

Planning tree species diversification in Kenya based on differences in tree species composition between farms. II. Analysis of tree niches.

R. Kindt^{1,*}, P. Van Damme², A.J. Simons¹ and H. Beeckman³

¹World Agroforestry Centre (ICRAF), PO Box 30677-00100, Nairobi, Kenya; ²Ghent University, Laboratory of Tropical and Subtropical Agriculture and Ethnobotany, Coupure Links 653, B-9000 Gent, Belgium; ³Royal Museum for Central-Africa, Leuvensesteenweg 13, B-3080 Tervuren, Belgium; *Author for correspondence (e-mail: R.Kindt@CGIAR.org; phone: 254-20-524000 or 1-650-8336645; fax: 254-20-524001 or 1-650-8336646)

Received 13 May 2004; accepted in revised form 30 September 2004

Key words: Diversification, Domestication, Hellinger distance, Landscape, Ordination, Redundancy analysis

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. Novel types of ordination using the Hellinger ecological distance and polynomial Redundancy Analysis indicated wide heterogeneity between farms with respect to tree species composition of five niches, including homestead, cropland, fallow, woodlot, and external boundary ($p \leq 0.05$). Multiple regression analysis confirmed the ordination results using the abundance of dominant species as the response variable. The relationship between location and species composition differed with those of two previous surveys. Methodological differences in sampling intensity, locations and time of sampling between these surveys could have caused the difference. The maps of spatial distribution of compositional types provided in the previous surveys were not confirmed, whereas villages were found to contain several farms with a species composition that was not typical of their village. Meaningful results about the species composition of a landscape should include several farms per village and use a sampling grid finer than $5 \times 5 \text{ km}^2$.

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 may be pursued, including objectives of: (i) diversification of the incomes from tree products for resource-poor farmers (Leakey and Simons 1998; Garrity 2004; Simons and

Leakey 2004; Kindt et al. 2004, 2006; (ii) increasing the stability and productivity of the agroecosystem (Loreau et al., 2001; Di Falco and Perrings 2003; Dorward et al. 2003; Kindt et al. 2006; and (iii) enhancing the conservation of biodiversity in landscape mosaics (Kristensen and Balslev 2003; Atta-Krah et al. 2004; McNeely 2004). 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 (Leakey and Simons 1998; Kindt et al. 2004).

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, 2006). One of the objectives of the study described in this article was to verify spatial patterns in compositional differences (differences in identity and abundance of species) that were described in two earlier studies. These earlier studies provided information on compositional differences between tree niches (such as homesteads, cropland and external boundaries) for areas of western Kenya that included the Vihiga district. Confirmation of the findings of the previous studies would imply that district-wide landscape domestication strategies could be designed. The other objective was to establish whether significant differences in tree composition existed in the landscape. We used ordination and regression analysis methods to analyse differences in tree composition between farms and discussed the superiority of these methods for such analysis purposes.

Methods

Study area and recorded data

We made inventories of all trees (woody perennials) found on 201 farms in western Kenya, cover-

ing a total area of 114 ha. Beentje (1994) was used as the key reference for species identification. For each species encountered on a farm during walks with household informants, its abundance (number of trees) in each niche was recorded. The categories of on-farm niches were: trees in the homestead area, trees mixed in cropland, trees on contours in cropland, trees on external boundaries of the farm, trees on internal boundaries on the farm, trees in woodlots, and trees in fallows. In some cases the type of niche is mainly defined by how the trees are arranged: trees mixed in cropland, on internal boundaries, on contours and in woodlots all occur in the cropland section of the farm, but have a different spatial arrangement. These categories of niches were distinguished in previous surveys of the same area (Bradley et al. 1985; Bradley 1991; Kindt 1997), and satisfied the criteria of being largely self-explanatory, being easily identifiable in the field and from aerial photographs, and corresponding to local experience with tree establishment (Bradley 1991).

The study was conducted 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 above sea level, annual mean temperature ranges 18.1–20.4°C, and annual bimodal rainfall ranges 1600–2000 mm.

Four villages were randomly 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 1991). Ebuchiebe village was selected in stratum A1, which has 185 homesteads km⁻² and extremely small farms (0.53 ha). In this stratum, 60% of total groundcover of woody biomass is located on farms. Mutambi village is located in stratum A2, which is similar to A1 in terms of population density (195 homesteads km⁻²), farm sizes (0.49 ha), and percentage of on-farm woody biomass (67%). Madidi village is located in stratum E, which is characterised by average values (112 homesteads km⁻², farm size of 0.87 ha, 46% on-farm woody biomass). Shimutu village was selected in stratum F2, which has 74 homesteads km⁻², farm sizes of 1.12 ha, and significantly smaller on-farm woody biomass (18%). 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 Lauriks et al. (1999), Kindt et al. (2004) and references therein.

Species composition as provided in two previous surveys

Species composition was compared with the results of two previous surveys that were conducted on different farms in the same area.

The first survey provided information on species composition by giving the percentage of farms where a particular tree species occurred in three on-farm niches (Bradley et al. 1985; Bradley 1991). The percentage of farms was interpolated from stacked columns that were available for Ebusikhale (stratum A1, $n = 90$) and Kegoye (stratum A2, $n = 60$). Differences in the number of farms for the present survey were investigated by chi-square tests (Mathsoft 1999). Another method of comparing the survey results was provided by calculating 95% confidence intervals for species frequencies as:

$$p \pm 2\sqrt{\frac{p(1-p)}{n-1}}$$

(p : species frequency; n : number of farms) (Hayek and Buzas 1997).

The second survey provided information on species composition by listing species proportions (species abundance/total abundance) for five hedge types (Lauriks et al. 1999). The five hedge types were derived by cluster analysis of 63 samples each taken on a farm near the centre of the 25-km² cells of a grid covering the Vihiga and Siaya districts. The 63 samples were taken within a 100 m interval of hedge, by pooling six subsamples that covered an entire length of 20 m. Differences in species proportions with the present survey were analysed by calculating 95% confidence intervals based on clustered sampling as:

$$p \pm 2\sqrt{\frac{\sum_{i=1}^n m_i^2(p_i - p)^2}{(n-1)m^2n^{-1}}}$$

(m_i : total number of trees on farm i ; p_i : species proportion at farm i ; p : species proportion in the survey; n : number of farms; m : total number of trees) (Hayek and Buzas 1997)

Ordination and regression methods

The influence of farm characteristics on differences in species composition was investigated by ordination and regression analysis. These are complementary techniques since ordination graphs display patterns for a fraction of variance for all species, whereas regression analysis investigates patterns in total variance for a single species. We exploited this complementarity by reducing the number of regression analyses to the dominant species of the two-dimensional ordination graphs (Kindt et al. 2006). Explanatory variables for differences in species composition and species abundance among farms included village identity, type of household head, type of house (wealth indicator), farm size, number of cattle (wealth indicator), the years that the farm had been under the present head, the age of the household head, the number of resident children, and the level of schooling (Kindt et al. 2004, 2006).

As ordination technique, we used the modern constrained ordination analysis methods of Linear and Polynomial Redundancy Analysis (LRDA and PRDA) based on transformed farm – species abundance matrices (Makarek and Legendre 2002). A transformation was selected that resulted in representing the Hellinger distance between farms in ordination graphs. This ecological distance is among the better distances for expressing differences in species composition (Legendre and Legendre 1998; Legendre and Gallagher 2001). Only those species that occurred on more than four farms were considered for the ordination analyses. Regression coefficients and their significance were estimated by stepwise multiple regression analysis (Mathsoft 1999). Species abundances were $\ln(a + 1)$ transformed before the ordination and regression analyses.

Results

Comparison with previous surveys

Several differences in farm frequencies can be observed between the first and most recent survey (Table 1). The clearest differences for stratum A1 for the present survey are the higher frequency of farms with *Euphorbia tirucalli* and *Markhamia lutea* on hedges, higher frequencies for *Mangifera*

Table 1. Farm number and average frequency (95% confidence interval) for the species mentioned by Bradley et al. (1985) (B.) compared with the present survey (K.).

Species	A1 with Ebukihale (B., n = 90) and Ebuchiebe (K., n = 50)			A2 with Kegoye (B., n = 60) and Mutambi (K., n = 50)							
	Hedge	Homestead	Cropland	Hedge	Homestead	Cropland					
	B.	K.	B.	K.	B.	K.					
<i>Cupressus lusitanica</i>	21 (14-32%)	10 ns (9-31%)	0 -	6** (!) (3-21%)*	19 (12-30%)	3* (0-13%) ns	16 ns (19-45%) ns	3 ns (0-13%) ns	0 -	13 (11-32%)	0** *
<i>Eucalyptus saligna</i>	18 (12-28%)	13 ns (13-39%)	21 ns	17 ns (20-48%)	20 ns	22* (30-58%) ns	5 ns (1-19%) ns	11*** (10-34%)*	0 -	15 (14-36%)	18 ns (22-50%) ns
<i>Euphorbia tirucalli</i>	40 (34-55%)	46*** (84-100%)*	0 -	0 -	0 -	0 -	34*** (55-81%)*	0 -	0 -	0 -	0 ns -
<i>Lantana camara</i>	24 (17-36%)	7 ns (4-24%) ns	0 -	0 -	0 -	0 -	15*** (17-43%)*	0 -	2 ns (0-10%) ns	0 -	1 ns (0-6%) ns
<i>Mangifera indica</i>	0 -	3** (!) (0-13%) ns	25 (18-37%)	32*** (50-78%)*	0 -	16*** (19-45%)*	0 -	18 (18-42%)	12 ns (12-36%) ns	0 -	22*** (30-58%)*
<i>Markhamia lutea</i>	13 (7-22%)	29*** (44-72%)*	0 -	9*** (!) (7-29%)*	24 (17-36%)	29*** (44-72%)*	10*** (!) (9-31%)*	0 -	7** (!) (4-24%)*	1 (0-5%)	14*** (15-41%)*
<i>Persea americana</i>	0 -	5** (!) (1-19%)*	18 (12-28%)	27*** (40-68%)*	13 (7-22%)	30*** (46-74%)*	2 ns (12-36%)	22 (24-49%)	23 ns (32-60%) ns	39 (53-77%)	44* (79-97%)*
<i>Psidium guajava</i>	16 (10-26%)	22** (30-58%)*	0 -	21*** (28-56%)*	0 -	12*** (!) (12-36%)	13*** (13-39%)*	18 (18-42%)	28* (42-70%)*	21 (23% - 47%)	29* (44-72%) ns
<i>Sesbania sesban</i>	0 -	0 -	0 -	8*** (!) (6-26%)	24 (17-36%)	33*** (52-80%)*	1 ns (0-6%) ns	0 -	2 ns (0-10%) ns	0 -	17*** (20-48%)*

Coding of significance levels of differences: ns < 5% < * < 1% < ** < 0.1% < ***. !: result of the χ^2 test needs to be treated with care due to small expected frequencies.

indica, *Persea americana* and *Psidium guajava* in the homestead area, and higher frequencies for *Mangifera indica*, *Markhamia lutea*, *Persea americana* and *Sesbania sesban* on cropland. For stratum A2, the most obvious differences for the most recent survey are the higher frequencies of farms with *Psidium guajava* and *Lantana camara* on hedges, the lower frequency for *Euphorbia tirucalli* on hedges, higher frequencies for *Psidium guajava* and *Eucalyptus saligna* in the homestead, and higher frequencies for *Mangifera indica*, *Markhamia lutea*, *Persea americana* and *Sesbania sesban* in cropland. The general trend of Table 1 is that higher frequencies were observed in the more recent survey for most species. The results further indicate that most species belong to several niches.

Several differences in species composition (measured as species proportions) were observed between the second and the most recent survey (Table 2). Species that also contributed to the ordination graph for the external boundary (Figure 1 and Table 4: *Buddleja davidii* and *Cupressus lusitanica*) were added to those species that described differences in hedge species composition (Lauriks et al. 1999). The major differences were the lower proportion of *Euphorbia tirucalli* (the dominant species for hedge type 1) in Mutambi and Madidi and the higher proportion of the same species in Ebuchiebe in the most recent survey. In the first two villages, the effect is mainly caused by the higher proportion of *Buddleja davidii*, *Cupressus lusitanica* and *Dracaena steudneri*.

Ebuchiebe had a lower proportion of *Lantana camara*, which is the dominant species of hedge type 2. The comparisons between the two surveys were hindered by the fact that confidence intervals for the species proportions could not be estimated for the second survey (this requires the original dataset).

The 25-km² cells east, south-east, south-west and west to the Ebuchiebe cell belonged to hedge type 1, which is a hedge type that describes the Ebuchiebe species composition very well (Lauriks et al. 1999). The coordinates of the farm sampled in the first survey (UTM 677.25 E, 12.71 N; Lauriks 1996) indicates that its position was about 3 km south-west from Ebuchiebe (UTM 677.71 E, 10.15 N). The hedge types that corresponded best to the composition recorded in Mutambi (696.27 E, 8.00 N), Madidi (686.08 E, 6.74 N) and Shimutu (701.98 E, 20.79 N) was hedge type 3 containing a mixture of *Dracaena steudneri*, *Euphorbia tirucalli*, *Lantana camara* and *Tithonia diversifolia*. Cells north-west (for the first village) and south of the cells (for the other villages) to which these villages belonged had this species composition.

Differences between hedge types were mainly between proportions of species, not between occurrences of these species (Table 2). Hedge types 1 and 2 share six of the seven species that were used to categorize differences between these types. At least five of these species were encountered in each village of the recent survey.

Table 2. The species composition in hedges (expressed as percentage of total) provided by Lauriks (1996) and Lauriks et al. (1999) for hedge types (L.), for hedge composition of individual 25 km² cells containing the village of the present survey (L. (village abbreviation)), and for village in the present survey (K. (village)) with 95% confidence interval.

Species	Area with hedge type 1					Area with hedge type 2		
	L.	L. (Mu.)	L. (Ma.)	K. (Mu.)	K. (Ma.)	L.	L. (Eb.)	K. (Eb.)
<i>Aloe</i> sp.	0.0	0.0	0.0	0.0	0.0	3.3	0.0	0.0
<i>Buddleja davidii</i>	–	–	–	27.5 (13.9–41.2)*	14.6 (7.3–22.0)*	–	–	0.0
<i>Cupressus lusitanica</i>	–	0.0	0.0	26.5 (11.2–41.8)*	26.8 (16.5–37.0)*	–	0.0	4.4 (0.0–11.6)
<i>Dracaena steudneri</i>	2.3	19.6	2.5	12.6 (2.9–22.3)	17.7 (10.0–25.4)*	0.8	0.0	0.3 (0.0–0.8)
<i>Euphorbia tirucalli</i>	64.3	62.7	91.5	18.4 (11.3–25.6)*	17.4 (11.0–23.7)*	5.2	2.1	74.0 (62.8–85.2)*
<i>Lantana camara</i>	13.8	0.0	0.0	9.0 (2.4–15.6)	12.5 (7.0–18.0)*	52.9	84.1	5.6 (0.4–10.7)*
<i>Markhamia lutea</i>	6.0	0.0	0.0	0.7 (0.2–1.1)*	3.9 (0.0–7.8)*	5.1	0.0	7.4 (2.8–12.0)*
<i>Psidium guajava</i>	3.0	0.0	4.5	0.8 (0.3–1.3)*	1.6 (0.6–2.6)*	2.1	4.8	3.7 (1.3–6.1)
<i>Tithonia diversifolia</i>	2.6	0.0	0.0	0.7 (0.0–1.8)*	0.0	6.1	0.0	0.0
Other species	8.1	17.7	1.5	3.8 (0.6–6.9)*	5.5 (1.3–9.8)	24.4	11.2	4.7 (2.1–7.2)*

* indicates that the confidence interval for the recent survey did not include the frequency of the hedge type and the cell of the previous survey.

Results of constrained ordination and multiple regression

All ordinations for different niches were significant ($p \leq 0.05$, Table 3), except LRDA and PRDA for contour planting and PRDA for the internal boundary, although there was some evidence for pattern for the internal boundary ($p = 0.06$; only 100 randomisation tests were made due to computational difficulties). Farm characteristics therefore do not explain differences in species composition for contour planting. PRDA explained substantially more variation than LRDA (respective averages are 47 and 14%) and could thus be preferred for ordination diagrams. The variation explained by the first axis of PRDA was on average only 49% of that of the first PCA axis, however. The fact that more variation is observed for PCA (an unconstrained ordination technique) than for RDA (a constrained ordination technique) means that some differences in species composition could not be explained by the explanatory variables (Kindt et al. 2006).

Several species dominated several niches, although the majority of species (55%) only dominated a single niche (Table 4). The species that dominated the highest number of niches were *Psidium guajava* (5 niches) and *Markhamia lutea* (4 niches). The dominant species of contour planting were a subset of the dominant species that were mixed in cropland and in fallows. The dominant species of external boundaries were a subset of the dominant species of external boundaries.

As examples of interpreting the ordination and regression results, the results for the external boundary and homestead area are discussed. The main differences in species composition for the external boundary can be observed between villages (Figures 1 and 2). Arrows in the ordination graphs indicate the directions for species and environmental variables where farms have larger values, whereas abbreviations in the figures show the typical position or centroid of farms of a certain category. Ebuchiebe contains more *Euphorbia tirucalli* (regression coefficient after stepwise regression = $r_s = 1.9$, $p < 0.0001$) and *Markhamia lutea* ($r_s = 0.7$, $p = 0.003$). Shimutu contains more *Lantana camara* ($r_s = 0.7$, $p = 0.05$), *Tithonia diversifolia* ($r_s = 1.6$, $p < 0.0001$) and *Psidium guajava* ($r_s = 0.5$, $p = 0.04$). Mutambi and Madidi can not be

clearly differentiated in the graph as the centroids were placed close together. Both villages contain more *Buddleja davidii* (negative coefficients for the other villages). *Cupressus lusitanica* was typical for Madidi (negative regression coefficients for the other villages). *Dracaena fragrans* is more dominant in Shimutu and Madidi (negative coefficients for the other villages). The ordination graph shows exceptional farms for each village that are plotted far from the centroid of the particular village.

The age of the household head was positively related to *Cupressus lusitanica* ($r_s = 2.0$, $p = 0.002$), as was a permanent house which is an indicator of wealth ($r_s = 2.2$, $p < 0.0001$). Larger farms contained more *Psidium guajava* ($r_s = 1.5$, $p = 0.005$), whereas farms with younger heads contained fewer trees of this species ($r_s = -1.0$, $p = 0.02$). Thatch-roofed houses contained less *Buddleja davidii* ($r_s = -0.5$, $p = 0.04$). Other patterns that were suggested by the ordination diagram were not confirmed by regression analysis.

Only three species are well represented in the ordination graphs for the homestead area (Figures 3 and 4). *Croton macrostachyus* is typical for Shimutu ($r_s = 1.0$, $p < 0.0001$). *Mangifera indica* is more typical of Ebuchiebe and Madidi (negative coefficients for the other villages). *Psidium guajava* had lower presence on farms with thatch-roofed houses ($r_s = -0.4$, $p = 0.009$). The species also occurred more in households with *de jure* female heads (female household heads that are not married) ($r_s = 0.3$, $p = 0.09$). *Croton macrostachyus* occurred more on farms with heads that had been longer in charge ($r_s = 0.5$, $p = 0.06$) and that had more cattle of local race ($r_s = 1.1$, $p = 0.003$) and less on farms with thatch-roofed houses ($r_s = -0.3$, $p = 0.02$). Of species that were not well represented, *Markhamia lutea* was associated with permanent houses ($r_s = 0.3$, $p = 0.03$) and the time the household head was in charge of the farm ($r_s = 0.8$, $p = 0.001$).

Discussion

Comparison with previous surveys

The results from the recent survey differed in various aspects with those obtained from two previous studies, while other patterns were con-

Table 3. Summary information on the ordination of the prevalent niches of trees on farms in western Kenya.

Function	All species	Number of dominant species	Variance		Significance		Variance		Significance		Variance	
			PRDA (all axes)	LRDA (all axes)	PRDA (all axes)	LRDA (all axes)	PCA (first axis)	PRDA (first axis)	LRDA (first axis)			
Cropland	61	10	39.70	17.12	0.01	0.0001	12.68	7.45	0.0001	6.48		
Homestead	62	7	40.41	13.79	0.01	0.0001	11.63	5.63	0.0001	4.01		
Woodlot	54	5	37.88	12.10	0.05	0.0006	33.78	10.21	0.0006	4.12		
Tree fallow	51	7	53.87	10.36	0.02	0.0184	31.98	13.43	0.0184	5.61		
External boundary	50	8	42.70	22.01	0.01	0.0001	20.63	13.31	0.0001	10.28		
Internal boundary	34	2	46.70	12.97	0.06	0.0002	24.59	10.36	0.0002	5.48		
Contour planting	19	2	66.18	8.76	0.44	0.1637	22.98	13.37	0.1637	3.84		

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%).

firmed. The specific locations where samples were taken were different in all surveys. The samples were also taken at different times and at different intensity, with only one sample per cell in the second survey and minimum 50 farms per village in the other surveys. Lauriks et al. (1999) (the second survey) also mentioned that some of their findings contrasted with those of Bradley (1991) (the first survey). For example, *Lantana camara* occurred on approximately a quarter of farms in stratum A1 in the first survey (Table 1), whereas this species dominated the same area in the second survey (hedge type 2, Table 2).

Since farms that belong to the same village can differ substantially in species composition as shown in the first and third survey, the sampling method of taking only one sample per village of the second survey could have sampled farms with hedge types that were less frequent in the village or cell. The regional distribution of hedge types in Vihiga depicted by the second survey could thus have been an artifact created by particular selections of farms. For instance, hedge type 3 was only sampled three times and types 4 and 5 twice – their distribution could be more an indication of being less frequent in the whole area than their occurrence in specific locations. An indication for potential sampling effects was also encountered during the second survey: the farm in the cell north of Madidi was sampled twice and contained a hedge more similar to type 1 and a hedge more typical of type 2.

Whereas the results obtained in various surveys did not agree completely with each other, they confirmed a spatial partitioning of species composition since typical species compositions for each village could be determined. The general spatial pattern of species composition over the survey area is therefore expected to be either patchier than described by the first and second survey (consisting of a smaller-scale mosaic of dominant compositional types), or to consist of large areas that have more irregular boundaries. To investigate which pattern of species composition provides a better description, a more detailed spatially distributed sampling scheme could be used (e.g., 1-km² cells or transect surveys). The sampling scheme should include various samples per village to investigate differences among farms at the smallest scale. Lauriks et al. (1999) had concluded that their method of using 25-km² cells was not detailed enough to reveal intra-district differences.

Table 4. Species that significantly contributed to ordination diagrams for the prevalent niches of trees on farms.

Species	Mixed in cropland	Homestead	Fallow	Woodlot	Crop contour	External boundary	Internal boundary
<i>Azadirachta indica</i>		+					
<i>Bridelia micrantha</i>			×	×			
<i>Buddleja davidii</i>						×	×
<i>Cajanus cajan</i>	+						
<i>Camellia sinensis</i>	+						
<i>Carica papaya</i>	+						
<i>Coffea arabica</i>	+						
<i>Croton macrostachyus</i>		×	×	×			
<i>Cupressus lusitanica</i>						×	×
<i>Dracaena fragrans</i>						+	
<i>Eucalyptus saligna</i>	×	×		×			
<i>Euphorbia tirucalli</i>						+	
<i>Harungana madagascariensis</i>			×	×			
<i>Lantana camara</i>						+	
<i>Mangifera indica</i>	×	×		×			
<i>Markhamia lutea</i>	×	×	×			×	
<i>Persea americana</i>	+						
<i>Psidium guajava</i>	×	×	×		×	×	
<i>Senna didymobotrya</i>			+				
<i>Sesbania sesban</i>	×		×		×		
<i>Syzygium cumini</i>		+					
<i>Tithonia diversifolia</i>						+	

×: Species that dominate several niches; +: Species that only dominate one particular niche.

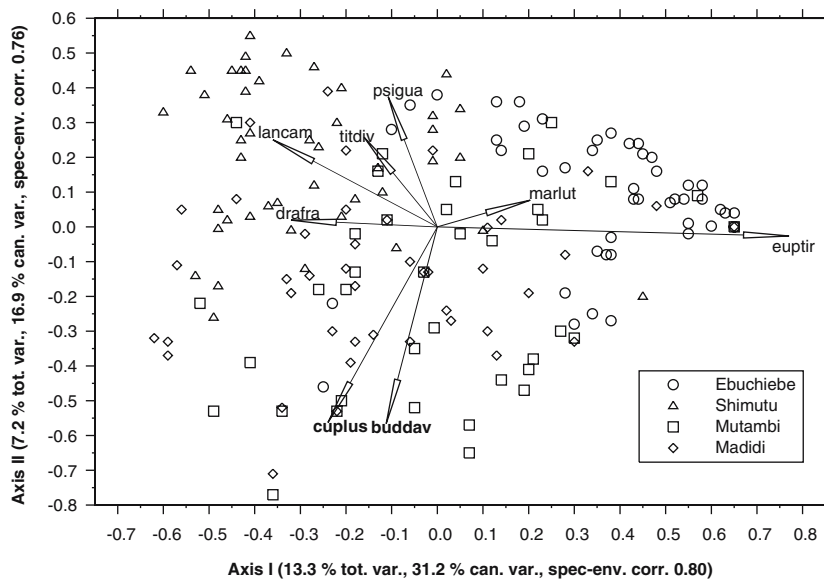


Figure 1. Ordination plot for trees on the external boundary provided by polynomial Redundancy analysis, showing positions of farms and species vectors for ordination axes 1 and 2 for the western Kenyan study. Abbreviations on the axes: % tot. var.: percentage of total variance explained on the axis; % can. var.: percentage of canonical variance explained on the axis; spec-env. corr.: species-environment correlation. Full species names can be found in Table 4.

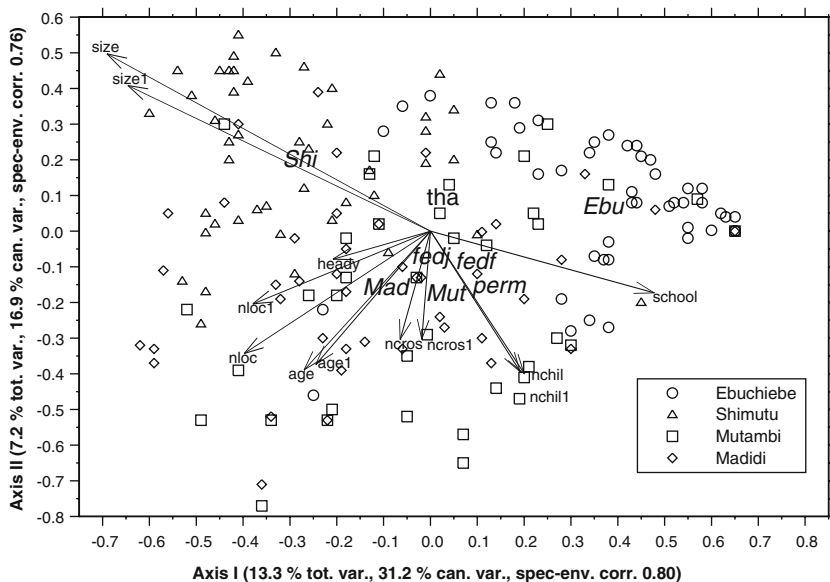


Figure 2. Ordination plot for trees on the external boundary provided by polynomial Redundancy analysis, showing positions of farms and explanatory variables for ordination axes 1 and 2 for the western Kenyan study. Abbreviations as in Figure 1.

Choice of method of investigating differences in species composition

A variety of methods was used in this article to investigate hypotheses that differences in species composition can be explained by some explana-

tory variables (such as the spatial subsection to which the site belongs). Some of these methods are compared in this section. One of the approaches involves calculating differences for different subsections of the survey area for each species separately (e.g., Bradley 1991; Harmand et al. 2003;

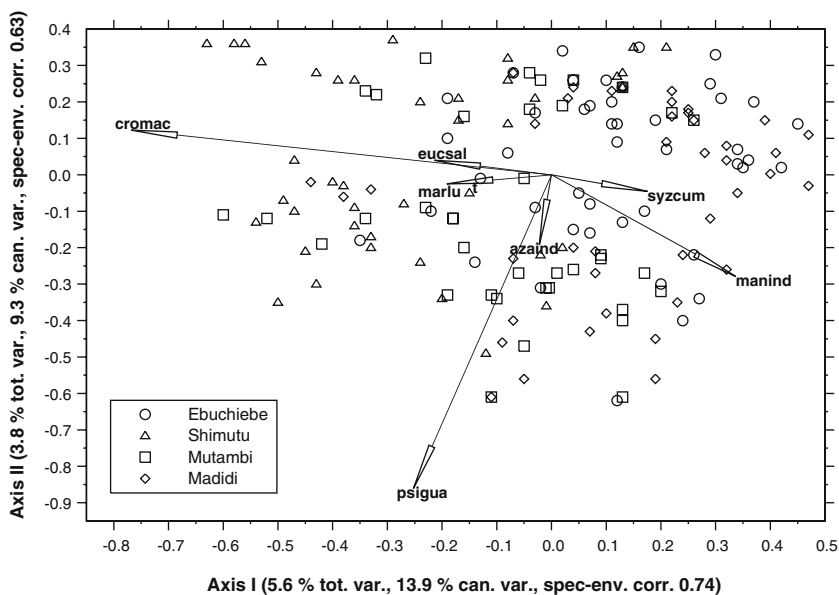


Figure 3. Ordination plot for trees in the homestead area provided by polynomial Redundancy analysis, showing positions of farms and species vectors for ordination axes 1 and 2 for the western Kenyan study. Abbreviations as in Figure 1.

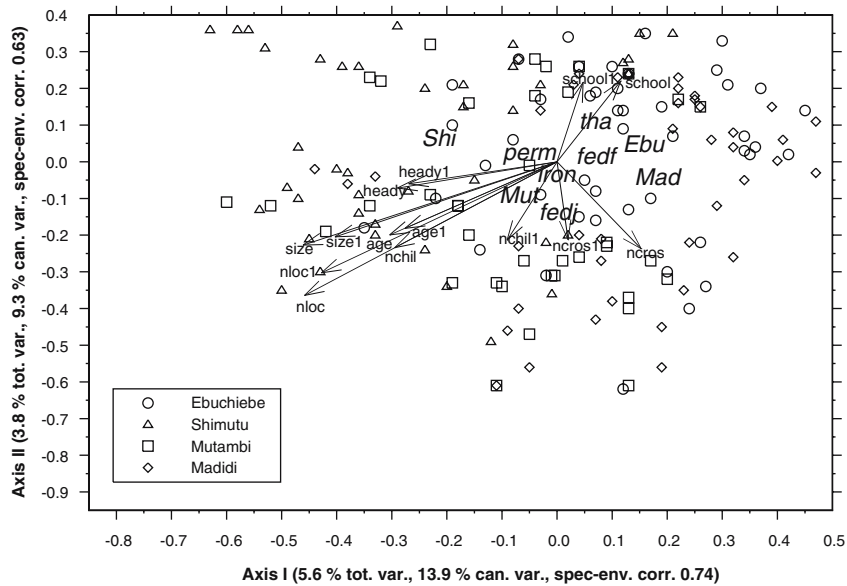


Figure 4. Ordination plot for trees in the homestead area provided by polynomial Redundancy analysis, showing positions of farms and explanatory variables for ordination axes 1 and 2 for the western Kenyan study. Abbreviations as in Figure 1.

Wezel and Bender 2003). For each subsection and species, researchers calculated the percentage of farms where the species was present, the average abundance, or the average proportion of trees. This approach has the advantage that the differences for each species are well documented. However, the approach will only be useful in describing differences among the subsections if it is coupled by a statistical test that investigates whether differences are significant. Such test can be provided by confidence intervals, χ^2 tests or multiple regression analysis. A disadvantage of this approach is that a separate analysis needs to be conducted for each species, which may not be practical. As an example, 62 species were listed for the homestead (Table 3), which would require 62 separate tests. Multiple testing may also result in a large type-I error.

A second approach calculates a distance matrix, where a pair-wise ecological distance is provided between all sites or between groupings of sites (e.g., Harvey and Haber 1999; Huang et al. 2002; Kaya et al. 2002). The advantage of this approach is that this truly is a multivariate analysis that is not based on multiple investigations for each species separately, so that this method does not suffer from the problems associated with multiple testing. One disadvantage of the method is that no

statistical test is provided to test whether differences are significant. Another disadvantage of the method is that no indication is given on the identities of the most important species for differences in species composition. Clustering methods (e.g., Lauriks et al., 1999; Rojas et al. 2001) can be described as approaches that summarise the information provided in the distance matrix. Due to this fact, clustering methods share the disadvantages of distance matrices. Another disadvantage is that some information is lost on the pair-wise differences among sites – in being a summary method, clustering methods show differences among clusters and not necessarily among individual members. Yet another disadvantage of clustering is that various algorithms exist that can lead to different results in case no clear clustering pattern exists in the data (Legendre and Legendre 1998; Quinn and Keough 2002). The advantage over distance matrices is that a summary of the results can be communicated in a single graph, and that it is possible to classify sites in a user-defined number of clusters.

Ordination methods overcome several of the shortcomings of the distance matrix and cluster approaches, especially by providing significance testing and information on important species for differences in species composition. Since ordination

results are usually interpreted by two-dimensional ordination graphs that only show a fraction of the total variance in differences in species composition, some information on patterns for individual species are lost. For this reason, ordination analysis can be combined with regression analysis, which provides statistical tests on differences in species composition, graphs that summarize differences in species composition, identities of the species that contribute most to significant differences in species composition, and analysis of the entire variance related to the most important species.

Results from the ordinations

As we observed for ordinations for functions of trees (Kindt et al. 2006), ordinations and multiple regressions did not explain all variation in the data. Although not all variation of Hellinger-distances among farms could be explained by RDA, the consistency between RDA and multiple regression results allows considering farm characteristics to guide diversification. Exceptional farms with species composition that was different from most farms of the same type resulted in low percentages of variation that was explained, but do not prevent targeting of interventions towards farms of the same type.

The regression results for the external boundary and the homestead area corresponded well to the results for all the trees found on the farm (Kindt et al. 2006). For example, the same patterns were observed in the distribution over villages of *Buddleja davidii*, *Croton macrostachyus*, *Cupressus lusitanica*, *Dracaena fragrans*, *Euphorbia tirucalli*, *Lantana camara*, *Markhamia lutea*, *Psidium guajava* and *Tithonia diversifolia*. The similarity in patterns is caused by the fact that most of the trees of the component species are found in the two niches.

A general process to determine the degree of tree diversity at farm and landscape levels and steps to increase it – based on the selection of areas of lowest diversity and analysis of compositional differences to identify species that can be distributed more widely – was discussed elsewhere (Kindt et al. 2006). Such process may lead to efforts that target contour planting and internal boundaries for diversification, since these niches had the lowest species richness of the average

farm (average species richness on contour planting is 0.2 [1.9 when only including those farms where the niche is represented], internal boundaries 0.9 [2.4], fallows 1.3 [6.3], woodlots 2.9 [3.9], external boundaries 4.5 [4.8], homesteads 5.0 [5.6], and cropland 6.5 [6.6]) and these niches only had two dominant species in ordination graphs (Table 4). Contour planting and internal boundaries could therefore be the focal niches for diversification. However, when we analyse niche diversity by the proportion of the dominant species (the Berger-Parker diversity index: Magurran 1988; Kindt et al. 2004), internal boundaries and contour planting are the most diverse with proportions of 16 and 19%, respectively (proportions for homesteads were 22%, external boundaries 30%, tree fallows 41%, cropland 56%, and woodlots 65%). The choice of the niche with the lowest diversity thus depends on the criterion used. Choice of the niche to diversify will in reality also depend on the importance attributed to its diversification by farmers, and is very likely to be related to the functions of the trees that are found in a particular niche.

Analysis for niches is different to analysis for functions (such as fruit, boundary demarcation or timber; Kindt et al. 2004, 2006), since the same tree can have several functions but can only occur in one niche. Diversification of one niche will therefore have no effect on the diversity of the other niches, whereas diversification of one function may also influence the diversity of other functions. This difference complicates the analysis of diversification for functions, and may require the collection of new information on utilisation patterns of trees when diversity has changed in one function. Although the investigation of diversification of niches is less complex, we expect that discussions with farming communities on diversification pathways will be more informative if such discussions focus on the different functions of trees rather than different niches of trees. Despite the facts that some information on the distribution of niches can be collected by remote sensing (the possibility of detecting a niche on aerial photographs is an important characteristic of good niche typologies; Bradley 1991) and that comparisons can be made with previous surveys, the facts that information on functions may be more relevant for planning diversification and that on-farm

sampling is required to identify species identities within niches (with the possible exception of some invasive nitrogen-fixing species; Asner and Vitousek 2005) leads us to the recommendation to always collect information on tree uses in tree diversity surveys. The collection of such ethnobotanical information should not dramatically increase the time needed for data collection, since information can be collected on a species-by-species basis rather than a tree-by-tree basis, although the latter option will offer the most detailed description of on-farm tree utilisation patterns (Kindt et al. 2004, 2006).

Conclusions

The present survey differed in various aspects with the findings of two earlier surveys. Because of these differences, larger-scale planning of tree diversification efforts requires more intense sampling than provided by the three studies, i.e., using a finer grid than the 5×5 km² grid of the second survey. Several farms should be sampled within the same village since not all farms within a village are expected to have similar species composition.

We recommend that future studies of tree diversity collect information on tree utilisation patterns if these studies are to be used for planning diversification. Ordination analysis combined with regression analysis as demonstrated in this paper offers several advantages that are not provided by other analysis methods of compositional differences, and should be considered for future studies of compositional differences.

Acknowledgements

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

- Asner G.P. and Vitousek P.M. 2005. Remote analysis of biological invasion and biogeochemical change. *Proc. Natl. Acad. Sci.* 102: 4383–4386.
- 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, 722pp.
- Bradley P.N. 1991. *Woodfuel, Women and Woodlots*, Vol. 1. Macmillan Education Ltd, London and Basingstoke, 338pp.
- Bradley P.N., Chavangi N. and Van Gelder A. 1985. Development research and energy planning in Kenya. *Ambio* 14: 228–236.
- Di Falco S. and Perrings C. 2003. Crop genetic diversity, productivity and stability of agroecosystems. A theoretical and empirical investigation. *Scot. 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.
- Harmand J.M., Donfack P. and Njiti C.F. 2003. Tree-root systems and herbaceous species-characteristics under tree species introduced into grazing lands in subhumid Cameroon. *Agroforest. Syst.* 59: 131–140.
- Harvey C.A. and Haber W.A. 1999. Remnant trees and the conservation of biodiversity in Costa Rican pastures. *Agroforest. Syst.* 44: 37–68.
- Hayek L.-A.C. and Buzas M.A. 1997. *Surveying Natural Populations*. Columbia University Press, New York. xvi, pp. 563.
- Huang W., Luukkanen W., Johanson S., Kaarakka V., Raisanen S. and Vihemaki H. 2002. Agroforestry for biodiversity conservation of nature reserves: functional group identification and analysis. *Agroforest. Syst.* 55: 65–72.
- Kaya M., Kammesheidt L. and Weidelt H.-J. 2002. The forest garden system of Saparua island, Central Maluku, Indonesia and its role in maintaining tree species diversity. *Agroforest. Syst.* 54: 225–234.
- Kindt R. 1997. Local Perceptions on Tree Propagation and Domestication – Results from a Survey in Western Kenya. *Int. Centre Res. Agroforest.*, Nairobi, 45pp.
- Kindt R., Van Damme P., Simons A.J. and Beeckman H. 2006. Planning tree species diversification in Kenya based on differences in tree species composition between farms I. Analysis of tree uses. *Agroforestry Systems* 67: 215–228.
- 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.
- Lauriks R. 1996. *Onderzoek naar veranderingen in landgebruik in West-Kenya met behulp van luchtfoto-interpretatie*. Ghent University, Ghent, 125pp.
- Lauriks R., De Wulf R., Carter S.E. and Niang A. 1999. A methodology for the description of border hedges and the analysis of variables influencing their distribution: a case study in western Kenya. *Agroforest. Syst.* 44: 69–86.

- 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, 853pp.
- 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.
- Magurran A.E. 1988. *Ecological Diversity and its Measurement*. Princeton University Press, Princeton, pp. x + 179.
- Makarenkov V. and Legendre P. 2002. Nonlinear redundancy analysis and canonical correspondence analysis based on polynomial regression. *Ecology* 83: 1146–1161.
- McNeely J.A. 2004. Nature vs. nurture: managing relationships between forests, agroforestry and wild biodiversity. *Agroforest. Syst.* 61: 155–165.
- Mathsoft I. 1999. *S-Plus 2000 User's Guide*. MathSoft Inc., Seattle, pp. xii + 558.
- Quinn G. and Keough M. 2002. *Experimental Design and Data Analysis for Biologists*. Cambridge University Press, Cambridge, 520pp.
- Rojas L., Godoy C., Hanson P., Kleinn C. and Hilje L. 2001. Hopper (Homoptera: Auchenorrhyncha) diversity in shaded coffee systems of Turrialba, Costa Rica. *Agroforest. Syst.* 53: 171–177.
- Simons A.J. and Leakey R.R.B. 2004. Tree domestication in tropical agroforestry. *Agroforest. Syst.* 61: 167–181.
- Wezel A. and Bender S. 2003. Plant species diversity of homegardens of Cuba and its significance for household food supply. *Agroforest. Syst.* 57: 39–49.