

Environmental and neighbourhood effects on tree fern distributions in a neotropical lowland rain forest

Jones, Mirkka M.^{1*}; Olivas Rojas, Paulo²; Tuomisto, Hanna^{1,3} & Clark, David B.⁴

¹Department of Biology, University of Turku, 20014 Turku, Finland; ²Department of Biological Sciences; Florida International University, Miami, FL 33199 USA; E-mail poliv001@fiu.edu; ³E-mail hanna.tuomisto@utu.fi;

⁴Department of Biological Sciences; University of Missouri-St. Louis, 8001 Natural Bridge Road, St. Louis, MO 63121, USA; E-mail dbclark@sloth.ots.ac.cr; *Corresponding author; Fax +358 23335730; E-mail mirkka.jones@utu.fi

Abstract

Questions: To what extent are the distributions of tropical rain forest tree ferns (*Cyatheaceae*) related to environmental variation, and is habitat specialization likely to play a role in their local coexistence?

Location: Lowland rain forest at La Selva Biological Station, Costa Rica.

Methods: Generalized linear (GLM) and generalized additive (GAM) logistic regression were used to model the incidence of four tree fern species in relation to environmental and neighbourhood variables in 1154 inventory plots regularly distributed across 6 km² of old-growth forest. Small and large size classes of the two most abundant species were modelled separately to see whether habitat associations change with ontogeny.

Results: GLM and GAM model results were similar. All species had significant distributional biases with respect to microhabitat. Environmental variables describing soil variation were included in the models most often, followed by topographic and forest structural variables. The distributions of small individuals were more strongly related to environmental variation than those of larger individuals. Significant neighbourhood effects (spatial autocorrelation in intraspecific distributions and non-random overlaps in the distributions of certain species pairs) were also identified. Overlaps between congeners did not differ from random, but there was a highly significant overlap in the distributions of the two most common species.

Conclusions: Our results support the view that habitat specialization is an important determinant of where on the rain forest landscape tree ferns grow, especially for juvenile plants. However, other factors, such as dispersal limitation, may also contribute to their local coexistence.

Keywords: Coexistence; Costa Rica; *Cyatheaceae*; Habitat; La Selva; Logistic regression; Microhabitat; Niche; Species distribution.

Nomenclature: Moran & Riba (1995).

Abbreviation: AUC = Area under the curve.

Introduction

The extent to which tropical rain forest plants have specialized towards different ecological niches remains contentious. Some researchers have proposed that their local coexistence is enabled by narrow niche specialisation to the physical environment, such as soils, or the biotic environment, such as pollinators (e.g. Ashton 1969; Grubb 1977; Svenning 1999; Clark et al. 1999). Others have argued that tropical plants may be divisible into just a few guilds, within which species are generalists (e.g. Hubbell & Foster 1986; Harms et al. 2001; Hubbell 2001; Valencia et al. 2004).

If plant species coexistence is enabled by specialization to different microhabitats, then species distribution patterns should not only be significantly related to environmental gradients, but also differ among species (MacArthur & Levins 1967; Abrams 1983; Pacala & Tilman 1994; Svenning 1999; Harms et al. 2001).

Rain forest plant species have been documented to be associated with particular edaphic, topographic or successional conditions in many studies (e.g. Clark et al. 1995, 1998; Lieberman et al. 1995; Davies et al. 1998; Svenning 1999, 2000, 2001; Webb & Peart 2000; Harms et al. 2001; Debski et al. 2002; Tuomisto et al. 2002; Cannon & Leighton 2004; Hall et al. 2004; Valencia et al. 2004; Paoli et al. 2006; Tuomisto 2006). However, significant microhabitat associations have been identified in an extremely variable proportion of species. Even where habitat associations have been identified for many species, some authors (e.g. Harms et al. 2001; Valencia et al. 2004) have concluded that habitat specialization was unlikely to be the main coexistence mechanism in their study sites, because considerable overlap in patterns of habitat association across species was observed.

The degree of habitat association detected may vary between plant groups and sites, and as a function of the length of the environmental gradient sampled (Duque et al. 2002; Tuomisto et al. 2002; Jones et al. 2006). Environmental data quality is also important. Many studies

used only a limited set of qualitative variables, such as soil type or topographic position, whereas access to a comprehensive set of quantitative variables, including soil chemical properties, might have allowed the identification of finer microhabitat partitioning (Clark et al. 1999; Hall et al. 2004).

Clumped distributions, at least over short distances, are common in rain forest plants (He et al. 1997; Condit et al. 2000; Plotkin et al. 2000). At least two causes may be invoked: patchy distributions of microhabitats, and dispersal limitation (including both vegetative reproduction and the tendency of most seeds or spores to fall close to the mother plant). Spatial autocorrelation in environmental variables and plant distributions needs to be accounted for in analyses to avoid inflating the apparent importance of habitat specialization (Smith 1994; Legendre & Legendre 1998; Harms et al. 2001).

In the current study, we analyse the mesoscale distributions of a group of plant species that are phylo-

genetically closely related and similar in life form, the tree ferns (*Cyatheaceae*). Four common species belonging to two genera were found in our 6-km² old-growth rain forest study site at the La Selva Biological Station, Costa Rica.

Our approach was to (1) calculate whether the distributional overlaps of the four species differed from random expectation and (2) model their distribution patterns using a range of environmental data, to see whether there is evidence for microhabitat specialization, and if so, whether the specialization patterns differ among species. Interspecific and intraspecific neighbourhood variables were also added to the habitat models to identify species associations and to model spatial autocorrelation. For the two most abundant species, we investigate whether their distributions among microhabitats change over the course of the species' lifetimes by comparing separate models for small and large individuals.

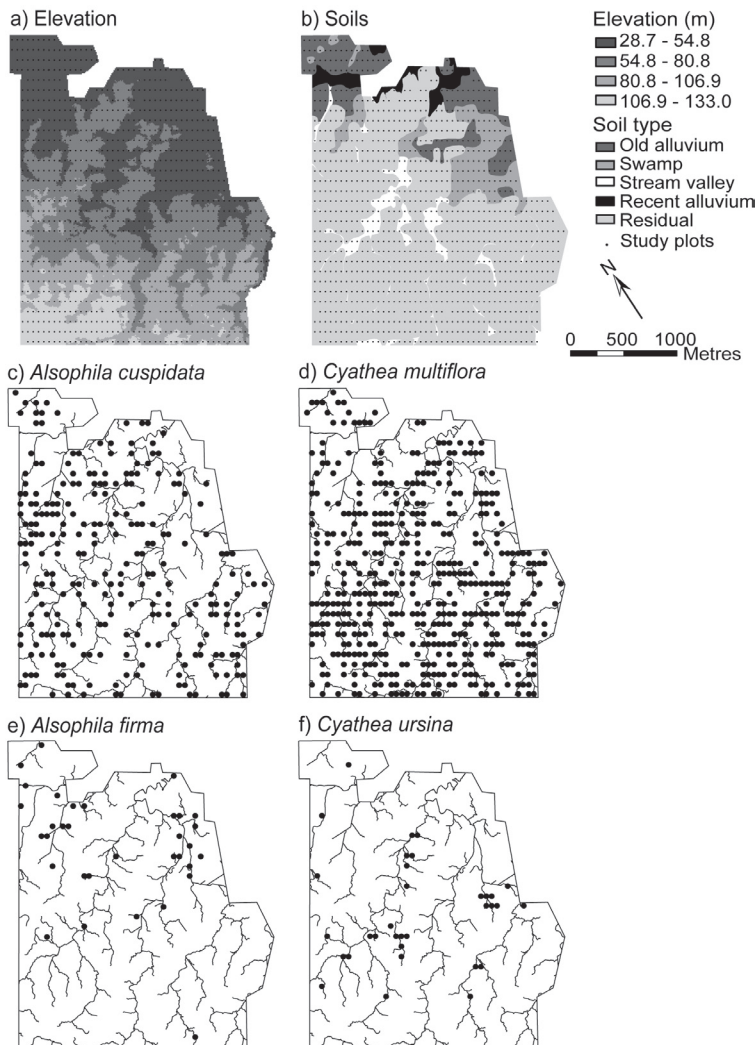


Fig. 1. Locations of the sample plots overlaid on (a) elevation and (b) soil maps of the study area, and distributions of the four study species (c-f) overlaid on a stream map.

Material and Methods

Study site

Fieldwork was carried out in the Caribbean lowlands of northeastern Costa Rica in ca. 6 km² of old growth rain forest at the La Selva Biological Station of the Organization for Tropical Studies (OTS). The site has a mean monthly temperature of 26 °C and mean annual precipitation of just over 4000 mm (OTS unpubl. data). On an average year, rainfall peaks in July and December and is lowest from January to April. Elevation varies from ca. 30 - 130 m a.s.l. (Fig. 1a). The area includes five main soil types (Fig 1b): alluvial terraces formed by both recent and historic flooding, swamps, residual soils formed by in-situ weathering of ancient lava flows, and stream valleys with infertile colluvial soils (Clark et al. 1999). The site is covered by a grid of permanent markers at a 50 m × 100 m spacing.

Field data

Tree ferns

The study area was sampled systematically for tree ferns between July 2001 and July 2002. A total of 1154 circular plots, each 100 m² in area and centred on a grid intersection (Fig. 1), was surveyed for all tree ferns with frond length > 10 cm. Voucher specimens can be found in the herbaria CR (Herbario Nacional de Costa Rica), USJ (Universidad de Costa Rica), TUR (University of Turku, Finland) and LSCR (the on-site herbarium of La Selva). Representative collection numbers with duplicates in all four herbaria (of the collection series of Mirkka Jones) are: *Alsophila cuspidata* (MJ 486), *Alsophila firma* (MJ 449), *Cyathea multiflora* (MJ 520), and *Cyathea ursina* (MJ 195).

The number of individuals of each species in each plot was recorded in eight size classes based on maximum frond length (in centimetres 10-20, 20-50, 50-100, 100- 150, 150-200, 200-300, 300-400 and 400+). Each separate stem arising from the ground was counted as one individual, although some *A. firma* stems were probably ramets of a single clone. Frond length was chosen to delimit size classes, because it was considered to be a more reliable correlate of tree fern age than trunk height. After falling over, *A. cuspidata*, at least, can resume vertical growth, in which case it produces large fronds on a disproportionately short trunk (M. Jones pers. observ.).

Soils

Each plot was allocated to one of the five qualitative soil classes (Fig. 1a). Additionally, quantitative soil data were available on pH, total C and P, and exchangeable

K, Ca, Mg and Mn from soil samples collected between March 1998 and May 1999 (D.B. Clark unpubl. data).

To reduce the dimensionality of the quantitative soil data, a principal components analysis (PCA) was run. The variables were all first standardized to give them equal weight. The first three PCA axes, which represented 86 % of the variation in the data, were used for modelling. Soil pH, Mg and Ca had negative loadings on Axis 1, which seemed to represent a drainage gradient: poorly drained or inundation-prone sites (swamp, recent alluvial and some stream valley soils) received low values on this axis. Axis 2 seemed more generally related to soil fertility, with all cations, P and pH having negative loadings. Axis 3 represented a soil carbon content gradient (positive loading of C), and its values were often high at low topographic positions and sometimes on flat ground in higher-lying areas, but were generally low in recent alluvial terraces.

Topographic data

Three topographic variables (topographic position, elevation and slope) were defined for the plots, based on data collected by Clark et al. (1999). Topographic position was defined as one of five classes: riparian, base of slope/flat low ground, mid-slope, upper slope, or flat high ground. Riparian plots were those overlapping with a stream. Elevation data came from two sources. For 1128 plots, elevation was measured directly in the field by optical survey. For the remaining 26 plots, elevation was based on Lidar measurements (Light detection and ranging data, cf. Hofton et al. 2002). Slope was measured in the steepest direction with a clinometer.

Canopy openness and forest structure

The size of the largest canopy gap visible from the plot centre was recorded on a scale of 0-25 using the canopy-scope method (Brown et al. 2000). These values were then reduced into 5 ordered classes: 0 = closed canopy, 1 = very small gap, 2-4 = small gap, 5-9 = medium gap, and ≥ 10 = large gap.

Forest structural data, collected from 1993 to 1995, included the number of large trees (DBH ≥ 10 cm) and their basal area (m².ha⁻¹) in each plot (Clark et al. 1999).

Distributional modelling

A simple randomization test was first run to check whether the observed level of co-occurrence between tree fern species deviated from random expectation. The observed occurrences of each species were randomly re-allocated over the 1154 plots, and the resulting number of co-occurrences of each species pair was counted. This was repeated 1000 × (using the program Resampling Stats

Table 1. Summary of the environmental and neighbourhood variables entered as potential explanatory variables of tree fern incidence in the GLM and GAM logistic regression models. Species in the neighbourhood variable abbreviations refers to the modelled tree fern species, i.e. either *Cyathea multiflora* (Cmult), *Alsophila cuspidata* (Acusp), *Alsophila firma* (Afirm) or *Cyathea ursina* (Cursi).

Variable type	Abbreviation	Scale	Description
Soil chemistry	Soil Axis 1	Continuous	Soil PCA Axis 1 (First axis of a PCA on seven soil chemical variables. See Methods for details)
	Soil Axis 2	Continuous	Soil PCA Axis 2
	Soil Axis 3	Continuous	Soil PCA Axis 3
Soil type	SV	Binary	Stream valley soils
	SW	Binary	Swamp soils
	RE	Binary	Residual soils
	RA	Binary	Recent alluvial soils
	OA	Binary	Old alluvial soils
Forest structure	Canopy openness	Categorical	Ordered five-class index of the size of the largest canopy gap visible (See Methods for details)
	Tree ba	Continuous	Basal area of trees ≥ 10 cm DBH in $\text{m}^2\cdot\text{ha}^{-1}$ (mean 0.3, range 0 – 2.7)
	N Trees	Continuous	Number of trees ≥ 10 cm DBH (mean 4.5, range 0 – 12)
Topography	Topographic position	Categorical	Ordered five-class index of relative topographic position (See Methods for details)
	Elevation	Continuous	Elevation in m above sea level (mean 67.4, range 33.9 – 130.0)
	Slope	Continuous	Slope in degrees (mean 12.1, range 0.0 – 44.0)
Neighbourhood	Species > 100	Binary	Incidence of large individuals (frond length > 100 cm) of the modelled species in a sample plot. (Included in models of small individuals only)
	Species 50 m	Binary	Incidence of individuals of the modelled species in a neighbouring plot (50 m away)
	Cmult	Binary	Incidence of <i>C. multiflora</i> in a sample plot. (Included in models of the other three species only)
	Acusp	Binary	Incidence of <i>A. cuspidata</i> in a sample plot. (Included in models of the other three species only)
	Afirm	Binary	Incidence of <i>A. firma</i> in a sample plot. (Included in models of the other three species only)
	Cursi	Binary	Incidence of <i>C. ursina</i> in a sample plot. (Included in models of the other three species only)

4.1) to create 95% confidence intervals for the degree of overlap between each species pair.

Generalized linear and generalized additive logistic regression models (GLM and GAM, respectively) were then produced to predict the incidence (presence/absence) of each of the four tree fern species using a combination of the environmental variables and biotic neighbourhood variables listed in Table 1. The biotic variables were derived from the tree fern dataset. All variables were standardized before testing. The maximum Pearson correlation between any pair of variables was 0.61 (between the residual soil type and elevation), so no variables were excluded on account of excessive co-linearity.

The GLMs (e.g. Nicholls 1989) were run as simple linear combinations of the explanatory variables. GAMs also constitute linear combinations of explanatory variables, but the relationship of continuous-scale variables to species incidence is represented by a data-driven smooth function (here, a thin plate regression spline, Wood & Augustin 2002; Wood 2003). GAM modelling is more flexible and may also be more realistic (Yee & Mitchell 1991).

Models were first built for tree ferns of all size classes together. For the two most abundant species, *C. multiflora* and *A. cuspidata*, additional models were built for small and large plants separately to see whether predictive performance and the variables selected depend on plant size. Small plants were those with fronds < 20 cm in length for *C. multiflora* and < 50 cm for *A. cuspidata*, and large plants were those with fronds > 100 cm in length for both

species. The size class limits were defined such that the small and large size classes had a similar frequency of occurrence across sample sites.

In model building, first the environmental variables were included alone. Backward elimination was applied to select the most parsimonious set of environmental explanatory variables using Akaike's Information Criterion (AIC, Akaike 1977).

To study neighbourhood effects, several biotic variables were then added to the final environmental models. These were the incidence of each of the three other tree fern species in a study plot, and the incidence of conspecifics in a neighbouring plot (50 m away). When all plants or small plants alone were modelled, these variables were based on individuals of all sizes. When large plants alone were modelled, only individuals of a similar or larger size were included. When small plants alone were modelled, a fifth neighbourhood variable, indicating the incidence in the same plot of large conspecific individuals (> 100 cm frond length), was also added. The interspecific neighbourhood variables were entered to model either positive or negative associations in the distributions of the modelled species. The intraspecific neighbourhood variables were entered to represent spatial autocorrelation in species distributions, due to intrinsic factors such as limited dispersal, or spatial dependence on any unmeasured, spatially autocorrelated habitat variables (Smith 1994; Wu & Huffer 1997). All the neighbourhood variables were added together, then backward elimination based on AIC was employed to

remove any variables that did not significantly improve overall model performance.

Zero occurrences of a modelled species in any category of a categorical explanatory variable yield incorrect numerical outcomes in logistic regression (Hosmer & Lemeshow 2000). To avoid this problem, categories can be combined in some reasonable way to eliminate the zero cell. Thus, because *C. ursina* did not occur in all soil types or topographic categories, the recent and old alluvial soil types, and the hilltop and upper slope categories of topographic position, were combined for models of this species. For models of the large size class of *C. multiflora* the neighbourhood variable *C. ursina* was omitted because large individuals of *C. multiflora* never co-occurred with individuals of *C. ursina*. Finally, for models of the large size class of *A. cuspidata*, the medium and large gap classes of canopy openness were combined.

The models were run using R-language (Anon. 2006) functions `glm` and `step.gam` in the package `stats` (v. 2.3.1) and `gam` in `mgcv` (v. 1.3-17).

Model evaluation

Model predictive performance was evaluated by calculating the area under the curve (AUC) of receiver-operating characteristic (ROC) plots. AUC is much less affected by species prevalence than other common measures of model performance, such as the correct classification rate, sensitivity, specificity or Cohen's κ , at least at intermediate prevalences (Manel et al. 2001; McPherson et al. 2004).

Evaluation was initially based on testing the ability of each model to correctly discriminate between species presences and absences in the full 1154 plot dataset. However, calibrating and testing models on the same data is known to result in over-optimistic estimates of prediction success (Fielding & Bell 1997). Therefore, the GLM models were also evaluated using the 0.632+ bootstrap approach (see Wintle et al. 2005 and link therein for R code). The original dataset was re-sampled 500 × to produce a training set and a test set. Each time, a GLM model was fitted using the training set, and two values of AUC were calculated, one using the training set and the other using the test set. The difference in these two AUC values was averaged over the 500 bootstrap runs to produce an estimate of optimism. This estimate was subtracted from the original AUC values to give a final, bootstrapped AUC value.

Finally, the standardised (Pearson) residuals of the all size classes models were examined for spatial autocorrelation relative to that observed in the raw species presence-absence data (Bio et al. 2002). This was done by producing correlograms of Moran's I (in the R Package

v. 4.0, Casgrain & Legendre 2001), which summarize spatial autocorrelation in the data at increasing distances (here at ca. 50 m intervals). The occurrence of significant spatial autocorrelation in the model residuals over short distances would suggest that the model fails to account for spatial autocorrelation due to limited dispersal or for spatial dependence on some autocorrelated environmental factor (Austin 2002; Bio et al. 2002; Schwarz et al. 2003).

Results

General observations

Tree ferns were a frequent component of the understorey plant community, being present in 53% of sample plots.

Cyathea multiflora was the most common species, both in absolute abundance (1410 individuals) and frequency (present in 42% of plots). Its mean abundance in occupied plots was 2.9 individuals. The largest individual had fronds just over 3 m in length. Small individuals (with fronds 10-20 cm in length) were present in 18% of plots, and large individuals (with fronds > 100 cm) were present in 16% of plots.

Alsophila cuspidata was the second most common species (681 individuals, 23% of plots) and its mean abundance in occupied plots was 2.6 individuals. This was the largest species, with fronds up to 4.4 m long. Small individuals (with fronds 10-50 cm in length) were present in 13% of plots, and large individuals (with fronds > 100 cm) were present in 12% of plots.

Cyathea ursina was much less abundant overall (148 individuals, 3% of plots), but it reached very high local densities at some sites and its mean abundance in occupied plots was 4.4 individuals. This species differs in life form from the others, being almost trunkless (maximum observed trunk height 33 cm, usually < 10 cm). Its fronds are also smaller than those of the other species (max. length 2.5 m).

Alsophila firma was almost as abundant as *C. ursina* (128 individuals in 3% of plots), with an average of 3.8 individuals in occupied plots. The maximum frond length found in any individual of this species was 3.7 m.

Tests of random distributional overlap

The degree of distributional overlap between the two most common species, *Cyathea multiflora* and *Alsophila cuspidata*, was quite high. They co-occurred in 156 plots, which represent 33% of the incidences of *C. multiflora* and 60% of the incidences of *A. cuspidata*. Randomization tests showed that this degree of overlap was higher

Table 2. Variables selected for the Environmental (E) and Environmental + Neighbourhood effects (E+N) models to explain the incidence of the four modelled tree fern species by GLM logistic regression modelling. Canopy openness and Topographic position were entered as ordered categorical variables (polynomial contrasts among the categories (Venables & Ripley 2002) were used to look for a significant linear, quadratic, cubic or 4th order trend in the relationship of species occurrence to these two variables). See Table 1 for explanations of variable abbreviations. The partial regression coefficients of the standardized variables are given, and their statistical significance is indicated as follows: $^{\circ} p < 0.1$; $* p < 0.05$; $** p < 0.01$; $*** p < 0.001$. AUC values indicate overall model predictive performance, with bootstrapped AUC values in brackets. An AUC value of 0.5 represents random expectation.

	<i>Alsophila cuspidata</i> (prevalence = 23%)		<i>Alsophila firma</i> (3%)		<i>Cyathea multiflora</i> (42%)		<i>Cyathea ursina</i> (3%)	
	E	E+N	E	E+N	E	E+N	E	E+N
Soil Axis 1	-0.37***	-0.36**			+0.30**	+0.37**	+0.81***	+0.42 $^{\circ}$
Soil Axis 2	-0.35***	-0.38***	-0.56***	-0.53**			+1.18***	+1.09***
Soil Axis 3	+0.15 $^{\circ}$	+0.22**			-0.13 $^{\circ}$	-0.15*		
SV	+0.32***	+0.29***					+0.45*	
SW			+0.47***	+0.42**			+0.47*	
RE	+0.18	+0.20						
RA	-0.20 $^{\circ}$				-0.42*	-0.42*		
Tree ba	-0.20*	-0.22*	+0.47**	+0.46**				
N Trees			-0.73**	-0.72**				
Topographic position					-linear***	-linear***	-linear***	-linear***
Elevation			-1.12*	-1.05*	+0.11			
Slope	+0.71***	+0.68***			+0.23**			
Species 50 m				+0.18 $^{\circ}$		+0.27***		+0.67***
Cmult		+0.42***						
Acusp						+0.34***		+0.52**
Afirm						+0.00		
Cursi		+0.19*						
AUC	0.73	0.76	0.89	0.89	0.67	0.70	0.88	0.90
(Bootstrap estimate)	(0.70)	(0.73)	(0.77)	(0.85)	(0.63)	(0.68)	(0.82)	(0.86)

than expected at random. When the distributions of the small individuals and the large individuals were considered separately, they too were found to overlap more often than expected. The observed overlaps between the other species pairs (all size classes together) fell within the boundaries of random expectation.

Habitat distributions of the species

All four tree fern species had distributional biases with respect to microhabitat. Tables 2 and 3 list the explanatory variables selected by GLM modelling. The results of the GAM models were so similar that for simplicity we focus on the GLM results alone. When all size classes were analysed together, between five and eight environmental variables were selected as predictors of the distribution of each species (Table 2). Soil variables were selected most often, followed by topographic and forest structural variables. The following distributional descriptions focus on the individually significant predictor variables ($p < 0.05$).

Alsophila cuspidata, *Cyathea multiflora*, and *C. ursina* were all associated with low topographic positions such as riparian sites and stream valley soils. However, they had a low predicted incidence, or did not occur at all (*A. cuspidata*), in sites with low values on Soil Axis 1, and also on Soil Axis 2 for *C. ursina*. Thus, they were

biased away from the most poorly drained sites with high soil Ca, Mg and pH. In contrast, *A. firma* was biased towards low-lying, swampy or occasionally flooding sites near larger streams with relatively nutrient-rich soils of high pH. *Cyathea multiflora* and especially *A. cuspidata* were associated with sloping sites, whereas *A. firma* and *C. ursina* were not.

The environmental models indicated that *Cyathea ursina* was positively associated with certain relatively less fertile stream valley and swamp soils, but when neighbourhood variables were added to the model these variables were removed. Unlike the other three species, *C. ursina* almost always colonized fallen tree trunks or branches and often lacked a root connection to the ground (M. Jones pers. obs.).

Forest structural variables were selected as significant predictors of the incidence of both *Alsophila* species. *Alsophila firma* was associated with sites with a high tree basal area, but visual inspection of the data suggested this to be mainly due to the distribution of small individuals. *Alsophila cuspidata* was associated with sites with a low tree basal area. These variables were not significant predictors of the incidence of either *Cyathea* species. Canopy openness was not incorporated into any of the models.

Separate models of the small and large size classes of *C. multiflora* and *A. cuspidata* indicated that small

Table 3. Variables selected for the Environmental (E) and Environmental + Neighbourhood effects (E+N) models to explain the incidence of small vs. large individuals of *A. cuspidata* and *C. multiflora* by GLM logistic regression modelling. Canopy openness and Topographic position were entered as ordered categorical variables (polynomial contrasts among the categories (Venables & Ripley 2002) were used to look for a significant linear, quadratic, cubic or 4th order trend in the relationship of species occurrence to these two variables). See Table 1 for explanations of variable abbreviations. The regression coefficients of the standardized variables are given, and their statistical significance is indicated as follows: $^{\circ} p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. AUC values indicate overall model predictive performance, with bootstrapped AUC values in brackets. An AUC value of 0.5 represents random expectation.

	<i>Alsophila cuspidata</i> (small)(prevalence = 13%)		<i>Alsophila cuspidata</i> (large)(12%)		<i>Cyathea multiflora</i> (small)(18%)		<i>Cyathea multiflora</i> (large)(16%)	
	E	E+N	E	E+N	E	E+N	E	E+N
Soil Axis 1	-0.33*	-0.51***	-0.21 $^{\circ}$	-0.20 $^{\circ}$	+0.26*	+0.24*		
Soil Axis 2	-0.53***	-0.56***	-0.28**	-0.31**				
Soil Axis 3								
SV							+0.26**	+0.23*
SW	-0.29*	-0.30*						
RE					-0.38**	-0.31*	+0.31**	+0.33**
RA	-0.34*	-0.29*	-0.34*	-0.30 $^{\circ}$	-0.28 $^{\circ}$	-0.22		
OA			-0.16		-0.34**	-0.26*		
Tree ba	-0.19	-0.18					-0.17 $^{\circ}$	-0.16
N Trees			-0.18 $^{\circ}$	-0.17 $^{\circ}$				
Canopy openness					+4 th order	+cubic		
Topographic position	-linear**							
Elevation								
Slope	+0.62***	+0.62***	+0.48***	+0.50***	+0.47***	+0.41***	+0.14	
Species > 100		+0.37***				+0.31***		
Species 50 m				+0.16 $^{\circ}$		+0.15 $^{\circ}$		+0.24**
Cmult		+0.36***		+0.19*				
Acusp						+0.13		+0.16*
Afirm								
Cursi		+0.24**		+0.20**		+0.11		
AUC	0.77	0.81	0.68	0.70	0.67	0.70	0.60	0.64

individuals of *C. multiflora*, and both size classes of *A. cuspidata*, were biased towards sloping sites. However, small individuals of *C. multiflora* were biased towards higher values on Soil Axis 1, indicating better drainage, whereas *A. cuspidata* was biased towards lower values. Small individuals of *A. cuspidata* were generally biased towards lower topographic positions, but away from the swamp and recent alluvial soils. Some bias away from recent alluvial soils was also evident in large *A. cuspidata*. Large individuals of *C. multiflora* were biased towards residual and stream valley soils, whereas small individuals were biased away from old alluvial and residual soils. Forest structural variables were included in the models slightly more often than in the all size-class analyses, but they were never statistically significant individually.

The inclusion of interspecific neighbourhood effects in the models (Tables 2 and 3) showed that *A. cuspidata* had a significantly overlapping distribution with both *Cyathea* species: the presence of *A. cuspidata* was a significant predictor of the presence of both *C. multiflora* and *C. ursina*, and vice versa. Significant negative associations were not detected for any species pair.

When all size classes were considered, the main differences between the two most common species

were their opposite distributional biases with respect to Soil axis 1, with *C. multiflora* being biased towards better drained sites, and to a lesser extent with respect to Soil axis 3. Within suitable soils, *C. multiflora* was significantly associated with lower topographic positions, whereas *A. cuspidata* was not, and *A. cuspidata* was more strongly associated with steeply sloping sites than *C. multiflora*. *Alsophila cuspidata* was positively associated with stream valley soils, and negatively associated with Soil axis 2, whereas *C. multiflora* was apparently indifferent to these variables.

The inclusion of intraspecific neighbourhood effects (Tables 2 and 3) showed there to be significant local spatial autocorrelation in species distributions, which was unaccounted for by the modelled environmental variables. The incidence of large conspecifics within a plot was a significant predictor of the incidence of small individuals in both *C. multiflora* and *A. cuspidata*. In addition, the incidence of conspecifics in neighbouring plots (50 m away) was included in predictive models of all four species, but was only significant in the two *Cyathea* species (except in small *C. multiflora*).

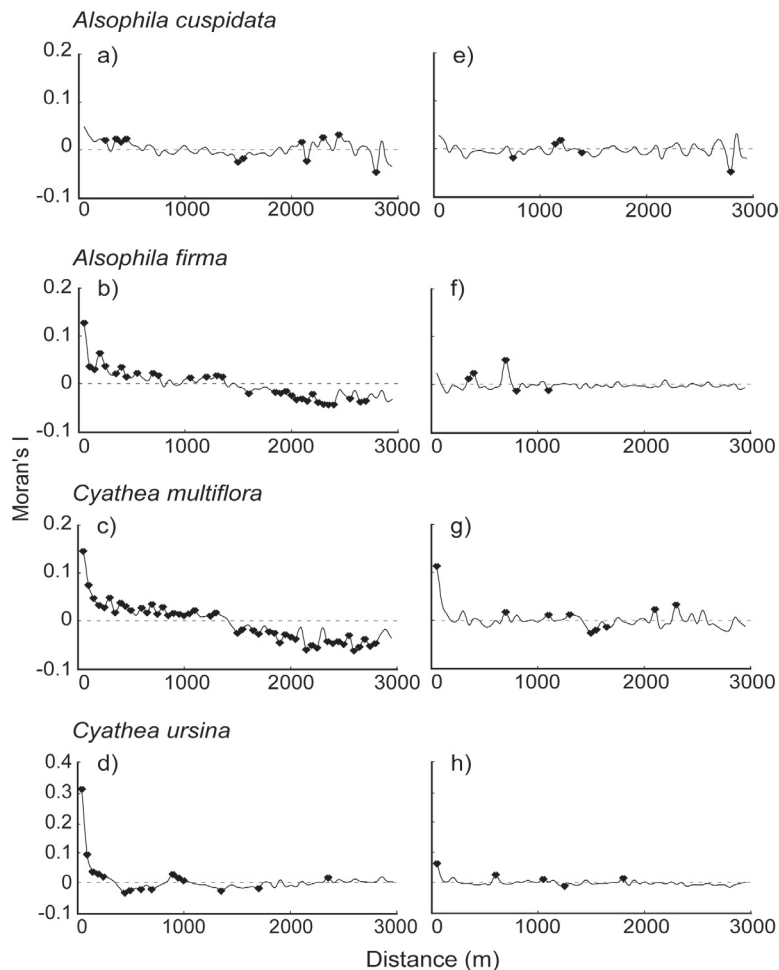


Fig. 2. Correlograms of Moran's I illustrating the patterns of spatial autocorrelation observed at ca. 50 m distance intervals in: (a-d) the species incidence data (all size classes) and (e-h) the environmental GLM model residuals for the four modelled tree fern species. Significant spatial autocorrelation ($p < 0.05$) is indicated by the black symbols.

Model evaluation

Model performance, as measured with the original AUC values, was either equal for the corresponding GLM and GAM models, or marginally better (on average 3% higher) for the GAM models. For the environmental models of *A. firma* and *C. ursina* the AUC values obtained with both modelling methods were identical. The greatest difference, an increase in AUC from 0.68 to 0.73, was observed in the environmental model for large individuals of *A. cuspidata*. However, this increase may also be explained by the fact that two more variables were included in the GAM than the GLM model. In general, however, the subset of variables selected for the GLM and GAM models was very similar. Running GAM models presumably did not markedly improve predictive performance because most of the smoothers were close to linear. In the environmental GAM models, the most flexible smoother was that for the responses of *C. multiflora* and *A. cuspidata* to Soil axis 1, whereby

peaks in their probabilities of occurrence at intermediate values on this axis were indicated.

Prediction success of the environmental GLM models (Table 2), based on the final, bootstrap-estimated AUC values, was moderate to good for the two rarest species, *A. firma* (AUC = 0.77) and *C. ursina* (AUC = 0.82) and moderate for *A. cuspidata* (AUC = 0.70). The models for *C. multiflora* performed less well (AUC = 0.63). With the addition of neighbourhood effects, model predictive power increased for all species. The best performing final model was that for *C. ursina* (AUC = 0.86).

Prediction success was better for small individuals and worse for large individuals of *A. cuspidata* than when all size classes were modelled together (Table 3). In the case of *C. multiflora* the results for small individuals and all size classes were similar, and again better than those for large individuals. The bootstrapped AUC value (0.52) indicated that the prediction success of the environmental model for large individuals of *C. multiflora* was scarcely better than random.

Comparing the original AUC values with the bootstrap-estimated values indicated there to be more uncertainty in environmental model specification and predictive success than in the models including neighbourhood effects. The differences between the original and bootstrapped AUC values were greatest for the environmental models of *A. firma* (13.5%), and the large size classes of *C. multiflora* (13.3%) and *A. cuspidata* (10.3%).

Correlograms of Moran's *I* (Fig. 2) show differences in the pattern of spatial aggregation among the tree fern species, and in their environmental model residuals, at distances ranging from c. 50 m to 3 km. The occurrences of *A. cuspidata* did not show significant local spatial autocorrelation. In contrast, those of *A. firma* and of both *Cyathea* species showed significant spatial aggregation over short distances, especially for sites 50–100 m apart. In *A. firma*, much of the local aggregation was accounted for by spatial pattern in environmental variables, as significant spatial autocorrelation was no longer present in the environmental model residuals. In the environmental model residuals for *C. multiflora* and *C. ursina*, Moran's *I* was much reduced relative to that in the raw distribution data, indicating that a part of the spatial aggregation was also explainable by induced (exogenous) spatial dependence rather than endogenous autocorrelation. However, significant spatial autocorrelation at the 50 m scale remained, which may reflect dispersal limitation.

Discussion

Our results support the view that habitat specialization is an important determinant of where on the rain forest landscape tree ferns grow. All four species had distributional biases with respect to habitat, with several environmental variables significantly related to their occurrence. The significant environmental variables were partly different, but quantitative soil chemical variables entered models of every species. Several earlier studies on rain forest plants have only considered qualitative classes of soil type or topographic position (e.g. Clark et al. 1999; Harms et al. 2001; Cannon & Leighton 2004; Valencia et al. 2004), which probably restricted their ability to identify microhabitat associations (Clark et al. 1999; Hall et al. 2004).

Small individuals of both *C. multiflora* and *A. cuspidata* were better modelled than large individuals. There are at least two possible explanations for this pattern. First, the recruitment success of these species may be determined by environmental suitability to a larger degree than later survival, and the latter may be more affected by random factors, such as physical damage

(cf. Clark & Clark 1989). Second, large individuals can vary considerably in age and some may be quite old, so the environmental data may correspond better to the conditions that prevailed during the lifetimes of the small individuals. Comparable studies are limited, but our results contrast with those of two earlier studies on trees in Borneo (Webb & Peart 2000; Paoli et al. 2006) in which more habitat associations were found in adults than in juveniles, suggesting that the effect of habitat filtering on species distributions builds up during their ontogeny. Whether this reflects real differences between trees and tree ferns, or localities, remains to be clarified.

Both randomization tests and adding neighbourhood variables to the habitat models showed that neither pair of congeners co-occurred more frequently than expected at random. In contrast, the two most common species, *C. multiflora* and *A. cuspidata*, co-occurred more often than expected at random, and they were highly significant positive predictors of each others' incidence in the logistic regression models. This suggests that their habitat preferences overlap, and that the presence of one species indicates the presence of such environmental conditions that are suitable for the other species but were not well modelled by the available environmental data. To a lesser extent, *Cyathea ursina* and *A. cuspidata* also overlapped in their site-wide habitat distributions, but segregated within plots. Both species had a bias towards stream valleys, but the microhabitat of *C. ursina* was unique at the within-plot scale, as this was the only species that usually grew on fallen tree trunks and branches. It is possible that greater separation between the distributions of the two most common species would have been found if the environmental variables had been measured at the exact location of each tree fern individual, if more environmental variables had been included, or if information on species abundances had been used (Stubbs & Wilson 2004; Tuomisto 2006).

An important ecological process that was probably inadequately modelled here is gap dynamics. Available space, released nutrients and suitable light conditions for establishment and growth created by tree and branch falls may be a key element influencing understorey plant distributions (Grubb 1977; Denslow 1987; Svenning 2000; Palmer et al. 2000; Hubbell 2001). Gap formation is also likely to play a stochastic role in the probability of survival of individuals (Clark & Clark 1989). However, the variables representing these effects were static, being an index of canopy openness at the time of inventory, and two measures of forest structure from several years earlier. Forest structure changes continuously, so the information these variables conveyed of forest successional processes during the lifetimes of the tree fern individuals will have been very limited. Forest structural variables were included in models of all species except *C. ursina*,

but were not among the most significant predictors except in models of *A. firma*. Soil chemistry at individual sample sites is also likely to change with time (Sollins 1998), but we assume that differences between sites representing the extremes of gradients in soil drainage and fertility are likely to be more constant.

In addition to being dependent on model quality, the predictability of species-habitat relationships is affected by the fact that all suitable habitat is unlikely to be occupied at a given time (Fielding & Bell 1997; Hanski 1998). Nonetheless, AUC values indicated moderate to good performance for models of all four tree fern species, except when large individuals alone were modelled. Several recent studies have suggested that GAM often performs better than GLM (e.g. Bio et al. 2002; Segurado & Araújo 2004), but in the present study the differences were marginal. This may be because the study area is relatively small, and the species responses along the relatively short environmental gradients can be adequately represented with a linear model. The habitat distributions of the two rarer species appeared to be better predicted than those of the two common species. However, comparison of the AUC values among species is problematic, because model evaluation is more uncertain for species of low incidence (McPherson et al. 2004). The bootstrap analyses indicated particularly high uncertainty in the environmental model results for one of the two rarer species, *A. firma*, as well as for large individuals of both common species.

As well as influencing model predictability, a sparse distribution pattern could, in itself, decrease the necessity for niche differentiation among species (Grubb 1986; Wright 2002). Although their distributions overlapped significantly, the two most common species in this study also often occurred separately (in 67% of plots occupied by *C. multiflora*, and in 40% of plots occupied by *A. cuspidata*), and average densities per occupied 100 m² plot were less than three for both species, with most individuals belonging to the smallest size classes. Therefore, crown and root overlap in these species is likely to be quite uncommon, and they may only rarely compete for light and soil resources. Other elements of the flora may be more important competitors (Wright 2002; Tuomisto 2006). However, even if microhabitat differences are not obvious at local spatial scales, niche differences may still exist at broader scales (cf. Tuomisto 2006).

Significant spatial autocorrelation, especially over distances of 50 – 100 m, was evident in three of the species, particularly in the *Cyathea* species. This is in agreement with several recent studies that have identified significant aggregation in rain forest plants (e.g. He et al. 1997; Condit et al. 2000; Plotkin et al. 2000; Svenning 2001). A failure to account for spatial autocorrelation in statistical analyses can inflate the apparent importance of

environmental variation for species distributions (Legendre & Legendre 1998; Harms et al. 2001). Nonetheless, in the present study adding neighbourhood variables to the initial environmental models had only a minor impact on the set of environmental variables selected. In the case of *A. firma*, examination of the model residuals showed that local-scale aggregation was almost entirely accounted for by the modelled environmental variables, whereas in the *Cyathea* species it was not, and the remainder may partly reflect an effect of dispersal limitation (Schwarz et al. 2003). Stronger dispersal limitation in *Cyathea* can be expected on morphological grounds, because both *Cyathea* species have short trunks in comparison with the *Alsophila* species (pers. obs.; Riba 1995; Moran 1995). The positive relationship between the incidence of large and small conspecifics found in our study could be due to either dispersal limitation or to temporal persistence of suitable establishment conditions (Svenning 2001; Hamill & Wright 1986). More long term studies would be needed to distinguish among these alternatives.

Acknowledgements. We thank E. Heegaard for modelling advice, R. Gonzalez for help during the tree fern survey, and J. Mackensen and E. Veldcamp for help with soil sampling. La Selva Biological Station of the Organization for Tropical Studies provided logistic support, and D. and M. Lieberman kindly allowed us to collect data in their restricted access study areas. The work was funded by grants from the Academy of Finland (to H. Tuomisto) and the Andrew W. Mellon Foundation (to D.B. and D.A. Clark). Research permits were granted by the Costa Rican Ministry of the Environment and Energy (MINAE).

References

- Anon. 2006. *R: A language and environment for statistical computing*. R Development Core Team, R Foundation for Statistical Computing, Vienna, AT, URL: <http://www.R-project.org>.
- Abrams, P.A. 1983. The theory of limiting similarity. *Annu. Rev. Ecol. Syst.* 14: 359-376.
- Akaike, H. 1977. On entropy maximization principle. In: Krishnaiah, P.R. (ed.) *Applications of statistics*, pp. 27-41. North-Holland Publishing Company, Amsterdam, NL.
- Ashton, P.S. 1969. Speciation among tropical forest trees: Some deductions in light of recent evidence. *Biol. J. Linn. Soc.* 1: 155-196.
- Austin, M.P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol. Model.* 157: 101-118.
- Bio, A.M.F., de Becker, P., de Bie, E., Huybrechts, W. & Wassen, M. 2002. Prediction of plant species distribution in lowland river valleys in Belgium: modelling species response to site conditions. *Biodivers. Conserv.* 11: 2189-2216.
- Brown, N., Jennings, S., Wheeler, P. & Nabe-Nielsen, J.

2000. An improved method for the rapid assessment of forest understorey light environments. *J. Appl. Ecol.* 37: 1044-1053.
- Cannon, C.H. & Leighton, M. 2004. Tree species distributions across five habitats in a Bornean rain forest. *J. Veg. Sci.* 15: 257-266.
- Casgrain, P. & Legendre, P. 2001. *The R Package for Multivariate and Spatial Analysis, Version 4.0*. Département de Sciences Biologiques, Université de Montréal, Montréal, CA.
- Clark, D.B. & Clark, D.A. 1989. The role of physical damage in the seedling mortality regime of a neotropical rain forest. *Oikos* 55: 225-230.
- Clark, D.A., Clark, D.B., Sandoval M., R. & Castro C., M.V. 1995. Edaphic and human effects on landscape-scale distributions of tropical rain forest palms. *Ecology* 76: 2581-2594.
- Clark, D.B., Clark, D.A. & Read, J.M. 1998. Edaphic variation and the mesoscale distribution of tree species in a neotropical rain forest. *J. Ecol.* 86: 101-112.
- Clark, D.B., Palmer, M.W. & Clark, D.A. 1999. Edaphic factors and the landscape-scale distributions of tropical rain forest trees. *Ecology* 80: 2662-2675.
- Condit, R., Ashton, P.S., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N., Hubbell, S.P., Foster, R.B., Itoh, A., LaFrankie, J.V., Lee, H.S., Losos, E., Manokaran, N., Sukumar, R. & Yamakura, T. 2000. Spatial patterns in the distribution of tropical tree species. *Science* 288: 1414-1418.
- Davies, S.J., Palmiotto, P.A., Ashton, P.S., Lee, H.S. & LaFrankie, J.V. 1998. Comparative ecology of 11 sympatric species of *Macaranga* in Borneo: tree distribution in relation to horizontal and vertical resource heterogeneity. *J. Ecol.* 86: 662-673.
- Debski, I., Burslem, D.F.R.P., Palmiotto, P.A., LaFrankie, J.V., Lee, H.S. & Manokaran, N. 2002. Habitat preferences of *Aporosa* in two Malaysian forests: implications for abundance and coexistence. *Ecology* 83: 2005-2018.
- Denslow, J.S. 1987. Tropical rainforest gaps and tree species diversity. *Annu. Rev. Ecol. Syst.* 18: 431-451.
- Duque, A., Sánchez, M., Cavelier, J. & Duivenvoorden, J.F. 2002. Different floristic patterns of woody understorey and canopy plants in Colombian Amazonia. *J. Trop. Ecol.* 18: 499-525.
- Fielding, A.H. & Bell, J.F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* 24: 38-49.
- Grubb, P.J. 1977. The maintenance of species-richness in plant communities: The importance of the regeneration niche. *Biol. Rev.* 52: 107-145.
- Grubb, P.J. 1986. Problems posed by sparse and patchily distributed species in species-rich plant communities. In: Diamond, J & Case T.J. (eds.) *Community ecology*, pp. 207-225. Harper & Row, New York, NY, US.
- Hall, J.S., McKenna, J.J., Ashton, P.M.S. & Gregoire, T.G. 2004. Habitat characterizations underestimate the role of edaphic factors controlling the distribution of *Entandrophragma*. *Ecology* 85: 2171-2183.
- Hamill, D.N. & Wright, S.J. 1986. Testing the dispersion of juveniles relative to adults: a new analytic method. *Ecology* 67: 952-957.
- Hanski, I. 1998. Metapopulation dynamics. *Nature* 396: 41-49.
- Harms, K.E., Condit, R., Hubbell, S.P. & Foster, R.B. 2001. Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. *J. Ecol.* 89: 947-959.
- He, F., Legendre, P. & La Frankie, J.V. 1997. Distribution patterns of tree species in a Malaysian tropical rain forest. *J. Veg. Sci.* 8: 105-114.
- Hofton, M.A., Rocchio, L.E., Blair, J.B. & Dubayah, R. 2002. Validation of vegetation canopy lidar sub-canopy topography measurements for a dense tropical forest. *J. Geodyn.* 34: 491-502.
- Hosmer, D.W. & Lemeshow, S. 2000. *Applied logistic regression*. 2nd ed. John Wiley & Sons, New York, NY, US.
- Hubbell, S.P. 2001. *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, NJ, US.
- Hubbell, S.P. & Foster, R.B. 1986. Biology, chance, and history and the structure of tropical rain forest tree communities. In: Diamond, J. & Case, T.J. (eds.) *Community ecology*, pp. 314-329. Harper & Row, New York, NY, US.
- Jones, M.M., Tuomisto, H., Clark, D.B. & Olivas, P. 2006. Effects of mesoscale environmental heterogeneity and dispersal limitation on floristic variation in rain forest ferns. *J. Ecol.* 94: 181-195.
- Legendre, P. & Legendre, L. 1998. *Numerical ecology*. 2nd ed. Elsevier, Amsterdam, NL.
- Lieberman, M., Lieberman, D., Peralta, R. & Hartshorn, G.S. 1995. Canopy closure and the distribution of tropical forest tree species at La Selva, Costa Rica. *J. Trop. Ecol.* 11: 161-177.
- MacArthur, R. & Levins, R. 1967. Limiting similarity convergence and divergence of coexisting species. *Am. Nat.* 101: 377-387.
- Manel, S., Williams, H.C. & Ormerod, S.J. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *J. Appl. Ecol.* 38: 921-931.
- McPherson, J.M., Jetz, W. & Rogers, D.J. 2004. The effect of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *J. Appl. Ecol.* 41: 811-823.
- Moran, R.C. 1995. *Cyathea*. In: Moran, R.C. & Riba, R. (eds.) *Flora Mesoamericana Volumen 1 Psilotaceae a Salviniaceae*, pp. 93-103. Universidad Nacional Autónoma de México, Mexico, MX.
- Moran, R.C. & Riba, R. (eds.) 1995. *Flora Mesoamericana Vol. 1 Psilotaceae a Salviniaceae*, pp. 93-103. Universidad Nacional Autónoma de México, Mexico, MX.
- Nicholls, A.O. 1989. How to make biological surveys go further with generalised linear models. *Biol. Conserv.* 50: 51-75.
- Pacala, S.W. & Tilman, D. 1994. Limiting similarity in mechanistic and spatial models of plant composition in heterogeneous environments. *Am. Nat.* 143: 222-257.
- Palmer, M.W., McAlister, S.D., Arévalo, J.R. & DeCoster, J.K. 2000. Changes in the understorey during 14 years following catastrophic windthrow in two Minnesota forests. *J. Veg. Sci.* 11: 841-854.

- Paoli, G.D., Curran, L.M. & Zak, D.R. 2006. Soil nutrients and beta diversity in the Bornean Dipterocarpaceae: evidence for niche partitioning by tropical rain forest trees. *J. Ecol.* 94: 157-170.
- 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. *J. Theor. Biol.* 207: 81-99.
- Riba, R. 1995. *Alsophila*. In: Moran, R.C. & Riba, R. (eds.) *Flora Mesoamericana Volumen 1 Psilotaceae a Salviniaceae*, pp. 88-90. Universidad Nacional Autónoma de México, Mexico, MX.
- Schwarz, P.A., Fahey, T.J. & McCulloch, C.E. 2003. Factors controlling spatial variation of tree species abundance in a forested landscape. *Ecology* 84: 1862-1878.
- Segurado, P. & Araújo, M.B. 2004. An evaluation of methods for modelling species distributions. *J. Biogeogr.* 31: 1555-1568.
- Smith, P.A. 1994. Autocorrelation in logistic regression modelling of species' distributions. *Global Ecol. Biogeogr.* 4: 47-61.
- Sollins, P. 1998. Factors influencing species composition in tropical lowland rain forest: does soil matter? *Ecology* 79: 23-30.
- Soltis, P.S., Soltis, D.E., Savolainen, V., Crane, P.R. & Barraclough, T.G. 2002. Rate heterogeneity among lineages of tracheophytes: Integration of molecular and fossil data and evidence for molecular living fossils. *Proc. Natl. Acad. Sci.* 99: 4430-4435.
- Stubbs, W.J. & Wilson, J.B. 2004. Evidence for limiting similarity in a sand dune community. *J. Ecol.* 92: 557-567.
- Svenning, J.-C. 1999. Microhabitat specialization in a species-rich palm community in Amazonian Ecuador. *J. Ecol.* 87: 55-65.
- Svenning, J.-C. 2000. Small canopy gaps influence plant distributions in the rain forest understory. *Biotropica* 32: 252-261.
- Svenning, J.-C. 2001. Environmental heterogeneity, recruitment limitation and the mesoscale distribution of palms in a tropical montane rain forest (Maquipucuna, Ecuador). *J. Trop. Ecol.* 17: 97-113.
- Tuomisto, H. 2006. Edaphic niche differentiation among *Polybotrya* ferns in Western Amazonia: implications for coexistence and speciation. *Ecography* 29: 273-284.
- Tuomisto, H., Ruokolainen, K., Poulsen, A.D., Moran, R.C., Quintana, C., Cañas, G. & Celi, J. 2002. Distribution and diversity of pteridophytes and Melastomataceae along edaphic gradients in Yasuni National Park, Ecuadorian Amazonia. *Biotropica* 34: 516-533.
- Valencia, R., Foster, R.B., Villa, G., Condit, R., Svenning, J.-C., Hernández, C., Romoleroux, K., Losos, E., Magård, E. & Balslev, H. 2004. Tree species distributions and local habitat variation in the Amazon: large forest plot in eastern Ecuador. *J. Ecol.* 92: 214-229.
- Venables, W.N. & Ripley, B.D. 2002. *Modern Applied Statistics with S*. 4th ed. Springer Science+Business Media, Inc., New York, NY, US.
- Webb, C.O. & Peart, D.R. 2000. Habitat associations of trees and seedlings in a Bornean rain forest. *J. Ecol.* 88: 464-478.
- Wintle, B.A., Elith, J. & Potts, J.M. 2005. Fauna habitat modelling and mapping: A review and case study in the Lower Hunter Central Coast region of NSW. *Aust. Ecol.* 30: 719-738.
- Wood, S.N. 2003. Thin-plate regression splines. *J. R. Stat. Soc. B* 65: 95-114.
- Wood, S.N. & Augustin, N.H. 2002. GAMs with integrated model selection using penalized regression splines and applications to environmental modeling. *Ecol. Model.* 157: 157-177.
- Wright, S.J. 2002. Plant diversity in tropical forests: a review of mechanisms of species coexistence. *Oecologia* 130: 1-14.
- Wu, H. & Huffer, F.W. 1997. Modelling the distribution of plant species using the autologistic regression model. *Environ. Ecol. Stat.* 4: 49-64.
- Yee, T.W. & Mitchell, N.D. 1991. Generalized additive models in plant ecology. *J. Veg. Sci.* 2: 587-602.

Received 19 June 2006;

Accepted 14 September 2006;

Co-ordinating Editor: J. Franklin.