

Planning tree species diversification in Kenya based on differences in tree species composition between farms. I. Analysis of tree uses

R. Kindt^{1,*}, P. Van Damme², A.J. Simons¹ and H. Beeckman^{3,4}

¹ICRAF, P.O. Box 30677–00100, Nairobi, Kenya; ²Ghent University, Laboratory of Tropical and Subtropical Agriculture and Ethnobotany, Coupure Links 653, B-9000 Gent, Belgium; ³Ghent University, Laboratory of Plant Ecology, Coupure Links 653, B-9000 Gent, Belgium; ⁴Royal Museum for Central-Africa, Leuvensesteenweg 13, B-3080 Tervuren, Belgium; *Author for correspondence (e-mail: R.Kindt@CGIAR.org; phone: (254-2)-524000 or (1-650)-8336645; fax: (254-2)-524001 or (1-650)-8336646)

Received 13 May 2004; accepted in revised form 3 August 2005

Key words: Diversification, Hellinger distance, Landscape, Ordination, Redundancy analysis, Species composition

Abstract

Concerns exist about the limited diversity of tree species in agricultural landscapes. Complete tree inventories were carried out on 201 farms from four villages in western Kenya to establish whether significant differences in tree species composition existed between farms, and if so their magnitude and implications for new introductions and plantings. Tree species composition was interpreted to encompass elements of both tree identity and abundance. Tree identity was viewed from both taxonomic and function (e.g. fruit, timber, medicine) perspectives. Novel types of ordination using the Hellinger ecological distance and polynomial Redundancy Analysis indicated wide heterogeneity between farms with respect to tree species composition. For the 12 most prevalent functions of trees, the analyses showed significant differences ($p < 0.05$). Partitioning of variance identified that village location explained much of the differences between farms suggesting that farmers share tree species within villages more than between villages. Differences between farms were assessed on two-dimensional ordination graphs. For five important tree functions, including beverage, charcoal, construction, fodder and medicine, two species dominated the compositional differences. For these functions, diversification can be achieved by village-to-village sharing even in the absence of any new species introductions. A general process to determine the degree of tree diversity at farm and landscape levels and steps to increase it are discussed.

Introduction

Concerns exist about the limited diversity of tree species in agricultural landscapes. Whereas the majority of species on a farm may be indigenous taxa, introduced exotic taxa account for many of the trees on a farm (Simons and Leakey 2004). This observation has led to the concept of domesticating a landscape, which involves ame-

liorating tree species diversity within and between functional uses (those products and services that trees provide in a particular landscape) (Simons and Leakey 2004).

There are various reasons why landscape domestication maybe be pursued. Some of the objectives of landscape domestication may include: (i) diversification of the incomes from tree products for resource-poor farmers; (ii) increase

the stability and productivity of the agroecosystem; and (iii) enhance the conservation of biodiversity in landscape mosaics.

Tree domestication is farmer-driven and seeks to match the intra-specific diversity of many locally important tree species to the needs of subsistence farmers, including their need to create marketing possibilities for a wide range of products (Simons and Leakey 2004). Although history has provided many examples of the development of monocultures, tree domestication will not always lead to the establishment of monocultures (Leakey and Simons 1998). It seems that the domestication of a wide range of Agroforestry Tree Products (AFTPs as defined by Simons and Leakey 2004) offers the best pathway towards their commercialisation, since at the one extreme the incentives for domestication are insufficient for self-use by farmers only, whereas at the other extreme large-scale production in monocultures for international markets may sweep away the benefits for small-scale farmers (Leakey and Simons 1998; Garrity 2004). Several precedents exist of farmers opting to grow a wide diversity of tree species for local markets instead of focusing on a single tree species to produce for international markets (Leakey and Simons 1998; Kindt et al. 2004). Evidence is building that farm diversification increases the mean income and reduces the variance of income of farmers (Di Falco and Perrings 2003; Dorward et al. 2003).

A second reason to seek landscape domestication and diversification is linked to improving ecosystem functioning. Although, studying the effect of biodiversity on ecosystem functioning requires complex ecological experiments or models, studies have shown that there is a positive, but conditional relationship between species diversity and ecosystem stability and/or productivity (e.g., Loreau et al. 2001; Norberg et al. 2001; Tilman et al. 2001; Whittaker et al. 2001). The conditions include that species have diversity in traits and that the ecosystem has environmental heterogeneity. Therefore, not all mixtures of higher diversity will have beneficial effects on ecosystem functioning. One of the challenges of targeting enhancement of ecosystem functioning by increasing biodiversity is therefore to select the identities and traits of the component species. The challenge is also to match the variation in environmental conditions at a specific scale in the landscape with a mixture of

species with the appropriate traits for the environmental conditions. However, the studies have shown that on average higher diversity will have beneficial effects. Also on average, having more species in the landscape means that redundancy is higher: losing one species from a diverse landscape will have smaller consequences than losing one species from a species-poor landscape. One of the methods of achieving landscape diversification is to promote wider distribution of species that are already present somewhere in the landscape. Reasons to adopt such approach could be related to species conservation in a conservation-through-use setting (Kindt and Lengkeek 1999; Kristensen and Balslev 2003; Kindt et al. 2004), biosafety precautions, ecological suitability, or fitness of purpose of the species.

A third possible reason to domesticate landscape is to increase the value of the landscape for biodiversity conservation. Agroforestry is uniquely suited to provide ecoagriculture solutions that successfully combine objectives for increased food security and biodiversity conservation gains, especially by promoting greater use of native tree species in agroforestry systems (Atta-Krah et al. 2004; Garrity 2004; McNeely 2004; Simons and Leakey 2004). Diversification could therefore result in improved conservation, although the links between development and conservation goals need to be explored carefully (Adams et al. 2004).

The objective of the study described in this article was to establish whether significant differences in tree composition existed in a landscape in western Kenya, and if so their magnitude and implications for landscape diversification. We used ordination and regression analysis methods to analyse differences in tree composition between farms. This study is part of a larger one examining options for landscape diversification based on information on current tree diversity levels in a particular landscape (Kindt et al. 2004).

Methods

Study area

We randomly-stratified selected 201 farms (taken to mean all land managed by a single household) in the Vihiga and Kakamega districts of western Kenya. The study area is inhabited predominantly

by the Luhya (Luyia) ethnic group and belongs to the same agroecological zone where altitude ranges 1500–1800 m a.s.l., annual mean temperature ranges 18.1–20.4 °C, and annual bimodal rainfall ranges 1600–2000 mm. Four villages were selected within the area, each located in a different stratum that mainly differed in farm sizes and arrangement of woody biomass in the landscape (Bradley et al. 1985). The selection of villages coincided with a gradient towards the species-rich Kakamega Forest National Reserve. A more detailed description of the survey area can be found in Kindt et al. (2004) and references therein.

Information recorded on tree species and farm characteristics

All trees (woody perennials) were inventoried using Beentje (1994) as the key reference. For each tree species encountered on a farm, its abundance (the total number of trees on the farm) and on-farm functions (defined for this survey as all products or services provided by the species to the farming household; other studies have used the terms ‘functional uses’ or ‘products and services’, we used ‘functions’ for brevity) were recorded by farm walks and participatory interviews with household informants. Because it is possible that some informants forgot to mention some functions during the interview, we made the assumption that a species had a particular function on all farms if more than 50% of farmers and minimum five farmers with the species mentioned the function. For example, although only 145 of the 201 farmers with *Mangifera indica* listed fruit as a function, we assumed that this species had fruit as a function on all 201 farms. The criteria of minimum 50% and 5 were set arbitrarily to avoid overestimation of functions. The adjustment procedure resulted in an increase of the species-farm-use combinations from 6859 to 7526 (10%).

Information recorded on farm characteristics revealed that a fair socio-economic cross-section of the population was interviewed (Kindt et al. 2004). The farm characteristics that were collected included: (i) type of household head (categorical variable with categories male-headed, female-headed de jure = not married, and female-headed de facto = husband with off-farm occupation who is absent most of the time from the farm); (ii) type

of house (categorical variable and indicator of wealth, with poorest households having thatched-roofed houses and wealthiest households having permanent houses); (iii) farm size (continuous variable), (iv) the number of cattle (continuous variable and indicator of wealth); (v) the years that the farm had been under the present household head (continuous variable), (vi) the age of the household head (continuous variable), (vii) the number of resident children (continuous variable); and (viii) the level of schooling (continuous variable, measured by minimum number of years to reach the level). In total, 183 farms out of the 201 surveyed had complete information on farm characteristics and the analyses were restricted to this subset. Summary statistics for these variables and a verification that they are not strongly correlated (strong correlation could complicate constrained ordination and regression methods) are provided in Kindt et al. (2004).

Ordination methods

Differences in species composition can be measured by ecological distance coefficients (Jongman et al. 1995; Legendre and Legendre 1998; Quinn and Keough 2002). We used the Hellinger distance, which was shown to be among the better ecological distance measures for compositional difference (Legendre and Gallagher 2001). The Hellinger distance between site 1 and 2 is calculated by:

$$\text{Hellinger distance}(1, 2) = \sqrt{\sum_{j=1}^S \left[\sqrt{\frac{n_{1j}}{n_{1+}}} - \sqrt{\frac{n_{2j}}{n_{2+}}} \right]^2}$$

where n_{aj} is the abundance of species j at site a ; n_{a+} the total abundance of site a ; and S the total number of species for the two sites combined. Only species occurring on more than five farms were included in the calculations, since the Hellinger distance is biased towards species with low abundances. The abundance of the remaining species was $\ln(a + 1)$ transformed to avoid that very abundant species would dominate the calculation of the Hellinger distance.

Ordination methods estimate positions of sites in multidimensional space so that the distances

between the positions of the sites approximate the difference in ecological distance between the sites. Constrained ordination methods provide a direct way of investigating the influence of site characteristics on differences in species composition, and provide the significance and magnitude of compositional differences. Ordination results are typically presented in two dimensional ordination graphs that maximise the variance that can be displayed, taking advantage of the fact that many species respond to the same gradients (Økland 1996). We combined two novel ordination techniques: (i) transforming species matrices prior to Redundancy Analysis so that Euclidean distances between sites were equal to the Hellinger distance (Legendre and Gallagher 2001); and (ii) polynomial Redundancy Analysis (Makarenkov and Legendre 2002; URL <http://www.bio.umontreal.ca/legendre/indexEnglish.html> last accessed 6–2005). The novel polynomial form of RDA (PRDA) allows for second-order relationships between ordination axes and environmental characteristics, whereas the linear form of RDA (LRDA) only includes first-order relationships.

Variance explained by the first axis of RDA was compared to the variance explained by the first axis in Principal Components Analysis (PCA); this is an expression of the level of constraint introduced by constrained ordination (RDA is the constrained ordination form of PCA). Partial Redundancy Analysis calculated the variance that was explained exclusively by village information, and the variance that was explained exclusively by farm characteristics (Borcard et al. 1992; Legendre and Legendre 1998). The number of species that significantly contributed to the ordination diagram was determined by the radius of $\sqrt{\frac{2}{n}}$ (n : number of ordination axes) of the equilibrium circle (Legendre and Legendre 1998). Further information on the interpretation of ordination results can be obtained from Jongman et al. (1995), Legendre and Legendre (1998), Makarenkov and Legendre (2002) and Quinn and Keough (2002) and in the results section.

Multiple regression analysis

Another statistical method by which the magnitude and significance of environmental characteristics on differences in species composition were

investigated was multiple regression analysis (Mathsoft 1999). Regression coefficients were obtained through stepwise multiple regressions by using the same explanatory variables that were used in the RDA. Species abundances were used as response variables, after transforming species abundance by $\ln(a + 1)$. Only species that were well represented in ordination graphs for the first two axes were investigated. The analyses presented here, therefore consisted of two steps of (i) first selecting the dominant species from the ordination graphs, and (ii) analysing regression models for the selected species.

Results

Summary of all ordination results

We summarised ordination results for the most prevalent functions, using the criterion of occurrence on a minimum of 40 farms (20%) to identify the common functions (Table 1).

Even when only species that occurred on more than five farms were considered, the majority of functions contained more than 10 species. This pattern illustrates the need for this study for techniques that reduce the number of dimensions of a dataset such as PCA or RDA, since it is complicated to compare the positions of sites on 10 or more axes (which would be required when using the species as ordination axes).

The equilibrium circle selected subsets of species from all the species that were included in a function before ordination. Because the first two ordination axes explain the largest percentages of variance, this technique thus allowed selecting the subset of species that contributed most to differences in species composition between farms.

All ordinations were significant ($p < 0.05$, based on 10,000 randomisations for LRDA and 100 randomisations for PRDA), except PRDA for the charcoal function. Obtaining a significant result means that the observed pattern is not likely to be an artefact of random distribution of species over farms. PRDA explained substantially more variance than LRDA (the respective averages are 44% and 18% of explained variance). For eight functions, this resulted in two axes of PRDA explaining more variance than all the variance explained by LRDA. Both methods, however,

Table 1. Summary information on the ordination of the prevalent functions of trees on farms in western Kenya. The number of dominant species was determined by adding the equilibrium circle to the ordination diagram for the first two ordination axes. Variance is expressed as percentage of total variance (PCA = 100%).

Function	All species	Number of dominant species	Variance PRDA (all axes)	Significance PRDA (all axes)	Variance LRDA (all axes)	Significance LRDA (all axes)	Variance PCA (first axis)	Variance LRDA (first axis)	Variance PRDA (first axis)
(All trees)	70	15	42.23	0.01	23.52	0.0001	12.17	8.62	9.60
Firewood	68	13	43.21	0.01	24.01	0.0001	12.49	8.68	9.53
Shade	50	8	42.21	0.01	12.27	0.0001	13.35	3.94	6.51
Timber	31	6	41.27	0.01	19.63	0.0001	23.77	8.23	11.42
Ornamental	31	3	47.60	0.01	11.78	0.0006	25.36	5.66	10.50
Boundary demarcation	27	6	48.44	0.01	30.30	0.0001	30.71	16.78	19.79
Medicine	26	2	43.51	0.01	10.59	0.0040	22.82	4.14	8.79
Soil fertility enhancement	22	6	45.84	0.01	17.17	0.0001	53.53	12.07	20.06
Charcoal	21	2	44.27	0.21	11.56	0.0108	50.47	6.88	14.92
Construction	18	2	42.49	0.02	13.67	0.0035	61.80	9.79	20.17
Fruit	13	4	36.01	0.01	14.97	0.0001	24.86	7.04	10.28
Fodder	7	2	48.41	0.01	14.14	0.0003	32.54	9.23	18.20
Beverage	2	2	50.33	0.01	34.26	0.0001	63.32	29.68	38.58

only explained a fraction of the total variance (PCA explains 100%), and the same pattern could be observed for the first ordination axis. The average variance explained on the first axis of PRDA was only 50% of that of the first PCA axis. The two first axes of PRDA explain on average 23% of variance, whereas two axes of PCA explain 50% of variance on average.

Variance partitioning provided evidence that differences between villages were significant for all functions (Table 2). For the charcoal, construction, fodder, timber, ornamental and soil fertility enhancement functions, using environmental characteristics alone did not result in significant relationships with LRDA ordination ($p < 0.07$). For the construction, timber and ornamental functions, using PRDA instead of LRDA made the ordination significant ($p < 0.07$). The reverse was the case for the medicine group ($p < 0.05$).

A total of 27 species dominated compositional differences in PRDA ordinations (Table 3). For several functions, most of the dominant species were a subset of the 15 species that dominated the ordination graph for all trees (firewood, timber, ornamental, boundary demarcation and beverage). Some functions had some species that only dominated their ordination graphs (shade, ornamental, medicine, soil fertility enhancement and fodder). In most cases, a species that dominated several ordination graphs occurred on the same subset of farms. Since this implies that the species had the same abundance distribution for several functions (the information on functions was captured on a species-by-species basis, not on a tree-by-tree basis), the same regression analysis is relevant for several functions. Shade, ornamental, soil fertility, charcoal and fodder were functions where species composition was different than the species composition of other functions, indicating that a smaller subset of farmers were using trees for these functions than for other functions.

Ordination and regression results for all trees

The horizontal axis of the ordination graphs followed the gradient of Ebuchiebe → Madidi → Mutambi → Shimutu towards Kakamega Forest (Figures 1 and 2). Especially differences in species composition between Ebuchiebe and the

Table 2. Variance partitioning for ordinations of the prevalent functions of trees on farms in western Kenya, with percentages of exclusive variance and significance for village and household differences in species composition.

Functions	Village Percentage (LRDA)	Village Percentage (PRDA)	Village Significance (LRDA and PRDA)	Household Percentage (LRDA)	Household Significance (LRDA)	Household Percentage (PRDA)	Household Significance (PRDA)
All trees	13.32	14.67	0.0001	7.30	0.0001	26.01	0.01
Firewood	13.55	15.00	0.0001	16.38	0.0001	35.58	0.01
Shade	4.57	10.86	0.0001	6.39	0.0059	36.33	0.01
Timber	13.94	12.53	0.0001	10.02	0.6370	31.66	0.01
Ornamental	5.29	12.29	0.0001	5.77	0.2854	41.59	0.06
Boundary demarcation	17.49	16.40	0.0001	8.19	0.0001	26.33	0.01
Medicine	0.2	12.79	0.0001	5.79	0.0001	38.71	0.21
Soil fertility enhancement	11.81	15.35	0.0001	5.21	0.6559	33.88	0.20
Charcoal	5.57	10.60	0.0001	5.59	0.4723	38.30	0.27
Construction	5.67	8.66	0.0007	7.53	0.1180	36.35	0.02
Fruit	7.44	11.22	0.0001	7.34	0.0674	28.38	0.06
Fodder	8.92	19.98	0.0001	5.80	0.7247	40.07	0.48
Beverage	21.07	17.07	0.0001	8.65	0.0014	24.72	0.01

other villages are highlighted, as Ebuchiebe is the only village with farms represented on the right-hand side of the graphs. The vertical axis highlights differences between Mutambi (upper-half) and Shimutu (lower-half). The third ordination axis featured differences for Madidi as it is the village that dominated the upper-left part of the graph (graph not included). The first three ordination axes therefore discriminate between the four villages. The typical species for each village can therefore be inferred from the graphs. Information from the ordination graph was cross-checked with the information from the stepwise regression analysis (Table 4).

Euphorbia tirucalli (finger euphorbia) *Markhamia lutea* (markhamia), *Sesbania sesban* (sesbania), *Grevillea robusta* (silky oak) and *Syzygium cumini* (Java plum) are typical species for Ebuchiebe. *Grevillea robusta* and *Markhamia lutea* were the only species where arrows pointed directly to the centroid for Ebuchiebe village. (The centroid positions show where farms of a certain category are predicted to occur; centroids for each categorical variable are indicated by italicised abbreviations in Figure 2. The arrows show the direction in the graph where sites with above average values for a species or a continuous variable are predicted to occur. The length of the arrow is an indication of the magnitude of differences in values). The position and length for the species arrow for *Euphorbia tirucalli* demonstrate clearly that this species dominates Ebuchiebe, however. For *Syzygium cumini*, regression analysis only indicated lower presence in Shimutu and Mutambi, but not higher presence in Ebuchiebe. The arrow for this species was the shortest, however.

Madidi is characterized by *Buddleja davidii* (butterfly bush), *Dracaena fragrans* (no common name), *Cupressus lusitanica* (cypress) and *Harungana madagascariensis* (blood tree). The only species that had negative stepwise regression coefficients for the other three villages was *Cupressus lusitanica*, however. *Buddleja davidii* features as an important species for Mutambi. Regression coefficients confirm its lower presence in Ebuchiebe and Shimutu.

Species that feature more in Mutambi are *Coffea arabica* (coffee) and *Camellia sinensis* (tea). *Buddleja davidii* is also typical for Madidi (see above).

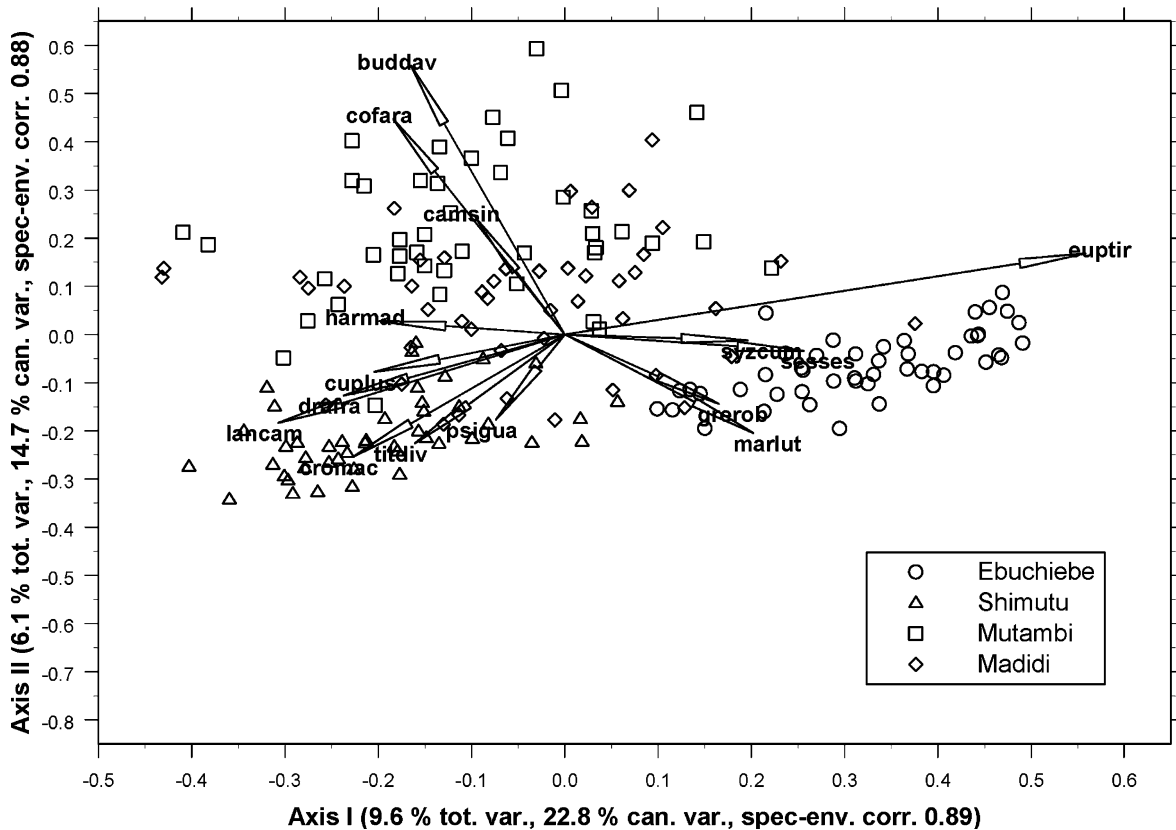


Figure 1. Ordination plot for all trees provided by polynomial Redundancy analysis, showing positions of farms and species arrows for ordination axes 1 and 2 for the western Kenyan study. Abbreviations on the axes: % tot. var.: percentage of total variance explained on the axis; % can. var.: percentage of canonical variance explained on the axis; spec-env. cor.: species-environment correlation. Full species names can be found in Table 3.

Typical for Shimutu are *Lantana camara* (lantana), *Croton macrostachyus* (broad-leaved croton), *Tithonia diversifolia* (tithonia), and *Psidium guajava* (guava). *Croton macrostachyus* and *Tithonia diversifolia* were the only species with positive regression coefficients for the village. For the other two species (*Lantana camara* and *Psidium guajava*), arrows pointed towards Madidi on the third ordination axis, whereas regression only confirmed lower presence in Ebuchiebe and Mutambi.

Not all relationships expected between species composition and farm characteristics could be confirmed. The relationships between farm size (the explanatory variable with longest arrow size in Figure 2) and *Lantana camara*, *Croton macrostachyus* and *Harungana madagascariensis* were confirmed (Table 4). The species with the smallest angle to the arrow representing farm size (*Tithonia diversifolia*) did not have a significant regression

coefficient for the variable ($r_a = -0.07$, $p = 0.90$), however.

The positive link between the number of local cattle and *Lantana camara* could not be confirmed by regression analysis. Neither could a positive correlation between *Harungana madagascariensis* and the number of years that the household was under the present head be confirmed. The relationships between *Euphorbia tirucalli*, *Sesbania sesban*, *Syzygium cumini*, and the number of children in the household were not confirmed by regression analysis. None of the species' arrows belonged to the same quadrant as the arrow for the number of children in all figures, however, and the arrow length was also quite short.

The centroids for the categorical explanatory variables describing the type of household head and the type of house (as indicator of wealth) were much closer to the origin than the centroids for villages. This pattern shows that differences in species

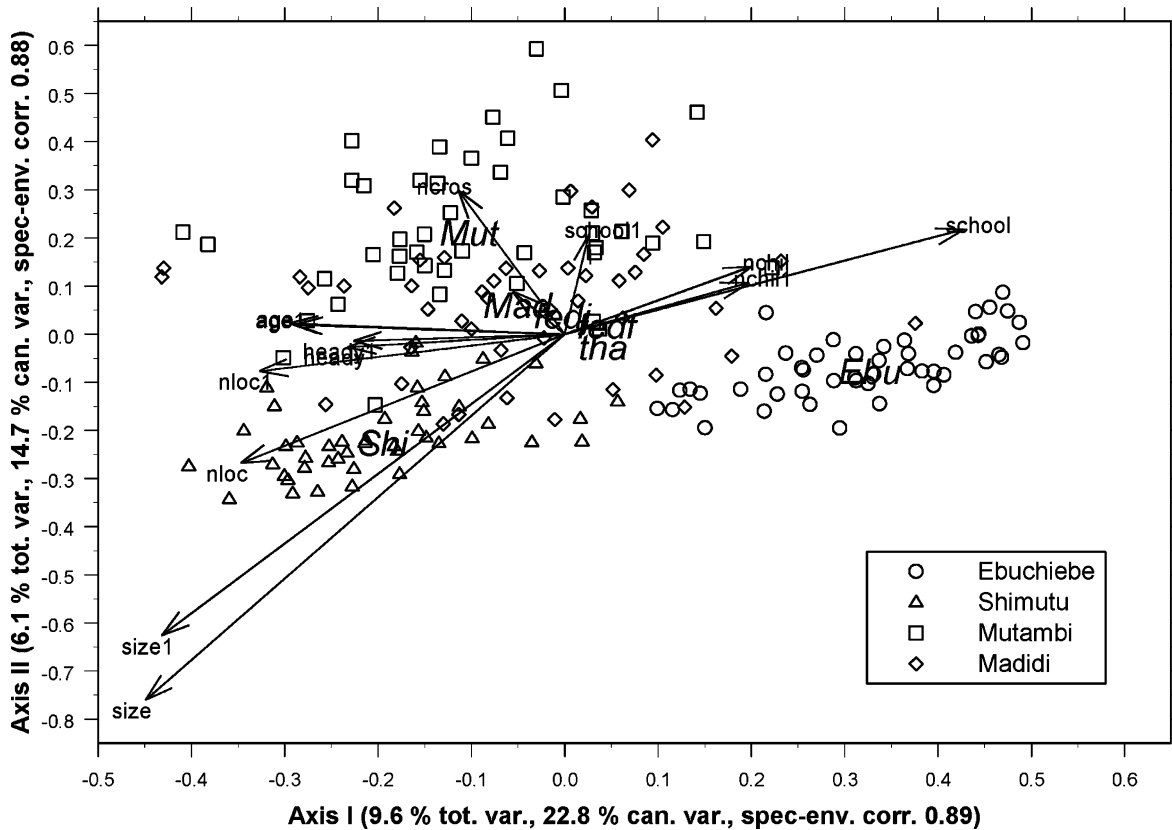


Figure 2. Ordination plot for all trees provided by polynomial Redundancy analysis, showing positions of farms and explanatory variables for ordination axes 1 and 2 for the western Kenyan study. Abbreviations as in Figure 1.

composition between these typologies of farms were smaller than spatial differences. By consequence, many patterns could not be confirmed, for example the relationship between *de facto* female-headed households (households where the husband was most of the time absent from the farm) and *Syzygium cumini* and *Sesbania sesban*. Some patterns could be confirmed for explanatory variables, however: the relationship between *de jure* female-headed households (households without a husband) and *Coffea arabica*, and the relationship between thatch-roofed houses and *Grevillea robusta*.

Discussion

Ordination methods

Two new methods of ordination analysis were combined for the ordination methods in this article: the transformation of the community

matrix so that differences in the Redundancy Analysis reflect the Hellinger distance, and the use of polynomial redundancy analysis. Both methods were developed to overcome some shortcomings in ordination methods that had been developed earlier (Legendre and Gallagher 2001; Makarenkov and Legendre 2002). Polynomial RDA provided a substantial increase in the variance that was explained, and hence in the ability to discover patterns in the data. Since non-linear patterns are typical between species and environmental variables, PRDA will provide a better approach in most situations by allowing for non-linear relationships (Økland 1999; Makarenkov and Legendre 2002).

Constrained methods such as the RDA used in this article are especially superior to unconstrained ordination methods in providing a significance test (Legendre and Legendre 1998; Makarenkov and Legendre 2002; Quinn and Keough 2002). Such significance test is similar to the significance tests of regression and ANOVA (an *F*-type statistic is

Table 4. Coefficients of village and household characteristics as explanatory variables on species abundance after stepwise linear regression. The significance of coefficients is provided between brackets below the regression coefficients. Total explained variance is expressed by the multiple correlation coefficient R^2 , with significance $p(F)$. The significance of regression coefficients was estimated by randomisation tests (Anderson and Legendre 1999; provided between brackets).

Species	$R^2(p(F))$	Intercept	Ebuechie-be village (farthest from forest)	Shimutu village (closest to forest)	Mutambi village (intermediate to forest)	Male headed	Female headed (no husband)	Farm size	Perma-nent house (most wealthy)	Thatch roofed house (least wealthy)	Number of cross-bred cattle (~wealth)	Number of local cattle (~wealth)	Years being head	Age head	Number of resident children	Level of schooling of head	
<i>Buddleja davidii</i>	0.394 (0.001)	2.33 (0.001)	-2.18 (0.0001)	-2.01 (0.0001)	-	-	0.44 (0.1546)	-	-0.73 (0.0446)	-0.46 (0.0776)	-	-	-	-0.87 (0.1135)	1.03 (0.0492)	-	
<i>Camellia sinensis</i>	0.217 (0.001)	-0.04 (0.870)	-	-	1.79 (0.0001)	-	-0.68 (0.0571)	1.42 (0.0449)	-	-	1.98 (0.0547)	-	-	-	-	-	-
<i>Coffea arabica</i>	0.491 (0.001)	-0.72 (0.041)	-0.37 (0.0748)	-	1.97 (0.0001)	-	0.81 (0.0010)	-	1.05 (0.0014)	-	-1.57 (0.0355)	-	-	1.13 (0.0162)	-	0.85 (0.0135)	-
<i>Croton macrostachyus</i>	0.475 (0.001)	-0.10 (0.742)	-	1.76 (0.0001)	0.42 (0.0216)	-	-	1.11 (0.0112)	-	-	-	-	-	1.26 (0.0018)	-0.57 (0.0942)	0.58 (0.0332)	-
<i>Cupressus lusitanica</i>	0.275 (0.001)	1.51 (0.001)	-1.78 (0.0001)	-0.74 (0.0487)	-1.30 (0.0006)	-	-	-	2.09 (0.0001)	-0.49 (0.1181)	-	-	-	1.91 (0.0038)	-	-	-
<i>Dracaena fragrans</i>	0.192 (0.001)	1.34 (0.001)	-1.88 (0.0001)	-	-1.23 (0.0003)	-	-	-	-	-	-	-	-	-	1.77 (0.0049)	-	-
<i>Euphorbia tirucalli</i>	0.266 (0.001)	1.95 (0.001)	1.96 (0.0001)	-1.22 (0.0009)	-	0.46 (0.1239)	-	-	-	-	-	-	-	-	-	-	-
<i>Grevillea robusta</i>	0.165 (0.001)	-0.32 (0.177)	0.40 (0.0033)	-	-0.27 (0.0485)	-	-	0.52 (0.0791)	-	0.37 (0.0042)	-	-	-	0.68 (0.0263)	-	0.40 (0.0469)	-
<i>Harungana madagascariensis</i>	0.301 (0.001)	0.23 (0.328)	-1.03 (0.0003)	-0.82 (0.0071)	0.72 (0.0077)	0.28 (0.1565)	-	3.57 (0.0001)	-	0.45 (0.0422)	-1.64 (0.0335)	-	-	-	-	-	-
<i>Lantana camara</i>	0.242 (0.001)	2.78 (0.001)	-1.50 (0.0001)	-	-0.60 (0.0737)	-	-	2.82 (0.0002)	-1.06 (0.0152)	-0.52 (0.0908)	-	-	-1.33 (0.0431)	-1.11 (0.1183)	-	-	-
<i>Markhamia lutea</i>	0.284 (0.001)	0.85 (0.001)	1.14 (0.0001)	-	-	-	-	2.91 (0.0001)	0.52 (0.1217)	-	-1.88 (0.0182)	-	0.83 (0.0580)	-	-	-	-
<i>Psidium guajava</i>	0.216 (0.001)	2.06 (0.001)	-0.71 (0.004)	-	-0.64 (0.0110)	-	-	2.82 (0)*	-	-	-	-	-	-	-	-	-
<i>Sesbania sesban</i>	0.195 (0.001)	1.32 (0.001)	0.55 (0.0364)	-0.97 (0.0022)	-0.41 (0.1218)	0.29 (0.1455)	-	1.46 (0.0091)	-	-	-	-	-0.80 (0.0762)	-	-	-0.60 (0.0780)	-
<i>Syzygium cumini</i>	0.287 (0.001)	0.42 (0)	-	-0.93 (0.0001)	-0.51 (0.0001)	0.18 (0.0434)	-	1.06 (0.0005)	-	-	-0.68 (0.0567)	-	-	-	-	-	-
<i>Tithonia diversifolia</i>	0.289 (0.001)	0.40 (0.029)	-	1.62 (0.0001)	-	-	-	-	-	-0.50 (0.0189)	-	-	-	-	-	-0.47 (0.1281)	-

calculated), and patterns should only be trusted when the tests are significant. For these reasons, results from unconstrained ordination methods (e.g. Gordon et al. 2003; Bouvet et al. 2004) should be interpreted with care.

Complexities of ordination algorithms result in normal percentages of variance explained by constrained ordination methods to range 20–50% (ter Braak and Smilauer 1998; Økland 1999). This makes it difficult to assess the efficiency of constrained ordination (RDA) in explaining patterns in species composition that are revealed by unconstrained ordination (PCA). By assuming that the first ordination axis of PCA showed the dominant gradient to which most species respond, we can analyse the performance of RDA, however. The average variance explained on the first axis of PRDA was only 50% of that of the first PCA axis. When we correlated the positions of the farms on the first PCA axis with the positions of the farms on the second PCA axis, we found strong correlations (minimum correlation 0.83 for ornamental, correlation >0.95 for eight functions). We can therefore conclude that the gradients expressed on the first PCA and RDA axes were largely similar. In combination with the fact that the majority of species that contributed to the ordination by PRDA also contributed to the PCA ordination (in general, more species contributed to the RDA than to the PCA ordination graphs), we can therefore conclude that the PRDA ordination graphs portrayed the main differences in species composition between farms.

Multiple regression explained low percentages of variance, with the highest amount of variance explained for *Coffea arabica* (49%). Low amounts of explained variance imply that single species were not strongly related to explanatory variables. The trends from ordination graphs were confirmed by multiple linear regression for most of the strong effects (where arrow lengths were longer, or where centroids were positioned far from the origin), however.

Jarvis et al. (2000) stated that gender, age, wealth or social status affect farmers' knowledge, actions and access to resources regarding the maintenance of crop diversity. Long et al. (2000) provide some examples of the influence of wealth, age and gender on crop diversity. Jarvis et al. (2000) also listed farm size, family size, and years of education as explanatory factors for the crop

varieties that farmers choose. An earlier analysis of explanatory characteristics of tree species diversity on farms using results from the same survey that was used here showed that the same explanatory variables could explain farm diversity (Kindt et al. 2004). These results show that variables that influence crop choices and tree diversity (wealth, age, gender, farm size, family size, length of tenure, and education) also influenced on-farm tree composition and differences in the abundance of separate species. The low percentages of explained variance imply that several exceptional farms can be found for each function, which means that one must be careful in making general statements about the species composition of farms of a certain typology.

Planning for agroecosystem diversification

This study was part of a larger one studying options for landscape domestication (Kindt 2004; Simons and Leakey 2004). Some of the reasons to opt for landscape domestication (defined as the amelioration of the tree species diversity within and between functional uses in a particular landscape) were provided in the introduction. Here we provide some examples of how the information presented in the results section could be utilised in planning for landscape domestication. It should be obvious that any landscape domestication strategy should involve farmers and stakeholders, and not be merely a desktop study of how ordination graphs could be changed.

As one example for the boundary demarcation function, the results provided evidence that the average farms of Ebuchiebe village were dominated by *Euphorbia tirucalli*. Although some farms of the same village already had a species composition that was more typical of other villages, one pathway of landscape domestication would be to stimulate wider planting of species that are more typical of other villages in the landscape, for example wider planting of *Cupressus lusitanica* or *Buddleja davidii*. A wider sharing of species that have a dominant presence in other villages of the same landscape could have several advantages, such as ensuring that the species has already been tested in farming conditions and allowing sharing between farmers of germplasm and knowledge. Such approach is a

landscape integration approach of distributing species more widely, in contrast to segregation approaches that distribute species only in particular subsets of a landscape (Van Noordwijk et al. 1997). For some local species, distribution of trees of the same species at a larger scale in a landscape may lead to those linkages of populations that are necessary to maintain species diversity necessary to avoid inbreeding depression and extinction of species (O'Neill et al. 2001; Atta-Krah et al. 2004). Care must be taken, however, not to overlook rare species that are rare because of overexploitation rather than being less fit to produce agroforestry tree products that farmers wish to commercialise than other more widely distributed species. As a hypothetical example, in case that *Agave sisalana* was rare for boundary demarcation only because of overexploitation of this species, then it would be more strategic to attempt to distribute this species more widely since this would have larger effects on increasing diversity.

In another study that focused on differences in tree species diversity (the average species richness of a farm) of the same functions, the construction, medicine, charcoal, beverage, fodder, ornamental and soil fertility groups were identified as having the lowest diversity using the criterion of less than two species (Kindt et al. 2004). Two species dominated the ordination graphs for the first five functions (with identities provided in Table 3). If priority was given to diversify the functions with the lowest alpha diversity (such approach would maximise the relative increment in diversity of adding a species), then the ordination results suggest one possible strategy to increase the average diversity of farms to two or more species: by distributing the two dominant species everywhere in the landscape. Implicit in this approach would also be to increase the evenness of species abundances in the selected functions, since compositional differences and diversity are not only influenced by species presence-absence but also by the proportions of each species. The study of alpha diversity and study of ordination graphs can therefore complement each other: (i) the study of alpha diversity can allow to prioritise functions of lowest diversity on the average farm; and (ii) the study of ordination graphs provides information on the dominant species that could be more widely distributed. Such approach could easily be implemented in other landscapes, since the collection of information on current tree diversity levels is

required to monitor the impacts of landscape domestication efforts and since the same information can be utilised to study differences in alpha diversity and differences in species composition as shown in this article.

Conclusions

Ordination graphs display patterns for a fraction of variance for all species, whereas regression analysis investigates patterns in total variance for a single species. Ordination and regression analyses are thus complementary techniques. We utilised this complementarity by selecting those species that contributed most to differences in species composition from the ordination graphs, and sequentially analysing to complete information that was available for these species. By doing so, we reduced the total number of regression analyses that were required (assuming that the same species were used for the same functions on all farms where the species occurred, the total number of regressions was reduced from 70 to 15 – in reality a larger reduction occurred). We can therefore recommend similar approaches for other studies on explanatory factors for differences in species composition.

We found evidence for differences in species composition between villages. As one pathway of increasing the tree diversity in the survey area, a wider sharing of species and information on their utilisation between villages is recommended. This study showed that it is possible to increase diversity at landscape levels without introducing new species. This approach provides several advantages, such as the fact that local knowledge on a species is available, that the species has already been tested in similar agroecosystem conditions, that information can be provided by farmer-to-farmer exchange visits, and that local biodiversity can be increased. Because of these advantages and because of growing evidence of the benefits of diversification and tree domestication, we expect that similar schemes can be implemented throughout the tropics.

Acknowledgements

We are very grateful for information provided by farmers, Jean-Marc Boffa, Richard Coe, Ian Dawson, Ard Lengkeek, John Weber and two

anonymous reviewers for inputs in the article and assistance provided by the ICRAF East and Central African regional programme especially through Stephen Ruigu and Amadou Niang. Roeland Kindt wants to thank VVOB and DFID for financial support, and Joseph Njeri for assistance during data collection.

References

- Adams W.M., Aveling R., Brockington D., Dickson B., Elliott J., Hutton J., Roe D., Vira B. and Wolmer W. 2004. Biodiversity eradication and the eradication of poverty. *Science* 306: 1146–1149.
- Anderson M.J. and Legendre P. 1999. An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. *J. Stat. Comput. Simul.* 62: 271–303.
- Atta-Krah K., Kindt R., Skilton J.N. and Amaral W. 2004. Managing biological and genetic diversity in tropical agroforestry. *Agroforest. Syst.* 61: 183–194.
- Beentje H.J. 1994. Kenya Trees, Shrubs and Lianas. National Museums of Kenya, Nairobi, pp. 722
- Borcard D., Legendre P. and Drapeau P. 1992. Partialling out the spatial component of ecological variation. *Ecology* 73: 1045–1055.
- Bradley P.N., Chavangi N. and Van Gelder A. 1985. Development research and energy planning in Kenya. *Ambio* 14: 228–236.
- Bouvet J.M., Fontaine C., Sanou H. and Cardé C. 2004. An analysis of the pattern of genetic variation in *Vitellaria paradoxa* using RAPD markers. *Agroforest. Syst.* 60: 61–69.
- Di Falco S. and Perrings C. 2003. Crop genetic diversity, productivity and stability of agroecosystems. A theoretical and empirical investigation. *Scott. J. Polit. Econ.* 50: 207–216.
- Dorward A., Poole N., Morrison J., Kydd J. and Urey I. 2003. Markets, institutions and technology: missing links in livelihoods analysis. *Dev. Policy Rev.* 21: 319–332.
- Garrity D.P. 2004. Agroforestry and the achievement of the Millennium Development Goals. *Agroforest. Syst.* 61: 5–17.
- Gordon J.E., Hawthorne W.D., Sandoval G. and Barrance A.J. 2003. Trees and farming in the dry zone of southern Honduras II: the potential for tree diversity conservation. *Agroforest. Syst.* 59: 107–117.
- Jarvis D.I., Myer L., Klemick H., Guarino L., Smale M., Brown A.H.D., Sadiki M., Sthapit B. and Hodgkin T. 2000. A training guide for in situ conservation on farm. Version 1. International Plant Genetic Resources Institute, Rome, pp. 161
- Jongman R.H., Braak C.J.F. and Van Tongeren O.F.R. 1995. Data Analysis in Community and Landscape Ecology. Cambridge University Press, Cambridge, pp. 299xxi +
- Kindt R. and Lengkeek A.G. 1999. Tree diversity on farm – use it or lose it. In: National workshop on agricultural biodiversity conservation, 27–29 January 1999, Nairobi, pp. 75–85. Intermediate Technology Development Group (ITDG).
- Kindt R., Simons A.J. and Van Damme P. 2004. Do farm characteristics explain differences in tree species diversity among western Kenyan farms? *Agroforest. Syst.* 63: 63–74.
- Kristensen M. and Balslev H. 2003. Perceptions, use and availability of woody plants among the Gourounsi in Burkina Faso. *Biodivers. Conserv.* 12: 1715–1739.
- Leakey R.R.B. and Simons A.J. 1998. The domestication and commercialization of indigenous trees in agroforestry for the alleviation of poverty. *Agroforest. Syst.* 38: 165–176.
- Legendre P. and Gallagher E.D. 2001. Ecologically meaningful transformations for ordination of species data. *Oecologia* 129: 271–280.
- Legendre P. and Legendre L. 1998. Numerical Ecology. Elsevier Science BV, Amsterdam, pp. 853
- Long J., Cromwell E. and Gold K. 2000. On-farm Management of Crop Diversity: An Introductory Bibliography. Overseas Development Institute for ITDG, London, pp. 42
- Loreau M., Naeem S., Inchausti P., Bengtsson J., Grime J.P., Hector A., Hooper D.U., Huston M.A., Raffaelli D., Schmid B., Tilman D. and Wardle D.A. 2001. Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science* 294: 804–807.
- McNeely J.A. 2004. Nature vs. nurture: managing relationships between forests, agroforestry and wild biodiversity. *Agroforest. Syst.* 61: 155–165.
- Makarek V. and Legendre P. 2002. Nonlinear redundancy analysis and canonical correspondence analysis based on polynomial regression. *Ecology* 83: 1146–1161.
- Mathsoft I. 1999. S-Plus 2000 User's Guide. MathSoft Inc, Seattle, pp. 558xii
- Norberg J., Swaney D.P., Dushoff J., Lin J., Casagrandi R. and Levin S.A. 2001. Phenotypic diversity and ecosystem functioning in changing environments: a theoretical framework. *Proc. Natl. Acad. Sci.* 98: 11376–11381.
- O' Neill G.A., Dawson I.K., Sotelo-Montes C., Guarino L., Current D., Guariguata M. and Weber J.C. 2001. Strategies for genetic conservation of trees in the Peruvian Amazon basin. *Biodivers. Conserv.* 10: 837–850.
- Økland R.H. 1996. Are ordination and constrained ordination alternative or complementary strategies in general ecological studies? *J. Veg. Sci.* 7: 289–292.
- Økland R.H. 1999. On the variation explained by ordination and constrained ordination axes. *J. Veg. Sci.* 10: 131–136.
- Quinn G. and Keough M. 2002. Experimental Design and Data Analysis for Biologists. Cambridge University Press, Cambridge, pp. 520
- Simons A.J. and Leakey R.R.B. 2004. Tree domestication in tropical agroforestry. *Agroforest. Syst.* 61: 167–181.
- ter Braak C.J.F. and Smilauer P. 1998. CANOCO reference manual and user's guide to Canoco for windows: software for canonical community ordination (version 4). Microcomputer Power, Ithaca, pp. 352
- Tilman D., Reich P.B., Knops J., Wedin D., Mielke T. and Lehman C. 2001. Diversity and productivity in a long-term grassland experiment. *Science* 294: 843–845.
- Van Noordwijk M., Tomich T.P., De Foresta H. and Michon G. 1997. To segregate or to integrate? The question

of balance between production and biodiversity conservation in complex agroforestry systems. *Agroforestry Today* 9: 6–9.

Whittaker R.J., Willis K.J. and Field R. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. *J. Biogeogr.* 28: 453–470.