Effects of topography on structuring species assemblages in a subtropical forest

Qinggang Wang[1,2], Ruwan Punchi-Manage[3,4], Zhijun Lu[1,*],
Scott B. Franklin[5], Zhiheng Wang[2], Yaoqi Li[2], Xiulian Chi[6],
Dachuan Bao[1], Yili Guo[1], Junmeng Lu[1], Yaozhan Xu[1,7],
Xiujuan Qiao[1] and Mingxi Jiang[1]

[1] *Key Laboratory of Aquatic Botany and Watershed Ecology, Chinese Academy of Sciences, Wuhan, Hubei 430074, PR China*
[2] *Department of Ecology and Key Laboratory for Earth Surface Processes of the Ministry of Education, College of Urban and Environmental Sciences, Peking University, Beijing 100871, PR China*
[3] *Department of Ecosystem Modeling, University of Göttingen, Büsgenweg 4, 37077 Göttingen, Germany*
[4] *Department of Ecological Modelling, UFZ Helmholtz Centre for Environmental Research-UFZ, PF 500136, 04301 Leipzig, Germany*
[5] *School of Biological Sciences, University of Northern Colorado, Greeley, CO 80639-0017, USA*
[6] *National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, PR China*
[7] *College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, PR China*
*Correspondence address. Key Laboratory of Aquatic Botany and Watershed Ecology, Wuhan Botanical Garden, Chinese Academy of Sciences, #1-1 Moshancun, Hongshan District, Wuhan, Hubei 430074, PR China. Tel: +86-27-87510986; Fax: +86-27-87510251; E-mail: luzj@wbgcas.cn

Abstract

Aims

Topography has long been recognized as an important factor in shaping species distributions. Many studies revealed that species may show species–habitat associations. However, few studies investigate how species assemblages are associated with local habitats, and it still remains unclear how the community–habitat associations vary with species abundance class and life stage. In this study, we analyzed the community–habitat associations in a subtropical montane forest.

Methods

The fully mapped 25-ha (500×500 m) forest plot is located in Badagongshan Nature Reserve in Hunan Province, Central China. It was divided into 625 (20×20 m) quadrats. Habitat types were classified by multivariate regression tree analyses that cluster areas with similar species composition according to the topographic characteristics. Indicator species analysis was used to identify the most important species for structuring species assemblages. We also compared the community–habitat associations for two levels of species abundances (i.e. abundant and rare) and three different life stages (i.e. saplings, juveniles and adults), while accounting for sample size effects.

Important Findings

The Badagongshan plot was divided into five distinct habitat types, which explained 34.7% of the variance in tree species composition. Even with sample size taken into account, community–habitat associations for rare species were much weaker than those for abundant species. Also when accounting for sample size, very small differences were found in the variance explained by topography for the three life stages. Indicator species of habitat types were mainly abundant species, and nearly all adult stage indicator species were also indicators in juvenile and sapling stages. Our study manifested that topographical habitat filtering was important in shaping overall local species compositions. However, habitat filtering was not important in shaping rare species' distributions in this forest. The community–habitat association patterns in this forest were mainly shaped by abundant species. In addition, during the transitions from saplings to juveniles, and from juveniles to adults, the relative importance of habitat filtering was very weak.

Keywords: multivariate regression tree analysis, habitat filtering, rare species, Badagongshan

Received: 9 November 2015, Revised: 10 May 2016, Accepted: 11 May 2016

INTRODUCTION

How species coexist in species-rich communities is one of the central questions in ecology (Hubbell 2001; Hutchinson 1961). Abiotic assembly rules suggest that habitat filtering preventing the establishment or persistence of species is an important driver in shaping species communities (Kraft *et al.* 2015). Biotic assembly rules suggest that species coexistence is maintained through competition for similar resources or shared enemies (e.g. Comita *et al.* 2010). In contrast, the neutral theory emphasizes that species are functionally equivalent, and species diversity in a community can be explained by speciation, dispersal limitation and demographic stochasticity (Hubbell 2001). Topography is often used as a surrogate of habitats for plant species, which is correlated with environmental conditions such as soil moisture (Daws *et al.* 2002; Engelbrecht *et al.* 2007), soil nutrients (John *et al.* 2007) and light availability (Svenning 2001) that may directly influence the demography (i.e. survive, growth and recruitment) of plants. Species-level analyses of fully mapped forest plots have revealed that many species showed species–habitat associations, and that the number of habitat specialists increases with topographic heterogeneity (e.g. Harms *et al.* 2001; Gunatilleke *et al.* 2006; Valencia *et al.* 2004). Moreover, due to changes in regeneration niche (Grubb 1977; Webb and Peart 2000) or physiological requirements such as light (Comita *et al.* 2007; Poorter *et al.* 2005), habitat associations of many species may not be consistent throughout their life stages (Comita *et al.* 2007; Webb and Peart 2000).

However, such species–habitat association analyses provide little information on how the species assemblages are related to local habitats, as they mainly focus on the relationships between individual species' distribution and topographic habitats (Harms *et al.* 2001). In addition, rare species are typically removed from these analyses due to sample size limitations. One approach that reveals habitat associations from the community perspective is to categorize the areas with distinct species composition into different habitat types according to local topographic properties using a multivariate spatial method (Kanagaraj *et al.* 2011; Legendre *et al.* 2009). The amount of variance in species composition explained by the habitat types (the degree of community–habitat association) could be related to habitat filtering processes. Furthermore, comparing the associated habitat types and/or the degree of community–habitat associations across species abundance classes (e.g. abundant species and rare species) could shed light on the role of habitat filtering in shaping species rarity. Abundant species could occupy nearly all their suitable habitats, as they usually have relative strong dispersal abilities (e.g. Dalling *et al.* 2002; Gaston 1994). In contrast, rare species that have weak dispersal abilities are confined in relative small areas and could not reach their potentially suitable habitats (Wang *et al.* 2012). Therefore, habitat filtering tends to play a weaker role in shaping distributions of rare species than abundant species. At community level, we would expect community–habitat association of rare species to be weaker than that of abundant species. In addition, by examining the changes in associated habitat types and/or the changes in the degree of community–habitat associations across life stages, we can better understand the relative importance of habitat filtering during the transition from early life stages to later ones. More specifically, if habitat filtering dominates community assembly processes, we would expect that species would become more strongly associated with their habitats as they age, and thus the community–habitat associations for later-stage assemblages would be enhanced. In contrast, if competition governed community assembly processes, we would expect weaker community–habitat associations for later-stage assemblages (Baldeck *et al.* 2013a; Kanagaraj *et al.* 2011). If neutral mortality plays a more important role, or the habitat filtering and competition processes counterbalanced each other, we would expect that community–habitat associations would not change with life stage (Baldeck *et al.* 2013a; Punchi-Manage *et al.* 2013).

Previous studies have indicated that the number of individuals changes significantly with life stages. Normally, the number of adult individuals is usually smaller than that of seedlings and saplings in natural tree communities. However, recent studies (e.g. Baldeck *et al.* 2013b; Punchi-Manage *et al.* 2014) pointed out that the degrees of community–habitat associations increased with the number of individuals included in the community data. For example, Baldeck *et al.* (2013b) found that differences of community–habitat associations for different life stages were directly related to the number of individuals of the life stages. Further, they found that changes in degree of community–habitat associations among different life stages after controlling sample size were much smaller than those without controlling sample size (Kanagaraj *et al.* 2011). Thus, it is necessary to take sample size into account when comparing community–habitat associations across life stages and/or across species abundance classes.

Here, we analyzed the community–habitat associations using multivariate regression tree (MRT) analysis in a subtropical montane forest in Badagongshan (BDGS) Nature Reserve in Central China. We also compared the MRT results with that of the unconstrained clustering method. More specially, we investigated: (i) whether this forest shows distinct habitats and to what extent variance in local species composition can be explained by habitat differentiation; (ii) is the community–habitat association of rare species weaker than that of abundant species, especially after controlling for sample size? and (iii) is the community–habitat association generally consistent across life stages, especially after controlling for sample size? We expected that distinct habitat differentiation could be detected, because BDGS forest plot emerged as strong fine-scale topography heterogeneity with deep valleys and steep slopes. Since rare species are usually narrowly distributed and could not reach all their potentially suitable habitats, rare species would be less associated with habitats than abundant species. Thus, we expected that the community–habitat associations of rare species would be weaker than that of abundant species, even

after controlling sample sizes. In addition, although some individual species may change habitat preferences with life stages, different tree communities should formed in the early stage and their associated habitats should maintain with life stage. Thus, we expected that the community–habitat associations could be generally consistent across life stages and the relative importance of habitat filtering could not exceeded competition and/or stochastic processes as tree age.

METHODS

Study site

The 25-ha BDGS forest plot (29°46.04′N, 110°5.24′E) is located in the Badagongshan Nature Reserve in Central China. Annual rainfall averages 2105.4 mm. The forest type is a montane evergreen and deciduous subtropical forest with *Cyclobalanopsis multinervis* (evergreen species) and *Fagus Lucida* (deciduous species) as its dominant species. All free-standing trees with diameter at breast height (dbh) ≥1 cm in the BDGS plot were tagged, mapped and identified to species (Guo *et al.* 2013; Wang *et al.* 2014a; Wang *et al.* 2014b); 238 species (including two subspecies), 114 genera and 53 families were identified in the first census during 2010–11. The soil type is yellow-brown soil. The elevation in this plot ranges from 1369.6 to 1470.9 m. The plot has fine-scale topographic heterogeneity with deep valleys and steep slopes (Fig. 1).

Topographic variables

The 25-ha BDGS plot was divided into 625 (20 × 20 m) quadrats, and for each quadrat, six topographic variables were calculated: elevation, slope, convexity, aspect, topographic wetness index (TWI) and vertical distance from the channel network

(VDC) (online supplementary Fig. S1). Elevation was defined as the mean elevation of the four corners. Convexity was the mean difference between the elevation of the focal quadrat and the elevation of the eight adjacent quadrats. Slope was calculated as the average angular deviation from horizontal of each of the four triangular planes formed by connecting three of its four corners. Aspect is the direction a slope faces, and cos (aspect) and sin (aspect) were calculated to make aspect data usable in linear models. TWI was defined as the ratio of the area upslope from any given point on the landscape to the local slope at that point and calculated using Tarboton's Deterministic Infinity Method (Tarboton 1997). VDC was the vertical distance from the channel network. TWI and VDC are frequently employed indexes quantifying topographical control on hydrological processes (Kanagaraj *et al.* 2011; Punchi-Manage *et al.* 2013). We obtained all topographical variables using SAGA GIS (http://www.saga-gis.org).

Statistical methodology

Multivariate regression tree analysis

Multivariate regression trees technique (MRT; De'ath 2002; Larsen and Speckman 2004) was used to group areas with similar species composition (i.e. species assemblages) according to topographic variables. MRT is a constrained clustering method that can explore the relationships between multispecies data and environmental characteristics (De'ath 2002). MRT is based on a recursive algorithm. In our case, the root node consists of all 625 quadrats (20 × 20 m). The algorithm determines the environmental value threshold that splits the quadrats into two groups by minimizing the species dissimilarity within groups (Larsen and Speckman 2004). The Bray–Curtis dissimilarity index was used to represent the species
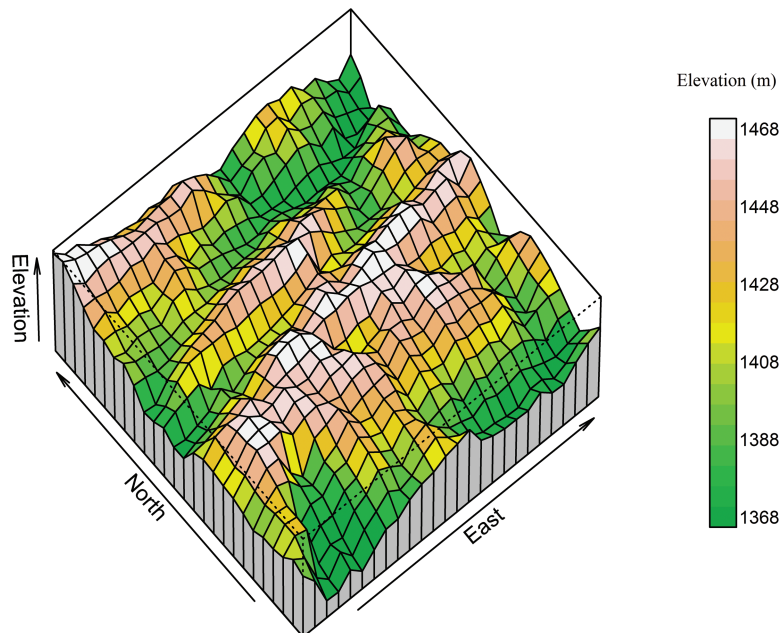

Figure 1: topographic map of the 500 × 500 m Badagongshan Forest Dynamic Plot (modified from Wang *et al.* 2014a).

dissimilarity between two quadrats j and k: $\Sigma_i (y_{ij} - y_{ik}) / \Sigma_i (y_{ij} + y_{ik})$, where y_{ij} is the abundance of the ith species in jth quadrat. The tree was pruned by removing successive pairs of child nodes that increased the deviance by a minimum. Tree size was selected by minimizing the cross-validated relative error (CVRE) with 1 SE rules in all the cases in order to avoid over-fitting the data. CVRE ranged from 0 (best predictors) to 1 (poorest predictors). The MRT analysis was performed using 'mvpart' package in R (R Development Core Team 2016).

Indicator species analysis

It is important to investigate which species are most important for habitat classification. We used indicator species analysis (Dufrene and Legendre 1997) to identify species that were statistically significant indicators of habitat types. Indicator value is defined as the product of the frequency and fidelity for each habitat type, ranging from 0 (species does not occur in a habitat) to 1 (species occurs in all quadrats of one habitat type and none of the quadrats of other habitat types). Species with high indicator values (>0.25 in this study) for particular habitat types are considered as indicator species. Indicator species analysis was performed using 'labdsv' package in R.

Definition of abundant and rare species

As abundant species and rare species are two relative concepts, it is difficult to obtain unified definitions for different communities and different scales. Hubbell and Foster (1986) classified those species with average density of <1 stem/ha as rare species. Gaston (1994) proposed the lower quartile of the frequency distribution of species abundances as rare species. However, the matrix derived from these two criteria produced a large proportion of zero values in rare species matrices. Here, we defined abundant species as the upper quartile (i.e. 25%) of the frequency distribution of species abundances, i.e. Magurran's quartile method (Magurran 2004), and the remaining species as rare species to avoid that only singletons or doubletons were grouped as rare species (Siqueira *et al.* 2012). Thus, we obtained 59 abundant species and 179 rare species accounting for 93.9% and 6.1% of total abundance, respectively. Preliminary analyses showed that the change of cut-off point (e.g. 20% and 30%) to define abundant and rare species would not change the general trends found in subsequent analyses.

Definition of saplings, juveniles and adult trees

Tree size (dbh) is frequently used to define tree life stages in many studies, as smaller trees are generally younger than the bigger ones. Ecologists often regarded the trees reaching about 3 cm dbh as established trees and trees with >10 cm dbh as adults (e.g. Baldeck *et al.* 2013b), i.e. adult (dbh ≥ 10 cm), juvenile (3 cm ≤ dbh < 10 cm) and sapling (dbh < 3 cm). In the BDGS plot, 12% of the trees were ≥10 cm dbh, 34% ≥3 and <10 cm dbh and 54% <3 cm dbh. In this study, to make the subassemblages of sapling, juvenile and adult have similar composition (online supplementary Fig. S2), we chose species-specific dbh to define saplings, juveniles and adults. For each species, the largest 12%, the second largest 34% and the remaining 54% were defined as adult, juvenile and sapling class, respectively. A total of 164 species with individuals ≥10 were included in our analysis.

Effects of sample size on topographic structuring

We found that stem counts of adults were smaller than those of saplings and juveniles, and stem counts of rare species were of course less than those of abundant species (online supplementary Table S1). To test whether the degree of community–habitat association changed with species abundance and life stages, we used fixed sample sizes for the three life stages and the two species abundance levels. Specifically, we thinned all trees from 10% of the total stems up to 90% of the stems by maintaining their population structures in increments of 10% (Baldeck *et al.* 2013b; Punchi-Manage *et al.* 2014). We used the 'spatstat' package in *R* to perform the random thinning. In MRT analyses, we sampled 199 times and for each sampling with 100 cross-validation trials. We calculated the variance explained by the topography and the CVRE for each MRT analysis.

Unconstrained cluster analysis

The MRT is not only a form of multivariate regression but also a constrained cluster method (De'ath 2002). The clusters can define assemblage types and the environmental values can define the associated habitat types. It is helpful to compare the results of MRT to unconstrained cluster analysis. In this study, we determined the habitat types using Partitioning Around Medoids methods (pam) based on topographical variables. The pam is a more robust version of *K*-mean clustering method, because it minimizes a sum of dissimilarities instead of a sum of squared distances (Reynolds *et al* 1992). Then, we tested whether this is a significant difference in species compositions between different habitat types using Multiple Response Permutation Procedures (MRPP). The best partition for different habitats was selected as the clusters with maximum Dunn Index. This was performed using 'clValid' function in 'clValid' package of R. The pam and MRPP were conducted in 'cluster' package and 'vegan' package, respectively.

RESULTS

MRT partitioned habitats of all trees in the BDGS plot into five types: (1) low plateau; (2) intermediate plateau; (3) high plateau; (4) shallow valley; and (5) deep valley (Fig. 2; online supplementary Fig. S3f). Habitat types were mainly determined by four topographic variables (i.e. convexity, elevation, TWI and VDC). Convexity determined the first split for habitat types (break point = −0.033), accounting for 16% of species variance. Elevation, TWI and VDC determined the second (break point = 1443), third (break point = 9.2) and fourth (break point = 7.7) splits of habitat types, respectively. CVRE
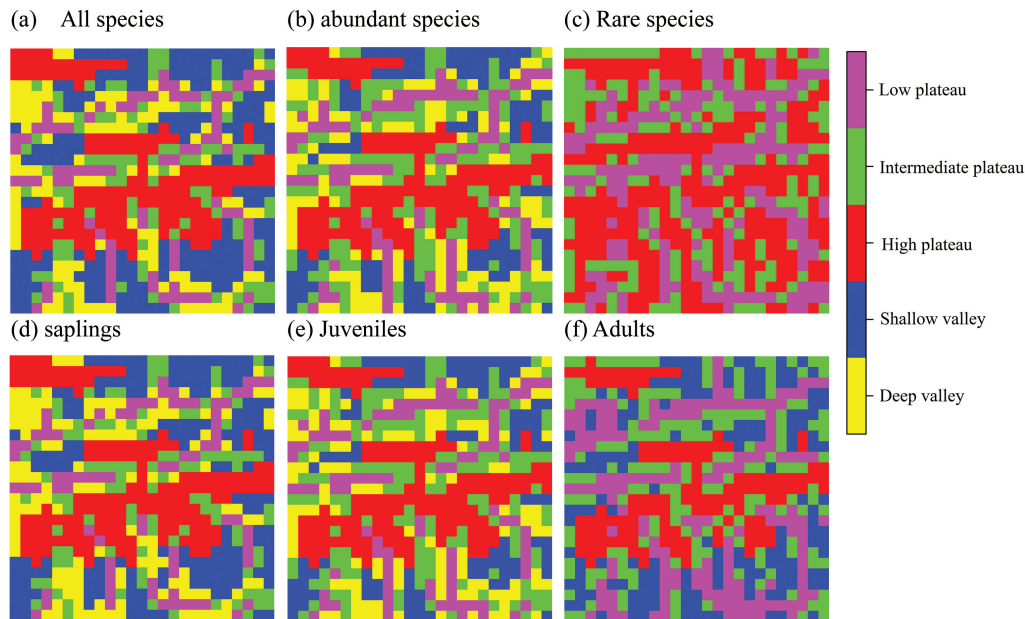
Figure 2: results of the multivariate regression tree (MRT) analyses for all trees, two species abundance levels (abundant and rare) and three life stages (sapling, juvenile and adult) in the 25-ha Badagongshan (BDGS) Forest Dynamic Plot (FDP) based on 20 × 20 m quadrats. (a) all species, (b) abundant species, (c) rare species, (d) saplings, (e) juveniles (f) adults. The five habitat types are as follows: (1) low plateau; (2) intermediate plateau; (3) high plateau; (4) shallow valley; and (5) deep valley.

of MRT was 0.719, and the species variance explained by MRT was 34.7% (Table 1). The corresponding average numbers of indicator species per plot for habitat types 1–5 were 0.143, 0.153, 0.270, 0.153 and 0.190. Indicator values were strongly positively correlated with species abundances (Fig. 3). MRT analyses for all trees showed that the proportions of variance explained by topographical variables increased, while the CVRE decreased with the increase of sample size (online supplementary Fig. S4).

Community–habitat associations for rare and abundant species

Habitats of abundant species were divided into five types, which was similar to the results for the whole community. In contrast, habitats of rare species were only grouped into three types: (1) low plateau, (2) intermediate and high plateau and (3) valley (Fig. 2; online supplementary Fig. S3d and e). The variance of species explained by habitat types was 33.8% for abundant species and 6.5% for rare species, and the corresponding CVREs were 0.71 and 0.95, respectively (Table 1). When the differences in sample sizes between abundant and rare species were accounted for, variances explained and the CVREs were still significantly different (Table 2; Fig. 4a). The average number of indicator species per habitat for abundant species and rare species was 22.6 and 4.3, respectively.

Community–habitat associations across life stages

MRT partitioned the habitats of saplings, juveniles and adults into five, five and four types, respectively (Fig. 2; online

supplementary Fig. S3a–c). The habitat types generated for different life stages were relatively similar, except that the habitats of low plateau identified at sapling and juvenile stages were merged with deep valley at the adult stage. For all three life stages, convexity was the first split for habitat types, accounting for 10.1–19.1% of species variance. For saplings, juveniles and adults, the species variances explained by MRT were 32.8%, 26.4% and 16.6%, respectively, and CVREs were 0.75, 0.78 and 0.87, respectively (Table 1). The number of indicator species per habitat for saplings, juveniles and adults is 16.8, 14 and 9.4, respectively. When the differences in sample sizes between saplings, juveniles and adult stages were accounted for, the variance explained by the topography was lower, and the CVRE was higher, albeit neither significantly, for the transition from saplings to juveniles. However, for the transition from juveniles to adults, the variance explained was significantly lower and the CVRE non-significantly higher (Table 2; Fig. 4b). Adults shared all indicator species with saplings and juveniles with very few exceptions (Table 1).

Unconstrained cluster analysis

The pam analyses divided the whole plot into six habitat types based on the maximum Dunn Index (online supplementary Fig. S5). The six habitat types were high plateau, intermediate plateau, low plateau, shallow valley, intermediate valley and deep valley (online supplementary Fig. S6). The MRPP tests showed that there were significant differences between six habitat types for all trees, abundant species, rare species, saplings, juveniles and adults ($P = 0.001$, online supplementary Table S2).

Table 1: results of MRT analyses and indicative species for all trees, three life stages (sapling, juvenile and adult) and two abundance levels (common and rare) in the 25-ha Badagongshan Forest Dynamic Plot, Central China

	Total no. of stems	Total no. of species	CVRE	Variance Explained	Total species	Total stems	Habitat type	Breakpoints	No. of sites	No. indicator Species
All species	186388	238	0.719 (±0.020)	0.347	194	32854	1 low plateau	Con ≥ −0.033, Elv < 1443, VDC < 7.7	112	3
					202	60228	2 intermediate plateau	Con ≥ −0.033, Elv < 1443, VDC ≥ 7.7	183	10
					182	54093	3 high plateau	Con ≥ −0.033, Elv ≥ 1443	148	19
					191	23755	4 shallow valley	Con < −0.033, TWI < 9.2	98	1
					181	15419	5 deep valley	Con < −0.033, TWI ≥ 9.2	84	4
Abundant species	175461	59	0.709 (±0.021)	0.338	59	36516	1 low plateau	Con ≥ −0.011, Elv < 1443, VDC < 13	120	3
					59	34957	2 intermediate plateau	Con ≥ −0.011, Elv < 1443, VDC ≥ 13	111	10
					59	50225	3 high plateau	Con ≥ −0.011, Elv ≥ 1443	142	17
					59	37355	4 shallow valley	Con < −0.011, TWI < 9.2	157	1
					59	16374	5 deep valley	Con < −0.011, TWI ≥ 9.2	95	7
Rare species	10927	179	0.955 (±0.001)	0.065	151	4557	3 plateau	Con ≥ −0.008	279	2
					148	3013	4 shallow valley	Con < −0.004, Con ≥ −0.034	167	0
					149	3352	5 deep valley	Con < −0.008, Con < −0.034	179	0
Saplings	100305	164	0.748 (±0.009)	0.328	160	19755	1 low plateau	Con ≥ −0.033, Elv < 1443, VDC < 8.7	128	1
					150	29975	2 intermediate plateau	Con ≥ −0.033, Elv < 1443, VDC ≥ 8.7	167	8
					141	31002	3 high plateau	Con ≥ −0.033, Elv ≥ 1443	148	16
					152	12686	4 shallow valley	Con < −0.033, TWI < 9.2	103	0
					142	6869	5 deep valley	Con < −0.033, TWI ≥ 9.2	79	3
Juveniles	63358	164	0.777 (±0.008)	0.264	151	13775	1 low plateau	Con ≥ −0.011, Elv < 1443, VDC < 13	124	2
					136	11877	2 intermediate plateau	Con ≥ −0.011, Elv < 1443, VDC ≥ 13	107	7
					128	17056	3 high plateau	Con ≥ −0.011, Elv ≥ 1443	142	10
					157	13987	4 shallow valley	Con < −0.011, TWI < 9.2	157	0
					147	6648	5 deep valley	Con < −0.011, TWI ≥ 9.2	95	3
Adults	22455	164	0.868 (±0.009)	0.167	139	6395	2 intermediate plateau	Con ≥ −0.007, Elv < 1443	162	2
					106	4651	3 high plateau	Con ≥ −0.007, Elv ≥ 1443	128	5
					142	6478	4 shallow valley	Con < −0.007, TWI < 8.3	172	3
					144	4927	5 deep valley	Con < −0.007, TWI ≥ 8.3	163	2

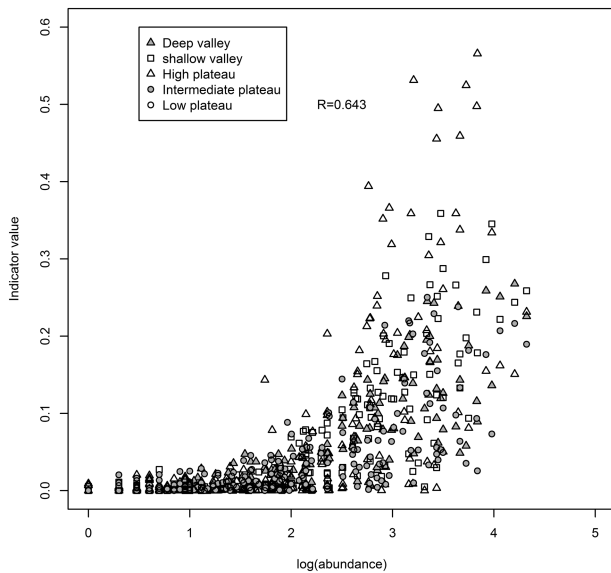Abbreviations: Con = convexity, Elv = Elevation.

Figure 3: the relationship between indicator values and species abundances in the 25-ha Badagongshan plot. The five habitat types are as follows: (1) low plateau; (2) intermediate plateau; (3) high plateau; (4) shallow valley; and (5) deep valley.

Table 2: the differences in variance explained by topographical habitats and the differences in Cross Validate Relative Error (CVRE) for MRT analyses between rare and abundant species; saplings and juveniles; and juveniles and adults

	Rare–Abundant	Sapling–Juvenile	Juvenile–Adult
Sample size	10 927	22 455	22 455
The variance difference (%)	**6.722**	−2.211	**−2.646**
CVRE difference	**−0.051**	0.019	0.015

The variance difference means the variance explained for abundant species (or the latter life stage) minus the variance explained for rare species (or the early life stage). CVRE difference means CVRE for abundant species (or the latter life stage) minus CVRE for rare species (or the early life stage). Bold text indicates that 95% of difference values computed from random samples are all larger (or all smaller) than zero, which is considered to be significant.

DISCUSSIONS

Topographic ridge-valley gradients (catenas) play a fundamental role in shaping plant species distributions (Chuyong *et al.* 2011; Valencia *et al.* 2004; Webb and Peart 2000). The BDGS plot can be generally divided into five topographic habitat types using MRT similar to the results of other tropical and subtropical forests. Using similar methods, Legendre *et al.* (2009), Kanagaraj *et al.* (2011) and Punchi-Manage *et al.* (2013) also divided the habitats into five types in Gutianshan, Barro Colorado Island (BCI) and Sinharaja forest plots, respectively. In general, the habitat classification presented by unconstrained cluster analyses coincided with the result of MRT in BDGS plot. They all detected the five or six ridge-valley gradients at this plot.

Convexity and elevation were the two most important topographical variables related to local species composition, accounting for 25.4% of local species variance in the BDGS plot; they are also related to hydrological processes and disturbance regimes (Moeslund *et al.* 2013). Two local hydrological variables, VDC and TWI that were also contained into the model, provided indirect evidence of the importance of a hydrological niche in shaping local species composition in mountain subtropical forests (Silvertown *et al.* 2015). Convexity was the first split of MRT in BDGS, in contrast to elevation that was the first split, while convexity was not included in the MRT for BCI and Sinharaja. Similarly, Brown *et al.* (2013) found a strong correlation between variance in pairwise spatial associations and topographic heterogeneity across 14 forest plots with an area of 16–52 ha.

The species assemblage variance explained by topographical variables in the BDGS plot (34.7%) was higher than that in both Sinharaja (22.73%) and BCI (~20.8%) plots. The higher explanatory power of topography habitats in the BDGS plot may suggest stronger driving forces of topography for the assembly of tree communities than in other forests (Baldeck *et al.* 2013b). The proportion of variance unexplained by topography (65.3%) may be attributed to stochasticity or habitat-independent spatial processes such as dispersal limitation. For example, several studies showed that habitat-independent spatial processes accounted for ca., a quarter of assemblage composition variance in subtropical or tropical forests (e.g. Legendre *et al.* 2009; Punchi-Manage *et al.* 2014). Such unexplained variance could also be due to unaccounted environmental variables, such as edaphic variables and light availability. For example, recent studies showed the importance of soil properties on the tree assemblages (Baldeck *et al.* 2013a; John *et al.* 2007). Topographical habitat filtering and other habitat-independent processes combinedly shaped local species compositions in this forest.

Abundance and community–habitat associations

Most of the articles that tested habitat associations at species levels showed weaker species–habitat associations for rare species than for abundant species (Gunatilleke *et al.* 2006). This is not surprising because habitat associations for individual species strongly depend on species abundances, and species–habitat associations for rare species are difficult to analyze. In this regard, a reasonable way to understand species–habitat associations for rare species is to pool rare species into a community (abundant species were also pooled, but they are less sensitive). Our results using MRT analysis showed that abundant and rare species had different habitat-driven assemblages. As expected, rare species demonstrated much weaker community–habitat associations than abundant species with and without controlling for sample sizes. Previous studies found that rare species tended to be more spatially aggregated than abundant species (Condit *et al.* 2000; Guo *et al.* 2013), and rare species with low abundance may be unable to distribute in all potential habitats. This suggests that habitat filtering was relatively unimportant
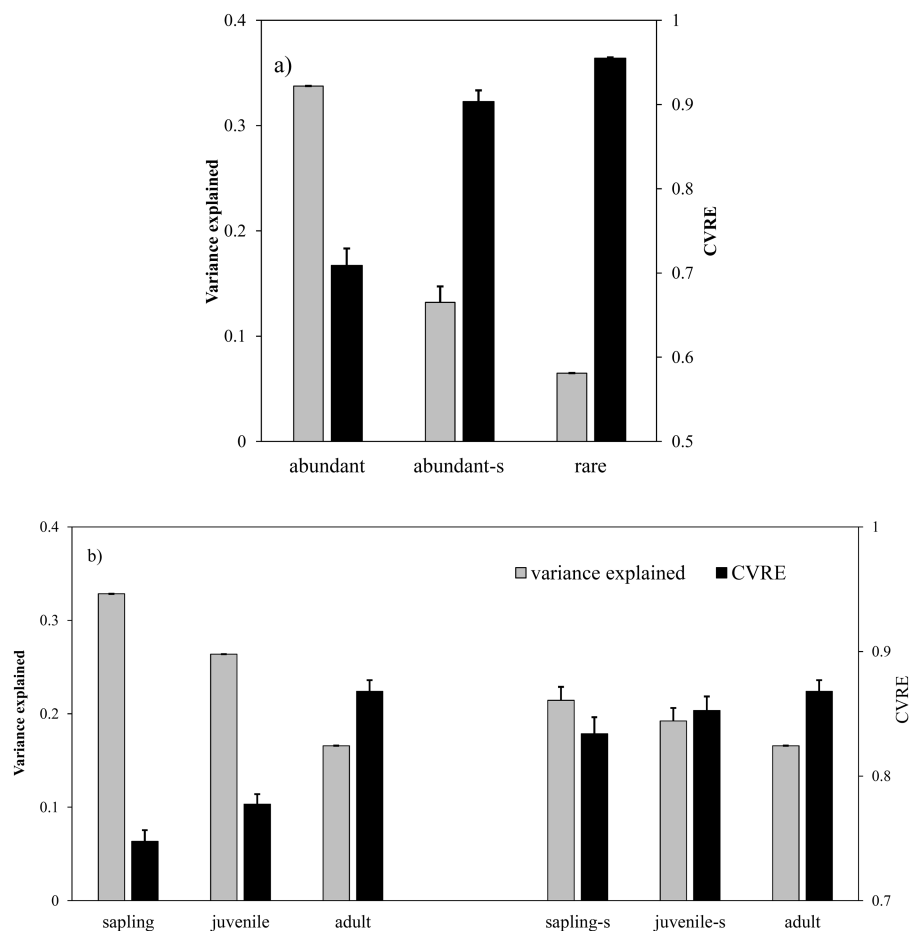
Figure 4: the degree of community -habitat association (species variance explained by habitats), cross-validated relative error (CVRE) and sampling size effects for two species abundance levels and three life stages. (a) two species abundance level, (b) three life stages. The symbols sapling-s and juvenile-s indicate saplings and juveniles assigned to the total number of adult individuals; abundant-s indicates abundant species assigned to the total number of rare specie individuals.

in shaping rare species' distributions in this forest. Other processes such as dispersal limitation may be the main driving force in shaping rare species' distributions. For example, Rossetto *et al.* (2008) found that dispersal limitation, rather than habitat specificity, restricted the distribution of the rare and endemic *Elaeocarpus sedentarius* rainforest tree in Australia. However, it is not clear how unaccounted variables, especially soil properties, could explain the rare species–habitat associations in this forest plot. On the other hand, topography is a good proxy for capturing small microclimatic changes, and hence inclusion of soil properties may not increase the understanding of rare species distributions of this forest plot.

We also noted that variance explained by the topography and CVRE was very similar for abundant species and all species. This may suggest that community–habitat associations are mainly driven by a handful of abundant species rather than a mass of rare species. This idea is further supported by indicator values that correlated strongly and positively with species abundance, and the abundant species had more indicator species than rare species. It is difficult to perform

analyses in an unbiased way for rare species, since their distributions are masked by the dominant species that shape the community. When using the rare species as indicators of all five habitats that were based on all species, we only detected zero indicators (all indicator values are smaller than 0.25). According to the definition of rare species by Gaston (1994), rare species may use spatially restricted resources. In this study, the whole forest plot grouped into five or six types of habitats. Therefore, this kind of analyses may bias toward abundant species besides the sample size effects.

The community–habitat associations across life stages

We found that the variance explained by habitats decreased and the CVRE increased (more 'noise') during the transition from saplings and juveniles to adults. Under the circumstances that there were no intense disturbances, this suggested that competition may weaken habitat associations over time (Kanagaraj *et al.* 2011). In addition, this could be due to that the regeneration niche is more restricted than adult niche

(Grubb 1977). However, the differences in the variance explained and the CVRE were very small, which may reflect that during the transition from saplings to juveniles and from juveniles to adults, stochastic effects outweighed habitat filtering and/or competition effect or habitat filtering and competition processes counterbalanced each other. Baldeck *et al.* (2013a) found that changes in variance explained by environmental variables among life stages were small and inconsistent. They interpreted it as the consequence of neutral mortality rather than habitat filtering. Limited community–habitat associations allude for the BDGS plot that stochasticity outweighs deterministic effects (e.g. habitat filtering) due to unpredictable neighborhoods in species-rich forest.

We found that the topographic habitat types were relatively stable for different life stages, which was similar to the results from the Sri Lankan Sinharaja plot (except depression and NE hilltop habitats; Punchi-Manage *et al.* 2013) but different from the BCI plot in Panama (Kanagaraj *et al.* 2011). The BCI forest differs with respect to other forest plots as a result of frequent visits of primates and large birds that cause ontogenetic shifts of species–habitat associations. Meanwhile, we found that individual species may shift habitat associations during their life stages change as shown in previous studies (Comita *et al.* 2007; Webb and Peart 2000). The constancy in habitat types across life stage may be because species suffer strong seed dispersal limitations that cause juveniles to locate adults nearby (Webb and Peart 2000). In this forest plot, we found that almost all indicator species for the habitats defined by adults were also indicator species for the habitats defined by juveniles and saplings.

Effects of sample size on community–habitat associations

We found weaker community–habitat associations for adult trees than for sapling and juvenile trees before controlling for sample size. In the BDGS plot, the number of stems in saplings, juveniles and adult cohorts was 100 305, 63 358, and 22 455, respectively. We found that degree of community–habitat associations for saplings was 2-fold higher than that for juvenile and adult trees when the sample sizes were not controlled for, but relatively stable across the three cohorts after controlling for sample size. Our results corroborate those where species variance explained by environmental variables increases with sample size (Baldeck *et al.* 2013b; Jones *et al.* 2008). Punchi-Manage *et al.* (2013) found that adult trees and recruitment showed weak habitat associations compared to juveniles and all trees at the Sinharaja forest plot. However, the number of individuals at adult and recruitment stages was much smaller than the juvenile stage. Thus, Baldeck *et al.* (2013b) and Punchi-Manage *et al.* (2014) suggested that community–habitat associations at different life stages were directly related to the number of individuals of the specific stages.

Since rare species have a much lower sample size than abundant species, such sample size effects were also obvious for the comparison analysis between rare and abundant species. Mazaris *et al.* (2013) suggested that after accounting for sample size, the contribution of abundant and rare species to the overall species richness pattern was much less. Our analyses showed that when taking sample sizes into account, the differences in the species variance explained by topography and CVREs for abundant and rare species were both much lower.

CONCLUSIONS

In this study, we assessed the relative contribution of topographic variables to tree community assembly in a subtropical forest with strong topographic heterogeneity. Over one-third of community composition variance could be explained by topographic habitats in this plot. However, the degree of community–habitat association for rare species was weaker than that of abundant species, suggesting that habitat filtering was not as important in shaping rare species distributions in this species-rich subtropical forest. In addition, the differences in degree of community–habitat associations among sapling, juveniles and adults were very small, suggesting that the relative importance of habitat filtering was weak during the transitions from saplings to juveniles and from juvenile to adults.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of Plant Ecology* online.

FUNDING

ACKNOWLEDGEMENTS

REFERENCES

Baldeck CA, Harms KE, Yavitt JB, *et al*. (2013a) Soil resources and topography shape local tree community structure in tropical forests. *Proc R Soc Lond B Biol Sci* **280**:1753.

Baldeck CA, Harms KE, Yavitt JB, *et al*. (2013b) Habitat filtering across tree life stages in tropical forest communities. *Proc R Soc B Biol Sci* **280**:1766.

Brown C, Burslem D, Illian JB, *et al.* (2013) Multispecies coexistence of trees in tropical forests: spatial signals of topographic niche differentiation increase with environmental heterogeneity. *Proc R Soc Lond B Biol Sci* **280**:1764.

Chuyong GB, Kenfack D, Harms KE, *et al.* (2011) Habitat specificity and diversity of tree species in an African wet tropical forest. *Plant Ecol* **212**:1363–74.

Comita LS, Condit R, Hubbell SP (2007) Developmental changes in habitat associations of tropical trees. *J Ecol* **95**:482–92.

Comita LS, Muller-Landau HC, Aguilar S, *et al.* (2010) Asymmetric density dependence shapes species abundance in a tropical tree community. *Science* **329**: 330–2.

Condit R, Aston P, Baker P, *et al.* (2000) Spatial patterns in the distribution of tropical tree species. *Science* **288**:1414–8.

Dalling J, Muller-Landau H, Wright S, *et al.* (2002) Role of dispersal in the recruitment limitation of neotropical pioneer species. *J Ecol* **90**:714–27.

Daws MI, Pearson TRH, Burslem D, *et al.* (2002) Topographic position affects the water regime in a semideciduous tropical forest in Panama. *Plant Soil* **238**:79–90.

De'ath G (2002) Multivariate regression trees: a new technique for modeling species-environment relationships. *Ecology* **83**:1105–17.

Dufrene M, Legendre P (1997) Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecol Monographs* **67**:345–66.

Engelbrecht BMJ, Comita LS, Condit R, *et al.* (2007) Drought sensitivity shapes species distribution patterns in tropical forests. *Nature* **447**:80–2.

Gaston KJ (1994) *Rarity*. London: Chapman and Hall.

Grubb PJ (1977) The maintenance of species richness in plant communities: the importance of the regeneration niche. *Biol Rev Camb Philos Soc* **52**:107–45.

Gunatilleke CVS, Gunatilleke IAUN, Esufali S, *et al.* (2006) Species-habitat associations in a Sri Lankan dipterocarp forest. *J Trop Ecol* **22**:371–84.

Guo Y, Lu J, Franklin SB, *et al.* (2013) Spatial distribution of tree species in a species-rich subtropical mountain forest in central China. *Can J For Res* **43**:826–35.

Harms KE, Condit R, Hubbell SP, *et al.* (2001) Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. *J Ecol* **89**:947–59.

Hubbell SP (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. New Jersey: Princeton University Press.

Hubbell SP, Foster RB (1986) Biology, chance and history and the structure of tropical rain forest tree communities. In Diamond JM, Case TJ (eds). *Community Ecology*. New York: Harper and Row, 314–29.

Hutchinson GE (1961) The paradox of the plankton. *Am Nat* **95**:137–45.

John R, Dalling JW, Harms KE, *et al.* (2007) Soil nutrients influence spatial distributions of tropical tree species. *Proc Natl Acad Sci USA* **104**:864–69.

Jones MM, Tuomisto H, Olivas PC (2008) Differences in the degree of environmental control on large and small tropical plants: just a sampling effect? *J Ecol* **96**:367–77.

Kanagaraj R, Wiegand T, Comita LS, *et al.* (2011) Tropical tree species assemblages in topographical habitats change in time and with life stage. *J Ecol* **99**:1441–52.

Kraft NJ, Adler PB, Godoy O, *et al.* (2015) Community assembly, coexistence and the environmental filtering metaphor. *Funct Ecol* **29**:592–9.

Larsen DR, Speckman PL (2004) Multivariate regression trees for analysis of abundance data. *Biometrics* **60**:543–9.

Legendre P, Mi X, Ren H, *et al.* (2009) Partitioning beta diversity in a subtropical broad-leaved forest of China. *Ecology* **90**:663–74.

Magurran AE (2004) *Measuring Biological Diversity*. Oxford: Blackwell Publishing.

Mazaris AD, Tsianou MA, Sigkounas A, *et al.* (2013) Accounting for the capacity of common and rare species to contribute to diversity spatial patterns: is it a sampling issue or a biological effect? *Ecol Indic* **32**:9–13.

Moeslund JE, Arge L, Bøcher PK, *et al.* (2013) Topography as a driver of local terrestrial vascular plant diversity patterns. *Nord J Bot* **31**:129–44.

Poorter L, Bongers F, Sterck FJ, *et al.* (2005) Beyond the regeneration phase: differentiation of height–light trajectories among tropical tree species. *J Ecol* **93**:256–67.

Punchi-Manage R, Getzin S, Wiegand T, *et al.* (2013) Effects of topography on structuring local species assemblages in a Sri Lankan mixed dipterocarp forest. *J Ecol* **101**:149–60.

Punchi-Manage R, Wiegand T, Wiegand K, *et al.* (2014) Effect of spatial processes and topography on structuring species assemblages in a Sri Lankan dipterocarp forest. *Ecology* **95**:376–86.

R Development Core Team (2016) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Reynolds AP, Richards G, de la Iglesia B, *et al.* (1992) Clustering rules: a comparison of partitioning and hierarchical clustering algorithms. *J Math Model Algorithms* **5**:475–504.

Rossetto M, Kooyman R, Sherwin W, *et al.* (2008) Dispersal limitations, rather than bottlenecks or habitat specificity, can restrict the distribution of rare and endemic rainforest trees. *Am J Bot* **95**:321–9.

Silvertown J, Araya Y, Gowing D (2015) Hydrological niches in terrestrial plant communities: a review. *J Ecol* **103**:93–108.

Siqueira T, Bini LM, Roque FO, *et al.* (2012) Common and rare species respond similar niche processes in macroinvertebrate metacommunities. *Ecography* **35**:183–92.

Svenning JC (2001) On the role of microenvironmental heterogeneity in the ecology and diversification of neotropical rain-forest palms (Arecaceae). *Bot Rev* **67**:1–53.

Tarboton DG (1997) A new method for determination of flow directions and upslope areas in grid digital elevation models. *Water Resour Res* **33**:309–19.

Valencia R, Foster RB, Villa G, *et al.* (2004) Tree species distributions and local habitat variation in the Amazon: large forest plot in eastern Ecuador. *J Ecol* **92**:214–29.

Wang Q, Bao D, Guo Y, *et al.* (2014a). Species associations in a species-rich subtropical forest were not well-explained by stochastic geometry of biodiversity. *PLOS ONE* **9**:e97300.

Wang Z, Fang J, Tang Z, *et al.* (2012) Geographical patterns in the beta diversity of China's woody plants: the influence of space, environment and range size. *Ecography* **35**:1092–102.

Wang Q, Xu Y, Lu Z, *et al.* (2014b). Disentangling the effects of topography and space on distributions of dominant species in a subtropical forest. *Chin Sci Bull* **59**:5113–22.

Webb CO, Peart DR (2000) Habitat associations of trees and seedlings in a Bornean rain forest. *J Ecol* **88**:464–78.