Journal of Ecology 2003
91, 757–775

Habitat association among Amazonian tree species: a landscape-scale approach

OLIVER L. PHILLIPS, PERCY NÚÑEZ VARGAS†‡, ABEL LORENZO MONTEAGUDO‡, ANTONIO PEÑA CRUZ§, MARIA-ELENA CHUSPE ZANS‡, WASHINGTON GALIANO SÁNCHEZ‡, MARKKU YLI-HALLA¶ and SAM ROSE*

School of Geography, Earth and Biosphere Institute, University of Leeds, Leeds, UK, †Biodiversidad Amazónico, Cusco, Peru, ‡Herbario Vargas, Universidad Nacional, San Antonio Abad del Cusco, Peru, §Jardin Botánico de Missouri, Oxapampa, Peru, and ¶MTT, Agrifood Research Finland, Jokioinen, Finland

Summary

1 Unravelling which factors affect where tropical trees grow is an important goal for ecologists and conservationists. At the landscape scale, debate is mostly focused on the degree to which the distributions of tree species are determined by soil conditions or by neutral, distance-dependent processes. Problems with spatial autocorrelation, sparse soil sampling, inclusion of species-poor sites with extreme edaphic conditions, and the difficulty of obtaining sufficient sample sizes have all complicated assessments for high diversity tropical forests.

2 We evaluated the extent and pervasiveness of habitat association of trees within a 10 000 km^2 species-rich lowland landscape of uniform climate in south-west Amazonia. Forests growing on two non-flooded landscape units were inventoried using 88 floristic plots and detailed soil analyses, sampling up to 849 tree species. We applied single-species and community-level analytical techniques (frequency-distributions of presence records, association analysis, indicator species analysis, ordination, Mantel correlations, and multiple regression of distance matrices) to quantify soil/floristic relationships while controlling for spatial autocorrelation.

3 Obligate habitat-restriction is very rare: among 230 tree species recorded in ≥ 10 localities only five (2.2%) were always restricted to one landscape unit or the other.

4 However, many species show a significant tendency to habitat association. For example, using Monte Carlo randomization tests, of the 34 most dominant species across the landscape the distributions of 26 (76.5%) are significantly related to habitat. We applied density-independent and frequency-independent estimates of habitat association and found that rarer species tend to score higher, suggesting that our full community estimates of habitat association are still underestimated due to the inadequate sampling of rarer species.

5 Community-level floristic variation across the whole landscape is related to the variation in 14 of 16 measured soil variables, and to the geographical distances between samples.

6 Multiple regression of distance matrices shows that 10% of the floristic variation can be attributed to spatial autocorrelation, but even after accounting for this at least 40% is attributable to measured environmental variation.

7 Our results suggest that substrate-mediated local processes play a much more important role than distance-dependent processes in structuring forest composition in Amazonian landscapes.

Key-words: forest, floristics, soils, spatial autocorrelation, tree species composition.

Journal of Ecology (2003) **91**, 757–775

Correspondence: Oliver Phillips, School of Geography, Earth and Biosphere Institute, University of Leeds, Leeds LS2 9JT, UK (e-mail o.phillips@geog.leeds.ac.uk).
*Present address: Volunteer Services Overseas, London, UK.

Introduction

A fundamental goal of plant ecology is to understand the relative importance of the environment and of chance in structuring plant communities. Tropical forests include the earth's most diverse communities, and niche differentiation with respect to resources is frequently invoked as a means by which this exceptional diversity is maintained (e.g. Grubb 1977; Denslow 1987). Several studies have demonstrated demographic patterns and life-history variation along successional light-intensity gradients that may help explain coexistence of tropical tree species (e.g. Clark & Clark 1992; Condit *et al*. 1996; Condit *et al*. 1999; Rees *et al*. 2001). However, the extent to which tropical trees specialize with respect to edaphically and topographically determined resources – such as soil nutrients and moisture availability – is the subject of an especially persistent and vigorous debate. Some studies have found no strong effect of soil conditions on lowland tropical tree floristic patterning (e.g. Poore 1968; Wong & Whitmore 1970; Knight 1975; Newbery & Proctor 1983; Hubbell & Foster 1986; Newbery *et al*. 1988; Hart *et al*. 1989; Pitman *et al*. 1999). Where researchers have reported that the floristic composition of tropical lowlands responds to edaphic differences, the studies rarely allow us to draw general conclusions about the degree of habitat association in tree species across species-rich tropical landscapes.

The large practical difficulties inherent in collecting and identifying thousands of non-fertile collections representing hundreds of species in poorly mapped landscapes have encouraged researchers to develop alternative approaches. Thus, for example, studies have often focused on family level patterns (e.g. Gentry 1988a; Terborgh & Andresen 1998; ter Steege *et al*. 2000), the impact of extreme soil conditions (such as white-sand podsols, floodplain entisols, swamp histosols, or limestone mollisols, e.g. Newbery & Proctor 1983; Kalliola *et al*. 1991; Vásquez & Phillips 2000, and cf. Sollins 1998), the importance of microhabitat topographic variation (e.g. Sabatier *et al*. 1997; Svenning 1999; Vormisto *et al*. 2000), relatively depauperate regions (e.g. Swaine 1996), distributions of particular tree taxa (e.g. Debski *et al*. 2002) or particular non-tree elements of the flora (e.g. Clark *et al*. 1995; Poulsen 1996; Fleck & Harder 2000; Tuomisto & Poulsen 2000; Tuomisto *et al*. 2003a,b), or do not make allowance for spatial aggregation of species' populations, a near-ubiquitous feature of tropical forests (cf. Condit *et al*. 2000).

Recently it has been shown that the importance of factors such as habitat variation in structuring tropical forest communities may vary greatly with scale and locality (Condit *et al*. 2002). This implies that our understanding of tropical community pattern and process is heavily influenced by the geography of ecological research, which is typically done at small scales in a limited number of localities. In the Neotropics, for example, the intensity of research in Central American forests outscores that in Amazonian forests by a factor of > 50 on a per area basis (based on a survey of ISI publications between January 2000 and June 2002 reporting primary ecological research from Panama, Costa Rica and Amazonia). Most studies that do clearly demonstrate substrate-mediated habitat association in tropical tree species are central American (e.g. Clark *et al*. 1998, 1999; Condit *et al*. 2000, 2002; Harms *et al*. 2001; Pyke *et al*. 2001; Faith & Ferrier 2002). In Amazonia, encompassing more than half the world's lowland rain forest, the importance of species' habitat associations in controlling community diversity and composition remains particularly contentious (e.g. Salo *et al*. 1986; Gentry 1988a; ter Steege *et al*. 1993; Duivenvoorden 1995; Tuomisto *et al*. 1995; Ruokolainen *et al*. 1997; Pitman *et al*. 1999; Condit *et al*. 2002). According to one view, large areas of the Amazon feature low beta-diversity and a substantially homogeneous flora. Pitman *et al*. (2001), for example, have shown that a similar set of tree species dominates Andean foreland landscapes > 10^3 km apart in Peru and Ecuador. Alternatively, the appearance of homogeneity could be exaggerated by the difficulty in mustering sufficient statistical power to detect heterogeneity in species distributions within landscapes (Ruokolainen *et al*. 1997).

Western Amazonia supports the world's most tree species-rich forests (e.g. Gentry 1988b) and the scarcer taxa that may be the most important for conservation are also the most difficult to characterize environmentally. With some important exceptions (e.g. Tuomisto *et al*. 1995; Kalliola *et al*. 1998) vegetation mapping in Amazonia has produced relatively uniform results, with only a handful of different environmentally determined vegetation formations broadly controlled by regional climatic and geomorphological features (e.g. UNESCO 1980; Sombroek 2001). Can these maps be relied on to reflect floristically defined communities, or is the floristic composition of forests a more finely grained affair, with distributions substantially mediated by localized soil formations? The question is not only important for plant ecologists but also for land-use planners and conservationists who need reliable vegetation maps to help identify forest communities and species at particular risk.

A few edaphically and hydrologically determined vegetation types *are* recognized within lowland Amazonia, particularly palm swamps, floodplain forests, and forests on podzolized white sand soils (e.g. Anderson 1981; Pires & Prance 1985; Kahn & Mejia 1990; Duque *et al*. 2002), but these occupy minor areas compared with the unflooded forest on non-podzolized soil, so-called 'terra firme' forest, which covers the majority of lowland Amazonia (UNESCO 1980). Ethnobotanical evidence (Fleck & Harder 2000; Shephard *et al.* 2001; Shephard *et al.* 2003) shows that indigenous people recognize numerous distinct habitats within terra firme, characterized by congruous variation in

soil, topography, vegetation structure, and characteristic species of flora and fauna, indicating that current academic understanding of the floristic landscape is still inadequate. Ecologists' attempts to characterize tree species variation across whole Amazonian landscapes are hampered by the difficulty of achieving sufficient sampling within individual sampling units while also realizing spatial replication across the landscape. Both features can generate high sampling error, and this renders it difficult to identify the signal of floristic patterns in floras of 1000 tree species or more. For example, Pitman *et al*. (1999) and Condit *et al*. (2002) sampled at least 19 000 Amazonian trees, but the limited soil data and small number of discrete sample plots may make detecting species-level habitat association difficult; Duivenvoorden (1995) used more plots (95) but sampled fewer stems. These studies concluded that species turnover from one substrate to another (β – diversity) is low in non-flooded forests.

Here we bring a new, large single-landscape Amazon ecofloristic data set to bear on the question of the extent to which tree species distributions are determined by edaphic factors within a tropical landscape. Specifically, we ask: (i) What proportion of species are completely confined to different terra firme habitats? (ii) Is a tendency to habitat association limited to a few individual species, or is it a widespread property of the Amazonian flora? (iii) Is habitat association of species independent of measures of ecological success such as density and frequency? (iv) What is the relative importance of stochastic and environmental factors in controlling floristic patterns across an Amazonian landscape? Each question was asked for all sampled species attaining a stem diameter of ≥ 2.5 cm, and then for the subset of species that attain a diameter of ≥ 10 cm.

Methods

FIELD WORK

Our study area in Madre de Dios, south-eastern Peru, was defined as a rectangle with *c*. 100 km edges, centred on the town of Puerto Maldonado. The region is still > 90% forested and is dominated by past (Pleistocene to Holocene) and present fluvial activity (Räsänen *et al*. 1992).

Sampling approach

The region has near uniform elevation (200–260 m a.m.s.l) and a seasonal tropical climate, with mean annual rainfall of 2200–2400 mm, 3 months a year averaging less than 100 mm, and a mean annual temperature of > 25 °C (Duellman & Koechlin 1991; Malhi *et al*. 2002).

Forests are found on three distinct geomorphological units (Salo *et al*. 1986; Räsänen *et al*. 1990, 1991, 1992; Salo & Kalliola 1990; Osher & Buol 1998): irregularly flooded areas of the contemporary floodplains of the Tambopata, Heath, Las Piedras and Madre de Dios rivers (*c*. 9% of the study area), no-longer flooded terraces of the Holocene floodplain of these rivers (*c*. 20%), and ancient Pleistocene alluvial terraces at least 40 000 years old (58%). Swamp forests are found in small histosol patches within each unit, particularly the contemporary floodplains, but most forest (89%) grows on better drained ultisols that dominate the Pleistocene surfaces (Osher & Buol 1998) or inceptisols and ultisols on the Holocene surface (Malhi *et al*. 2004). Samples were located in old-growth forests in 13 community territories and protected areas (Fig. 1, Appendix 1 and 2 in Supplementary Material), stratified by geomorphology on the basis of visual assessment of a Landsat image (path 002 row 069) ground-truthed over a 2-year period by checking relative topographical positions and detailed mapping with local residents. Our focus here is on patterns within the ultisols and inceptisols of the terra firme forests, so we deliberately exclude all samples in swamp forests and contemporary floodplains from the analysis. The region lacks other important and distinctive tropical soils, such as highly weathered oxisols or white sand soils (spodosols), or those derived from basaltic, volcanic ash, limestone or serpentine substrates. Therefore, our approach provides a rather conservative evaluation of the extent of species association within mature tropical forests, and even within mature unflooded tropical forests. Holocene and Pleistocene surfaces are well distributed throughout, so by sampling both surfaces across the landscape we aimed to disentangle the roles of spatial and environmental factors.

Sample units

Our approach is a modification of an inventory approach first used extensively in tropical forests by Gentry (e.g. Gentry 1982, 1988a; Enquist & Niklas 2001; Phillips & Miller 2002; Phillips *et al*. 2003). Forests at each location were sampled in 1998 and 1999 by 10 2×50 m subplots, totalling 0.1 ha, and located within a 100×180 m sampling grid so as to systematically subsample 1.8 ha of forest. Each non-scandent plant rooted within the transect area and with a stem diameter of ≥ 2.5 cm at 1.30 m height (= diameter at breast height, d.b.h) was included in the sample, and every plant measured and identified or recorded as a unique 'morphospecies'. Voucher collections were made for each unique species and whenever there was any uncertainty to its identity. Repeated collections of sterile plants were frequently needed to reliably distinguish morphospecies. A full set of duplicates is deposited in Peru at CUZ, where vouchers were identified and cross-referenced. A partial set is also maintained at USM.

We collected soil samples (0–15 cm below the organic material layer) from each inventory location by auguring each of the 2×50 subplots at one or more randomly

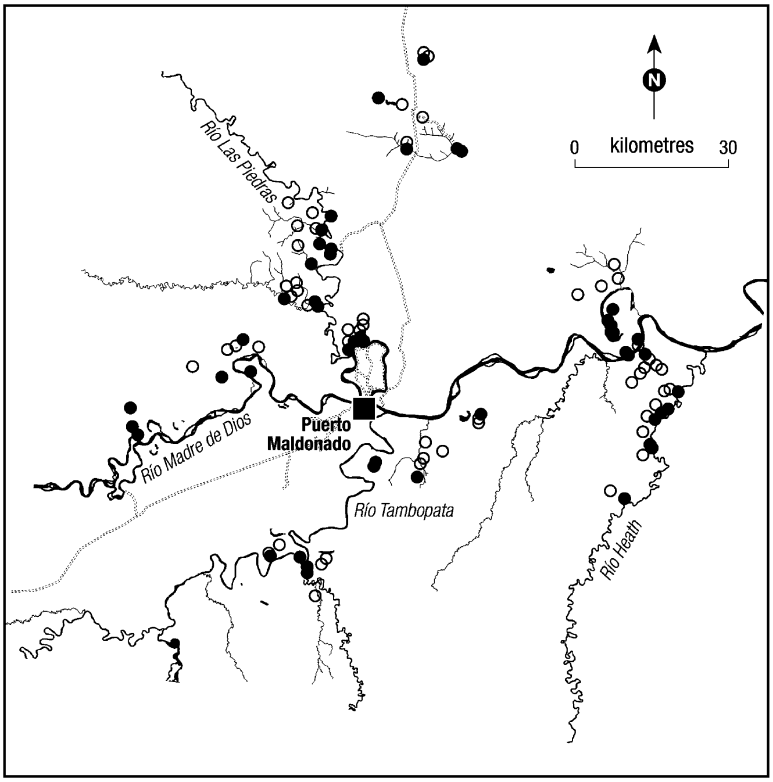

Fig. 1 Map of study area showing distribution of sample locations. Closed circles represent Holocene samples; open circles represent Pleistocene samples. All samples are located between 12° S and 13° S and 68°30′ W and 69°30′ W.

chosen points and then bulking the subsamples for each inventory. Composite samples were air-dried, cleaned by removing macroscopic organic material, and subsampled. Drainage conditions were assessed visually on a scale of 1 (permanently water-logged) to 10 (excessively drained) for each subplot and mean values derived for each inventory. Chemical and physical soil properties were analysed at the Agricultural Research Center in Finland following ISRIC protocols (Van Reeuwijk 1995): soil pH was measured in a 1-M KCl suspension; exchangeable Ca, Mg, K and Na were extracted with 1 M ammonium acetate (pH 7.0); exchangeable Al was extracted with 1 M KCl; plant-available P was determined by the Bray 1 method (0.03 M NH_4F–0.025 M HCl extraction); clay (< 2 μm), silt (2–63 μm) and sand (0.63–2 mm) content was determined after pre-treatment with citrate – dithionite – bicarbonate; and loss of weight on ignition (LOI) was determined by heating the dried soils at 420 °C for 6 h. Effective cation exchange capacity (ECEC) was calculated as the sum of cation charge, expressed in cmol(+) kg^{-1}.

ANALYSES

Eighty-eight mature forest samples were included in the data set. All records of unidentified morphospecies were removed to eliminate the possibility that inconsistent cross-referencing of vouchers within and between communities might bias the results. Each species in this data set was checked for maximum diameter within the 88×0.1-ha samples and against an independent data set of trees in 16×1.0-ha permanent sample plots in eastern Madre de Dios (Gentry 1988a; Phillips *et al*. 1998, Phillips, Vásquez Núñez, & Monteagudo, unpublished data) to identify the plants that attain ≥ 10 cm stem diameter as self-supporting adults. Analyses that follow were applied both to species ≥ 10 cm d.b.h. ('trees'), and to all species ≥ 2.5 cm d.b.h. in our data set.

We evaluate habitat association within terra firme forests using two distinct approaches. In the first, we simply divide our samples into Holocene and Pleistocene surfaces and examine the extent to which individual species are generally confined to one or other landscape unit. In the second approach, we examine the extent to which the entire floristic composition of forests varies with both physical and spatial factors, using ordination, Mantel tests, and multiple regression on distance matrices.

Treatment of spatial autocorrelation

Patterns of habitat association can be confounded by spatial autocorrelation, as geographically proximate samples are more likely to share species due to stochastic processes than geographically distant samples (e.g. Condit *et al*. 2000; Harms *et al*. 2001). Our analyses

either specifically parcel out and quantify the components of floristic variation potentially due to spatial autocorrelation and due to environmental attributes (cf. Condit *et al*. 2002), or use permutation procedures to compare patterns against null expectations generated by hypotheses of no habitat effect. Traditional chi-squared association analysis of density patterns is also used for comparison.

ANALYTICAL APPROACH 1: SPECIES DISTRIBUTIONS ACROSS LANDSCAPE UNITS

First, we examine whether species are sampled in only one habitat ('narrowly restricted', following Pitman *et al*. 1999), or not ('widespread'). By this definition the probability of a species appearing to be 'narrowly restricted' will vary artifactually as an inverse function of the frequency with which it is encountered, so we compare the empirically observed pattern with a null expectation of frequency distributions generated by a binomial distribution assuming no habitat restriction. This is the simplest technique for evaluating habitat association, as it takes account of presence/absence data but makes no allowance for differences in relative densities between habitats. Secondly, we relax the test of habitat association from 'narrow restriction' to 'tendency'. To assess the tendency with which species are favoured by one habitat or another we compare for each species both their frequency-distribution patterns and density-distribution patterns against null expectations, and evaluate overall differences in species frequencies and densities between habitats using chi-squared tests of association. Thirdly, we use indicator species analysis (Dufrene & Legendre 1997) to account for both relative abundance and relative frequency of each species across the landscape, by testing the degree of habitat association at the level of species against individually parameterized null models.

The null hypothesis is that the highest indicator value for each species in one or other habitat (*IVmax*) is no larger than would be expected if the species was distributed at random across the two landscape units. Significance is estimated by a Monte Carlo procedure that reassigns species densities and frequencies to habitats 1000 times. The probability of type I error is based on the proportion of times that the *IVmax* score for each species from the randomized data set equals or exceeds its *IVmax* score from the actual data set (McCune & Mefford 1999). Finally, to address the question of whether habitat association is associated with species density in the landscape we developed simple density- and frequency-independent indices of habitat association for each species, based on the relative difference between the probability of being encountered in Pleistocene samples and Holocene samples, and we relate both these indices and *IVmax* scores to species density values.

Thus, the density-independent index of habitat association,

A_{Dx} = the absolute value of $\{(P_D)/(P_D + H_D) - (H_D)/(P_D + H_D)\}$

where, for each species x:

(P_D) = the density of that species in the Pleistocene samples;
(H_D) = the density of that species in the Holocene samples.

A_{Dx} varies between 0 (= no habitat association with equal density in each landscape) and 1 (= every stem of that species found in one habitat or the other).

Similarly, the frequency-independent index of habitat association,

A_{Fx} = the absolute value of $\{(P_F)/(P_F + H_F) - (H_F)/(P_F + H_F)\}$

where, for each species x:

(P_F) = the frequency of that species in the Pleistocene samples;
(H_F) = the frequency of that species in the Holocene samples.

A_{Fx} varies between 0 (= no habitat association) and 1 (= every record of that species being from samples in one habitat or the other).

We used the entire data set for the first and second set of analyses, following established practice. However, we needed to eliminate the possibility of spatial autocorrelation affecting our other analyses (the indicator species analysis, and calculations of density-independent and frequency-independent habitat association scores). Our procedure was as follows. First, we examined at what spatial scale within-habitat sample clustering occurred (i.e. over what intersample pair distances was there a greater than expected probability of both samples representing the same habitat). Over intersample distances of $< c.$ 5000 m and especially $< c.$ 1500 m sample pairs are more likely than not to represent the same habitat. Then, we determined the number of within-habitat pairwise combinations that had to be eliminated to bring the odds to 1 in 2 of a random sample pair being from the same habitat over all distances < 5000 m. Finally, we progressively removed samples from the data base, starting with the shortest distance of same-habitat sample pairs; for each sample pair we selected the one sample that also had the second nearest same-habitat neighbour. In practice there are few samples very close to one another (of 3829 potential pairwise combinations, only 14 have distances < 1000 m), so the removal of relatively few samples from the data set (nine out of 88) eliminates the problem. Of the original 38 Holocene samples we needed to remove five (LAT7, PNB1, PTA9, SAB9 and SAB10);

of the original 50 Pleistocene samples we needed to remove four (BOC5, PTA5, PTA6 and SJC3).

ANALYTICAL APPROACH 2: FLORISTIC COMPOSITION AS A FUNCTION OF ENVIRONMENTAL AND GEOGRAPHICAL DISTANCE

Construction of distance matrices

We explore potential relationships between plant species composition and soil on the basis of site-to-site comparisons. Distance between all possible pairs of sites was measured for each of three classes of variable – (i) plant species composition, (ii) each of 15 soil variables, and (iii) geographical distance among the sites – to produce a dissimilarity matrix between all possible sample-pairs for all variables.

We measured floristic distance between sample pairs using the Sørensen (Bray-Curtis) index of similarity, widely used in community analyses because it retains sensitivity to the data structure without giving undue weight to outliers (e.g. Ludwig & Reynolds 1988). Distance matrices on the basis of soil characteristics (Ca, K, Mg, Na, sum of base cations, Al, ECEC, Al/ECEC, P, pH, LOI, sand, silt, clay, fraction < 0.063 mm), and geographical distances are based on Euclidean distance, i.e. the difference in the values of the each sample pair. Before calculating distances we transformed variables using Tukey's ladder of powers (Table 1); this corrected positively skewed distributions and ensures that absolute differences between low values receive greater analytical weight than differences between high values. This treatment is more likely to reflect environmental and spatial differences experienced by plants than would non-transformation of raw variables. For example, Hubbell's (2001) neutral theory predicts non-linear distance decay in similarity, and Condit *et al*. (2000, 2002) have shown this result empirically in tropical forest landscapes.

Table 1 Soil and Distance variable transformations

Variable	Transformation
pH	$-(x - 3.2)^{-1}$
Ca	$\ln(x)$
K	$\ln(x)$
Mg	$\log_{10}(x)$
Na	$-x^{-0.5}$
Al	x^{-2}
ECEC	$-x^{-0.5}$
Al/ECEC	x
P	$-x^{-0.5}$
Dry matter	$(x - 89.5)^3$
Loss on ignition	$-x^{-0.5}$
Sand	x^{-2}
Silt	x^2
Clay	$\ln(x)$
Fraction < 0.063 mm	x
Mean drainage	x^2
Geographical distance	$\ln(x)$

Analyses using the distance matrices

We used the distance matrices in four different analyses: first, by plots of sample-pair similarity against distance to quantify spatial decay; secondly, a non-metric multidimensional distance scaling ordination (NMDS, Kruskal 1964) to explore the compositional patterns; thirdly, Mantel's test on matrix correlation (Mantel 1967) to test for interdependence among key variables; and finally, multiple regression on distance matrices (Legendre *et al*. 1994), to model the full floristic variation in terms of the spatial and environmental factors.

NMDS is an ordination method that arranges the samples in a user-defined limited number of dimensions so that the rank order of distances is as similar as possible to the rank order in the original data. The solution is found iteratively to find the best number k of dimensions for a given data matrix. We used the program PC-ORD for producing the NMDS ordinations, running 400 iterations each starting with a random configuration, adopting $k = 6$, and applying an instability criterion of 10^{-5}. The Mantel test involves computing the Pearson correlation coefficient between the values of two matrices (Smouse *et al*. 1986), and using a Monte Carlo procedure to estimate the probability of error. We distinguished a 0.1% probability of error by comparing observed distributions of r against the distribution of random values generated from permuting one of the matrices and recalculating r 999 times. A partial Mantel test was used to evaluate how correlations between floristic composition and environmental variables changed after controlling for the effect of geographical distance. The Mantel tests were performed with PC-ORD and with the R-Package (Legendre & Vaudor 1991) for all pairwise variable combinations.

In the multiple regression method on distance matrices the variation in one dependent matrix is expressed in terms of variation in a set of independent matrices. The computational procedure mimics that of normal multiple regression, except that the significance of parameters is estimated by the Monte Carlo permutation procedure described above (Legendre *et al*. 1994). We developed models to describe the variation of the floristic data set in terms of environmental and spatial factors, by undertaking a multiple regression of the floristics distance matrix against the variable distance matrices. Both forward selection and backwards elimination methods were used. Here our aim is to ascertain the contribution of environmental factors to floristic pattern within the terra firme landscape, having accounted for the spatial fraction. The three-stage model generation and selection procedure reflects this: (i) we built the largest model possible in which each factor contributes significantly (at $P = 0.01$); (ii) we eliminated independent variables with negative b-coefficients, until all remaining b-coefficients were positive (negative b-coefficients are not interpretable as they imply that as values of environmental variables in samples become closer the differences in species

composition between them become greater); finally (iii), we applied backward elimination with Bonferroni corrected probability levels (estimated on the basis of 999 permutations) to decide eliminations, holding the intersite spatial distance matrix as the last factor to be excluded so that we could differentiate all floristic variation that had a spatial structure. The procedure was conducted first for the whole data set, and variables in the best all-species model were evaluated for tree species in order to be able to compare the two data sets directly. All multiple regressions on distance matrices were performed with the Permute! 3.4 program (Legendre *et al*. 1994).

Results

OVERVIEW: SOIL VARIATION

Pleistocene and Holocene substrates differ significantly in all but one measured soil parameter (Table 2), with the former having on average lower pH, lower cation concentrations, lower CEC, more sand, less silt and clay, and better drainage.

Overview: tree species diversity and stem density

The 88 samples inventory a range of 132–357 individuals per 0.1-ha sample (mean = 233.3) and between 63 and 136 species and morphospecies (mean = 92.9) (Table 3). Our sample contains 20 528 plants, 692 tree species, and 157 tree morphospecies. This compares with an estimate of 1004 tree species known from independent collections to attain at least 10 cm diameter in all Madre de Dios (Pitman *et al*. 2001). That figure spans a much larger geographical area and includes species restricted to floodplain, swamp and montane habitats, so we surmise that our inventories encountered 85–100% of tree species in the lowland terra firme landscape we sampled.

Treatment of spatial autocorrelation

One objective of the sampling design was to ensure that Pleistocene and Holocene samples were well-mixed across the landscape, so that the straight-line distance between samples does not affect the probability that they will be of the same habitat type. In practice, as described in the Methods section, this was not quite achieved; over short distances sample pairs were more likely to be from the same habitat. However, the average dispersal of samples is equivalent for pairs of samples of both habitat types and for pairs of samples from different habitat types (mean and median Pleistocene-Pleistocene distance = 42.9 km, 44.5 km, respectively; mean and median Holocene-Holocene distance = 39.9 km, 43.9 km; mean and median Pleistocene-Holocene distance = 41.1 km, 43.4 km). Additionally, as discussed below, partial Mantel tests show no geographical effect of distance on sample edaphic

Table 2 Soil chemistry, particle size distribution and estimated drainage in 88 Pleistocene and Holocene landscape sites. DM = dry matter; LOI = loss on ignition; Dr = Drainage

	n	pH (1 M KCl)	Ca (mg kg^{-1})	K (mg kg^{-1})	Mg (mg kg^{-1})	Na (mg kg^{-1})	Al/ ECEC (Cmol^{+} kg^{-1})	Al (mg kg^{-1})	ECEC (%)	P (mg kg^{-1})	DM 420 °C	LOI 420 °C	Sand (%)	Silt (%)	Clay (%)	< 0.063 mm (%)	Dr (1–10 scale)
Mean	51	3.79	40.8	55.1	27.1	3.6	210	3.1	69.5	2.1	98.9	2.89	17.7	64.3	17.9	50.8	6.7
(range) Pleistocene		(3.51–4.38)	(2.9–229)	(21.2–114)	(7.1–63.3)	(0.9–9.5)	(10–622)	(1.0–7.4)	(8.9–92.9)	(0.96–13.1)	(92.3–100)	(1.04–6.00)	(1.0–54.1)	(39.4–84.8)	(4.9–37.5)	(21.8–83.5)	(4.4–8.0)
Mean	37	4.05	1082	89.5	265.4	15.4	165	10.0	26.8	4.1	98.0	3.73	2.7	71.0	26.3	84.2	4.3
(range) Holocene		(3.57–5.39)	(10.2–3402)	(32.7–197)	(10.8–789)	(2.8–109.8)	(3–835)	(1.8–22.5)	(0.3–93.6)	(1.1–14.4)	(89.6–99.7)	(1.96–6.72)	(0.0–28.9)	(50.8–89.2)	(6.6–49.1)	(26.3–99.6)	(2.0–7.0)
Mann–Whitney test, W		1950	1459	1751	1441	1469	2500	1450	2965	1744	2785	1874	3055	1927	1864	1486	3069
		**	***	***	***	***	(*)	***	***	***	***	***	***	**	***	***	***

(*) $P < 0.10$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 3 Plant life-form representation in different subsets of the Madre de Dios floristic datasets. Analyses in this paper are based on the subsets listed under column 2(b) (terra firme forests, fully identified species only)

	(1) 101 samples, including swamps and contemporary floodplains				(2) 88 samples, excluding swamps and contemporary floodplains			
	(a) All morphospecies		(b) Fully identified species only		(a) All morphospecies		(b) Fully identified species only	
Adult form	Individuals	Species	Individuals	Species	Individuals	Species	Individuals	Species
Either shrub or herb	268	9	268	9	233	9	174	8
Obligate liana	170	24	126	16	135	23	50	17
Obligate shrub	1 309	99	1 049	89	754	93	648	85
Unknown: either shrub or tree	235	38	0	0	147	30	0	0
Tree (non-scandent stem ≥ 10 cm d.b.h)	21 063	896	18 438	701	19 259	849	18 229	692
SUM	23 045	1066	19 881	817	20 528	1004	19 101	802
Trees (%)	91.4%	84.1%	92.7%	85.8%	93.8%	84.6%	95.4%	85.7%

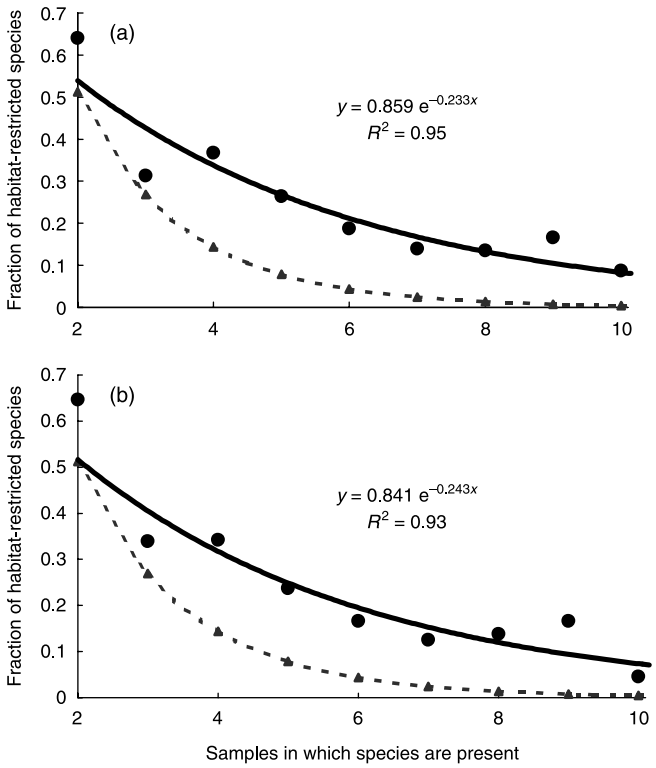

Fig. 2 Proportion of habitat restricted species vs. frequency. Circles: actual proportion of species that are habitat-restricted. Solid line: best-fit exponential model for proportion of species that are habitat-restricted. Triangles and dotted line: expected proportion of habitat-restricted species under null expectations of no habitat association. (a) All species. (b) Trees only.

distance. We therefore infer that the degree of spatial proximity between samples in this landscape does not substantially affect the probability of their sharing similar soils.

ANALYTICAL APPROACH 1: SPECIES DISTRIBUTIONS ACROSS LANDSCAPE UNITS

Holocene vs. Pleistocene landscapes: are species restricted to one or the other?

The number of species restricted to a single habitat can be described as a negative exponential function of the number of localities from which it was sampled (Fig. 2). Very few frequent species are completely restricted to a single habitat. Of all species present in at least 10 localities, only five tree species out of 230 trees and 242 taxa ≥ 2.5 cm d.b.h. are restricted to one habitat or another. However, for both the tree species and the whole data set the number of completely habitat-restricted species modelled by a binomial probability distribution significantly exceeds null predictions irrespective of species' frequencies.

Holocene vs. Pleistocene landscapes: are species favoured by one or the other?

Frequency Binomial tests of association between species frequency and habitat class show that out of 651 species ≥ 2.5 cm d.b.h. occurring in at least two samples, 235 are significantly associated with one or the other habitat (36%, all tests at $P < 0.05$). Similarly, of 563 trees occurring in ≥ 2 samples, 199 are significantly associated with one or other habitat (35%). The proportion of species for which significant habitat association was detected increases with increasing frequency of species. Thus, out of 249 species ≥ 2.5 cm d.b.h. occurring in ≥ 10 samples, 117 are associated with one or other habitat (47%), and of 234 trees occurring in ≥ 10 samples, 113 are significantly associated with one or the other (48%). Similarly, of the 114 species ≥ 2.5 cm d.b.h and 112 tree species occurring in ≥ 20 samples, 60 are associated with one or the other habitat (53% and 54%, respectively). Contingency analyses also suggest a widespread tendency to habitat association (values of χ^2 range from 102 (trees in at least 20 samples) to 1871 (species ≥ 2.5 cm d.b.h. in at least two samples), $P < 0.001$).

Density Of 348 species ≥ 2.5 cm d.b.h. with at least 10 individuals, 223 are significantly associated with one or the other habitat (64%, all tests at $P < 0.05$). Similarly, of 317 trees with ≥ 10 individuals, 212 were significantly associated with one or the other (67%). With increasing density of species the proportion for which significant habitat association is detectable increases. Out of 204 species ≥ 2.5 cm d.b.h. with at least 20 individuals, 144 are associated with one or other habitat (71%), and of 194 trees with ≥ 20 individuals, 138 (71%) have significant habitat association. Similarly, of the 39 tree species and 40 species ≥ 2.5 cm d.b.h. with ≥ 100 individuals, 90% of each are associated with one or other habitat. Contingency analyses also suggest a strong tendency to habitat association (values of χ^2 range from 961 (trees with ≥ 100 individuals) to 38 759 (all species with > 2 individuals), $P < 0.001$).

Frequency and *Density* Indicator species analysis shows that most abundant species in the landscape have a significant tendency to one or other of the terra firme habitats and have value as habitat indicators even after removing nine samples to account for any potential effects of spatial autocorrelation (Table 4). However, the proportion of habitat indicators falls off rapidly with decreasing population density (Fig. 3).

In sum, results from analyses of (i) species entirely restricted to one habitat, (ii) distributions by habitat of species' frequencies, (iii) distributions by habitat of species' densities, and (iv) indicator species, all show that more species are associated with individual habitats than expected by chance alone. Among the more frequent and dense species, more than half are significantly

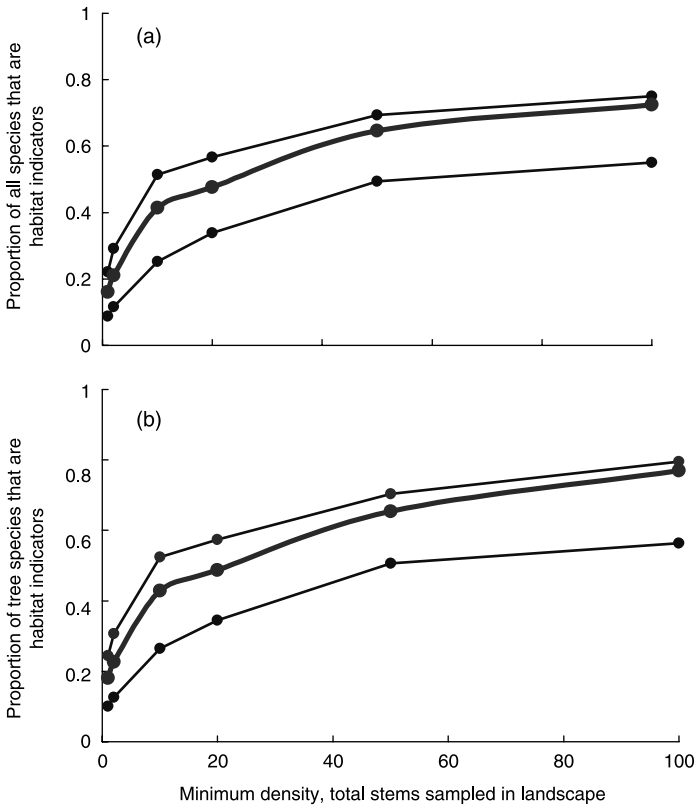

Fig. 3 Proportion of all species that are habitat indicators, as a function of species' densities (minimum cutoff values). Upper line, at $P < 0.1$; middle line, at $P < 0.05$; lower line, at $P < 0.01$. (a) All species. (b) Trees only.

Journal of Ecology, **91**, 757–775

Table 4 Indicator Species Analysis. Scores and habitat tendencies for each species with ≥ 100 stems in the landscape sample. IV (Indicator Values) expressed as a percentage, with a maximum theoretical value of 100. Species sorted by *P*-value and IV

Family	Genus and species	Habit	Density (*n* stems)	Frequency (*n* samples)	Habitat tendency	Observed IV	Randomised mean IV	Randomised SD IV	*P*
Monimiaceae	*Siparuna decipiens* (Tul.) A. DC.	T	521	67	Pleistocene	82.6	46.8	3.65	0.001
Monimiaceae	*Mollinedia killipii* JF Macbr.	T	147	31	Pleistocene	44.4	24.9	4.25	0.001
Chrysobalanaceae	*Hirtella racemosa* Lam.	T	352	61	Pleistocene	73.3	43.8	4.39	0.001
Sterculiaceae	*Theobroma cacao* L.	T	133	37	Holocene	52.0	29.0	4.41	0.001
Annonaceae	*Unonopsis floribunda* Diels	T	105	38	Holocene	67.0	29.6	4.71	0.001
Cecropiaceae	*Pourouma minor* Benoist	T	129	41	Pleistocene	80.2	31.7	4.74	0.001
Monimiaceae	*Siparuna cristata* (Poepp. & Endl.) A. DC	T	362	49	Pleistocene	60.4	36.7	4.74	0.001
Arecaceae	*Socratea exorrhiza* (Mart.) H. Wendl.	T	102	42	Holocene	58.9	32.5	4.83	0.001
Chrysobalanaceae	*Hirtella triandra* Sw.	T	131	39	Pleistocene	55.9	30.5	4.85	0.001
Meliaceae	*Guarea macrophylla* ssp. *pachycarpa* (C. DC.) Pennington	T	129	32	Holocene	64.3	27.1	4.99	0.001
Myristicaceae	*Iryanthera juruensis* Warb.	T	197	53	Pleistocene	70.7	39.7	5.05	0.001
Moraceae	*Sorocea pileata* W.C. Burger	T	136	48	Holocene	53.1	35.5	4.27	0.002
Moraceae	*Pseudolmedia laevis* (Ruiz & Pav.) JF Macbr.	T	378	63	Holocene	56.6	45.0	4.27	0.002
Fabaceae	*Tachigali bracteosa* (Harms) Zarucchi & Pipoly	T	225	52	Pleistocene	54.0	39.0	4.96	0.002
Bombacaceae	*Quararibea witii* K. Schum. & Ulbr.	T	109	25	Holocene	43.0	22.0	5.07	0.002
Apocynaceae	*Aspidosperma tambopatense* A.H. Gentry	T	135	44	Pleistocene	50.7	33.4	4.58	0.003
Chrysobalanaceae	*Hirtella excelsa* Standl. ex Prance	T	106	46	Pleistocene	48.3	34.5	4.39	0.005
Strelitziaceae	*Phenakospermum guyanense* (Rich.) Endl.	T	871	22	Pleistocene	37.4	19.5	4.73	0.006
Piperaceae	*Piper pseudoarboreum* Yuncker	T	240	43	Pleistocene	37.2	20.5	4.80	0.008
Arecaceae	*Bactris concinna* Mart.	S	151	42	Holocene	51.6	33.6	5.75	0.009
Nyctaginaceae	*Neea macrophylla* Poepp. & Endl.	T	118	39	Pleistocene	42.1	30.4	4.67	0.021
Myristicaceae	*Iryanthera laevis* Markgr.	T	289	63	Pleistocene	56.0	45.0	4.29	0.022
Fabaceae	*Tachigali polyphylla* Poeppig	T	168	49	Pleistocene	48.8	36.8	4.91	0.026
Violaceae	*Leonia glycycarpa* Ruiz & Pav.	T	225	64	Holocene	55.0	45.4	4.12	0.027
Myristicaceae	*Virola calophylla* Warb.	T	171	58	Pleistocene	51.6	41.9	4.42	0.033
Monimiaceae	*Siparuna cuspidata* (Tul.) A. DC.	T	120	31	Pleistocene	36.8	25.6	4.94	0.035
Arecaceae	*Iriartea deltoidea* Ruiz & Pav.	T	432	60	Holocene	53.7	44.1	5.10	0.052
Meliaceae	*Guarea gomma* Pulle	T	105	44	Pleistocene	38.2	33.7	4.87	0.179
Euphorbiaceae	*Pausandra trianae* (Müll. Arg.) Baill.	T	306	22	Pleistocene	21.2	18.8	4.10	0.242
Violaceae	*Rinorea viridifolia* Rusby	T	355	27	Holocene	24.9	22.9	4.90	0.292
Burseraceae	*Protium neglectum* Swart.	T	237	53	Pleistocene	41.8	40.0	5.04	0.315
Arecaceae	*Oenocarpus mapora* H. Karst.	T	281	68	Holocene	45.3	48.2	4.08	0.753
Arecaceae	*Euterpe precatoria* Mart.	T	387	72	Pleistocene	46.8	50.4	3.98	0.823
Meliaceae	*Trichilia quadrijuga* Kunth.	T	106	39	Holocene	24.5	30.7	4.73	0.987

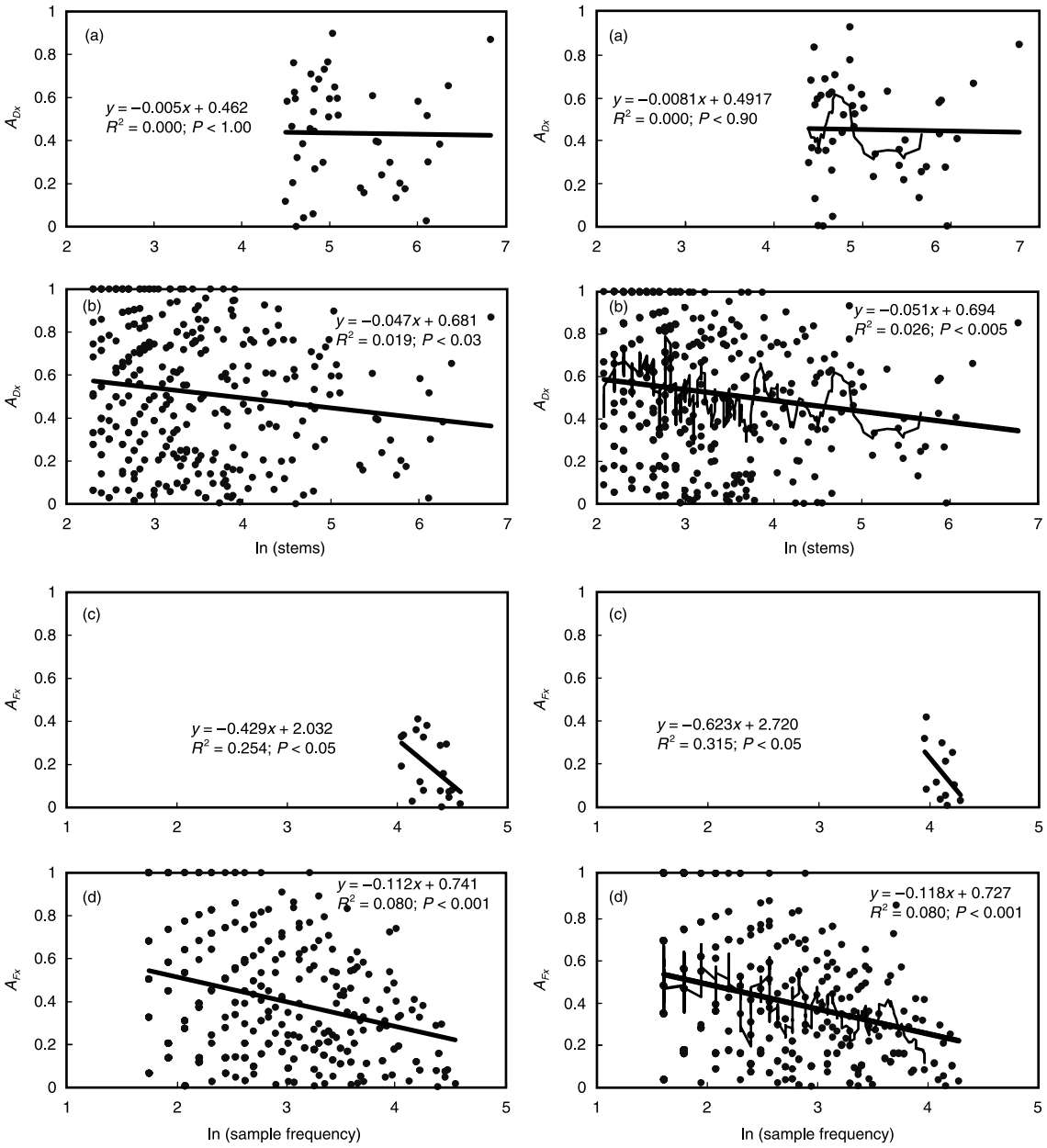

Fig. 4 Habitat association indices vs. density and frequency across the landscape. Left side, all species; right side, trees only. A_{Dx} and A_{Fx} vary between 0 (= no habitat association) and 1 (= all stems found in one habitat or the other); see text for details. (a) Density–independent habitat association vs. ln (density) for tree species; all species with ≥ 10 stems per hectare. Lines depict a linear regression and a 10-point moving average. (b) Density–independent habitat association vs. ln (density) for tree species; all species with ≥ 1 stem per hectare. Lines depict a linear regression and a 10-point moving average. (c) Frequency–independent habitat association vs. ln (frequency); all tree species present in 50 or more samples. Lines depict a linear regression. (d) Frequency–independent habitat association vs. ln (frequency), all tree species present in five or more samples. Lines depict a linear regression and a 10-point moving average.

associated with one habitat. These patterns are equivalent for the larger data set of all sampled species and the slightly smaller subset of tree species. For both groups the rate of habitat association also falls off with decreasing population density. This may simply be an artifact of incomplete sampling of the rare species, or it might reflect a real underlying pattern of rarer species being more randomly distributed across the landscape.

To help evaluate whether species' habitat association is in fact related to rarity in the landscape, we investigated how the density-independent and frequency-independent indices of each species' habitat association (A_{Dx}, A_{Fx}) vary with stem density and frequency across all samples. Among the most dense species, values of A_{Dx} are invariate with stem density, but once lower density species are included there is a weak but significant negative relationship between a species density and its degree of habitat association (Fig. 4a,b). Frequency-independent habitat association (A_{Fx}) is negatively correlated with species' frequencies in samples across

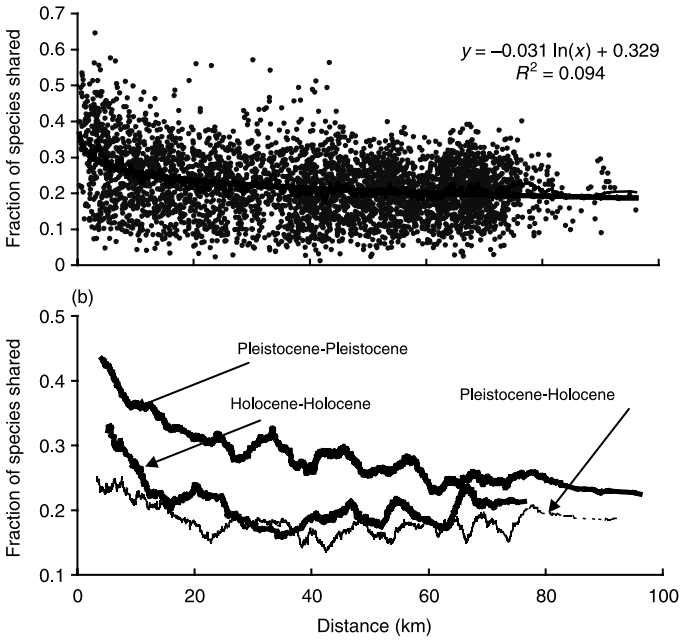

Fig. 5 Spatial decay in floristic similarity between sample pairs. Sørensen Index (proportion of shared species) is plotted as a function of straight-line distance between plot centroids. Graphs show (a) floristic comparisons between all sample pairs, and (b) comparisons separately for Pleistocene, Holocene, and Pleistocene-Holocene pairs, with 50-sample moving averages for tree species. Results for all species (not shown) are essentially identical. Floristic similarity can be described by a log-function of distance, i.e. Pleistocene pairs: S.I. = $-0.0537 \times \ln(\text{distance}) + 0.846$, *R*-squared = 0.334; Holocene pairs: S.I. = $-0.0332 \times \ln(\text{distance}) + 0.557$, *R*-squared = 0.151; Pleistocene – Holocene pairs: S.I. = $-0.0210 \times \ln(\text{distance}) + 0.440$, *R*-squared = 0.060.

more than an order of magnitude range (Fig. 4c,d). Neither A_{Dx} or A_{Fx} are strongly related to habitat (comparing mean values for trees vs. other self-supporting plants, $t = 1.09$, $P = 0.28$, d.f. = 109; $t = 0.85$, $P = 0.36$, d.f. = 109, respectively, for A_{Dx} and A_{Fx}).

ANALYTICAL APPROACH 2: FLORISTIC COMPOSITION AS A FUNCTION OF ENVIRONMENTAL AND GEOGRAPHICAL DISTANCE

Spatial decay in floristic similarity

Distant sites are less likely to share species than close sites, but the increase in species turnover reaches an asymptote for distances > *c*. 20 km and overall the relationship is rather weak (Fig. 5a). If the first order of environmental variation is accounted for by only comparing sites within landscape units, then up to 33% of the residual variation can be accounted for by spatial decay (Fig. 5b). However, the pattern of spatial decay differs with environmental conditions, such that at short distances samples on Holocene surfaces are less similar than are samples on Pleistocene surfaces.

Non-metric multidimensional distance scaling ordination

This analysis resulted in a two-dimensional solution, which posthoc analysis showed accounted for 74% (axis 1 = 25%, and axis 2 = 49%) of the tree compositional variation across the terra firme landscape. The main axis of floristic variation correlates with an edaphic gradient from low to high exchangeable cation concentrations (Table 5), broadly reflecting the gradient from Pleistocene to Holocene substrates (Fig. 6). As might be expected, the set of species that is most strongly associated with the main NMDS axis (Table 6) strongly overlaps with the set of species that has greatest indicator value for the simple Pleistocene/Holocene habitat dichotomy (Table 4). This set of species is completely dominated by trees, but our tree and non-tree species are indistinguishable across the landscape in terms of mean axis 2 scores (for all species with at least two individuals, mean value of *tau* for trees = 0.137, mean value of *tau* for other self-supporting plants = 0.123, test of difference: $t = 1.50$, $P = 0.16$, d.f. = 125).

Mantel's tests

We used the Mantel's test on matrix correlation to explore the relationship between environmental variables and floristic variation, and conducted partial Mantel tests to control for effects of spatial autocorrelation. Variation in each environmental variable except for silt and sand concentrations is significantly correlated with floristic variation, irrespective of whether or not the spatial pattern of intersample distance was accounted for (Fig. 7), and was equivalent whether the analysis was restricted to trees or expanded to include

Table 5 Association of each environmental variable with NMDS axis 2 and axis 1

Variable	Axis 2 R^2	Axis 2 tau	Axis 1 R^2	Axis 1 tau
Drainage	0.318*	−0.339	0.002	−0.040
pH	0.292	0.380	0.097	0.211
Al	0.241	−0.366	0.281*	−0.366
Ca	0.785**	0.673**	0.266*	0.378
K	0.253	0.297	0.018	−0.055
Mg	0.703**	0.587**	0.110	0.202
Na	0.376*	0.412	0.019	0.145
Al/ECEC	0.692**	−0.658**	0.286*	−0.375
ECEC	0.433*	0.393	0.030	0.110
P	0.217	0.282	0.133	0.238
LOI percentage	0.056	0.140	0.003	0.094
Dry matter percentage	0.080	−0.235	0.001	−0.011
Clay	0.077	0.148	0.000	0.046
Silt	0.060	0.145	0.001	−0.064
Sand	0.350*	−0.398	0.001	−0.035
> 0.063 fraction	0.452*	0.443	0.014	0.079

*$P < 0.05$; **$P < 0.01$.

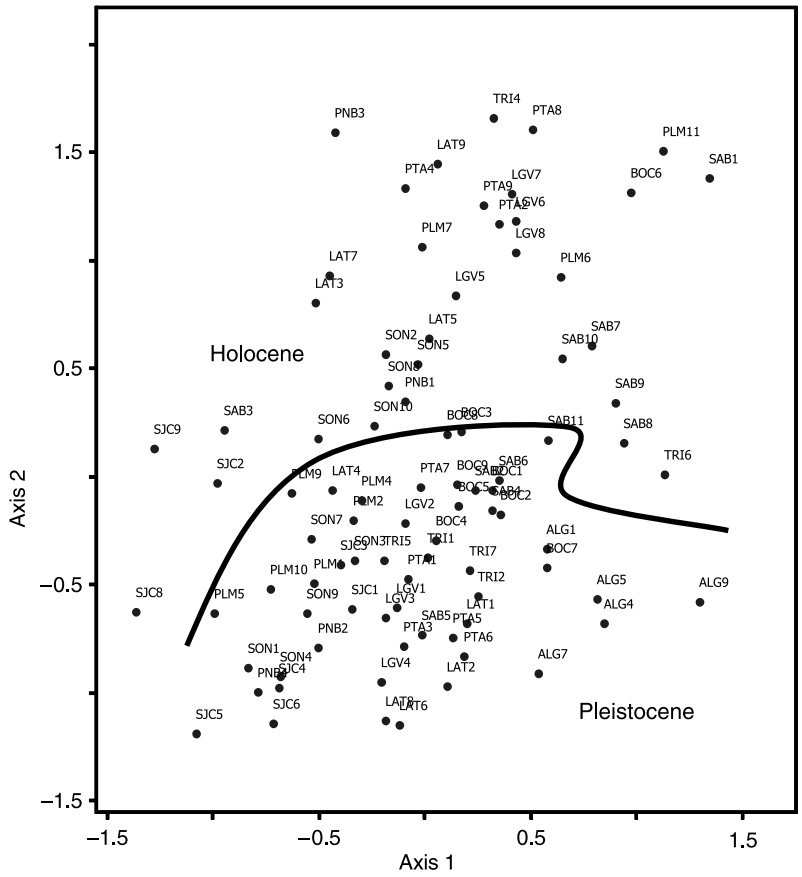

Fig. 6 Non-metric multidimensional sample scores (first two axes), showing relative position of sample plots in tree species space. Results for all species (not shown) are essentially identical. Curve corresponds to the division between Holocene and Pleistocene sites, except for two Holocene sites slightly below the line: BOC9 and BOC7.

all species. The partial Mantel tests show that distance between samples is weakly correlated with only one environmental variable (Na), consistent with our aim of sampling both geomorphic units across the whole landscape.

Multiple regression of distance matrices

Irrespective of whether analyses are performed on the entire floristic data set or on the tree subset, distance between sites accounts for 10% of the variation, and

Table 6 20 most influential species in generating NMS structure, as measured by correlation with axis 2 and sorted by axis 2 score. All species are trees

Family	Genus, Species	nms axis 2, score	Axis 2, r^2	Axis 2, tau
Cecropiaceae	*Pourouma minor*	–0.636	0.265	–0.511
Bignoniaceae	*Jacaranda copaia* ssp. *spectabilis*	–0.540	0.205	–0.392
Rubiaceae	*Amaioua corymbosa*	–0.516	0.267	–0.461
Myristicaceae	*Iryanthera juruensis*	–0.514	0.206	–0.439
Monimiaceae	*Siparuna decipiens*	–0.450	0.473	–0.563
Moraceae	*Pseudolmedia laevis*	0.461	0.304	0.454
Moraceae	*Sorocea pileata*	0.605	0.412	0.413
Annonaceae	*Unonopsis floribunda*	0.645	0.344	0.543
Sterculiaceae	*Theobroma cacao*	0.676	0.330	0.446
Arecaceae	*Astrocaryum murumuru*	0.686	0.227	0.452
Annonaceae	*Oxandra acuminata*	0.718	0.322	0.417
Arecaceae	*Attalea butyracea*	0.745	0.234	0.458
Bombacaceae	*Quararibea witii*	0.905	0.224	0.373
Myrsinaceae	*Stylogyne ambigua*	0.941	0.208	0.354
Meliaceae	*Guarea macrophylla* ssp. *pachycarpa*	0.978	0.234	0.495
Polygonaceae	*Triplaris americana*	1.025	0.280	0.396
Flacourtiaceae	*Hasseltia floribunda*	1.055	0.268	0.397
Sapotaceae	*Pouteria ephedrantha*	1.111	0.236	0.383
Phytolaccaceae	*Gallesia integrifolia*	1.123	0.213	0.372
Capparaceae	*Capparis macrophylla*	1.136	0.221	0.334

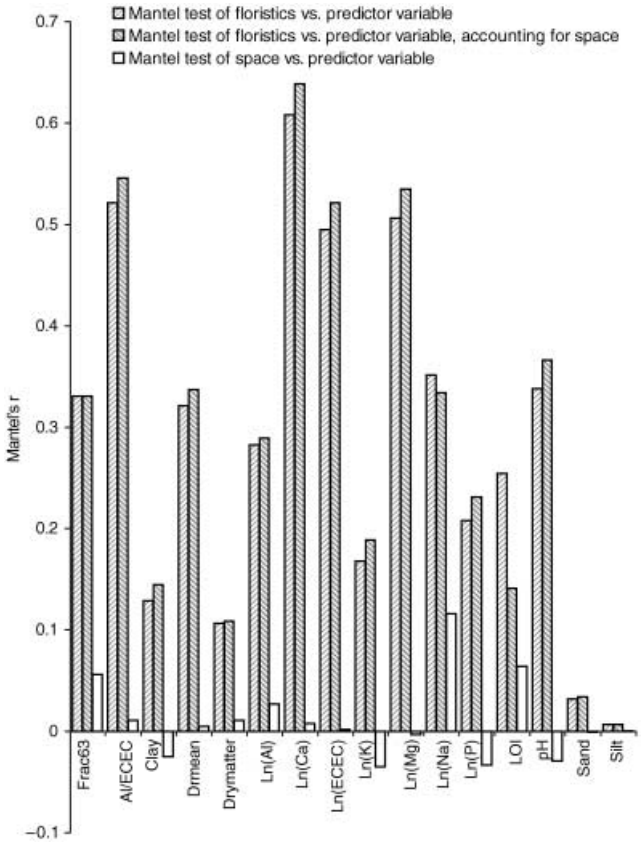

Fig. 7 Mantel tests on matrix correlation, showing relationship between environmental variables and tree species floristic variation, before and after accounting for effects of spatial autocorrelation, and between environmental variables and distance. Results for all species (not shown) are essentially identical. Values of $r > 0.075$ are significant at $P = 0.05$; values > 0.095 are significant at $P = 0.01$; values > 0.130 are significant at $P = 0.001$.

pure environmental factors account for more than 40% (Table 7). Of the floristic variation related to soil factors, a large portion is related to the soil's chemical attributes (cations), but soil drainage factors significantly improve the overall model.

Discussion

The discussion that follows applies primarily to tree species, which contribute > 86% of species and > 95% of individuals to our fully identified samples (Table 3),

Table 7 Multiple regression models of floristic variation based on environmental and spatial factor distance matrices, established by a combination of forward and backward elimination (see text for details)

(a) Final models for floristic variation

	All species		Trees only	
Variable	b, partial regression coefficient	*P*-value in final model	b, partial regression coefficient	*P*-value in final model
Drainage	0.147	< 0.001	0.149	< 0.001
ln (Ca)	0.323	< 0.001	0.298	< 0.001
ln (ECEC)	0.148	< 0.001	0.152	< 0.001
Al/ECEC	0.178	< 0.001	0.184	< 0.001
ln (space)	0.311	< 0.001	0.305	< 0.001

(b) Effect of removing the variable on model *r*-squared, when variables removed by backwards elimination

	All species		Trees only	
Removal of variable	Δr^2	*P*-value in the model that includes that variable and all listed below	Δr^2	*P*-value in the model that includes that variable and all listed below
drainage	–2.0	< 0.001	–2.1	< 0.001
ln (Ca)	–0.8	< 0.001	–0.7	< 0.001
ln (ECEC)	–11.9	< 0.001	–12.1	< 0.001
Al/ECEC	–26.1	< 0.001	–26.4	< 0.001
ln (space)	–10.1	< 0.001	–9.8	< 0.001

and therefore dominate the results. Our data set of non-trees (110 species, 872 individuals) is too small to analyse separately, and we therefore draw no firm conclusions about their behaviour. However, we note that including them did not substantially affect any analytical result.

Several (but not all) previous studies on non-flooded tropical forests have shown correlations between plant species and soil conditions, at a range of spatial scales. However, the degree to which the contemporary environment controls the landscape-scale distribution of tropical trees remains contentious, especially in Amazonia, which is dominated by oxisols and ultisols (e.g. Sombroek 2001). This study has sought to address this question by tackling three problems that have traditionally complicated interpretation: first, the difficulty of obtaining sufficient replication and spatial coverage of species-rich forest communities; secondly, inadequate characterization of the edaphic environment; and thirdly the lack of explicit treatment and quantification of spatial autocorrelation.

SINGLE SPECIES ANALYSIS

Our results suggest that very few species are habitat-specific when this is defined most narrowly as 'habitat-restricted'. This confirms recent findings (e.g. Pitman *et al*. 1999) that only a small proportion of Amazon tree species are confined to specific unflooded habitats. This result in itself is not necessarily inconsistent with the paradigm of forest communities being substantially structured by associations between species and local soil conditions; most forest landscapes are patchy mosaics of varying physical conditions and so few species are likely to be entirely immune from 'mass effects' (Shmida & Wilson 1985). For most species, seed dispersal between habitat patches may be a frequent event and therefore only very strong habitat-specific competition or preferences in growing conditions would prevent occasional regeneration of a species in the 'wrong' habitat.

When the definition of habitat association is relaxed, analyses of frequency- and density-distributions of populations demonstrate that a large fraction of species have a significant tendency to be favoured by one habitat or another, and that this is not an artifact of the spatial structure of the landscape. The proportion of species with significant habitat associations approaches 80% for those relatively well-sampled species with > 100 individuals, a greater proportion than in any other study we know of in a lowland tropical forest landscape. This could be a result of either an exceptional degree of habitat specialization in the species we studied, unlikely given the lack of extreme soil types in our analysis, or simply of relatively effective sampling at many sites across the landscape creating reasonable statistical power.

Most of the species *significantly* associated with one habitat or another are relatively densely distributed across the landscape, such that species averaging > 10 individuals ha^{-1} are nearly five times more likely to be habitat-indicators than species averaging 0.1 trees ha^{-1} (Fig. 3), but the *tendency* to habitat association is weakly *negatively* correlated with species overall population

densities and frequencies (Fig. 4). We draw two conclusions from this. First, the overall habitat association observed is not driven by a relatively small subset of more dense and frequent species, but appears to be a general property of many species across the landscape. Secondly, there is certainly no evidence of a decline in habitat association for species with low population densities: the low proportion of low-density and low-frequency species proven to be associated with one habitat or another reflects inadequate sampling across the landscape for these species rather than a genuine biological phenomenon.

Pitman *et al*. (1999, 2001) have shown that relatively few tree species dominate forests in west Amazonian landscapes, suggesting that dominant species may be relatively indifferent to environmental heterogeneity. Our data appear to offer only weak support for this view, as although more common species do have lower than average degrees of habitat association, there are several very common species that associate strongly with one or other of our terra firme habitats. Examples include *Phenakospermum guyanense* (Rich.) Endl., *Siparuna decipiens* (Tul.) A. DC. and *S. cristata* (Poepp. & Endl.) A. DC, all of which are strongly associated with Pleistocene landscapes and present at densities of > 40 individuals ha^{-1}. However, we note that these exceptions are understory or subcanopy species (and in the case of *Phenakospermum*, an over-sized herb) that represent relatively little biomass or productivity and so are hardly 'dominants'. Thus, the regional pattern that emerges from our data helps to reconcile two apparently contradictory views of how Amazon forests may be put together: on the one hand that they are largely homogeneous, even simple, communities, insensitive to edaphic changes across landscapes and consistently dominated by an oligarchy of widely successful species; and on the other that they are complex and strongly patchy systems affected by even subtle changes in environmental conditions.

It appears that a relatively small group of species are able to physically dominate forests almost regardless of soil conditions, and yet that most tree species are associated with soil factors. Our estimates of the degree of species-level habitat association in our landscape may be still too low, as statistical power to detect relationships is limited by insufficient sample size among the low-density species. Our data do not allow us to comment definitively on whether habitat-specificity is higher still for smaller non-trees such as understory palms (Ruokolainen & Vormisto 2000), melastomataceous shrubs (Ruokolainen *et al*. 1997) and terrestrial pteridophytes (Tuomisto & Poulsen 2000; Vormisto *et al*. 2000). However, species distributions in these groups are known to be highly non-random elsewhere in western Amazonia (Tuomisto *et al*. 2003a,b).

COMMUNITY ANALYSES ALONG GRADIENTS

Is it possible to reconcile our findings with a third, non-equilibrial, view of how forests are put together (Hubbell *et al*. 1999; Hubbell 2001), which holds that forests are dominated by stochastic processes such that composition is substantially affected by the species that happen to be able to disperse into and recruit in a locality? In our second set of analyses, we examined how landscape-scale species composition varied along environmental gradients, and attempted to identify the combined community-level effects of largely spatial processes such as dispersal limitation. Tree species composition is affected by a general edaphic gradient from nutrient poor sandy-clay soils to relatively fertile silty-clay soils (Table 5). The gradient reflects the broad Pleistocene/Holocene landscape categories, but the actual variation in soil and floristic conditions (Table 6, Fig. 7) is more complex than that simple dichotomy might suggest. A substantial proportion (10%) of the variation in forest composition is explained by the distance between samples: this could be caused by 'classical' distance decay, whereby populations are somewhat patchy due to distance-dependent population processes (reflecting individual species' dispersal kernels and the historical happenstance of where populations have been in the past), or an interaction with edaphic factors that themselves in principle could show distance decay. However, as we have shown that spatial and environmental factors are mostly unrelated to one another at the landscape scale, this portion of the floristic variation is likely to be attributable to distance-dependent processes such as dispersal limitation.

These processes do not operate equally everywhere: between pairs of Holocene samples no spatial decay in composition is evident over a distance of $> c.$ 10 km, suggesting efficient mixing of species, yet between Pleistocene pairs similarity continues to decrease gently with distance across the whole study region, indicating that species may be poorly dispersed within this geomorphic unit. Additional multiple regression analyses of the whole data set using un-transformed raw distance matrices (not shown) were only able to attribute 3% of floristic variation to space, supporting the inference that the effects of dispersal limitation are only substantially felt over short distances. These interpretations are complicated by the fact that we showed that the pattern of spatial decay in the two landscape units is dissimilar. There is also a possibility that some additional, undetected, spatial effects might operate at finer scales within habitats.

How do these results compare with other forests? Recent work by Condit *et al*. (2002) showed that distance-decay is much greater in a Panamanian landscape than in two Amazonian landscapes. Our results confirm that in parts of western Amazonia, at least, beta-diversity is less than in central Panama. However, as the Panamanian landscape has strongly spatial climatic gradients it was difficult to quantify the likely drivers of high beta-diversity, which could be related to climate, geology, dispersal, or all three (Pyke *et al*. 2001; Duivenvoorden *et al*. 2002). In our data set, beta-diversity is greatest,

and the relationship between distance and composition weakest, for Pleistocene-Holocene sample pairs, illustrating the dominant effect of this environmental contrast in structuring tree species composition. Multiple regression of distance matrices confirm this, showing that a much greater proportion (40%) of landscape-scale floristic variation is explained purely by soil conditions than by the effects of spatial autocorrelation (10%). The chemical composition of the soil, particularly the concentrations of base cations and Aluminium, appears to be the most important factor, with a small role for drainage conditions. The soil structural variables that we measured (relative contribution of clay, sand and silt) do not contribute to the best multiple regression models, suggesting that any effect they might have on species composition is mediated by soil chemistry and drainage.

There are several reasons to suspect that the combined effects of substrate on landscape-scale forest floristics have been underestimated in our analysis. First, although we sampled a wide range of soil variables, we may not have captured all the substrate variation that may be biologically meaningful. For example: (i) our samples were confined to the A horizon; (ii) we did not measure nitrogen or various biologically active trace elements; (iii) we did not quantify the mycoflora; (iv) our estimate of drainage conditions was necessarily superficial; and (v) we have not evaluated possible temporal variations in nutrient supply (cf. Newbery *et al*. 1988). Secondly, the extent of each sample is 1.8 ha, yet a portion of environmental control is known to be mediated at a much finer spatial scale than this (e.g. Vormisto *et al*. 2000; Harms *et al*. 2001), so we did not achieve a perfect microtopographical representation of the conditions in which each plant was growing. Thirdly, our samples include all individuals as small as 2.5 cm diameter and are dominated by juveniles of trees. Environment-mediated effects such as competition may continue to operate up to much larger size-classes, possibly leading to better mapping onto substrates by adult populations than juvenile populations (e.g. Webb & Peart 2000). Finally, our samples are subject to substantial sampling error as we only inventoried a fraction of species present in each locality.

Regardless of these concerns, our results confirm that tree species' composition within a lowland Amazon landscape respond strongly to comparatively small variations in soil, and evidence suggests that the distributional patterns of most species segregate by geomorphic unit even within a landscape with comparatively little soil variation. We emphasize that finding pervasive and deterministic habitat association among Amazon tree species does not prove pervasive physiological *specialization* in particular soil environments: other biological processes could contribute to the empirical patterns observed in the realized niches of these species. Indicator species of Pleistocene terraces, for example, may simply tolerate the low-nutrient, high-aluminium conditions, rather than depend on them, and their lower densities in the richer soil habitats may be due to the effects of interspecific competition. Dispersal limitation appears to operate mostly over relatively short distances, confirming findings from elsewhere. The overall weak effect of geographical proximity at the landscape scale suggests that these species-rich Amazon forest communities are structured more by *in situ* processes mediated in a deterministic way by substrate conditions, than they are by spatial processes. Clearly, experimental approaches will be needed to tease apart the biological mechanisms by which the distributions of Amazon tree species are non-randomly controlled.

Acknowledgements

We are grateful to several colleagues for their assistance during the field inventories, and especially to Fernando Cornejo, César Chacón, Alejandro Farfan and Wilfredo Ramirez. We thank Peruvian Safaris S.A., Bahuaja Lodge, and the people of the communities of La Torre, Palma Real, Alegría, Tres Islas, Sandoval, Jorge Chávez, Loero, Sonene, Puerto Arturo, Lago Valencia, Boca Pariamanu and Sabaluyo for their generous hospitality. We acknowledge invaluable help given by Fernando Cornejo in identifying some collections, and helpful discussions with Tim Baker, Pippa Chapman, Jon Lloyd, Kalle Ruokolainen and David Wood, as well as the contributions of three anonymous referees. Mark Newcombe helped develop Fig. 1. We thank INRENA for providing the necessary permits for conducting the work and UNSAAC and its herbarium for support. This work was funded by grant ERP-196 to the University of Leeds from the UK Department for International Development (Environment Research Department), and a UK Natural Environment Research Council Research Fellowship to OP.

Supplementary material

The following material is available from http://www.blackwellpublishing.com/products/journals/suppmat/JEC/JEC815/JEC815sm.htm

Appendix S1 Location and geomorphology of each forest sample in eastern Madre de Dios.

Appendix S2 Soil chemistry and structure of each forest sample in eastern Madre de Dios.

References

Anderson, A.B. (1981) White-sand vegetation of Brazilian Amazonia. *Biotropica*, **13**, 199–210.

Clark, D.A. & Clark, D.B. (1992) Life history diversity of canopy and emergent trees in a neotropical rain forest. *Ecological Monographs*, **62**, 315–344.

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, R.M. & Castro, 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.

Condit, R., Ashton, P.S., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N. *et al.* (2000) Spatial patterns in the distribution of tropical tree species. *Science*, **288**, 1414–1418.

Condit, R., Ashton, P.S., Manokaran, N., LaFrankie, J.V., Hubbell, S.P. & Foster, R.B. (1999) Dynamics of the forest communities at Pasoh and Barro Colorado: comparing two 50-ha plots. *Philosophical Transactions of the Royal Society of London Series B. Biology Sciences*, **354**, 1739–1748.

Condit, R., Hubbell, S.P. & Foster, R.B. (1996) Assessing the response of plant functional types to climatic change in tropical forests. *Journal of Vegetation Science*, **7**, 405–416.

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

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**, 2005–18.

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

Duellman, W.E. & Koechlin, J.E. (1991) The Reserva Cuzco Amazónico, Peru: biological investigations, conservation, and ecotourism. *University of Kansas Museum of Natural History Occasional Papers*, **142**, 1–38.

Dufrene, M. & Legendre, P. (1997) Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs*, **67**, 345–366.

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., Cavelier, J. & Duivenvoorden, J.F. (2002) Different floristic patterns of woody understorey and canopy plants in Colombian Amazonia. *Journal of Tropical Ecology*, **18**, 499–525.

Enquist, B.J. & Niklas, K.J. (2001) Invariant scaling relations across tree-dominated communities. *Nature*, **410**, 655–660.

Faith, D.P. & Ferrier, S. (2002) Linking beta diversity, environmental variation, and biodiversity assessment. *Science dEbate* http://www.sciencemag.org/cgi/eletters/295/5555/636.

Fleck, D.W. & Harder, J.D. (2000) Matses Indian rainforest habitat classification and mammalian diversity in Amazonian Peru. *Journal of Ethnobiology*, **20**, 1–36.

Gentry, A.H. (1982) Patterns of neotropical plant species diversity. *Evolutionary Biology*, **15**, 1–84.

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

Gentry, A.H. (1988b) Tree species richness of upper Amazonian forests. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 156–159.

Grubb, P.J. (1977) The maintenance of species-richness in plant communities: the importance of the regeneration niche. *Biology Reviews*, **52**, 107–145.

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.

Hart, T.B., Hart, J.A. & Murphy, P.G. (1989) Monodominant and species-rich forests of the humid tropics: causes for their co-occurrence. *American Naturalist*, **133**, 613–633.

Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, New Jersey.

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

Hubbell, S.P., Foster, R.B., O'Brien, S.T., Harms, K.E., Condit, R., Wechsler, B. *et al.* (1999) Light gap disturbances, recruitment limitation, and tree diversity in a neotropical rainforest. *Science*, **283**, 554–557.

Kahn, F. & Mejia, K. (1990) Palm communities in wetland forest ecosystems of Peruvian Amazonia. *Forest Ecology and Management*, **33**, 169–179.

Kalliola, R., Puhakka, M., Salo, J., Tuomisto, H. & Ruokolainen, K. (1991) The dynamics, distribution and classification of swamp vegetation in Peruvian Amazonia. *Annales Botanicae Fennici*, **28**, 225–239.

Kalliola, R., Ruokolainen, K., Tuomisto, H., Linna, A. & Mäki, S. (1998) Mapa geoecológico de la zona de iquitos y variación ambiental. *Geoecología y Desarrollo Amazónico: Estudio Integrado en la Zona de Iquitos, Perú*. Annales Universitatis Turkuensis Series A II 114 (eds R. Kalliola & S. Flores Paitán), pp. 443–457. University of Turku, Finland.

Knight, D.H. (1975) A phytosociological analysis of species-rich tropical forest on Barro Colorado Island, Panama. *Ecological Monographs*, **45**, 259–284.

Kruskal, J.B. (1964) Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, **29**, 115–129.

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

Legendre, P. & Vaudor, A. (1991) *The Royal-Package: Multidimensional Analysis, Spatial Analysis*. Département de Sciences Biologiques, Université de Montréal, Montreal.

Ludwig, J.A. & Reynolds, J.F. (1988) *Statistical Ecology*. John Wiley and Sons, New York.

Malhi, Y., Baker, T.R., Phillips, O.L., Almeida, S., Alvarez, E., Arroyo, L. *et al.* (2004) The above-ground wood productivity and net primary productivity of 104 neotropical forests. *Global Change Biology*, in press.

Malhi, Y., Phillips, O.L., Baker, T.R., Almeida, S., Frederiksen, T., Grace, J. *et al.* (2002) An international network to understand the biomass and dynamics of Amazonian forests (RAINFOR). *Journal of Vegetation Science*, **13**, 439–450.

Mantel, N. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.

McCune, B. & Mefford, M.J. (1999) *PC-ORD. Multivariate Analysis of Ecological Data, Version 4*. MjM Software Design, Gleneden Beach, Oregon.

Newbery, D.McC., Alexander, I.J., Thomas, D.W. & Gartlan, J.S. (1988) Ectomycorrhizal rain-forest legumes and soil phosphorous on Korup National Park, Cameroon. *New Phytologist*, **109**, 433–450.

Newbery, D.McC. & Proctor, J. (1983) Ecological studies in four contrasting lowland rain forests in Gunung Mulu national park, Sarawak. IV. Associations between tree distribution and soil factors. *Journal of Ecology*, **72**, 475–493.

Osher, L.J. & Buol, S.W. (1998) Relationship of soil properties to parent material and landscape position in eastern Madre de Dios, Peru. *Geoderma*, **83**, 143–166.

Phillips, O.L., Malhi, Y., Higuchi, N., Laurance, W.F., Nuñez, V.P., Vásquez, M.R. *et al.* (1998) Changes in the carbon balance of tropical forest: evidence from long-term plots. *Science*, **282**, 439–442.

Phillips, O.L. & Miller, J.S. (2002) *Global Patterns of Forest Diversity: the Dataset of Alwyn Gentry*. *Monographs in Systematic Botany*, *Volume 89*. Missouri Botanical Garden, St Louis, Missouri.

Phillips, O.L., Vásquez Martínez, R., Núñez Vargas, P., Lorenzo Monteagudo, A., Chuspe Zans, M.-E., Galiano Sánchez, W. *et al.* (2003) Efficient plot-based tropical forest floristic assessment. *Journal of Tropical Ecology*, in press.

Pires, J.M. & Prance, G.T. (1985) The vegetation types of the Brazilian Amazon. *Key Environments: Amazonia* (eds G.T. Prance & T.E. Lovejoy), pp. 109–145. Pergamon, Oxford.

Pitman, N.C., Terborgh, J., Silman, M.R. & Núñez, P.V. (1999) Tree species distributions in an upper Amazonian forest. *Ecology*, **80**, 2651–2661.

Pitman, N.C., Terborgh, J.W., Silman, M.R., Núñez, P.V., Neill, D.A., Ceron, C.E. *et al.* (2001) Dominance and distribution of tree species in upper Amazonian terra firme forests. *Ecology*, **82**, 2101–2117.

Poore, M.E.D. (1968) Studies in Malaysian rain forest. I. The forest on Triassic sediments in Jengka forest reserve. *Journal of Ecology*, **56**, 143–196.

Poulsen, A.D. (1996) Species richness and density of ground herbs within a plot of lowland rainforest in north-west Borneo. *Journal of Tropical Ecology*, **12**, 177–190.

Pyke, C.R., Condit, R., Aguilar, S. & Lao, S. (2001) Floristic composition across a climatic gradient in a neotropical lowland forest. *Journal of Vegetation Science*, **12**, 553–566.

Räsänen, M., Neller, R., Salo, J. & Jungner, H. (1992) Recent and ancient fluvial deposition systems in the Amazonian foreland basin, Peru. *Geological Magazine*, **129**, 293–306.

Räsänen, M., Salo, J. & Jungner, H. (1991) Holocene floodplain lake-sediments in the Amazon – C-14 dating and paleoecological use. *Quaternary Science Reviews*, **10**, 363–372.

Räsänen, M., Salo, J., Jungner, H. & Romero Pittman, L. (1990) Evolution of the western Amazon lowland relief: impact of Andean foreland dynamics. *Terra Nova*, **2**, 320–332.

Rees, M., Condit, R., Crawley, M., Pacala, S. & Tilman, D. (2001) Long-term studies of vegetation dynamics. *Science*, **293**, 650–655.

Ruokolainen, K., Linna, A. & Tuomisto, H. (1997) Use of Melastomataceae and pteridophytes for revealing phytogeographic patterns in Amazonian rain forests. *Journal of Tropical Ecology*, **13**, 243–256.

Ruokolainen, K. & Vormisto, J. (2000) The most widespread Amazonian palms tend to be tall and habitat generalists. *Basic and Applied Ecology*, **1**, 97–108.

Sabatier, D., Grimaldi, M., Prévost, M.-F., Guillaume, J., Godron, M., Dosso, M. *et al.* (1997) The influence of soil cover organization on the floristic and structural heterogeneity of a Guianan rain forest. *Plant Ecology*, **131**, 81–108.

Salo, J. & Kalliola, R. (1990) River dynamics and natural forest regeneration in the Peruvian Amazon. *Rain Forest Regeneration and Management*. Volume 6. Man in the Biosphere Series (eds A. Gómez-Pompa *et al.*), pp. 245–256. UNESCO and Cambridge University Press, Cambridge.

Salo, J., Kalliola, R., Hakkinen, I., Makinen, Y., Niemela, P., Puhakka, M. *et al.* (1986) River dynamics and the diversity of Amazon lowland forest. *Nature*, **322**, 254–258.

Shephard, G.N. Jr, Yu, D.W., Lizarralde, M. & Italiano, M. (2001) Rain forest habitat classification among the Matsigenka of the Peruvian Amazon. *Journal of Ethnobiology*, **21**, 1–38.

Shephard, G.N. Jr, Yu, D.W. & Nelson, B.W. (2003) Ethnobotanical ground-truthing and forest diversity in the Western Amazon. *Advances in Economic Botany*, in press.

Shmida, A. & Wilson, M.V. (1985) Biological determinants of species diversity. *Journal of Biogeography*, **12**, 1–20.

Smouse, P.E., Long, J.C. & Sokal, R.R. (1986) Multiple regression and correlation extensions of the mantel test of matrix correspondence. *Systematic Zoology*, **35**, 627–632.

Sollins, P. (1998) Factors influencing species composition in tropical lowland rain forest: does soil matter? *Ecology*, **79**, 23–30.

Sombroek, W. (2001) Spatial and temporal patterns of Amazon rainfall – consequences for the planning of agricultural occupation and the protection of primary forests. *Ambio*, **30**, 388–396.

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.

ter Steege, H., Sabatier, D., Castellanos, H., Van Andel, T., Duivenvoorden, J., De Oliveira, A.A. *et al.* (2000) An analysis of the floristic composition and diversity of Amazonian forests including those of the Guiana Shield. *Journal of Tropical Ecology*, **16**, 801–828.

Svenning, J.C. (1999) Microhabitat specialization in species-rich palm community in Amazonian Ecuador. *Journal of Ecology*, **87**, 55–65.

Swaine, M.D. (1996) Rainfall and soil fertility as factors limiting forest species distributions in Ghana. *Journal of Ecology*, **84**, 419–428.

Terborgh, J. & Andresen, E. (1998) The composition of Amazonian forests: patterns at local and regional scales. *Journal of Tropical Ecology*, **14**, 645–664.

Tuomisto, H. & Poulsen, A.D. (2000) Pteridophyte diversity and species composition in four Amazonian rain forests. *Journal of Vegetation Science*, **11**, 383–396.

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.

UNESCO. (1980) *Mapa de la Vegetación de America del Sur, 1 : 5 000 000*. Institut de la Carte Internationale du Tapis Vegetal, Toulouse.

Van Reeuwijk, L.P. (1995) *Procedures for Soil Analysis*. Technical Paper 9. International Soil Reference and. Information Centre, Wageningen, The Netherlands.

Vásquez, M.R. & Phillips, O.L. (2000) Floristics and ecology of a high-diversity forest at Allpahuayo, Amazonian Peru. *Annals of the Missouri Botanical Garden*, **87**, 499–527.

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.

Webb, C.O. & Peart, D.R. (2000) Habitat associations of trees and seedlings in a Bornean rain forest. *Journal of Ecology*, **88**, 464–478.

Wong, Y.K. & Whitmore, T.C. (1970) On the influence of soil properties on species distribution in a Malayan lowland dipterocarp forest. *Malaysian Forester*, **33**, 42–54.

Received 10 September 2002
revision accepted 20 June 2003