) [Computer software]. https://CRAN.R-project.org/package=ggplot2

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*,*3*(4), 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x