

# Effects of mesoscale environmental heterogeneity and dispersal limitation on floristic variation in rain forest ferns

MIRKKA M. JONES, HANNA TUOMISTO, DAVID B. CLARK\* and PAULO OLIVAS†

*Department of Biology, University of Turku, 20014 Turku, Finland, \*Department of Biology, University of Missouri-St Louis, 8001 Natural Bridge Road, St Louis, Missouri 63121, USA, and †Department of Biological Sciences, Florida International University, Miami, Florida 33199 USA*

## Summary

**1** Field studies to evaluate the roles of environmental variation and random dispersal in explaining variation in the floristic composition of rain forest plants at landscape to regional scales have yet to reach a consensus. Moreover, only one study has focused on scales below 10 km<sup>2</sup>, where the effects of dispersal limitation are expected to be easiest to observe.

**2** In the present study, we estimate the importance of differences in some key environmental variables (describing canopy openness, soils and topography) relative to the geographical distances between sample plots as determinants of differences in pteridophyte (ferns and fern allies) species composition between plots within a *c.* 5.7 km<sup>2</sup> lowland rain forest site in Costa Rica.

**3** To assess the relative importance of environmental vs. geographical distances in relation to the length of environmental gradient covered, we compared the results obtained over the full range of soil types, including swamps, with those for upland soils alone.

**4** Environmental variability was found to be a far stronger predictor of changes in floristic differences than the geographical distance between sample plots. In particular, differences in soil nutrient content, drainage and canopy openness correlated with floristic differences.

**5** The decline in mean floristic similarity with increasing geographical distance was stronger than proposed by the random dispersal model over short distances (up to *c.* 100 m), which is probably attributable to both dispersal limitation and environmental changes. The scatter around the mean was large at all distances.

**6** Our initial expectation was that the effects of dispersal limitation (represented by geographical distance) on observed patterns of floristic similarity would be stronger, and those of environmental differences weaker, than at broader spatial scales. Instead, these results suggest that the niche assembly view is a more accurate representation of pteridophyte communities at local to mesoscales than the dispersal assembly view.

*Key-words:* Costa Rica, dispersal limitation, environmental heterogeneity, floristic variation, La Selva, mesoscale, niche assembly, pteridophytes, variation partitioning, vegetation-environment relationships

*Journal of Ecology* (2006) **94**, 181–195  
doi: 10.1111/j.1365-2745.2005.01071.x

## Introduction

The relative importance of stochastic and deterministic processes in structuring rain forest plant communities has long been a subject of debate (e.g. Fedorov 1966;

Ashton 1969; Hubbell & Foster 1986), with the corresponding extreme views of community structuring now often being termed niche assembly and dispersal assembly (Hubbell 2001).

Under niche assembly, plant communities are expected to develop towards an equilibrium at which their species abundances and composition are closely determined by site environmental characteristics. There is general

agreement that large environmental differences between sites produce noticeable effects on vegetation structure and plant species composition. Such obvious differences within rain forests include seasonally inundated or permanently waterlogged forests as opposed to well-drained upland forests (*tierra firme*) and forests with extremely nutrient-poor white sand soils as opposed to those with less extreme soils (Prance 1989). However, opinions vary widely on whether environmental differences within these broad forest types, especially in *tierra firme*, are of sufficient ecological importance to be reflected in floristic patterns (e.g. Gentry 1988; Duivenvoorden 1995; Tuomisto *et al.* 1995; Condit 1996; Clark *et al.* 1999; Pitman *et al.* 1999, 2001; Tuomisto, Ruokolainen & Yli-Halla 2003).

Hubbell (2001) advocated a predominantly dispersal assembly view, at least for *tierra firme* rain forests. He suggested that dispersal limitation could decouple the control of species composition from resource-based niches. Thus, the composition of a local community will be less dependent on the ecological properties of the site itself than on the composition of the surrounding metacommunity and the immigration rate of its constituent species. In the extreme case, where community composition is entirely controlled by dispersal assembly, floristic similarity between sites will decrease with increasing geographical distance and will be independent of any environmental differences between sites. Condit *et al.* (2002) further suggested that the decline in similarity should be linear with the logarithm of geographical distance. Thus, especially in communities with many rare species (Hubbell 2001), this decline should be most obvious over short distances.

The extent to which these two alternative views of rain forests are supported by field data has recently been tested at landscape to regional scales (> 10 km<sup>2</sup>) in tropical America and south-east Asia. The results of these studies have varied. In western Amazonia, Terborgh *et al.* (1996) found mature floodplain forests to converge in tree species composition, despite intervening distances of some tens of kilometres, and interpreted this as evidence that random dynamics with dispersal limitation are not the major force determining the floristic composition of tropical forests. Condit *et al.* (2002) found the tree species composition of Panamanian forests to be more heterogeneous and of Amazonian forests to be more homogeneous than expected under the random dispersal model, and re-analyses of their Panamanian data showed that environmental distances explained a large part of the variation in floristic differences between sites (Duivenvoorden *et al.* 2002; Ruokolainen & Tuomisto 2002). A number of other recent studies have tried to quantify the extent to which environmental and geographical distances together explain floristic differences between sites, as well as to quantify their relative roles. Both have invariably been shown to be of consequence, but their relative importance has differed between studies. The environmental model has provided consistently better predictive power for

Amazonian pteridophytes and Melastomataceae than the geographical model (Ruokolainen *et al.* 1997; Tuomisto, Poulsen, Ruokolainen, Moran, Quintana & Cañas 2003; Tuomisto, Ruokolainen, Aguilar & Sarmiento 2003; Tuomisto, Ruokolainen & Yli-Halla 2003). The results for trees and palms have been more variable. Some studies (Ruokolainen *et al.* 1997; Phillips *et al.* 2003) found environmental factors to be more important, whereas others (Potts *et al.* 2002; Vormisto, Svenning, Hall & Balslev 2004) suggested that the spatial separation of sites was more important. Several studies have noted that when the sampled environmental gradient is long, environmental variables have more explanatory power than when the gradient is made shorter by excluding, for example, inundated or white sand sites (e.g. Ruokolainen *et al.* 1997). In fact, Duque *et al.* (2002) found a stronger effect of environmental factors than spatial separation when all their landscape units were analysed, but found the reverse when only upland soil units were considered.

At more restricted spatial scales, while environmental factors have received considerable attention, studies that explicitly separate the relative role of environmental and geographical distance effects are conspicuously lacking in tropical rain forests (but see Svenning & Skov 2002 and Gilbert & Lechowicz 2004 for temperate examples). A mesoscale study in Panama (Svenning *et al.* 2004) evaluated the relative roles of spatial patterns and environmental variables on species composition, but did not evaluate the effect of distances as such. In studies focusing on environmental variation alone, the factors that have most commonly been found to affect plant distributions within tropical rain forests at local (< 1 km<sup>2</sup>) to meso-scales (1–10 km<sup>2</sup>) include gap dynamics (Denslow 1987; but see Lieberman *et al.* 1995), topographic position and edaphic characteristics such as soil fertility, texture and drainage (e.g. ter Steege *et al.* 1993; Tuomisto & Ruokolainen 1994; Svenning 1999; Vormisto *et al.* 2000; Harms *et al.* 2001; Péllissier *et al.* 2001; Vormisto, Tuomisto & Oksanen 2004; Valencia *et al.* 2004). Potts *et al.* (2004) analysed niche overlap in terms of the occurrence of tree species over altitudinal gradients at two sites in Malaysia, using altitude as a surrogate for more detailed environmental data. They found a lower mean species overlap than expected from null models designed to minimize any habitat associations, especially at the more environmentally heterogeneous of the two sites.

Within our mesoscale study site in Costa Rica, environmental effects on species composition have been demonstrated for some plant groups but the role of geographical distances has not been investigated previously. Lieberman *et al.* (1985) found the species composition of large (> 10 cm d.b.h.) trees, palms, tree ferns and lianas to vary continuously with altitude on upland soils, but to be patchy under poorly drained swamp conditions. Clark *et al.* (1995, 1998, 1999) studied the distribution and abundance of palms and trees more extensively. They concluded that at least 30% of the

species had edaphic or topographic biases in their distributions; a clear floristic gradient was observed in response to soil type and, within the soil types, in response to topographic position. However, the environmental variables were only able to explain a small proportion of the floristic variation. This was suggested to reflect the lack of soil chemical data at that time, the impacts of other unmeasured environmental variables and noise due to small sample size in the floristic data. These methodological issues have been addressed as far as possible in the current study, through the inclusion of soil chemical and canopy openness data, and through our focus on a common understorey plant group, the pteridophytes (ferns and fern allies). Spatially and taxonomically representative data are more easily collected for pteridophytes than for trees and, even though pteridophytes may be more efficient long-distance dispersers than trees, both groups show leptokurtic dispersal (most propagules fall close to the mother plant; Wolf *et al.* 2001).

To our knowledge, this study is the first in which the effects of environmental and geographical distances on floristic distances are explicitly compared at the local-mesoscale in a tropical rain forest site. The effects of dispersal-limitation on observed patterns of floristic similarity are expected to be strongest at these scales (due to the logarithmic relationship; Hubbell 2001), and environmental effects less pronounced than at broader spatial scales.

Our specific aims here were:

1. To estimate the importance of differences in some key environmental variables (describing canopy openness, soils and topography) relative to the geographical distances between sample sites as determinants of variation in floristic composition within a *c.* 5.7 km<sup>2</sup> study area.
2. To assess the relative importance of environmental variation vs. geographical distances in relation to the length of the environmental gradient covered.

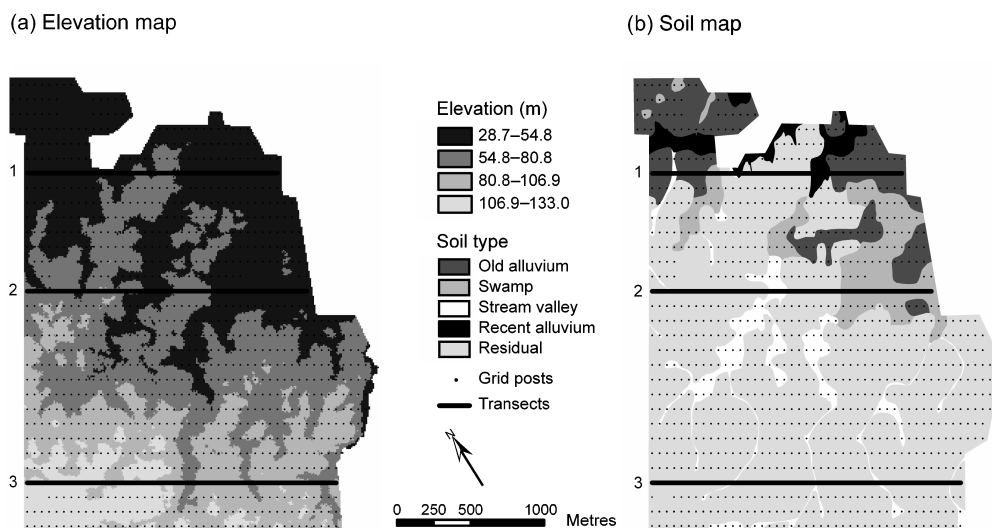
## Materials and methods

### FIELDWORK

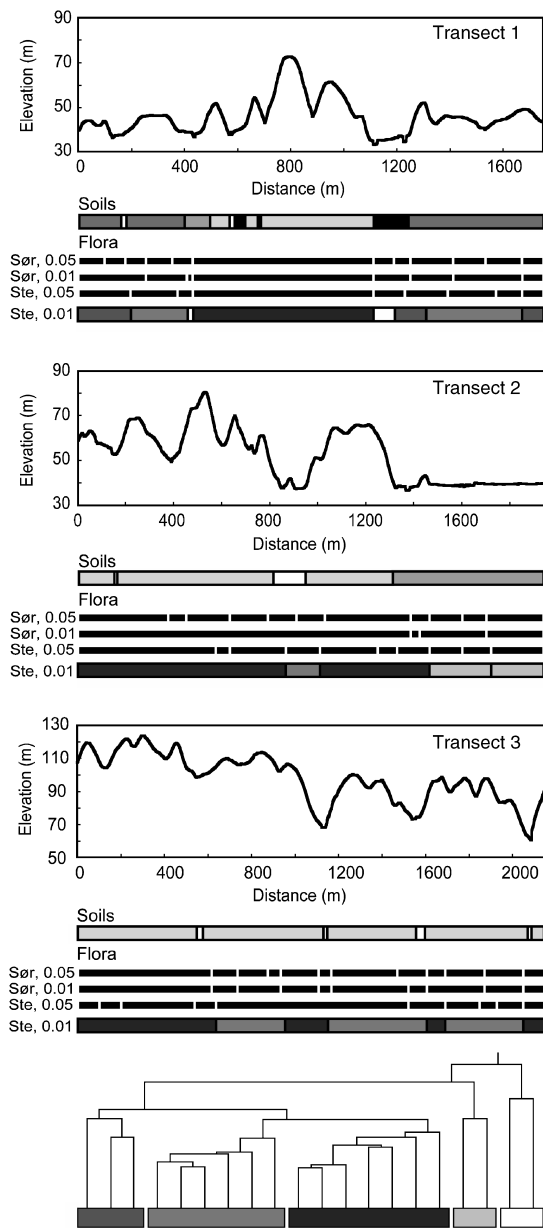
The study was carried out in a 573-ha tract of old growth forest belonging to La Selva Biological Station of the Organization for Tropical Studies from June to September 2001 (10°26' N, 84°00' W). The area is in the Caribbean lowlands of Costa Rica and is classified as tropical wet forest in the Holdridge system (Hartshorn & Hammel 1994). It has a mean monthly temperature of 25.8 °C and receives, on average, over 100 mm of rain each month, totalling about 4000 mm annually (Sanford *et al.* 1994). Elevation in the area is between *c.* 30 m and 130 m above sea level (Fig. 1a). The area covers a range of soil types, including alluvial terraces formed by both recent and historic flooding, swamps, 'residual' soils formed by the *in-situ* weathering of ancient lava flows, and stream valleys with infertile colluvial soils (Sollins *et al.* 1994; Clark *et al.* 1999; Fig. 1b).

The study area is covered by a grid with permanent intersection markers (grid posts) with 50 × 100 m spacing, placed to decimetre accuracy (Hofton *et al.* 2002). In the present study three continuous transects were established. Their locations were planned with reference to the topographic and soil maps (Fig. 1) so as to cover as wide a range of topographic and soil variation as possible, while adhering to the grid system. Each transect was 5 m wide, and their lengths (1750 m, 1940 m and 2150 m) were determined by the extent of the grid system. The transects were measured using a tape measure following the contours of the terrain, whereas the grid system has an exact horizontal spacing. Consequently, in hilly terrain the distances measured between grid posts were slightly longer than 50 m.

Clinometer (Suunto, Vantaa, Finland) readings were taken ( $\pm 1^\circ$ ) at all slope changes along the length of each transect to map topographic profiles (Fig. 2). Five



**Fig. 1** Map of the study area at La Selva biological station with the locations of the three study transects indicated on an elevation map (a) and a soil map (b).



**Fig. 2** Topographic profiles of the three study transects with corresponding soil and floristic classifications. Beneath each profile the first bar indicates the distribution of the five main soil types (as in Fig. 1b). The breaks in the three black lines and the vertical lines in the lowermost bar for each transect show where chronological clustering detected significant discontinuities in floristic composition. The analyses were conducted using two measures of floristic similarity, the Sørensen index (the first two lines), and the Steinhaus index (the next two lines) and two different alpha levels, 0.05 and 0.01. The chronological clustering results in the lowest bars (Steinhaus index, alpha = 0.01) illustrate the floristic similarities of those sections from all three transects. Five main classes were identified from these sections by hierarchical clustering (see UPGMA dendrogram at bottom of figure).

(aspect) to distinguish sites on either side of a north–south axis, and cosine (aspect) to distinguish those on either side of an east–west axis. Elevation was converted into relative elevations within each transect (= actual elevation – minimum elevation within that transect) to compensate for the fact that the general elevation of the terrain rose almost perpendicularly to the transects. Relative topographic position was defined as one of five ordered classes: (i) flat high ground (e.g. hilltop or raised terrace); (ii) upper slope; (iii) mid-slope; (iv) base of slope/flat, low ground; and (v) riparian. This is a modification of the classification used by Clark *et al.* (1999). Canopy openness was estimated at every grid post along the transects using the canopy-scope method of Brown *et al.* (2000) to estimate the size of the largest visible canopy gap on a scale of 0–25 (mean values of this index by soil type are in Table 1).

The occurrence of pteridophytes was recorded in 5 × 5 m contiguous subunits within the transects. In each subunit, all individuals with at least one green frond longer than 10 cm were identified and counted. Epiphytic and climbing individuals were considered only when they had such leaves less than 2 m above the ground. The inventory included terrestrial herbs and tree ferns, plants growing on fallen trees (including canopy species in cases where they had survived the fall), low trunk epiphytes and climbers. In species able to reproduce vegetatively, all apparently separate individuals were counted, although some may have been ramets of a single clone. Representative voucher specimens of each species were collected for later verification of species identifications. In addition, individuals whose reliable identification was not possible in the field were collected. Duplicates of the specimens are deposited in several herbaria in Costa Rica (CR, USJ and the herbarium of the La Selva Biological Station) and Finland (TUR). Unicates are in CR. Each species was classified *a posteriori*, based on field observations made both during this study and later, as predominantly either terrestrial, climbing, epiphytic, epilithic, or a combination of those categories (see Table S1 in Supplementary Material). The epiphytic class includes plants found growing on fallen trees and branches, as well as those growing on standing trees, provided they did not have a root connection to the ground. The epilithic plants were those growing on stones, especially at streamsides.

Soil chemical data were obtained from samples collected for a separate study on small-scale heterogeneity in soils (D. B. Clark, unpublished data, mean values by soil type in Table 1). Surface soil samples were taken to a depth of 10 cm at 1171 grid posts in the study area, and each consisted of 8–10 pooled subsamples collected within 2 m of the grid post. All subsamples were combined in the field, and any roots, organic layer and stones were removed. The samples were then dried, sieved with a 2 mm mesh and sent to the Institute of Soil Science and Forest Nutrition at Göttingen in Germany for chemical analysis. Soil pH was measured in 1 M KCl, exchangeable K, Ca, Mg, Mn, Fe and Al in 1 M NH<sub>4</sub>Cl,

topographic variables were defined for each grid post location, using data on slope, aspect, elevation and topographic position collected previously (Clark *et al.* 1999). Aspect was divided into two components, sine

**Table 1** Mean values ( $\pm$  SD) for the canopy openness index and soil chemical data, by soil type, as measured within the 25-m sample plots centred on grid posts

	Residual ( <i>n</i> = 71)	Old alluvium ( <i>n</i> = 16)	Stream valley ( <i>n</i> = 8)	Recent alluvium ( <i>n</i> = 4)	Swamp ( <i>n</i> = 13)
Canopy openness (0–25)	2.4 $\pm$ 3.2	2.6 $\pm$ 3.8	3.4 $\pm$ 2.8	9.0 $\pm$ 9.5	4.0 $\pm$ 2.5
pH	3.7 $\pm$ 0.1	3.7 $\pm$ 0.1	4.0 $\pm$ 0.3	4.2 $\pm$ 0.4	4.4 $\pm$ 0.3
Ca (mmol(+) kg <sup>-1</sup> )	5.7 $\pm$ 4.3	7.7 $\pm$ 5.9	20.7 $\pm$ 24.9	58.5 $\pm$ 39.7	114.3 $\pm$ 117.0
K (mmol(+) kg <sup>-1</sup> )	2.3 $\pm$ 0.9	3.0 $\pm$ 0.9	2.9 $\pm$ 1.7	3.3 $\pm$ 1.1	4.7 $\pm$ 2.4
Mg (mmol(+) kg <sup>-1</sup> )	4.0 $\pm$ 1.1	5.2 $\pm$ 2.7	8.5 $\pm$ 6.9	27.1 $\pm$ 22.6	57.8 $\pm$ 68.3
Mn (mmol(+) kg <sup>-1</sup> )	1.3 $\pm$ 1.3	2.4 $\pm$ 1.7	1.8 $\pm$ 1.2	3.4 $\pm$ 0.5	3.6 $\pm$ 2.0
Fe (mmol(+) kg <sup>-1</sup> )	2.9 $\pm$ 1.7	1.3 $\pm$ 1.2	4.8 $\pm$ 8.0	0.3 $\pm$ 0.2	2.2 $\pm$ 3.4
Al (mmol(+) kg <sup>-1</sup> )	69.9 $\pm$ 14.7	71.5 $\pm$ 12.3	53.6 $\pm$ 27.1	23.9 $\pm$ 27.5	20.1 $\pm$ 15.8
C (mg g <sup>-1</sup> )	56.2 $\pm$ 13.7	46.4 $\pm$ 7.6	84.1 $\pm$ 76.9	64.0 $\pm$ 4.2	113.7 $\pm$ 86.0
P (mg g <sup>-1</sup> )	1.1 $\pm$ 0.3	1.0 $\pm$ 0.2	1.3 $\pm$ 0.3	2.0 $\pm$ 0.2	1.9 $\pm$ 0.6

total C using combustion, and total P using HNO<sub>3</sub>-pressure extraction. More details on the methodology can be found in König & Fortmann (1996a,b).

Each grid intersection was also defined as belonging to one of five broad soil classes defined by Clark *et al.* (1998) (old alluvium, recent alluvium, residual, stream valley or swamp, see Fig. 1b).

#### DATA ANALYSIS

Each of the original 5 × 5 m subunits contained few pteridophyte individuals and species, so the numerical analyses were done on larger subunits to reduce the effect of noise and avoid the occurrence of subunits without any pteridophytes. Four or five contiguous subunits were pooled to obtain new subunits either 20 m or 25 m long (100 m<sup>2</sup> and 125 m<sup>2</sup>, respectively). The 20-m subunits (the minimum size needed to avoid empty subunits; *n* = 294) were contiguous along the transects, and were used in the analyses that only involved floristic data. The 25-m subunits (*n* = 112) were separated from each other by *c.* 25 m because they were centred on the grid posts at which the soil samples and canopy openness measurements had been taken (i.e. they included the 5-m subunit in which the grid post was situated, and two subunits from either side). Canopy openness and soil Fe data were missing for a few grid posts, so analyses involving these two variables include fewer subunits than the other analyses (analyses with canopy openness data, *n* = 108; analyses with Fe data, *n* = 105; analyses including both canopy openness and Fe data, *n* = 102).

#### Calculation of resemblance matrices

All the analyses were based on resemblance matrices, which express a measure of the similarity (or distance) of all the possible pairs of subunits in floristic, environmental or geographical terms, using an appropriate index for each of the data types. Some of the analyses required the conversion of similarity indices (S) to distance indices (D), or vice versa, in which case the formula  $S = 1 - D$  was used.

Floristic resemblance matrices were computed using two indices of floristic similarity, the Sørensen index and

the Steinhaus index, which are equivalent except that the former only uses species presence-absence data, whereas the latter also takes into account species abundance.

Environmental resemblance matrices were computed using Euclidean distances. Prior to the distance calculations, the quantitative soil variables (except pH) were transformed by taking their natural logarithm (ln). Logarithmic transformation of the soil chemical variables gives more weight to a unit difference in their concentrations when the overall content is low than when it is high. The five qualitative soil classes were coded as binary (dummy) variables. Separate distance matrices were computed using each of the environmental variables in turn, and additional distance matrices were computed using a combination of several variables. In the latter case, the variables were standardized to zero mean and unit variance before computing the distances in order to give each variable equal weight.

A geographical distance matrix was calculated using the X and Y coordinates of each grid post in the biological station's coordinate system. These distances were then ln-transformed. According to Hubbell's neutral theory (Hubbell 2001), a linear relationship is expected between the logarithm of geographical distance between sites and their floristic resemblance (Condit *et al.* 2002; Hubbell 2001).

#### Clustering

Two clustering methods were used to visualize spatial patterns in floristic similarity both within and between transects, and the outputs are plotted alongside the topography and soil types of each transect in Fig. 2. Both clustering methods use an agglomerative proportional-link linkage algorithm to group sites based on their floristic similarity. The connectedness was always set to 0.5, i.e. midway between single-link (nearest neighbour) and complete-link (furthest neighbour) linkage.

First, chronological clustering was used to identify where statistically significant changes in species composition occurred within each transect. Each transect

was analysed separately using floristic similarity matrices based on the contiguous 20-m subunits. Chronological clustering has a spatial constraint, which means that only adjacent subunits (or groups of subunits) can be grouped. A permutation test is performed at each step of the agglomeration procedure to test whether two groups whose fusion is proposed by the algorithm should be joined. The null hypothesis is that the two groups are not significantly different in their floristic composition and hence should be fused (Legendre *et al.* 1985; Legendre & Legendre 1998). Two rejection levels of the null hypothesis were compared:  $\alpha = 0.05$  and  $\alpha = 0.01$ . When the analysis recognizes a break, the probability that the two groups separated by the break actually belong to the same random population is smaller than the value of  $\alpha$ . Thus, breaks in the output indicate sites where significant floristic change occurs, and the intervening sections are internally more homogeneous.

The transect sections produced by chronological clustering, from all three transects, were then grouped by hierarchical clustering (shown at the bottom of Fig. 2) in order to visualize their floristic interrelationships. This was done for the sections produced using the Steinhaus index at  $\alpha = 0.01$  only, because this index of floristic similarity and  $\alpha$ -level resulted in the fewest and longest sections, which serve to generalize site-wide patterns in floristic composition. For the hierarchical clustering the Steinhaus index was re-calculated from the average species abundances per 100 m<sup>2</sup> in each transect section instead of their total abundances, to account for differences in length among sections.

#### *Ordination*

To compare floristic similarity patterns among sample sites with those based on quantitative soil chemistry, principal coordinates analyses (PCoAs) were carried out separately on the floristic and soil chemical data. The floristic ordinations were based on the Sørensen and Steinhaus indices. The environmental ordination was based on a Euclidean distance matrix combining all the soil nutrients (C, P, K, Ca, Mg, Fe, Mn, Al).

#### *Matrix correlations*

Mantel tests of matrix correspondence were then carried out to test for correlations between all the floristic, environmental and geographical distance matrices. These provided an overview of which of the measured variables could potentially explain the observed variation in floristic differences between plots. The standardized form of the Mantel statistic was used, which equals the Pearson correlation coefficient computed between the cell values of two resemblance matrices (Legendre & Legendre 1998). The Mantel test is especially appropriate for multivariate data because it allows several variables to be combined before the correlation is calculated, whereas a linear Pearson's correlation

can only be computed between two variables at a time. When interpreting Mantel test results, it is important to keep in mind that the correlations are those of the distances and not of the original variables; when both kinds of correlations are computed on the same data, the former are often considerably lower than the latter, although their statistical significance is usually similar (Legendre 2000). The analyses were repeated on five different data sets in order to understand better how the results are influenced by the length and type of environmental gradients included: (i) all transects ( $n = 112$  subunits); (ii) upland soils only (i.e. subunits on old alluvial and residual soils,  $n = 87$ ); (iii) transect 1 ( $n = 34$ ); (iv) transect 2 ( $n = 37$ ); and (v) transect 3 ( $n = 41$ ). Where canopy openness or Fe data were missing, the number of subunits included was slightly fewer than indicated here.

Simple Mantel tests were first carried out between floristic and environmental distance matrices, using the distance matrices based on one environmental variable at a time. Those environmental variables that yielded a significant matrix correlation with floristic distances in each of the three transects separately were then used to calculate a combined environmental distance matrix for each of the five data sets.

The second set of Mantel tests was aimed at assessing the relative strengths of the correlations between floristic and geographical vs. environmental distances, with the latter represented by the combined environmental distance matrices. To do this, both simple and partial Mantel tests were carried out. First, the correlations between floristic distances and environmental vs. geographical distances as such were compared. Secondly, we assessed how much the floristic-environmental correlations decreased when the effect of geographical distance was partialled out, and how much the floristic-geographical correlations decreased when the effect of environmental distance was partialled out.

The statistical significances of all Mantel correlations were assessed by Monte Carlo permutation tests with 999 permutations, which allows the statistical significance of the correlation coefficient to be tested at the  $P < 0.001$  level for each individual correlation.

#### *Multiple regression on distance matrices*

Multiple regressions on distance matrices were carried out in order to quantify the proportion of variation in floristic distances between sites that can be explained by their intervening geographical distance and by differences in the measured environmental variables. Multiple regressions allow a more accurate representation of environmental effects than Mantel tests, as each environmental variable is entered into the analysis independently, and the final solution uses the best combination of those variables. Thus, each environmental variable selected has a significant contribution to explaining the variance in the floristic distance matrix. In Mantel tests, all the environmental variables included

in a test are combined in a single distance matrix, which leads to a loss of information.

The reasoning follows that of Duivenvoorden *et al.* (2002) and Tuomisto, Ruokolainen & Yli-Halla (2003), who adapted the idea of variation partitioning, originally introduced by Borcard *et al.* (1992) in the context of canonical correspondence analysis, to the distance matrix context. The purpose of the multiple regression analysis is to partition the variation in the floristic distances between site pairs into four components: (i) variation explained by geographical distances between sites alone; (ii) variation jointly explained by environmental and geographical distances; (iii) variation explained by environmental distances alone; and (iv) variation that is not explained by either geographical distances or the measured environmental distances. The analysis was carried out separately on two data sets: (i) all transects ( $n = 102$ ); and (ii) upland soil subunits only ( $n = 82$ ). With both data sets, the analyses were run twice to compare the results obtained with presence-absence data (Sørensen index) and abundance data (Steinhaus index).

To select those environmental distance matrices with a statistically significant ( $P < 0.05$  after Bonferroni correction) contribution to explaining the variation in floristic distances, multiple regression with backward elimination was carried out using all the distance matrices based on a single environmental variable as independent matrices (following Legendre *et al.* 1994). The significance levels of the regression parameters ( $R^2$  and the partial regression coefficients) were obtained by Monte Carlo permutation with 999 permutations. Three further multiple regression analyses were then carried out. The independent distance matrices in these analyses were: (i) the geographical distance matrix; (ii) the environmental distance matrices selected by backward elimination; and (iii) all of these together. The  $R$ -squares of these regressions ( $R_G^2$ ,  $R_E^2$  and  $R_T^2$ ) were used to calculate the four components of the variation partitioning, i.e. the variation in floristic differences best explained by: geographical distances alone =  $R_T^2 - R_E^2$ ; geographical and environmental distances =  $R_G^2 + R_E^2 - R_T^2$ ; environmental distances alone =  $R_T^2 - R_G^2$ , or left unexplained =  $1 - R_T^2$ .

The regression analyses were carried out using the program PERMUTE v.3.4 and the other analyses using the R Package v.4.0 (Casgrain & Legendre 2001). Both programs are freely available from the web site <http://www.bio.umontreal.ca/legendre/indexEnglish.html>

#### *Distance decay in floristic similarity*

Finally, we calculated a linear and a spline regression of floristic similarity (Sørensen index) against the logarithm of geographical distance to check visually for evidence of dispersal limitation (using the program JMP vs. 3.1, SAS Institute Inc.). This was done using the original  $5 \times 5$  m subunits to allow us to look in detail for evidence of declining similarity over very short spatial scales.

## Results

### GENERAL FLORISTIC OBSERVATIONS

Within the total area surveyed (2.92 ha), we recorded 27 551 pteridophyte individuals, giving an average density of 9433 individuals per hectare. A total of 96 species and 39 genera were represented in the data set (see Table S1).

The four most species-rich genera were *Asplenium* (nine spp.), *Elaphoglossum* (eight spp.), *Thelypteris* (seven spp.) and *Tectaria* (six spp.). *Polypodium* and *Diplazium* shared fifth place, with five species each. However, the most frequently encountered genera reflect the abundances of the five most common species, these being *Danaea wendlandii* (27% of individuals), *Salpichlaena* sp. 1 (*aff. volubilis*) (15%), *Polybotrya villosula* (9%), *Lomariopsis vestita* (8%) and *Adiantum obliquum* s.l. (6%).

The majority (59%) of the species encountered in the survey were classed as mostly terrestrial, being tree ferns (Cyatheaaceae, 4%), climbers on the lower parts of tree trunks (4%) or other usually ground-rooting species (51%). Thirty-seven per cent of species were usually epiphytic, and 4% were usually epilithic (typically growing on rocks at stream banks). Of the individuals encountered, 88% belonged to species that we classed as mostly terrestrial (tree ferns 2%, low climbers 12%, other terrestrial 74%), whereas just 10% of individuals belonged to the typically epiphytic species. Typically epilithic species constituted only 2% of the individuals encountered, which reflects the extremely small area of rocky streamside habitat covered during the survey.

For both the full data set and the upland soil subset there was considerable variation in the floristic similarities computed between the 25-m transect subunits. In the full data set some subunits did not share any species at all (both Sørensen and Steinhaus similarity index values = 0), whereas subunits in the upland soils always shared at least one species (minimum value for the Sørensen index = 0.09, and for the Steinhaus index = 0.01). Other subunits had exactly the same species present (Sørensen index = 1), albeit in slightly different abundances (Steinhaus index = 0.93). Mean floristic similarity between subunits for the upland soil subset was somewhat higher than for the full data set, both when the Sørensen index was used (0.53 vs. 0.45), and when the Steinhaus index was used (0.35 vs. 0.28).

### CLUSTERING

Spatially constrained clustering of the floristic data (20-m subunits) resulted in several breaks along each transect, indicating statistically significant changes in species composition at those locations (Fig. 2). Some of the breaks corresponded to the transition from higher ground to valley bottoms, indicating that distinct floristic assemblages are found close to streams and in swamps as compared with the higher ground. However, there were also changes in the flora within the well-drained

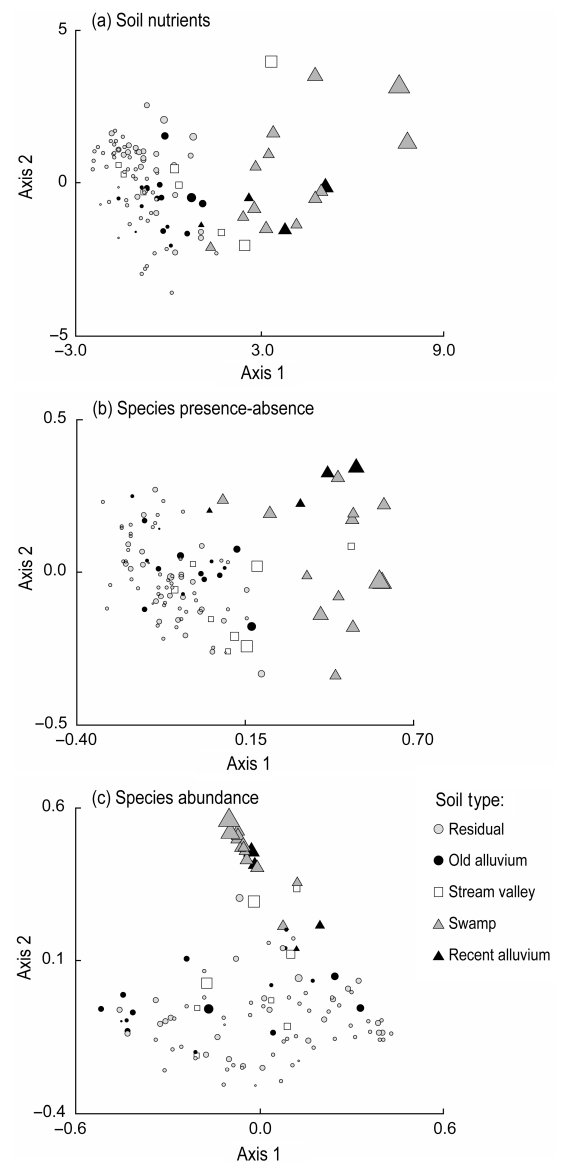
higher ground, as well as within large stream valleys and swamps. On the whole, significant changes in floristic composition were detected slightly more frequently with species presence-absence data (Sørensen index) than with abundance data (Steinhaus index). Increasing the alpha level of the cluster fusion test from  $\alpha = 0.01$  to  $\alpha = 0.05$  resulted in a clear increase in the number of accepted breaks in floristic composition. This was especially the case in transect 2, which was divided into just four or five floristic segments at  $\alpha = 0.01$  (depending on whether the Sørensen or Steinhaus index was used), but into 10 or 11 segments at  $\alpha = 0.05$ . The degree of correspondence between the breaks obtained with the two floristic similarity indices and  $\alpha$ -levels was variable. Three breaks corresponding to a stream valley and one to a shallow hill top position were recognized at the same location no matter which similarity index or alpha level was used. These breaks are likely to represent rather sharp changes in floristic composition, whereas those breaks that are more labile probably represent more gradual floristic changes.

The hierarchical clustering results illustrate broad floristic patterns among the study transects (Fig. 2). We decided to recognize five floristic classes, as there were five clear branches in the clustering dendrogram. One of these classes was only found in the extensive swamp of transect 2. Two of the classes were restricted to transect 1, where one of them was found in two sites of low topographic position (either on swamp soil or on recent alluvial soil), and the other in three sites with either recent or old alluvial soils. The two remaining classes were represented in all three transects, indicating that floristically similar sites occurred in different parts of the study area. These classes were mainly found on the residual and old alluvial soils.

#### ORDINATION

Ordination based on soil chemical data (Fig. 3a) shows that the sample sites covered a clear edaphic gradient from the residual soils at one extreme to the swamp soils at the other. The alluvial soils are intermediate, with old alluvium overlapping with residual soils, and recent alluvium overlapping with swamp soils. The stream valley soils are too widely scattered in the ordination diagram to be characterized as a uniform group along this gradient. The first ordination axis is most strongly defined by soil Ca, Mg, Al and K, as well as pH. All these variables correlate with the first axis with  $|r| > 0.7$ ; the correlations are positive with the exception of Al. Soil Mn and P also correlate positively ( $r > 0.6$ ) with the first axis. Consequently, axis 1 can be interpreted as a soil nutrient gradient, which is also clearly related to soil moisture and drainage conditions. Axis 2 correlates strongly with soil C ( $r = 0.76$ ). This gradient is apparent in both the upland and swamp soil types and is related to variation in soil organic matter.

The same basic pattern is seen in the floristic ordinations, whether based on presence-absence data (Fig. 3b)



**Fig. 3** Ordination diagrams (PCoA) where distances between sites were measured in terms of (a) several soil nutrients (Euclidean distance), (b) floristic composition using presence-absence data (Sørensen index) and (c) floristic composition using abundance data (Steinhaus index). The shaded symbols indicate which of the five qualitative soil classes each site belongs to. The height of each symbol is proportional to the natural logarithm of soil Mg content at that site.

or abundance data (Fig. 3c), although the ordination using presence-absence data more closely corresponds to the soil ordination. In the presence-absence ordination (Fig. 3b), axis 1 seems to be related to soil drainage. It correlates strongly and positively with soil Ca and Mg, as well as pH ( $r > 0.7$ ), and negatively with soil Al ( $r = -0.70$ ). Axis 2 does not correlate well with any of the soil variables (the strongest correlation is with soil Fe,  $r = -0.30$ ). In the ordination based on species abundance data (Fig. 3c), the strongest correlation of axis 1 is with soil Mn ( $r = 0.25$ ), and those of axis 2 are with soil Mg, Ca, pH and Al ( $|r| > 0.6$ ). Axis 2 seems clearly related to soil drainage conditions, whereas



axis 1 is probably determined by factors other than soil variation.

The swamp soil sites are very dispersed in the ordinations based on soil data and species presence-absence, but are very tightly associated on the first two axes of the ordination based on species abundance. However, in the latter ordination they are separated along the third ordination axis (not shown). Thus, they are rather variable in terms of both their soils and floristic composition. The two upland soil types appear to overlap relatively more in their soil nutrient levels. Floristic composition within them is, however, extremely variable when species abundances are taken into account.

#### MATRIX CORRELATIONS

The results of the Mantel tests between the floristic distances and geographical and environmental distances are summarized in Table 2. Generally, the correlations were somewhat higher when the Sørensen index (presence-absence data) was used to measure floristic distance than when the Steinhaus index (abundance data) was used, especially in the analyses including all transects. The correlations obtained with the different environmental variables varied widely. Soil pH, for example, yielded high and statistically significant matrix correlations with all floristic data sets, whereas the matrix correlations with aspect were always low and, with one exception, non-significant.

The results also differed clearly among the different data sets. In the analyses of the full data set (all transects together), highly significant correlations were found between floristic distances and most of the environmental distances, including all those based on quantitative soil properties. The correlation coefficients were especially high with distances based on soil Mg, Ca, Al, pH and the qualitative variable 'Swamp'. On the other hand, the correlations with aspect and the old alluvial and stream valley soil types were low and non-significant.

In analyses of the upland soil sites only (old alluvial and residual soils), the correlations between floristic and environmental distances were generally weaker than for the full data set. This was especially so when environmental distance was based on the soil variables. The canopy openness index was an exception, and gave the highest correlations with both floristic presence-absence data ( $r = 0.29$ ) and abundance data ( $r = 0.25$ ).

The results of the correlation tests were somewhat different when the data from each transect were analysed separately. Correlations between floristic and environmental distances were generally strongest in transect 2, and weakest in transect 3 (Table 2). This is likely to be because transect 2 covered the longest environmental gradient of all, running from a sizeable stretch of swamp to upland ridges, whereas transect 3 was the least heterogeneous in soil nutrient and drainage conditions (Figs 1 and 2).

Five of the environmental variables yielded a significant Mantel correlation with floristic distance in all

three transects separately: soil Ca, Mg and Al content, soil pH and topographic position (Table 2). These variables were included in a combined environmental distance matrix.

The correlations between the floristic and environmental distances, using the combined environmental distance matrix, were little affected by partialling out the effect of geographical distance (Table 2). The same was true of the correlations between floristic distances and geographical distances, except in the case of transect 2, where partialling out the effect of environmental distance resulted in a somewhat reduced correlation.

#### MULTIPLE REGRESSION ON DISTANCE MATRICES

The multiple regression results are illustrated in Table 3. The final environmental models included between two and six environmental distance matrices. Both qualitative and quantitative data were incorporated into the models. All the models contained data on canopy openness and topographic position, and all but one contained quantitative soil data in the form of soil Ca and, in one case, soil pH. Qualitative information on soil type (swamp and either recent alluvial or residual soils) was also included in the all sites models, but the qualitative distinction between the two upland soil types (residual vs. old alluvium) was not a significant explanatory factor within the upland soils.

The variation in floristic distances between sites was partitioned into components explained by the selected environmental variables alone, the ln-transformed geographical distances alone, and both kinds of distances together. For the full data set, a total of either 50% (presence-absence data) or 33% (abundance data) of the variation in floristic distances was explained overall (Fig. 4a,c). For the subset of sites on upland soils, the total proportion of explainable variation was either 18% (presence-absence data) or 10% (abundance data; Fig. 4b,d). Most of the explanatory power was purely environmental; the geographical distances and the mixed geographical and environmental fractions together explained just 0.7–1.1% of the total variation in floristic distances. The fraction uniquely explained by the geographical distances between sites was marginally larger in the upland soils (0.4–0.5%) than in the full data set (0.2–0.3%) with both floristic distance measures.

#### DISTANCE DECAY IN FLORISTIC SIMILARITY

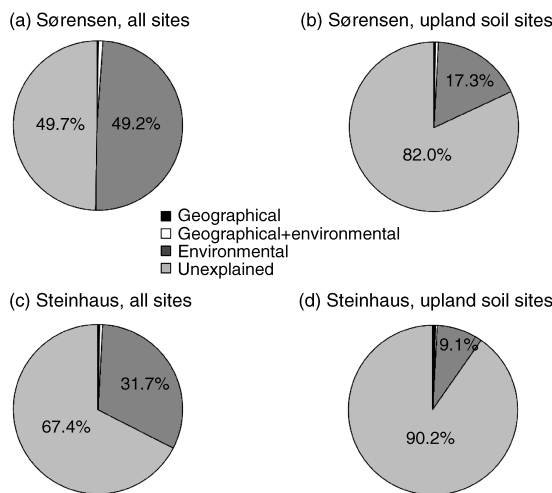
Fitting a spline function showed mean floristic similarity to decline very rapidly over short distances (< 100 m), then to decline much more gently and slightly irregularly up to the *c.* 3 km maximum distance covered in this study (Fig. 5). The initial decline indicated by the spline function was much steeper than the best linear fit to the data. Both the linear and spline regressions provided a poor fit to the data, explaining just 0.5% and 1.4% of the variation in floristic similarity with distance, respectively.

**Table 2** Matrix correlations (Mantel *r*) of floristic distances, based on the Sørensen or Steinhaus index, with the environmental or geographical distances between sample sites. The statistical significance of the correlations was determined by Monte Carlo permutation (999 permutations of the floristic distance matrix. \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001)

Environmental or geographical distance matrix		Floristic distance matrix														
		All transects			Upland soils			Transect 1			Transect 2			Transect 3		
		Sør	Site	Ste	Sør	Site	Ste	Sør	Site	Ste	Sør	Site	Ste	Sør	Site	Ste
<b>Simple Mantel tests</b>																
Soil chemistry (pH + all nutrients)		0.60***	0.52***	0.18**	0.23**	0.59***	0.46***	0.60***	0.62***	0.62***	0.62***	0.62***	0.62***	0.62***	0.30**	0.29**
pH		0.56***	0.46***	0.14*	0.11*	0.57***	0.44***	0.62***	0.62***	0.62***	0.62***	0.62***	0.62***	0.17*	0.21*	
Ca		0.61***	0.51***	0.09	0.19**	0.57***	0.37***	0.65***	0.65***	0.65***	0.65***	0.65***	0.65***	0.28**	0.27**	
K		0.28***	0.26***	0.14*	0.14*	0.13	0.09	0.34**	0.41***	0.41***	0.41***	0.41***	0.41***	0.23*	0.12	
Mg		0.64***	0.54***	0.10	0.14*	0.63***	0.41***	0.68***	0.68***	0.68***	0.68***	0.68***	0.68***	0.24**	0.22**	
Mn		0.15***	0.10**	0.02	0.10*	-0.06	-0.01	0.22**	0.20**	0.20**	0.20**	0.20**	0.20**	0.14*	0.08	
Fe		0.25***	0.20***	0.17**	0.17**	0.30***	0.25***	0.13*	0.19**	0.19**	0.19**	0.19**	0.19**	0.12	0.11	
Al		0.56***	0.45***	0.14**	0.15*	0.61***	0.47***	0.49***	0.49***	0.49***	0.49***	0.49***	0.49***	0.19*	0.24**	
C		0.23**	0.31***	0.12*	0.06	0.15	0.23**	0.32**	0.32**	0.32**	0.32**	0.32**	0.32**	0.10	0.06	
P		0.31***	0.23***	-0.01	-0.03	0.38***	0.24***	0.34***	0.34***	0.34***	0.34***	0.34***	0.34***	-0.04	0.02	
Soil type (old alluvium)		-0.08	-0.01	0.13*	0.06	0.07*	0.15**	NA	NA	NA	NA	NA	NA	NA	NA	
Soil type (recent alluvium)		0.31***	0.17*	NA	NA	0.45***	0.22*	NA	NA	NA	NA	NA	NA	NA	NA	
Soil type (residual)		0.30***	0.28***	0.13*	0.06	0.01	0.07	0.52***	0.50***	0.50***	0.50***	0.50***	0.50***	0.16	0.14	
Soil type (stream valley)		0.06	0.06	NA	NA	0.26*	0.15	-0.11	0.01	0.01	0.01	0.01	0.01	0.16	0.14	
Soil type (swamp)		0.53***	0.48***	NA	NA	0.19	0.24*	0.74***	0.69***	0.69***	0.69***	0.69***	0.69***	NA	NA	
Canopy openness		0.22**	0.22**	0.25**	0.29**	0.23*	0.23*	0.29**	0.34***	0.34***	0.34***	0.34***	0.34***	0.23*	0.13	
All topographic variables		0.28***	0.25***	0.15***	0.16**	0.20*	0.14*	0.39***	0.39***	0.39***	0.39***	0.39***	0.39***	0.25***	0.25**	
Slope		0.12*	0.14**	0.07	0.02	-0.01	0.08	0.26**	0.26**	0.26**	0.26**	0.26**	0.26**	0.11	0.10	
Sine (aspect)		-0.02	-0.02	0.02	0.04	-0.04	-0.04	-0.08*	-0.08*	-0.08*	-0.08*	-0.08*	-0.08*	0.08	0.08	
Cosine (aspect)		0.01	0.00	-0.01	-0.03	-0.04	-0.02	0.03	0.04	0.04	0.04	0.04	0.04	-0.07	-0.07	
(Relative) Elevation		0.19**	0.14**	0.07	0.09	0.24*	0.13	0.45***	0.40***	0.40***	0.40***	0.40***	0.40***	0.22*	0.22*	
Topographic position		0.35***	0.30***	0.17**	0.21***	0.31**	0.20**	0.41***	0.41***	0.41***	0.41***	0.41***	0.41***	0.20**	0.21**	
<b>Partial Mantel tests</b>																
Combined environmental		0.66***	0.56***	0.25***	0.31***	0.65***	0.45***	0.66***	0.66***	0.66***	0.66***	0.66***	0.66***	0.39***	0.42***	
Geographical		0.08***	0.08***	0.08***	0.08***	0.04	0.12**	0.39***	0.39***	0.39***	0.39***	0.39***	0.39***	0.09*	0.03	
Combined environmental – Geographical		0.66***	0.55***	0.25***	0.31***	0.65***	0.45***	0.62***	0.62***	0.62***	0.62***	0.62***	0.62***	0.39***	0.42***	
Geographical – Combined environmental		0.07***	0.07***	0.06**	0.06**	0.05	0.13**	0.24***	0.24***	0.24***	0.24***	0.24***	0.24***	0.08*	0.03	

**Table 3** Results of multiple regression analyses showing the geographical and environmental distance matrices with a statistically significant ( $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$  after Bonferroni correction) contribution to explaining variation in floristic distances. Floristic distance was based on either the Sørensen or the Steinhaus index. The analyses were run on all sample sites together vs. on upland soil sites only. The  $R^2$  values listed formed the basis of the variation partitioning procedure presented in the 'Materials and methods' section. The results of the variation partitioning are illustrated in Fig. 4

$R^2$	Explanatory distance matrices
All sites	
Sørensen	
$R_G^2$ 0.011***	Geographical
$R_E^2$ 0.500***	Environmental (Canopy openness + Recent alluvial soil + Swamp soil + pH + Ca + Topographic position)
$R_T^2$ 0.503***	Environmental + Geographical
Steinhaus	
$R_G^2$ 0.009***	Geographical
$R_E^2$ 0.324***	Environmental (Canopy openness + Residual soil + Swamp soil + Ca + Topographic position)
$R_T^2$ 0.326***	Environmental + Geographical
Upland soils	
Sørensen	
$R_G^2$ 0.007***	Geographical
$R_E^2$ 0.176***	Environmental (Canopy openness + Ca + Topographic position)
$R_T^2$ 0.180***	Environmental + Geographical
Steinhaus	
$R_G^2$ 0.007**	Geographical
$R_E^2$ 0.092***	Environmental (Canopy openness + Topographic position)
$R_T^2$ 0.098***	Environmental + Geographical

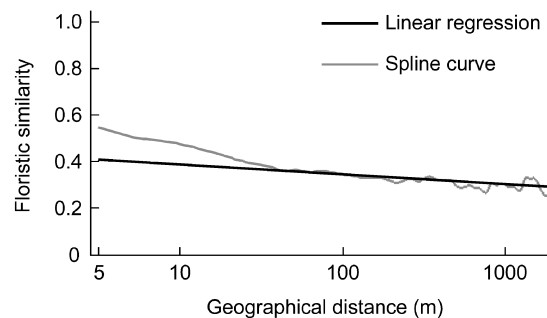


**Fig. 4** Diagrams showing the proportion of variation in floristic distances that can be explained by geographical and/or environmental distances. Floristic distances were based on either the Sørensen index (presence-absence data; a, b) or the Steinhaus index (abundance data; c, d). The analyses were run separately on the full data set (a, c) and the upland soil subset (b, d).

## Discussion

### RELATIVE ROLES OF DISPERSAL LIMITATION AND ENVIRONMENTAL VARIATION

In all the analyses, environmental differences were found to be far stronger predictors of site-wide changes in floristic differences than the geographical distances between sample plots. In transect 2, geographical distance effects were apparently stronger than in the other data



**Fig. 5** The relationship between floristic similarity (Sørensen index) and log-geographical distance based on the  $5 \times 5$  m subunits ( $n = 1141$ ). The spline curve ( $R^2 = 0.014$ ) illustrates how mean floristic similarity changes with distance. The linear regression line ( $R^2 = 0.005$ ,  $P < 0.001$ ) approximates the expectation for the relationship between floristic similarity and distance from neutral theory.

sets, which may be because the ends of this transect corresponded to two habitat extremes, whose effect was only partly accounted for by our environmental data. Otherwise the environmental and geographical effects on floristic composition appeared to be largely independent of each other, as partialling out the effect of one had little effect on the results for the other in the Mantel tests.

Our study plots covered an environmental gradient running from richer, poorly drained swamp and recent alluvial soils to poorer, well-drained upland soils. Floristic variation was seen to correspond clearly to this edaphic-topographic gradient. When based on species presence-absence data, variation partitioning showed

that the measured environmental differences alone explained 49% of the variation in floristic distances site-wide, and 17% when the environmental gradient was curtailed by including upland sites only. For species abundance data, the proportion of explained variation was lower: environmental differences explained 32% of site-wide variation in floristic distances and 9% in the upland soils. Geographical distances alone explained 0.5% or less of the variation in floristic distances in all cases. Restricting our analyses to the more environmentally homogeneous upland soils did not significantly strengthen the relationship between geographical distance and floristic variation (cf. Duque *et al.* 2002).

These findings are contrary to the dispersal assembly view of plant communities (Hubbell 2001), according to which floristic similarity between sites decreases as a monotonic function of the geographical distance between those sites. The results provide more support for the niche-assembly view, although half or more of the variation in floristic distances could not be accounted for by variation in either the measured environmental differences or geographical distances.

We reasoned in advance that as small areas harbour less environmental variation than large areas, we should find a smaller role for environmental variation at the mesoscale than has been found earlier at broader spatial scales. Similarly, as the effect of dispersal limitation is related to the logarithm of geographical distance (Hubbell 2001), per unit distance its effects should be more noticeable at small than at large geographical distances. These expectations were only partly borne out. Four studies in western Amazonia have recently estimated the relative roles of geographical and environmental distances at landscape to regional scales by regression analysis, and in these the unique contribution of geographical distances was 20–40% in the broadest-scale studies (Tuomisto, Ruokolainen & Yli-Halla 2003; Vormisto, Svenning, Hall & Balslev 2004) and 1.4–10% in the intermediate-scale studies (Phillips *et al.* 2003; Tuomisto, Ruokolainen, Aguilar & Sarmiento 2003). The contribution of less than 1% in the present study was hence much lower than expected. As to the unique contribution of environmental distances, the values reported in the four Amazonian studies ranged from 8% to 40%. This is similar to the range found in the present mesoscale study. Consequently, the only result that conformed to our prior expectations was that the contribution of environmental distances in the present study decreased when the environmental gradient was truncated by excluding the plots representing swamp and alluvial soils from the analyses. In comparing these studies, it should be acknowledged that they incorporated different aspects of environmental variation. All included some soil variables (except Tuomisto, Ruokolainen, Aguilar & Sarmiento 2003, who used satellite imagery as a surrogate for environmental data), but our study also included canopy openness and detailed topographic data, which probably allowed us to explain slightly more floristic variation than if we

had analysed soils alone. Phillips *et al.* (2003) designed their landscape-scale sampling of trees to minimize the correlation between environmental and geographical distances, but nevertheless found geographical distance to explain 10% of the variation in floristic differences. This is a much larger proportion than in our study.

We did find a rapid decline in mean floristic similarity at very short distances. However, mean floristic similarity at these short distances, as described by the spline curve, diverged considerably from a linear relationship with log-geographical distance (Fig. 5). Thus, this decline appears too steep to fit the predictions of neutral theory. At distances greater than 100 m the decline was much gentler and slightly irregular. Apart from these small irregularities, which are likely to be partly random noise and partly related to changes in environmental conditions with distance (cf. Tuomisto, Ruokolainen & Yli-Halla 2003), mean floristic similarity at these distances did not depart much from a linear relationship, which could be an indication of some dispersal limitation that fits neutral theory.

Spatial autocorrelation in floristic similarities over short distances is related to both biotic processes, such as a higher probability of dispersal to nearby sites, and to spatial autocorrelation in habitat variables (Legendre 1993). Environmental conditions at our study site are highly autocorrelated at the same local scales at which we observed steeply declining floristic similarity. A move of 100 m often represents a change in topographic position (Fig. 2), and distances much shorter than this represent changes from gap to understorey conditions (Sanford *et al.* 1986; M. M. Jones *et al.*, personal observations). Consequently, the decline we observed over short distances is confounded by environmental effects and cannot be attributed to dispersal limitation alone. Condit *et al.* (2002) came to a similar conclusion regarding the rapid decline in floristic similarity they observed at short distances (also to *c.* 100 m) in a study on canopy trees in Panama and Ecuador. They concluded that the trees were more aggregated than expected under Hubbell's neutral theory (Hubbell 2001) over these short distances, possibly due to light gap effects. In a temperate forest study carried out within an area of 10 km<sup>2</sup>, Gilbert & Lechowicz (2004) found that distance did not predict species turnover in six plant groups, and made only a slight contribution to turnover in a seventh. The genus *Carex*, for which floristic similarity fluctuated with distance, was chosen to illustrate deviations at their site from neutral expectations. They did not observe a decline in similarity even at the shortest distances included in their study, but this may have been because the minimum distance between any pair of sample plots was 135 m.

In conclusion, dispersal limitation is of little importance as a determinant of variation in pteridophyte composition at the mesoscale, but may be of increasing importance, alongside environmental differences, at landscape to regional scales. At La Selva, most fern species are continually fertile and produce large quantities of spores

(although spore shadows are generally leptokurtic, Wolf *et al.* 2001), whereas only a few species appear reliant upon vegetative reproduction (M. M. Jones *et al.*, personal observations). This, in conjunction with relatively small-scale environmental variation and strong environmental effects on the probability of germination and survival, is likely to account for the small effect of geographical distance we observed at scales of up to a few km.

#### RELATIVE ROLES OF DIFFERENT ENVIRONMENTAL VARIABLES

Floristic patterns were most strongly associated with the main drainage (topographic) and soil fertility gradient. This confirms earlier results for large trees and palms at the same site (Clark *et al.* 1995, 1998, 1999). The patterns were the same whether floristic similarity was calculated from species presence-absence or abundance data, although the environmental data were able to explain a higher proportion of the variation in presence-absence than in abundance data. In one earlier study (Tuomisto, Poulsen, Ruokolainen, Moran, Quintana & Cañas 2003), floristic differences based on abundance data were better predicted than differences based on presence-absence data, but in another, the results varied from between localities (Tuomisto & Poulsen 2000). The reasons for these differences are not clear but may have to do with the scales at which species abundances vary over the landscape relative to the size of the sampling unit.

Qualitative soil data were useful for predicting floristic differences when the full environmental gradient was modelled, specifically to distinguish the swamp and recent alluvial or residual soils from the other soil types. As can be seen from the ordination diagram (Fig. 3), these soil types represent the extremes of the main drainage and soil fertility gradient. The importance of soil drainage and topography for floristic variation is generally acknowledged (e.g. Lieberman *et al.* 1985; Svenning 1999; Webb & Peart 2000; Harms *et al.* 2001; Pélissier *et al.* 2001). Two of the identified broad floristic groups were associated with the more poorly drained recent alluvial and swamp soils. The other three groups included a combination of upland and stream valley soils. However, there were also floristically distinct areas within these broad groups, which did not clearly correspond to changes in soil type or topographic position.

Within the restricted environmental gradient of the upland soils, the qualitative distinction between old alluvial and residual soils was not useful for predicting floristic differences. Instead, as over the full environmental gradient, information on topographic position was informative. Streamside sites are floristically variable and distinct from other topographic positions. In addition, valleys and lower slopes are usually floristically distinct from ridges and upper slopes (see clustering results, Fig. 2).

An earlier paper by Lieberman *et al.* (1985) suggested that altitude was the main factor behind a continuous floristic gradient for large woody plants at our study

site. Our results suggest that relative topographic position is much more important than elevation *per se*. La Selva lies at the base of a mountain, so there is a c. 100 m rise in elevation over the site. Our transects were positioned almost at right angles to this gradient at different average elevations, so stream valley sites in the highest transect (transect 3) were found at elevations equivalent to ridge tops in the lowest transect (transect 1). Within each transect individually, elevation was a good proxy for topographic position and hence for environmental change among sample plots, and yielded significant correlations in the Mantel tests. However, site-wide there was no correlation between elevational distance and floristic distance, unless the relative elevations within each transect were used. Even then, the correlations with topographic position were clearly higher.

Quantitative soil chemistry was incorporated into all but one of the predictive models of floristic distances (Table 3), in the form of soil Ca and, in one case, pH. At the site-wide scale, the highest correlations with floristic distances were obtained with the distance matrices based on quantitative soil chemical data (soil Mg, Al, Ca or pH) (Table 2). As was suggested by Clark *et al.* (1999), quantitative soil data were found to provide more ecologically useful information for this site than qualitative data. Strong relationships between floristic distances and distances based on soil cation levels have also been observed at wider spatial and environmental scales in western Amazonia, both for pteridophytes and other plant groups (e.g. Ruokolainen *et al.* 1997; Phillips *et al.* 2003; Tuomisto, Poulsen, Ruokolainen, Moran, Quintana & Cañas 2003; Tuomisto, Ruokolainen, Aguilar & Sarmiento 2003; Tuomisto, Ruokolainen & Ylihalla 2003; Vormisto, Svenning, Hall & Balslev 2004).

Canopy openness was moderately correlated with floristic differences in all data sets, and was a significant predictor of floristic differences in all the regression models. Pteridophyte responses to gap-formation are multiple: the resulting disturbance and environmental change have positive effects on some species and negative effects on others (M. M. Jones *et al.*, personal observations). In addition, gap formation changed the suite of species sampled in a plot because a few species more typical of the canopy were found on fallen trees. However, the inclusion of canopy openness in the models cannot be interpreted unambiguously as a gap effect, because many streamside and some swamps also have relatively open canopies (Table 1).

The forest at La Selva is very dynamic, with a high tree stem turnover (2.3–2.8% year<sup>-1</sup> for canopy level trees; Clark *et al.* 2004) and frequency of gap formation (Sanford *et al.* 1986). This may partly account for the high proportion of variation in floristic distances (50–90%) that was unexplained by either geographical distance or our large environmental data set, especially in the upland soil subset. Forest successional dynamics have an obvious impact on pteridophyte communities at local scales, and it seems likely that the shorter the soil or drainage gradient considered, the more dominant

are species responses to gap formation and subsequent succession in contributing to floristic variation. Consequently, our inability to include information on past light and disturbance environments limits the degree to which we can predict floristic patterns at the mesoscale.

### Acknowledgements

We thank Jens Mackensen and Edzo Veldkamp for help with the soil sampling and chemical analyses, Rigoberto Gonzalez for assistance during the fern inventory and the La Selva Biological Station of the Organization for Tropical Studies for logistic support. Roger Jones, Detlev Kelm, Kati Salovaara and two anonymous reviewers provided useful comments on previous drafts. The work was funded by the Academy of Finland and the Andrew W. Mellon Foundation. Research permits were kindly granted by the Costa Rican Ministry of the Environment and Energy (MINAE).

### References

Ashton, P.S. (1969) Speciation among tropical forest trees: some deductions in the light of recent evidence. *Speciation in Tropical Environments* (ed. R.H. Lowe-McConnell), pp. 155–196. Academic Press, London.

Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology*, **73**, 1045–1055.

Brown, N., Jennings, S., Wheeler, P. & Nabe-Nielsen, J. (2000) An improved method for the rapid assessment of forest understorey light environments. *Journal of Applied Ecology*, **37**, 1044–1053.

Casgrain, P. & Legendre, P. (2001) *The R Package for Multivariate and Spatial Analysis, Version 4.0*. Département de Sciences Biologiques, Université de Montréal, Montréal, Canada.

Clark, D.B., Clark, D.A. & Read, J.M. (1998) Edaphic variation and the mesoscale distribution of tree species in a neotropical rain forest. *Journal of Ecology*, **86**, 101–112.

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., Palmer, M.W. & Clark, D.A. (1999) Edaphic factors and the landscape-scale distributions of tropical rain forest trees. *Ecology*, **80**, 2662–2675.

Clark, D.B., Soto Castro, C., Alfaro Alvarado, L.D. & Read, J. (2004) Quantifying mortality of tropical rain forest trees using high-spatial-resolution satellite data. *Ecology Letters*, **7**, 52–59.

Condit, R. (1996) Defining and mapping vegetation types in mega-diverse tropical forests. *Trends in Ecology and Evolution*, **11**, 4–5.

Condit, R., Pitman, N., Leigh, E.G., Chave, J., Terborgh, J., Foster, R. *et al.* (2002) Beta-diversity in tropical forest trees. *Science*, **295**, 666–669.

Denslow, J.S. (1987) Tropical rain forest gaps and tree species diversity. *Annual Review of Ecology and Systematics*, **18**, 431–451.

Duivenvoorden, J.F. (1995) Tree species composition and rain forest-environment relationships in the middle Caquetá area, Colombia, NW Amazonia. *Vegetatio*, **120**, 91–113.

Duivenvoorden, J.F., Svenning, J.-C. & Wright, S.J. (2002) Beta diversity in tropical forests. *Science*, **295**, 636–637.

Duque, A., Sánchez, M., Cavallier, J. & Duivenvoorden, J.E. (2002) Different floristic patterns in woody understorey

and canopy plants in Colombian Amazonia. *Journal of Tropical Ecology*, **18**, 499–525.

Fedorov, A.A. (1966) The structure of the tropical rain forest and speciation in the humid tropics. *Journal of Ecology*, **54**, 1–11.

Gentry, A.H. (1988) Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Annals of the Missouri Botanical Garden*, **75**, 1–34.

Gilbert, B. & Lechowicz, M.J. (2004) Neutrality, niches, and dispersal in a temperate forest understory. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 7651–7656.

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. *Journal of Ecology*, **89**, 947–959.

Hartshorn, G.S. & Hammel, B.E. (1994) Vegetation types and floristic patterns. *La Selva: Ecology and Natural History of a Neotropical Rain Forest* (eds L.A. Mc Dade, K.S. Bawa, H.A. Hespenheide & G.S. Hartshorn), pp. 118–157. University of Chicago Press, Chicago.

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. *Journal of Geodynamics*, **34**, 491–502.

Hubbell, S.P. (2001) *The Unified Theory of Biodiversity and Biogeography*. Monographs in Population Biology 32. Princeton University Press, Princeton.

Hubbell, S.P. & Foster, R.B. (1986) Biology, chance, and history and the structure of tropical rain forest tree communities. *Community Ecology* (eds J. Diamond & T.J. Case), pp. 314–329. Harper & Row, New York.

König, N. & Fortmann, H. (1996a) Probenvorbereitungs-, Untersuchungs- und Elementbestimmungs-Methoden des Umweltanalytik-Labors der Niedersächsischen Forstlichen Versuchsanstalt und des Zentrallabors II des Forschungszentrums Waldökosysteme. Teil 1: Elementbestimmungsmethoden A-M. *Berichte des Forschungszentrums Waldökosysteme, Reihe B, Bd. 46*.

König, N. & Fortmann, H. (1996b) Probenvorbereitungs-, Untersuchungs- und Elementbestimmungs-Methoden des Umweltanalytik-Labors der Niedersächsischen Forstlichen Versuchsanstalt und des Zentrallabors II des Forschungszentrums Waldökosysteme. Teil 2: Elementbestimmungsmethoden N-Z und Sammelanhänge. *Berichte des Forschungszentrums Waldökosysteme, Reihe B, Bd. 47*.

Legendre, P. (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology*, **74**, 1659–1673.

Legendre, P. (2000) Comparison of permutation methods for the partial correlation and partial Mantel tests. *Journal of Statistical Computation and Simulation*, **67**, 37–73.

Legendre, P., Dallot, S. & Legendre, L. (1985) Succession of species within a community: chronological clustering, with applications to marine and freshwater zooplankton. *American Naturalist*, **125**, 257–288.

Legendre, P., Lapointe, F.-J. & Casgrain, P. (1994) Modeling brain evolution from behavior: a permutational regression approach. *Evolution*, **48**, 1487–1499.

Legendre, P. & Legendre, L. (1998) *Numerical Ecology*, 2nd edn. Elsevier, Amsterdam.

Lieberman, M., Lieberman, D., Hartshorn, G.S. & Peralta, R. (1985) Small-scale altitudinal variation in lowland wet tropical forest vegetation. *Journal of Ecology*, **73**, 505–516.

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. *Journal of Tropical Ecology*, **11**, 161–177.

Pélissier, R., Dray, S. & Sabatier, D. (2001) Within-plot relationships between tree species occurrences and hydrological soil constraints: an example in French Guiana investigated through canonical correlation analysis. *Plant Ecology*, **162**, 143–156.

- Phillips, O.L., Núñez Vargas, P., Lorenzo Monteagudo, A., Peña Cruz, A., Chuspe Zans, M.-E., Galiano Sánchez, W. *et al.* (2003) Habitat association among Amazonian tree species: a landscape-scale approach. *Journal of Ecology*, **91**, 757–775.
- Pitman, N.C.A., Terborgh, J., Silman, M.R., Nuñez, V.P. (1999) Tree species distributions in an upper Amazonian forest. *Ecology*, **80**, 2651–2661.
- Pitman, N.C.A., Terborgh, J., Silman, M.R., Nuñez, V.P., Neill, D.A., Cerón, C.E. *et al.* (2001) Dominance and distribution of tree species in upper Amazonian terra firme forests. *Ecology*, **82**, 2101–2117.
- Potts, M.D., Ashton, P.S., Kaufman, L.S. & Plotkin, J.B. (2002) Habitat patterns in tropical rain forests: a comparison of 105 plots in northwest Borneo. *Ecology*, **83**, 2782–2797.
- Potts, M.D., Davies, S.J., Bossert, W.H., Tan, S. & Nur Supardi, M.N. (2004) Habitat heterogeneity and niche structure of trees in two tropical rain forests. *Oecologia*, **139**, 446–453.
- Prance, G.T. (1989) American tropical forests. *Ecosystems of the World 14B. Tropical Rain Forest Ecosystems: Biogeographical and Ecological Studies* (eds H. Lieth & M.J.A. Werger), pp. 99–132. Elsevier, Amsterdam, The Netherlands.
- Ruokolainen, K., Linna, A. & Tuomisto, H. (1997) Use of Melastomataceae and pteridophytes for revealing phylogeographical patterns in Amazonian rain forests. *Journal of Tropical Ecology*, **13**, 243–256.
- Ruokolainen, K. & Tuomisto, H. (2002) Beta-diversity in tropical forests. *Science*, **297**, 1439a.
- Sanford, R.L., Braker, H.E. & Hartshorn, G.S. (1986) Canopy openings in a primary Neotropical lowland forest. *Journal of Tropical Ecology*, **2**, 277–282.
- Sanford, R.L., Paaby, P., Luvall, J.C. & Phillips, E. (1994) Climate, geomorphology, and aquatic systems. *La Selva: Ecology and Natural History of a Neotropical Rain Forest* (eds L.A. McDade, K.S. Bawa, H.A. Hespenheide & G.S. Hartshorn), pp. 19–33. University of Chicago Press, Chicago.
- Sollins, P., Sancho, M.F., Mata, Ch.R. & Sanford, R.L. (1994) Soils and soil process research. *La Selva: Ecology and Natural History of a Neotropical Rain Forest* (eds L.A. McDade, K.S. Bawa, H.A. Hespenheide & G.S. Hartshorn), pp. 34–53. University of Chicago Press, Chicago.
- ter Steege, H., Jetten, V.G., Polak, A.M. & Werger, M.J.A. (1993) Tropical rain forest types and soil factors in a watershed area in Guyana. *Journal of Vegetation Science*, **4**, 705–716.
- Svenning, J.-C. (1999) Microhabitat specialization in a species-rich palm community in Amazonian Ecuador. *Journal of Ecology*, **87**, 55–65.
- Svenning, J.-C., Kinner, D.A., Stallard, R.F., Engelbrecht, B.M.J. & Wright, S.J. (2004) Ecological determinism in plant community structure across a tropical forest landscape. *Ecology*, **85**, 2526–2538.
- Svenning, J.-C. & Skov, F. (2002) Mesoscale distribution of understorey plants in temperate forest (Kalø, Denmark): the importance of environment and dispersal. *Plant Ecology*, **160**, 169–185.
- Terborgh, J., Foster, R.B. & Nuñez, V.P. (1996) Tropical tree communities: a test of the nonequilibrium hypothesis. *Ecology*, **77**, 561–567.
- Tuomisto, H. & Poulsen, A. (2000) Pteridophyte diversity and species composition in four Amazonian rain forests. *Journal of Vegetation Science*, **11**, 383–396.
- Tuomisto, H., Poulsen, A.D., Ruokolainen, K., Moran, R.C., Quintana, C. & Cañas, G. (2003) Linking floristic patterns with soil heterogeneity and satellite imagery in Ecuadorian Amazonia. *Ecological Applications*, **13**, 352–371.
- Tuomisto, H. & Ruokolainen, K. (1994) Distribution of Pteridophyta and Melastomataceae along an edaphic gradient in an Amazonian rain forest. *Journal of Vegetation Science*, **5**, 25–34.
- Tuomisto, H., Ruokolainen, K., Aguilar, M. & Sarmiento, A. (2003) Floristic patterns along a 43-km long transect in an Amazonian rain forest. *Journal of Ecology*, **91**, 743–756.
- Tuomisto, H., Ruokolainen, K., Kalliola, R., Linna, A., Danjoy, W. & Rodriguez, Z. (1995) Dissecting Amazonian biodiversity. *Science*, **269**, 63–66.
- Tuomisto, H., Ruokolainen, K. & Yli-Halla, M. (2003) Dispersal, environment, and floristic variation of Western Amazonian forests. *Science*, **299**, 241–244.
- Valencia, R., Foster, R.B., Villa, G., Condit, R., Svenning, J.-C., Hernández, C. *et al.* (2004) Tree species distributions and local habitat variation in the Amazon: large forest plot in eastern Ecuador. *Journal of Ecology*, **92**, 214–229.
- Vormisto, J., Phillips, O.L., Ruokolainen, K., Tuomisto, H. & Vásquez, R. (2000) A comparison of fine-scale distribution patterns of four plant groups in an Amazonian rainforest. *Ecography*, **23**, 349–359.
- Vormisto, J., Svenning, J.-C., Hall, P. & Balslev, B. (2004) Diversity and dominance in palm (Arecaceae) communities in terra firme forests in the western Amazon basin. *Journal of Ecology*, **92**, 577–588.
- Vormisto, J., Tuomisto, H. & Oksanen, J. (2004) Palm distribution patterns in Amazonian rainforests: what is the role of topographic variation? *Journal of Vegetation Science*, **15**, 485–494.
- Webb, C.O. & Peart, D.R. (2000) Habitat associations of trees and seedlings in a Bornean rain forest. *Journal of Ecology*, **88**, 464–478.
- Wolf, P.G., Schneider, H. & Ranker, T. (2001) Geographic distributions of homosporous ferns: does dispersal obscure evidence of vicariance? *Journal of Biogeography*, **28**, 263–270.

Received 13 April 2005

revision accepted 10 August 2005

Handling Editor: Gerhard Zotz

## Supplementary material

The following supplementary material is available online from [www.Blackwell-Synergy.com](http://www.Blackwell-Synergy.com)

**Table S1** Details of the pteridophyte species encountered during the inventory.