

Spatial aggregation of tropical trees at multiple spatial scales

Maxime Réjou-Méchain^{1,2*}, Olivier Flores¹, Nils Bourland³, Jean-Louis Doucet³, Richard F. Fétéké^{3,4}, Alexandra Pasquier² and Olivier J. Hardy⁵

¹UMR 5175 CEFE, CNRS, 1919 Route de Mende, 34293 Montpellier Cedex 5, France; ²CIRAD, 'Forest resources and public policies' Research Unit, Campus International de Baillarguet, TA 10C, BP 5035, Montpellier 34035, France; ³Laboratoire de Foresterie des Régions tropicales et subtropicales – Unité de Gestion des Ressources forestières et des Milieux Naturels. Gembloux Agro-Bio Tech, Université de Liège, Passage des Déportés 2, 5030 Gembloux, Belgium; ⁴Société d'Exploitation Forestière Pallisco, Cellule Inventaires et Aménagement, B.P. 394, Douala, Cameroun; and ⁵Evolutionary Biology and Ecology, Faculté des Sciences, Université Libre de Bruxelles, CP 160/12, 50 Av. F. Roosevelt, 1050 Brussels, Belgium

Summary

1. In tropical forests, species distribution patterns may be strongly context-dependent owing to local stochasticity of recruitment and/or to the specific history and environment of each site. Recent studies have reported, however, that the degree of spatial aggregation of tropical tree species is partly determined by some species traits irrespectively of site conditions, at least at a very local scale (< 200 m).

2. Here, we used standardized large-scale forest inventories of five Central African tropical forests (9670 0.5-ha plots spread over 5550 km²) to quantify the spatial aggregation of 106 tropical tree species at larger spatial scales. For this purpose, we developed a new statistic to quantify the respective contributions of different spatial scales to the aggregation patterns, and we tested whether patterns were consistent across sites. We finally asked whether species characteristics related to dispersal ability, to response to disturbances and to biogeographical range could significantly explain aggregation patterns.

3. Although aggregation patterns varied substantially among sites within each species, they displayed inter-site consistencies (21–24% of the total variance explained by species identity) at the local scale (0.2–1 km) and at the mesoscale (1–10 km) but not at the landscape scale (> 10 km). At the two former scales, upper taxonomical levels (family and/or order) significantly explained variation in the degree of species aggregation, while at the landscape scale, aggregation was entirely contingent on the site considered. Few species characteristics, except dispersal syndromes and wood density, were able to significantly explain aggregation patterns.

4. Synthesis. One of our most striking results is the high context dependence of species aggregation patterns, whatever the spatial scale considered. However, we showed that species distribution patterns can be predicted, to an extent, at spatial scales much larger than previously investigated in this context. Such patterns may be explained by traits displaying phylogenetic conservatism (such as dispersal syndrome), but further studies are necessary to clearly identify them.

Key-words: Central Africa, determinants of plant community diversity and structure, dispersal limitation, distribution patterns, niche differentiation, species traits, tropical forest, variance partitioning

Introduction

In order to understand the mechanisms controlling species spatial distribution, much effort has been devoted to the analysis of species turnover (beta diversity) among local communities,

attempting to assess the relative roles of abiotic filters, species interactions, dispersal limitation and/or biogeographical history (e.g. Condit *et al.* 2002; Tuomisto, Ruokolainen & Yli-Halla 2003; Hardy & Sonké 2004; Jones *et al.* 2006; Réjou-Méchain & Hardy 2011). A complementary approach consists of investigating the degree of spatial aggregation within each population individually (i.e. the extent to which

*Correspondence author. E-mail: maxime.rejou@gmail.com

conspecific individuals tend to occur in spatial clusters). This may help infer which processes are important in shaping species distribution patterns (Tilman & Kareiva 1997; Condit *et al.* 2000). Indeed, the species aggregation pattern is a major determinant of the species–area relationship (Plotkin *et al.* 2000) and could give insights on how a species uses resources (Condit *et al.* 2000; Montoya *et al.* 2009) or is affected by its dispersal ability (Hubbell 1979; Seidler & Plotkin 2006). Distribution of conspecific adult trees could also show a trend of overdispersion at small spatial scales following density- or distance-dependent juvenile mortality because of predation (Janzen 1970; Connell 1971; Bell, Freckleton & Lewis 2006) or competition (Kenkel 1988; Stoll & Bergius 2005).

Species aggregation patterns may thus reflect underlying processes, but as the spatial scale investigated changes, processes, and thus patterns, are also expected to change. Indeed, processes occurring at small scales (e.g. competition or predation) can differ from those prevailing at large scales (e.g. biogeographical processes; Wiens 1989; Levin 1992; Schluter & Ricklefs 1993; Huston 1999). Thus, identifying which patterns are preserved or lost as we move from one scale to another can enhance our understanding of the specific processes that control species distribution and their scale of influence (Levin 1992).

Recently, Montoya *et al.* (2009) showed that the degree of tree species aggregation in temperate forests was substantially underestimated by species distribution models

based on environmental predictors. They suggested that such high unexplained spatial structures may result from species-specific processes, irrespective of environmental conditions. In tropical forests, tree species distribution patterns may be strongly context-dependent, first because of strong local stochasticity in establishment and recruitment processes (Hubbell 2001) and secondly because species distribution is often strongly contingent to the specific characteristics of each site (e.g. past disturbances or environmental heterogeneity). Several studies have shown that the interaction between the environmental heterogeneity and the biological characteristics of species (i.e. species traits) can influence distribution patterns at various spatial scales (e.g. Muller-Landau 2004; ter Steege *et al.* 2006; Engelbrecht *et al.* 2007; Franklin & Rey 2007; Baltzer *et al.* 2008). Nevertheless, it has also been shown that the degree of spatial aggregation of species can be partly predicted by species traits that do not interact with the environment, which has, for example, been shown with dispersal syndrome at least at the local scale (0–200 m; Hubbell 1979; Flores 2005; Seidler & Plotkin 2006). Hence, such traits may also confer an endogenous property to species aggregation patterns, irrespective of the site context. Therefore, quantifying how aggregation patterns vary among species and/or according to specific biological characteristics at multiple spatial scales can help formulate hypotheses about the main processes driving species distribution.

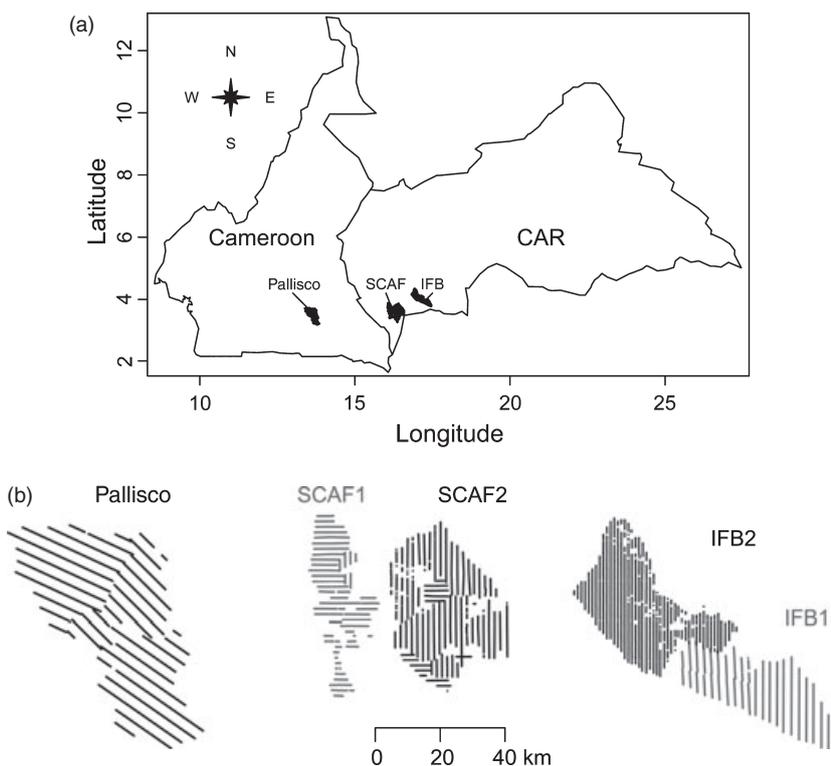


Fig. 1. The study sites. (a) Location of the sites in Cameroon and in the Central African Republic (CAR) (in black). (b) The sampling design of each site showing transects along which 0.5-ha plots ($20 \times 250 \text{ m}^2$ or $25 \times 200 \text{ m}^2$) were sampled for trees with d.b.h. ≥ 30 cm (2298 plots for Pallisco, 1429 for SCAF1, 3108 for SCAF2, 1277 for IFB1 and 1558 for IFB2).

Previous studies of tropical tree aggregation patterns have primarily concerned single sites and often a single (but see Plotkin, Chave & Ashton 2002; Wiegand *et al.* 2007) and relatively fine spatial scale (≤ 50 ha; Hubbell 1979; He, Legendre & LaFrankie 1997; Condit *et al.* 2000; Plotkin *et al.* 2000; Plotkin, Chave & Ashton 2002; Aiba, Kitayama & Takyu 2004; Franklin & Rey 2007; Picard *et al.* 2009). However, assessing whether aggregation patterns differ among species at multiple scales requires data on several large sites to obtain replicates. To our knowledge, such an approach has not yet been taken in tropical forests, probably because of the difficulties to generate an appropriate data set.

Here, we use a large and unique data set (284 587 trees with diameter at breast height (d.b.h.) ≥ 30 -cm) to assess the distribution of 106 species in five sites covering *c.* 5550 km² of Central Africa (Fig. 1), the second largest, but still poorly known, area of tropical rain forest in the world (Stocks *et al.* 2008). Data come from surveys conducted by timber companies in the south-east Cameroon and south-west Central African Republic to implement forest management plans (Réjou-Méchain *et al.* 2008, 2011). We developed a new statistic (A_s index) to quantify aggregation patterns independently at three different spatial scales (local: 0.2–1 km; meso: 1–10 km; landscape: 10–80 km) for each species in each site (Fig. 2). As a first objective, we asked whether species aggregation patterns are consistent across sites at each spatial scale.

Assuming an increasing influence of environmental and biogeographical processes on distribution patterns with larger spatial scales (Levin 1989, 1992; Wiens 1989), we expect that the consistencies of aggregation patterns across sites tend to fade when the spatial scale increases. Similarly, considering that closely related species tend to share similar biological characteristics (e.g. Casper, Heard & Apanius 1992; Cavender-Bares, Kitajima & Bazzaz 2004; Chave *et al.* 2006) that likely influence species aggregation patterns, we hypothesize that species of a same lineage share similar aggregation patterns at smaller scales.

As a second objective, we seek to determine whether differences in aggregation patterns can be explained by species characteristics related to dispersal ability and response to disturbance. In particular, we expect stronger aggregation in three situations: first, in species with unassisted dispersal compared with wind- and animal-dispersed species at the smaller scales, because seed dispersal distances are expected to be shorter (Hubbell 1979; Howe & Smallwood 1982; Flores, Gourlet-Fleury & Picard 2006; Seidler & Plotkin 2006); second, in light-demanding species, usually characterized by low wood density (Turner 2001; van Gelder, Poorter & Sterck 2006), than in shade-tolerant species at the smaller scales, as a consequence of pioneer recruitment in spatially structured gaps (Denslow 1980, 1987; Franklin & Rey 2007); and thirdly, in species with restricted phylogeographical range size, which

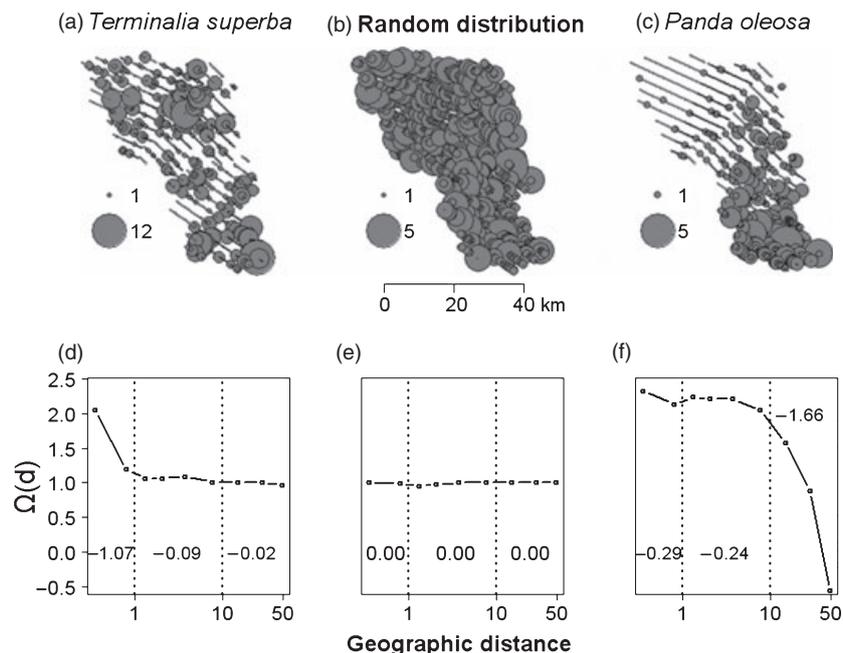


Fig. 2. Quantification of spatial aggregation of tree species. Three examples of species distribution at the Pallisco site are represented by the number of individuals per 0.5-ha plot, proportional to circle size (top), and the corresponding pair-correlation functions ($\Omega_s(d)$) against the logarithm of the distance (d) (bottom). $\Omega_s(d)$ is the mean density of conspecific individuals at a distance d from existing conspecifics, divided by the species' mean density. Dotted lines on the lower graphs delimit the distance ranges investigated (local scale: 0.2–1 km; mesoscale: 1–10 km; landscape scale 10–70 km) and the slopes of Ω_s against $\ln(d)$ within these three distance ranges are given inside the figures. The contribution of a given scale to the aggregation pattern, A_s , was quantified as minus the slope. (a, d) Aggregation in *Terminalia superba* (1333 individuals) occurs at the local scale while larger scales contribute little; (b, e) an artificial data set devoid of aggregation pattern by randomizing individuals of *Terminalia superba* among plots ($A_s = 0$ at all scales); (c, f) aggregation in *Panda oleosa* (486 individuals) occurs essentially at the landscape scale while smaller scales contribute little. See also Appendix S1 and Figs S1–S3 for additional illustration on simulated data.

is often correlated with recent speciation (Paul & Tonsor 2008), low dispersal ability (Edwards & Westoby 1996) and/or high habitat specialization (Gaston 2003).

Materials and methods

FLORISTIC DATA

The study sites are located in Cameroon and Central African Republic (CAR) in a humid tropical climate. According to the WORLDCLIM data base (Hijmans *et al.* 2005), annual rainfall ranges from 1487 to 1645 mm in the whole study area, which is covered by mixed moist semi-evergreen Guineo-Congolian rain forests according to White (White 1983). Floristic data were obtained from forest inventories conducted by three logging companies: Pallisco, Industrie Forestière de Batalimo (IFB) and Société Centrafricaine Forestière (SCAF). Extensive control resampling performed on previously sampled plots in CAR ($n = 1107$; Réjou-Méchain *et al.* 2011) showed that commercial forest inventories are well suited to detect patterns of floristic variation. The first study site is located in south-east Cameroon and the other ones in south-west CAR (Fig. 1a). Inventories were conducted in 2003 for Pallisco, 1993–1996 for IFB and 2005–2006 for SCAF. They were based on a systematic sampling design within rectangular plots of 0.5 ha (250×20 m for Pallisco and 200×25 m for IFB and SCAF) that were contiguous (along the small side of plots) on parallel transects separated by 2 or 3 km except for some part of the IFB site (IFB2) where plots were non-adjacent and transects were closer (Fig. 1b).

In most analyses, we separated the SCAF concession into two study areas, SCAF1 and SCAF2, which differ with respect to edaphic conditions: clay soils for SCAF1 and sandy soils for SCAF2. Plots located close to soil type limits were eliminated to avoid border effects and errors owing to imprecision of the soil map. We also separated the IFB concession into two sites, IFB1 and IFB2, which differed with respect to past disturbance. Unlike IFB1, IFB2 has experienced important disturbances (repeated fire events and logging) in the past, and some areas are currently in a transition phase towards savannas (Boulvert 1986; Réjou-Méchain *et al.* 2008). Finally, we eliminated plots containing < 10 individuals with d.b.h. > 30 cm. The total number of plots was 9670 (2298 for Pallisco, 188 000 ha; 1429 for SCAF1, 65 000 ha; 3108 for SCAF2, 143 000 ha; 1277 for IFB1, 70 000 ha; and 1558 for IFB2, 89 000 ha).

Census protocols, which were standardized among the three timber companies, facilitated comparison across the entire study region. All trees with a d.b.h. ≥ 30 cm were identified, when possible, in each plot at species level (229 identified taxa in SCAF, 287 in IFB and 442 in Pallisco). In tropical rain forests, identification of tree species can be doubtful, especially when inventories are implemented by timber companies, mostly interested in commercial species. In the case of the CAR, funded programmes have ensured an interactive control of the botanical identifications increasing the reliability of species identification (Réjou-Méchain *et al.* 2011). Furthermore, based on expert knowledge, we eliminated species for which botanical knowledge was limited and for which confusion with other species was possible. For each site, species with fewer than 100 individuals were also not considered to ensure sufficient precision of the measures of species aggregation. We thus focused our attention on 106 species (accounting for 70.7% of all individuals > 30 cm d.b.h.) for which botanical identifications were reliable (51 in Pallisco, 68 in SCAF1, 82 in SCAF2, 64 in IFB1 and 64 in IFB2) and which correspond to 92 genera, 31 families and 14 orders (Table S1 in Supporting Information). Taxonomic

classification at the order and family levels followed that of the Angiosperm Phylogeny Group (APG II 2003), and botanical synonyms of order and species were updated according to the African Flowering Plants Database (<http://www.ville-ge.ch/musinfo/bd/cjb/africa/>).

QUANTIFYING SPATIAL AGGREGATION

Aggregation can be defined as a decrease in density with spatial scale. Patterns of spatial aggregation of each species s were characterized by the pair-correlation function $\Omega_s(d)$ (Wiegand & Moloney 2004; Cousens, Dytham & Law 2008), also called 'relative neighbourhood density function' (Condit *et al.* 2000) or 'radial distribution function' in statistical mechanics. Applied to our data, $\Omega_s(d)$ is the average density of individuals of species s at a distance d from existing conspecific trees divided by the mean density of species s over the whole sampled area (Condit *et al.* 2000; Fig. 2). Using the plot data, $\Omega_s(d)$ was estimated by a numerical approach *sensu* Wiegand & Moloney (2004). First, the quantity $\Omega_{sij} = x_{si}x_{sj}/\bar{x}_s^2$ was computed for each pair of plots i, j , where x_{si} and x_{sj} are the abundances (number of individuals) of species s in plot i and j , respectively, and \bar{x}_s is the mean abundance of species s over the whole set of plots in the study site. $\Omega_s(d)$ can then be obtained by averaging Ω_{sij} over all pairs of plots separated by a distance d . In practice, one must consider distance classes and, for graphical representations (Fig. 2), Ω_{sij} was averaged over seven distance classes defined as (in km):]0, 0.5],]0.5, 1],]1, 1.7],]1.7, 2.5],]2.5, 5],]5, 10],]10, 20],]20, 40],]40, 90]. Under a random spatial distribution, $\Omega_s(d)$ is approximately equal to 1 for all d . When aggregation occurs, $\Omega_s(d) > 1$ for low d and $\Omega_s(d) < 1$ for large d (Fig. 2 and Figs S1–S3).

$\Omega_s(d)$ is closely related to Ripley's K statistic (Ripley 1981) but has the advantage of isolating specific distance classes because it is not a cumulative distribution as is Ripley's K , which confounds effects at large distances with effects at shorter distances (Condit *et al.* 2000; Wiegand & Moloney 2004). Nevertheless, $\Omega_s(d)$ does not provide *per se* the contribution of each spatial scale to the aggregation pattern because for low d , $\Omega_s(d)$ describes the degree of aggregation contributed by all scales included in the sampled area. For example, as illustrated in Figs 2 and S1, $\Omega_s(d) > 1$ at low d can occur for (i) a species distributed randomly within a set of small patches (Fig. S2a) or (ii) a species distributed at one end of the sampled area (Fig. S2c). However, aggregation is observed only at small scale in case (i) and at large scale in case (ii). These two cases can be well differentiated from the shape of the $\Omega_s(d)$ function because in case (i) $\Omega_s(d)$ drops at a short distance and then remains horizontal, while in case (ii) $\Omega_s(d)$ is near horizontal at a short distance and then drops at large distance. The distance at which the slope of $\Omega_s(d)$ is steepest thus indicates the scale contributing to the aggregation pattern. When two contrasting spatial scales contribute to aggregation (for example, if individuals occur in small patches distributed at one end of the sampled area (Fig. S2d)), the slope of $\Omega_s(d)$ is steepest both at short distance and at large distance, while it is flatter at intermediate distance. In other words, the slope of $\Omega_s(d)$ (i.e. its derivative with respect to distance) at distance d gives the contribution of scale d to the overall degree of aggregation. Hence, we quantified the contribution of a spatial scale (defined by a distance range) to the aggregation pattern by A_s , defined as minus the slope of $\Omega_s(d)$ on $\ln(d)$ within this distance range. This approach efficiently partitions aggregation pattern across spatial scales (Fig. S2d). Further investigations of the properties of $\Omega_s(d)$ and A_s using simulated data sets are presented in Appendix S1 and Figs S1–S3. Regressing $\Omega_s(d)$ on $\ln(d)$ rather than directly on d is

justified, first because $\Omega_s(d)$ decreases almost linearly with the logarithm of the distance when aggregation is because of limited dispersal (Appendix S1 and Fig. S2), and secondly because the slope is then mathematically dimensionless (i.e. it is not affected by the fact that distance is measured in m or km).

With our data, A_s was estimated as minus the slope of the regression of Ω_{sij} on $\ln(d_{ij})$ over all i, j pairs included in the distance range defining the spatial scale (A_s is thus independent from the arbitrarily defined distance classes used to represent $\Omega_s(d)$). For each site, we calculated one value of A_s for each species at each of three different spatial scales: a local scale (0.2–1 km), a mesoscale (1–10 km) and a landscape scale (10–70 km). To test whether a given species displayed significant aggregation, we performed 4000 complete random permutations of abundances among plots. In this way, any aggregation pattern occurring within our plots (i.e. at a scale not investigated because the positions of trees within plots are unknown) remained unaffected by the randomization procedure and did not, therefore, affect the patterns observed at larger scales. For each randomized data set, A_s was calculated and compared with the A_s obtained without randomization. The mean P -value is then given by the mean proportion of A_s obtained after randomization that are higher than or equal to the A_s value without randomization.

SPECIES CHARACTERISTICS

Data on diaspore type and range size [following the classification of White (1979)] were obtained from an extensive literature survey of regional flora (see Appendix S2). For dispersal modes, unassisted dispersal refers to species whose diaspores have no aril or pulp and which are not plumed or winged. Wind dispersal refers to species with plumed or winged diaspores and animal dispersal to species with diaspores with edible arils or pulp. In few cases, dispersal syndromes were reassigned based on expert knowledge (e.g. when diaspores devoid of pulp or aril are known to be dispersed by some animals). For shade tolerance, we completed and homogenized the literature survey according to our knowledge of species ecology following the classification of Hawthorne (1995): pioneer (P) species require gaps for establishment, non-pioneer light-demanding species (NPLD) can establish in shade but need a gap to reach the canopy, and shade-tolerant (ST) species establish and grow in shade. For all these species characteristics, consistency between references was checked in order to obtain reliable data (see Table S1). Wood densities (i.e. wood-specific gravities) were obtained from Gourlet-Fleury *et al.* (2011). Wood density was available for most species ($n = 65$). In absence of wood density information at the species level, we allocated the measurements at the genus level, the family level, the order level or the global level mean value following a set of rules described in Gourlet-Fleury *et al.* (2011).

STATISTICAL ANALYSES

We analysed the effects of species, sites and specific characteristics on A_s values using linear mixed models. This type of model is suited for testing overall effects of explanatory variables on the variability of the response variable, rather than comparing mean effects between modalities. Moreover, a major interest of mixed models is to account for disequilibrium and pseudoreplication in the design of the analyses (Bolker *et al.* 2009). In all these analyses, A_s values were $\ln(1 + x)$ -transformed to approach normality. Model fitting was performed by maximum likelihood using lme4 package (Bates & Sarkar 2006) in R 2.10.0 statistical package (R Development Core Team 2009) with a Gaussian error structure.

First, for each of the three spatial scales, we fitted a linear mixed model with a fixed intercept and species, sites and mean species abundance per plot (in each site) considered as random effects. This model allowed us to investigate the effects of species identity on A_s values while accounting for inter-site heterogeneity and species abundance, or *vice versa*. Controlling for species abundance is justified because in our simulations implementing dispersal limitation, we found that, for identical dispersal parameters, less abundant species have larger A_s values (Appendix S1). Note that the model only controls for a general effect of the sites on A_s values (i.e. considering that aggregation could be consistently different across sites because of particular environmental conditions and/or history). Second, we fitted a linear mixed model where hierarchically nested levels of taxonomy (order, family and species) and sites were considered as random effects. We did not consider the genus level because most genera were represented by a single species. This model allowed us to partition the variability observed at species level into higher taxonomical components while accounting for heterogeneity across sites. Finally, we investigated the link between species characteristics and A_s values, considering characteristics as fixed effects and species and sites as random effects. This model allowed us to estimate the particular effect of a given characteristic while controlling for species identity and site heterogeneity. The significance of each effect was assessed by likelihood ratio tests comparing the fitted model with reduced models where the effect was removed (Bolker *et al.* 2009).

Results

AGGREGATION AT DIFFERENT SPATIAL SCALES

A_s values, the contribution of a given spatial scale to the aggregation pattern, were significantly positive in most species at all three spatial scales, indicating that aggregation tends to occur at all scales. Over the five study sites, the percentages of sampled species exhibiting significantly positive A_s values (at 5% error level) were 48–65% at the local scale (0.2–1 km), 69–97% at the mesoscale (1–10 km) and 39–71% at the landscape scale (10–70 km), depending on the site investigated.

In all sites, A_s values significantly decreased at increasing spatial scales (Wilcoxon signed-rank test, $P < 0.05$): A_s values were significantly higher at the local- than at the mesoscale, and significantly higher at the meso- than at the landscape scale. Hence, most species displayed a concave curve form of the pair-correlation function $\Omega_s(d)$ against the logarithm of the distance (d) (e.g. *Terminalia superba* in Fig. 2).

The mean A_s values per species across sites at each spatial scale were significantly correlated across spatial scales (Spearman's rank correlation: local/meso: $r = 0.62$; meso/landscape: $r = 0.43$; local/landscape: $r = 0.36$; $P < 10^{-4}$ for all pairwise comparisons, Fig. S4). Hence, although A_s values differed substantially across the three spatial scales, species displayed consistent ranking at the three scales.

CONSISTENCY OF SPECIES AGGREGATION ACROSS SITES AND SITE EFFECT

A linear mixed model used to test the explanatory power of sites, species and mean species abundance per plot on transformed A_s values revealed significant species effects at the

local- and mesoscales, explaining *c.* 21% and 24% of the total variance, respectively (Table 1). At the landscape scale, the species effect was not significant. Hence, species aggregation displayed inter-site consistency at the local- and the mesoscales, although a large part of variation remained unexplained.

There was no significant effect of sites on A_s values at any spatial scale when the five study sites were considered (Table 1). Hence, the variability of aggregation patterns could not be explained by general differences among sites. However, when the adjacent SCAF1 and SCAF2 sites, which lie on contrasting soil types (clay versus sandy soils), were pooled together into a single data set (SCAF), the mean A_s value in SCAF was higher than in the three other sites (Pallisco, IFB1 and IFB2) at the landscape scale (mean A_s in SCAF = 0.18, Pallisco = 0.07, IFB1 = 0.10, IFB2 = 0.03, mixed model LRT test: $P < 10^{-4}$). Macrogeographical edaphic heterogeneity had thus a statistically significant influence on aggregation patterns at the landscape scale.

CONSISTENCY OF SPECIES AGGREGATION WITHIN TAXONOMIC LEVELS

Considering botanical order, family and species as nested random factors in a linear mixed model, we identified taxonomic levels that could significantly explain variation of aggregation patterns (Table 1). At the local scale, only the family level had a significant effect on A_s values. At the mesoscale, we found that A_s values could be significantly explained by variability among clades at all taxonomic levels. Finally, no taxonomic effect was detected at the landscape scale.

LINK BETWEEN SPECIES CHARACTERISTICS AND AGGREGATION PATTERNS

Only two of the species characteristics considered (i.e. diaspore type and wood density) showed a significant correlation with A_s values (Table 2, Fig. S5). The linear mixed models detected

a significant effect of diaspore type at the meso- and the landscape scales ($P < 0.05$), species with diaspores adapted to assisted dispersal (wind and animal dispersal) being less aggregated than species with unassisted dispersal. We also found that wood density was significantly negatively correlated to A_s values at the landscape scale, indicating that soft-wooded species tend to be more aggregated than hard-wooded species at this scale.

Discussion

To our knowledge, this is the first comparative study of spatial aggregation patterns in tropical tree species that covers a large range of spatial scales and where each species is studied in several sites. Most species showed spatial aggregation which, in the majority of cases, was most pronounced at smaller spatial scales. This is consistent with the results of Condit *et al.* (2000) whose study focussed on smaller spatial scales. Interestingly, despite this variation across scales, we detected a consistent ranking of species aggregation values at different scales. Different and/or complementary processes are thus likely to act in a similar way over spatial scales (Ricklefs 1987; Wiens 1989; Levin 1992; Huston 1999; Leibold *et al.* 2004). Consistent with the results of Condit *et al.* (2000) and with the simulated patterns under dispersal limitation (Appendix S1, Fig. S2 and S3), rare species tended to be more aggregated than common species at all spatial scales (Fig. S6). However, there was no significant effect of species abundance when species and sites were taken into account in our mixed model.

Although we observed substantial variation in species aggregation among sites, 21–24% of the variance in aggregation intensity was explained by a species effect at the local scale (0.2–1 km) and mesoscale (1–10 km). Hence, the level of spatial aggregation is partly determined by an emergent property of the species, indicative that biological characteristics generally affect species aggregation patterns. Particularly interesting here are the statistically significant effects of taxonomic levels

Table 1. Variance partitioning of aggregation patterns (A_s values) at three spatial scales using two mixed linear models: (1) Linear mixed model with a fixed intercept and species, sites and mean species abundance per plot considered as random effects. (2) Linear mixed model where hierarchically nested levels of taxonomy (order, family and species) and sites were considered as random effects. A_s values were $\ln(1 + x)$ -transformed to approach normality, and the significance of each factor was assessed with likelihood ratio tests (see Material and methods)

	<i>n</i>	Local		Meso		Landscape	
		Variance (%)	<i>P</i> -value	Variance (%)	<i>P</i> -value	Variance (%)	<i>P</i> -value
(1)							
Sites	5	0.0	NS (1.00)	1.7	NS (0.15)	0.8	NS (0.40)
Species	106	20.8	<0.01	23.9	<0.001	16.3	NS (0.14)
Species abundance	329	30.1	NS (0.12)	0.0	NS (1.00)	0.0	NS (0.13)
Residuals		49.1		74.4		82.9	
(2)							
Taxonomic levels							
Order	14	0.0	NS (0.25)	0.0	<0.01	0.0	NS (0.19)
Family	31	7.8	<0.01	8.5	<0.05	3.1	NS (0.37)
Species	106	10.3	NS (0.16)	15.9	<0.01	12.5	NS (0.22)
Sites	5	0.0	NS (0.99)	1.8	NS (0.12)	0.9	NS (0.36)
Residuals		81.9		73.8		83.5	

Table 2. Quantifying and testing the links between species aggregation (A_s values) and species characteristics. Estimates and standard deviation (SD) were calculated with mixed linear models. Each species characteristic was tested independently after removing species with missing data. In the model, A_s values were $\ln(1 + x)$ -transformed to approach normality. The significance of each factor was assessed with likelihood ratio tests (see Material and methods). See also Fig. S5 in Supporting Information

Species characteristics	<i>n</i>	Local			Meso			Landscape		
		Estimate	SD	<i>P</i> -value	Estimate	SD	<i>P</i> -value	Estimate	SD	<i>P</i> -value
Dispersal mode										
Unassisted	7	0.35	0.13	NS (0.76)	0.35	0.06	< 0.05	0.18	0.06	< 0.05
Wind	32	0.30	0.14		0.19	0.07		0.07	0.06	
Animal	67	0.35	0.13		0.19	0.06		0.12	0.06	
Shade tolerance										
Pioneers	38	0.37	0.05	NS (0.19)	0.24	0.03	NS (0.27)	0.12	0.02	NS (0.20)
NPLD	29	0.24	0.07		0.18	0.04		0.06	0.03	
Shade-tolerant	38	0.34	0.07		0.19	0.04		0.09	0.03	
Range size										
One domain	5	0.19	0.16	NS (0.47)	0.14		NS (0.52)	0.04		NS (0.15)
Two domains	28	0.40	0.17		0.21			0.12		
Region	53	0.31	0.16		0.19			0.09		
(Inter-)Continental	20	0.36	0.17		0.24			0.14		
Wood density	106	-0.27	0.24	NS (0.27)	-0.07	0.12	NS (0.56)	-0.20	0.09	< 0.05

NPLD, non-pioneers light demander.

above the species level, indicating that closely related species may share similar biological characteristics that influence species aggregation patterns at these scales. Similarly, Condit *et al.* (2000) found that some families (i.e. Dipterocarpaceae) were more aggregated than other families at small spatial scales in two Malaysian plots.

We showed that dispersal abilities could influence the spatial patterns of species up to the landscape scale (> 10 km). Indeed, species with unassisted seed dispersal were significantly more aggregated than wind- and animal-dispersed species at the meso- and the landscape scales. This result is in contrast to results from previous studies reporting significant correlation between dispersal mode and species aggregation patterns only at fine spatial scales [< 200 m, i.e. smaller than our smallest scale; (Hubbell 1979; Seidler & Plotkin 2006)]. Although these studies did not investigate aggregation patterns at scales larger than 1000 m, they concluded that above 200 m, aggregation patterns were caused by factors other than dispersal (Seidler & Plotkin 2006). Our results did not confirm these conclusions. A recent study had shown that species sharing the same seed dispersal mode showed wide variation in real dispersal abilities at local scale (< 1000 m; Muller-Landau *et al.* 2008). Hence, the local stochasticity of actual dispersal abilities and/or the low statistical power because of the small number of species showing unassisted dispersal (seven species in our study) could explain why we observed no relationship between dispersal mode and species aggregation patterns at the local scale. To a lesser extent, we also found that soft-wooded species, often fast-growing and short-lived pioneer species (Turner 2001; Muller-Landau 2004; van Gelder, Poorter & Sterck 2006), were significantly more aggregated than hard-wooded species at the landscape scale. This could result from past disturbances (e.g. forest logging, villages) leading to spatial clustering of

secondary forest species. However, the correlation was very low (Fig. S5), and the number of tests performed (four species characteristics at three spatial scales) may increase the type I error rate. Our index of shade tolerance was furthermore not correlated to A_s values. Hence, our interpretation must be considered with caution.

The overall low correlation observed between species characteristics and aggregation patterns at the smallest scales indicates that other relevant species characteristics may have been overlooked. Species characteristics may also affect aggregation by interacting in a complex way so that, if they generally show a phylogenetic signal, taxonomic or phylogenetic distance may better explain differences among species than a single characteristic. Another hypothesis is that the taxonomic signal detected at these scales could be explained by phylogenetic niche conservatism (Harvey & Pagel 1991), species of a same lineage being likely to occupy similar more or less spatially structured habitats (Queenborough *et al.* 2009). If true, the way forward is to use traits linked to drought sensitivity and/or nutrient use (Engelbrecht *et al.* 2007; Baltzer *et al.* 2008) to test whether they can explain species aggregation patterns.

And last, one of the most striking results of our study is the high residuals observed in our mixed models at the three spatial scales (Table 1). This indicates that species aggregation patterns are highly context-dependent. Nevertheless, variation among sites is not explained by general differences among sites, except when a steep environmental gradient occurs at the largest spatial scale. This high variability could result from high stochasticity in tree recruitment and establishment in tropical forests (Hubbell 2001), from different disturbance regimes or from site-specific effects (e.g. particular colonization history of a species in a site or different environmental heterogeneity

across sites, Montoya *et al.* 2009). Moreover, species spatial patterns were inconsistent across sites at the landscape scale (10–80 km), which confirms that species aggregation patterns may be solely context-dependant at this scale. A general site effect was indeed evident when pooling two adjacent sites located on contrasting soil types (SCAF1 and SCAF2): mean aggregation across species increased substantially at the landscape scale because, in this region, many species show contrasting distribution on different soil types (Réjou-Méchain *et al.* 2008; Gourlet-Fleury *et al.* 2011). By contrast, climatic factors, rainfall in particular, are unlikely to play a significant role in shaping aggregation patterns in the study region. Indeed, Réjou-Méchain *et al.* (2008) showed that rainfall had little explanatory power on the floristic variation in this region, which is partly because of the low rainfall gradient observed at the regional scale.

In conclusion, although species aggregation patterns may be mainly driven by stochastic factors and by the specific characteristics of the sites (history and environmental heterogeneity), we found that species aggregation displayed some inter-site and within-clade consistencies at the local- and the mesoscales. These results indicate that species aggregation patterns of trees could, to a certain extent, be predicted by species identity up to scales that were rarely investigated before in this context. However, further investigations that integrate complex combinations of species characteristics and how they explain aggregation at the local- and mesoscales in interaction with historical and environmental effects are necessary to better infer the underlying mechanisms controlling aggregation patterns. For this purpose, further theoretical studies are also needed to complete our preliminary results based on simulated data sets (Appendix S1) to better understand which processes generate which patterns.

Acknowledgements

We thank J. Chave, P.-O. Cheptou, L.S. Comita, R. Pélissier, A.S.L. Rodriguez, J.D. Thompson, the Handling Editor and two anonymous reviewers for their constructive comments on the manuscript, G. Dauby and O. Lacheneau for their help with identifying species that could be misidentified and O. Gimenez for help with mixed model analyses. We are grateful to the PARPAF project and the leading consortium CIRAD (Centre de Coopération Internationale en Recherche Agronomique pour le Développement) and FRM (Forêt Ressources Management), as well as Gembloux Agro-Bio Tech (Belgium) via Project PPR 10.000 and the Belgian Fund for Scientific Research (F.R.S.-FNRS) via grant F.4.519.10.F. We address special thanks to 'Ministère des Eaux, Forêts, Chasses et Pêches' of CAR and to the three forest companies that provided, albeit restricted, access to their inventory data for research purpose.

References

- Aiba, S., Kitayama, K. & Takyu, M. (2004) Habitat associations with topography and canopy structure of tree species in a tropical montane forest on Mount Kinabalu, Borneo. *Plant Ecology*, **174**, 147–161.
- APG II (2003) An update of the Angiosperm phylogeny group classification for the families of flowering plants. *Botanical Journal of the Linnean Society*, **14**, 399–436.
- Baltzer, J.L., Davies, S.J., Bunyavejchewin, S. & Noor, N.S.M. (2008) The role of desiccation tolerance in determining tree species distributions along the Malay-Thai Peninsula. *Functional Ecology*, **22**, 221–231.
- Bates, D. & Sarkar, D. (2006) Linear mixed-effects models using S4 classes (R package version 0.9975-10). Available at: <http://cran.r-project.org/web/packages/lme4/index.html> (accessed July 14, 2009).
- Bell, T., Freckleton, R.P. & Lewis, O.T. (2006) Plant pathogens drive density-dependent seedling mortality in a tropical tree. *Ecology Letters*, **9**, 569–574.
- Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H. & White, J.S.S. (2009) Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, **24**, 127–135.
- Boulvert, Y. (1986) *Carte phytogéographique de la République centrafricaine (feuille OUEST – feuille EST) à 1 : 1 000 000*. Editions de l'ORSTOM, Paris.
- Casper, B.B., Heard, S.B. & Apanius, V. (1992) Ecological correlates of single-seededness in a woody tropical flora. *Oecologia*, **90**, 212–217.
- Cavender-Bares, J., Kitajima, K. & Bazzaz, F.A. (2004) Multiple trait associations in relation to habitat differentiation among 17 Floridian oak species. *Ecological Monographs*, **74**, 635–662.
- Chave, J., Muller-Landau, H.C., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N. *et al.* (2000) Spatial patterns in the distribution of tropical tree species. *Science*, **288**, 1414–1418.
- Condit, R., Pitman, N., Leigh, E.G., Chave, J., Terborgh, J., Foster, R.B. *et al.* (2002) Beta-diversity in tropical forest trees. *Science*, **295**, 666–669.
- Connell, J.H. (1971) On the roles of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. *Dynamics of Populations* (eds P.J. den Boer & G.R. Gradwell), pp. 298–312. Center for Agricultural Publication and Documentation, Wageningen, Netherlands.
- Cousens, R., Dytham, C. & Law, R. (2008) *Dispersal in Plants: A Population Perspective*. Oxford University Press, Oxford.
- Denslow, J.S. (1980) Gap partitioning among tropical rainforest trees. *Biotropica*, **12**(Suppl.), 47–55.
- Denslow, J.S. (1987) Tropical rainforest gaps and tree species diversity. *Annual Review of Ecology and Systematics*, **18**, 431–451.
- Edwards, W. & Westoby, M. (1996) Reserve mass and dispersal investment in relation to geographic range of plant species: phylogenetically independent contrasts. *Journal of Biogeography*, **23**, 329–338.
- Engelbrecht, B.M.J., Comita, L.S., Condit, R., Kursar, T.A., Tyree, M.T., Turner, B.L. & Hubbell, S.P. (2007) Drought sensitivity shapes species distribution patterns in tropical forests. *Nature*, **447**, 80–82.
- Flores, O. (2005) *Déterminisme de la régénération chez quinze espèces d'arbres tropicaux en forêt guyanaise: les effets de l'environnement et de la limitation par la dispersion*. Université Montpellier II, Montpellier.
- Flores, O., Gourlet-Fleury, S. & Picard, N. (2006) Local disturbance, forest structure and dispersal effects on sapling distribution of light-demanding and shade-tolerant species in a French Guianian forest. *Acta Oecologica*, **29**, 141–154.
- Franklin, J. & Rey, S.J. (2007) Spatial patterns of tropical forest trees in Western Polynesia suggest recruitment limitations during secondary succession. *Journal of Tropical Ecology*, **23**, 1–12.
- Gaston, K.J. (2003) *The Structure and Dynamics of Geographic Ranges*. Oxford University Press, Oxford.
- van Gelder, H.A., Poorter, L. & Sterck, F.J. (2006) Wood mechanics, allometry, and life-history variation in a tropical rain forest tree community. *New Phytologist*, **171**, 367–378.
- Gourlet-Fleury, S., Rossi, V., Réjou-Méchain, M., Freycon, V., Fayolle, A., Saint-Andre, L. *et al.* (2011) Environmental filtering of dense-wooded species controls aboveground biomass stored in African moist forests. *Journal of Ecology*, **99**, 981–990.
- Hardy, O.J. & Sonké, B. (2004) Spatial pattern analysis of tree species distribution in a tropical rain forest of Cameroon: assessing the role of limited dispersal and niche differentiation. *Forest Ecology and Management*, **197**, 191–202.
- Harvey, P.H. & Pagel, M.D. (1991) *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford.
- Hawthorne, W.D. (1995) *Ecological Profiles of Ghanaian Forest Trees*. *Tropical Forestry Paper*, 29. Oxford Forestry Institute, Oxford, UK.
- He, F.L., Legendre, P. & LaFrankie, J.V. (1997) Distribution patterns of tree species in a Malaysian tropical rain forest. *Journal of Vegetation Science*, **8**, 105–114.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surface for global land areas. *International journal of climatology*, **25**, 1965–1978.
- Howe, H.F. & Smallwood, J. (1982) Ecology of seed dispersal. *Annual Review of Ecology and Systematics*, **13**, 201–228.
- Hubbell, S.P. (1979) Tree dispersion, abundance, and diversity in a tropical dry forest. *Science*, **203**, 1299–1309.

- Hubbell, S.P. (2001) *The Unified Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton.
- Huston, M.A. (1999) Local processes and regional patterns: appropriate scales for understanding variation in the diversity of plants and animals. *Oikos*, **86**, 393–401.
- Janzen, D.H. (1970) Herbivores and number of tree species in tropical forests. *American Naturalist*, **104**, 501–528.
- Jones, M., Tuomisto, H., Clark, D.B. & Olivas, P. (2006) Effects of mesoscale environmental heterogeneity and dispersal limitation on floristic variation in rain forest ferns. *Journal of Ecology*, **94**, 181–195.
- Kenkel, N.C. (1988) Pattern of self-thinning in jack-pine – testing the random mortality hypothesis. *Ecology*, **69**, 1017–1024.
- Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. & Gonzalez, A. (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, **7**, 601–613.
- Levin, S.A. (1989) Challenges in the development of a theory of ecosystem structure and function. *Perspective in Ecological Theory* (eds J. Roughgarden, R.M. May & S.A. Levin), pp. 242–255. Princeton University Press, Princeton.
- Levin, S.A. (1992) The problem of pattern and scale in ecology. *Ecology*, **73**, 1943–1967.
- Montoya, D., Purves, D.W., Urbieto, I.R. & Zavala, M.A. (2009) Do species distribution models explain spatial structure within tree species ranges? *Global Ecology and Biogeography*, **18**, 662–673.
- Muller-Landau, H.C. (2004) Interspecific and inter-site variation in wood specific gravity of tropical trees. *Biotropica*, **36**, 20–32.
- Muller-Landau, H.C., Wright, S.J., Calderon, O., Condit, R. & Hubbell, S.P. (2008) Interspecific variation in primary seed dispersal in a tropical forest. *Journal of Ecology*, **96**, 653–667.
- Paul, J.R. & Tonsor, S.J. (2008) Explaining geographic size by species age: a test using neotropical *Piper* species. *Tropical Forest Community Ecology* (eds W.P. Carson & S.A. Schnitzer), pp. 46–62. Blackwell Science, London.
- Picard, N., Bar-Hen, A., Mortier, F. & Chadoeuf, J. (2009) Understanding the dynamics of an undisturbed tropical rain forest from the spatial patterns of trees. *Journal of Ecology*, **97**, 97–108.
- Plotkin, J.B., Chave, J.M. & Ashton, P.S. (2002) Cluster analysis of spatial patterns in Malaysian tree species. *American Naturalist*, **160**, 629–644.
- Plotkin, J.B., Potts, M.D., Leslie, N., Manokaran, N., LaFrankie, J. & Ashton, P.S. (2000) Species-area curves, spatial aggregation, and habitat specialization in tropical forests. *Journal of Theoretical Biology*, **207**, 81–99.
- Queenborough, S.A., Burslem, D.F.R.P., Garwood, N.C. & Valencia, R. (2009) Taxonomic scale-dependence of habitat niche partitioning and biotic neighbourhood on survival of tropical tree seedlings. *Proceedings of the Royal Society Biological Sciences Series B*, **276**, 4197–4205.
- R Development Core Team (2009) *R: A Language and Environment for Statistical Computing*. (ed. R Foundation for Statistical Computing). Vienna, Austria.
- Réjou-Méchain, M. & Hardy, O.J. (2011) Properties of similarity indices under niche-based and dispersal-based processes in communities. *American Naturalist*, **177**, 589–604.
- Réjou-Méchain, M., Pélissier, R., Gourlet-Fleury, S., Couteron, P., Nasi, R. & Thompson, J.D. (2008) Regional variation in tropical forest tree species composition in the Central African Republic: an assessment based on inventories by forest companies. *Journal of Tropical Ecology*, **24**, 663–674.
- Réjou-Méchain, M., Fayolle, A., Nasi, R., Gourlet-Fleury, S., Doucet, J.L., Gally, M., Hubert, D., Pasquier, A. & Billand, A. (2011) Detecting large-scale diversity patterns in tropical trees: can we trust commercial forest inventories? *Forest Ecology and Management*, **261**, 187–194.
- Ricklefs, R.E. (1987) Community diversity – relative roles of local and regional processes. *Science*, **235**, 167–171.
- Ripley, B. (1981) *Spatial Statistics*. John Wiley, New York.
- Schluter, D. & Ricklefs, R.E. (1993) Convergence and the regional component of species diversity. *Species Diversity in Ecological Communities* (eds R.E. Ricklefs & D. Schluter), pp. 230–240. University of Chicago Press, Chicago.
- Seidler, T.G. & Plotkin, J.B. (2006) Seed dispersal and spatial pattern in tropical trees. *PLoS BIOLOGY*, **4**, 2132–2137.
- ter Steege, H., Pitman, N.C.A., Phillips, O.L., Chave, J., Sabatier, D., Duque, A., Molino, J.F., Prevoist, M.F., Spichiger, R., Castellanos, H., von Hildebrand, P. & Vasquez, R. (2006) Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature*, **443**, 444–447.
- Stocks, G., Seales, L., Paniagua, F., Maehr, E. & Bruna, E. (2008) The geographical and institutional distribution of ecological research in the tropics. *Biotropica*, **40**, 397–404.
- Stoll, P. & Bergius, E. (2005) Pattern and process: competition causes regular spacing of individuals within plant populations. *Journal of Ecology*, **93**, 395–403.
- Tilman, D. & Kareiva, P. (1997) *Spatial Ecology: the Role of Space in Population Dynamics and Interspecific Interaction*. Princeton University Press, Princeton, New Jersey.
- Tuomisto, H., Ruokolainen, K. & Yli-Halla, M. (2003) Dispersal, environment, and floristic variation of western Amazonian forests. *Science*, **299**, 241–244.
- Turner, I.M. (2001) *The Ecology of Trees in the Tropical Rain Forest*. Cambridge university press, Cambridge, UK.
- White, F. (1979) The Guineo-Congolian Region and its relationships to other phytochoria. *Bulletin du Jardin Botanique National de Belgique*, **49**, 11–55.
- White, F. (1983) *The Vegetation of Africa. A Descriptive Memoir to Accompany the UNESCO/AETFAT/UNSO Vegetation Map of Africa*. UNESCO, Paris.
- Wiegand, T. & Moloney, K.A. (2004) Rings, circles, and null-models for point pattern analysis in ecology. *Oikos*, **104**, 209–229.
- Wiegand, T., Gunatilleke, S., Gunatilleke, N. & Okuda, T. (2007) Analyzing the spatial structure of a Sri Lankan tree species with multiple scales of clustering. *Ecology*, **88**, 3088–3102.
- Wiens, J.A. (1989) Spatial scaling in ecology. *Functional Ecology*, **3**, 385–397.

Received 24 November 2010; accepted 28 June 2011

Handling Editor: Charles Canham

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Aggregation patterns and properties of A_s values using simulated datasets.

Appendix S2. References used in the literature survey of regional flora.

Table S1. The 106 studied tree species, their characteristics and A_s values.

Figure S1. Simulated aggregation patterns under habitat filtering with different spatial configurations of suitable habitats and without dispersal limitation.

Figure S2. Simulated aggregation patterns for two abundant species under dispersal limitation without habitat filtering.

Figure S3. Simulated aggregation patterns for two rare species under dispersal limitation without habitat filtering.

Figure S4. Correlation of A_s values between different spatial scales.

Figure S5. Differences in aggregation patterns (mean A_s values across sites) according to species characteristics at the local, meso- and landscape scales.

Figure S6. Aggregation patterns (A_s values) as a function of the abundance of each species (log scale).

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.