Genet Resour Crop Evol (2007) 54:1599–1612
DOI 10.1007/s10722-006-9172-7

RESEARCH ARTICLE

Spatial patterns of diversity and genetic erosion of traditional cassava (*Manihot esculenta* Crantz) in the Peruvian Amazon: An evaluation of socio-economic and environmental indicators

Louise Willemen · Xavier Scheldeman · Víctor Soto Cabellos · Simón Rafael Salazar · Luigi Guarino

Received: 6 February 2006 / Accepted: 15 November 2006 / Published online: 23 February 2007

Abstract This study evaluates quantitatively the suitability of the use of site-specific socio-economic and environmental data as indicators to rapidly assess patterns of diversity and genetic erosion risk in cassava. Socio-economic data as well as farmers' estimation of genetic erosion were collected in the study area, the Ucayali region of the Peruvian Amazon, through interviews with 285 cassava farmers in 50 communities, while diversity was assessed based on agromorphological characterization of 295 cassava accessions. Using multivariate regression analyses, 50 and 45% of the variation in respectively diversity and genetic erosion estimation could be explained by a selected set of socio-economic and environmental indicators. In both regression models four out of the total of 38 variables proved to contribute significantly (at $p < 0.10$ level). Additionally, the study revealed that farmers are a good direct source of information on the diversity present at community level, which can contribute to the development of methodologies to assess diversity more rapidly. The results of this study are valuable for the development of models to rapidly assess diversity dynamics in large areas.

Keywords Amazon · Genetic erosion · Genetic diversity · GIS · *Manihot esculenta*

Introduction

Remote indigenous settlements are often and stereotypically portrayed as refuges of agrobiodiversity. But is the cliché that equates picturesque landscapes and a patchwork of small fields with a multitude of traditionally cultivated crop varieties true?

Crop genetic diversity and its dynamics are believed to be a result of a complex process involving both human and environmental drivers. Globally, this process has led to a decrease in crop diversity in many farming systems, frequently because traditional varieties are being replaced by modern varieties (FAO 1997). However, as landraces are often well adapted to specific environments, they do have a clear advantage in marginal areas. Besides their direct use, these genetic resources have an important

---

L. Willemen · X. Scheldeman (✉)
Regional Office for the Americas, International Plant Genetic Resources Institute, Apartado Aéreo 6713, Cali, Colombia
e-mail: x.scheldeman@cgiar.org

L. Willemen
e-mail: Louise.Willemen@wur.nl

V. Soto Cabellos · S. Rafael Salazar
Instituto Nacional de Investigación y Extensión Agraria (INIEA), Av. La Molina 1981, La Molina 12, Peru

L. Guarino
Secretariat of the Pacific Community, Private Mail Bag, Suva, Fiji Islands

potential value in future breeding programs as well (IPGRI 2004), and therefore need to be conserved.

Observing this process of changing crop genetic diversity, Frankel (1970) already stated in 1970 that “it is now generally recognized that many of the ancient genetic reservoirs are rapidly disappearing”. Thereafter, the conservation of crop genetic diversity and a good understanding of the threats it faces have become an increasingly important issue in recent global initiatives such as the Global Plan of Action (GPA) for the Conservation and Sustainable Use of Plant Genetic Resources for Food and Agriculture, agreed by over 150 countries (FAO 1996) and the Strategic Plan of the Convention on Biological Diversity (Secretariat of the Convention on Biological Diversity 2002).

The GPA includes two key recommendations in the area of genetic erosion: under “Surveying and inventorying” (Recommendation 1) it urges to draw upon local knowledge to develop methodologies to assess diversity; under “Developing monitoring and early warning systems” (Recommendation 18) it calls for increased efforts to quantify, predict and mitigate the effects of crop genetic erosion. The Strategic Plan of the Convention on Biological Diversity identifies the lack of understanding on losses of biodiversity as a threat for effective implementation of conservation strategies, and should therefore be studied more extensively.

More easily accessible socio-economic and environmental data could serve as indicators for diversity or genetic erosion risk in cases where no diversity data are available (Hutchinson and Weiss 1999; Synnevåg et al. 1999). Researchers could thus carry out a rapid assessment of the current diversity and the threats it faces without using resource-intensive morphological or molecular analyses. By doing so, a much needed balance could be found between rapidly applicable but very general empirical models (as in Guarino 1995) and resource consuming but precise direct quantitative analyses (as in Brantestam et al. 2004). Different studies on the relationship between crop diversity and external factors have been carried out (Synnevåg et al. 1999; Morin et al. 2002; Oliveira and Martins 2002; Brush et al. 2003; Goa 2003; Guarino 1995) but, to our understanding, these studies have not truly evaluated in a quantitative manner the use of socio-economic and environmental factors as explanatory indicators of diversity. On a global level, identified threats to plant diversity mostly refer to wild species, such as those listed by the Millennium Ecosystem Assessment (2005), the Secretariat of the Convention on Biological Diversity (2001) and WWF (2004). These are, however, based on general assumptions and have only seldom been validated on the ground.

Responding directly to the above mentioned recommendations of the GPA and the Strategic Plan of the Convention on Biological Diversity, this paper evaluates quantitatively the suitability of the use of socio-economic data together with environmental data at community level to explain patterns of cassava diversity and genetic erosion risk in the Ucayali Department of the Peruvian Amazon. Most of the socio-economic and environmental data included in our analysis have been mentioned as potential diversity or genetic erosion indicators by other authors (Brush 1999; Guarino 1995). We hypothesize that not all these factors contribute equally to the diversity and genetic erosion estimation in the case study area; we aim, therefore, at presenting the most explanatory ones.

Although diversity and genetic erosion states were expected to be related, both were conceived as a result of distinct processes, under the assumption that, due to different local circumstances, not all areas presenting low diversity have been subjected to strong genetic erosion and vice versa.

The presented method should be able to rapidly indicate hotspots of genetic diversity as well as of genetic erosion risk, which would facilitate implementation of the necessary strategies to conserve and use the genetic resources present.

Material and Methods

Study area

The Amazonian region is known to be one of the most biodiverse on Earth, presenting also an important diversity within cultivated species (Coomes and Burt 1997; Vadez et al. 2004; Adin et al. 2004). The region is considered a centre of diversity of a number of crop species, including cassava (Vavilov 1951). It is well known that in the Amazon region natural habitats and the wild plants they contain are being lost rapidly. Human activities, which are an important cause of these losses, are also affecting the diversity of cultivated species (Vadez et al. 2004). The Ucayali region covers approximately 102,000 km^2, with some 315,000 inhabitants, of which the majority lives in the rapidly expanding city of Pucallpa (INEI 1993). The region is characterised by the Ucayali river and its many tributaries, including the Aguaytía river, and its intricate floodplain system bordered by slightly higher fluvial plains. Major road improvements in the region in the late 1960s facilitated the important connection between Pucallpa and Lima. Figure 1 shows an overview of the situation of the main study area characteristics. The natural vegetation of the entire region is rain forest. Two principal socio-cultural groups are found in the rural population: Amerindians, the indigenous dwellers of the region; and colonists, immigrants mainly coming from Lima, the Andes or other Amazonian provinces (in equal parts). The indigenous inhabitants are dedicated primarily to farming along the fertile floodplains of the rivers, complementing this activity by fishing and hunting. The colonists are farmers or cattle ranchers living primarily along the highway to Lima and secondary roads penetrating the adjacent forest.

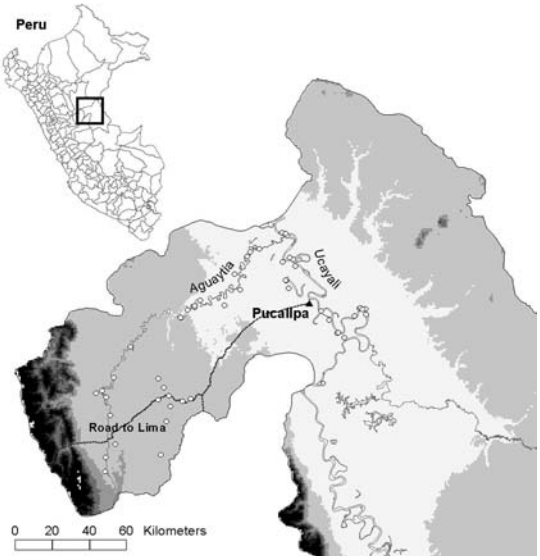

Fig. 1 Study area. Besides presenting the elevation intervals with grey-shades, the two main rivers, the city of Pucallpa, the main road to Lima, and the locations of all studied communities are displayed

Data

Between 2001-2002, a field survey was carried out interviewing 285 farmers in 50 communities in the Ucayali region, representative of the range of socio-economic and agro-ecological conditions present. A total of 10–15% of the households in each community were surveyed. The survey, based on the potential diversity indicators by Brush (1999) and the empirical model by Guarino (1995), included questions on the socio-economic situation of the farmers, their farming practices, cultural background and current agricultural trends. Additionally, farmers were asked about the number of distinct cassava varieties grown at the time of the survey and about their observation of losses of varieties in the recent past. Table 1 presents a complete overview of the survey data. All survey results were summarized per community; the mean value together with its standard deviation were calculated for the quantitative data, and in the case of qualitative multi-state questions the percentage for each class was presented.

Coordinates were taken for all the communities surveyed using a Global Positioning System (GPS) receiver. Spatial data for the study area were extracted from different digital sources. Mean annual temperature, annual precipitation and altitude data were obtained from Worldclim 1.3 (Hijmans et al. 2005), all at 30-second resolution (grid cell ± 1 km) grids. The Instituto de Investigaciones de la Amazonía Peruana (IIAP) provided thematic maps of the region with the main roads, rivers, urban areas, and geomorphology layers, the Centro Internacional de Agricultura Tropical (CIAT) a deforestation map, indicating the number of years since forest was cleared. With the aid of these digital maps, distance to the main road, rivers and Pucallpa were also derived for each community using a GIS (see Table 1).

Table 1 Summary of the survey and spatial data. For the numerical factors the mean and standard deviations over all communities are given. The categorical factors are summarized by the percentage of each class

Factor	Description		
Survey data			
Numerical	Mean	SD	
Amount of varieties planted by farmer	2.1	0.7	
Number of varieties known by farmer	2.8	0.8	
Farmer's age	44.4	6.2	
Number of families members	5.7	1.8	
Years of residence in community	18.9	10.3	
Farm area (ha)	10.4	9.4	
Area cultivated with cassava (ha)	1.03	0.6	
Annual cassava production (kg)	1113	812	
Categorical binomial (yes/no)	%		
Observation of variety losses (yes)	47.1		
Native ethnic group (yes)	37.3		
Use of fertilizer (yes)	8.3		
Use of herbicides (yes)	4.1		
Possession of livestock (yes)	95.8		
Use of contracted labour forces (yes)	25.0		
Contact with agricultural extension services (yes)	25.5		
Local origin of planted varieties (yes)	95.0		
Intention to cultivate more varieties (yes)	96.4		
Cultivation of commercial varieties (yes)	1.2		
Knowing other persons with more varieties (yes)	75.5		
Knowing other communities with more varieties (yes)	87.9		
Experiencing good market possibilities (yes)	38.4		
Experiencing problems in cultivating cassava (yes)	56.1		
Use of previous harvest as planting material (yes)	87.8		
Farmer's intention to continue growing cassava (yes)	93.3		
Categorical nominal (class order)	%	%	%
Cassava use (alimentation/processing/market)	81.1	16.3	2.6
Source of planting material (farm/relative/market)	71.0	23.7	5.3
Observation of trend in the under cassava cultivated area (decrease/stable/increase)	6.4	83.1	10.5
Observation of trend in number of planted cassava varieties (decrease/stable/increase)	22.3	69.2	8.5
Estimation of present diversity (low/medium/high)	40.6	37.4	22.0
Observation of trend in cassava price (decrease/stable/increase)	12.7	65.3	22.0
Spatial data			
Numerical	Mean	SD	
Altitude (m)	196	60	
Years after deforestation (yr)	21.3	7.1	
Mean annual temperature (°C)	26.0	0.3	
Annual precipitation (mm)	2826	1132	
Distance to Pucallpa (km)	60.7	39.5	
Distance to river (km)	0.9	1.5	
Distance to road (km)	20.1	10.9	
Categorical nominal (class order)	%	%	%
Geomorphology (mountains/ hills/ plains)	24.0	18.0	58.0

Additionally, cassava planting material was collected at the 50 communities in the study area. In total, 295 cassava samples, representative for the morphological diversity at community level, were collected. This germplasm was planted at research stations in the Pucallpa area and characterized *ex situ* using a cassava descriptor list (Fukuda and Guevara 1998). A total of 24 different quantitative morphological traits were measured and some 36 qualitative were classified. Germplasm is being maintained in the Peruvian national genebank, managed by the Instituto Nacional de Investigación y Extensión Agraria (INIEA).

Diversity and genetic erosion indices and distribution

To define cassava diversity at community level, the morphological characterisation data were combined into a single index, by grouping morphologically similar accessions using cluster analysis. Hair et al. (1998) recommend that to improve cluster analysis ''the selection of the variables to be included in the cluster analysis must be done with regard to theoretical and conceptual as well as practical considerations,'' indicating that only variables that characterize well the objects being clustered should be included. We included solely the most distinguishing descriptors, as listed by the Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA) in Fukuda and Guevara (1998).

All selected quantitative variables were standardized to reduce the effect on the clustering due to different measurement scales. As both quantitative and qualitative data were used in the same analysis, the DAISY algorithm for mixed data types (Kaufman and Rousseeuw 1990) was used to calculate dissimilarities between accessions. A hierarchical cluster analysis was carried out on the dissimilarity matrix using the average linkage method to realize the grouping. The Kelley-Gardner-Sutcliffe (Kelley et al. 1996) penalty function was applied to determine the optimal number of clusters for the hierarchical clustering. It was thus possible in the end to assign each cassava accession to a cluster group. The number of different clusters present in a community was used as the indicator for the cassava agromorphological diversity present, and this is referred to below as the ''diversity index''. All statistical operations in this paper were carried out using the statistical software package R 2.0 (R Development Core Team 2004).

Due to a lack of baseline and time series data, genetic erosion could not be measured directly in the field (although the data collected by this study now clearly forms an important baseline for future studies following trends in cassava diversity). Therefore, farmers' own observations on the loss of varieties within their communities were used to define a ''genetic erosion index''. The term ''genetic erosion'' in this paper is used to mean the total loss of crop varieties at a particular location within a certain time span, in this case the losses of traditional cassava varieties in different communities in the Ucayali department in the last decade. The presence or absence of genetic erosion was assessed by asking farmers whether they had noticed any disappearance of varieties in the last 5–10 years. The proportion of farmers per community confirming the loss of varieties was used as an index of genetic erosion for each community.

Both the diversity and genetic erosion index per community were plotted in a GIS using different circle sizes to indicate their value in a given community. To visualize general spatial patterns, community point data were converted to a raster format with grid cell of 0.15 degree (~16 km), using a 0.30 degree mean circular neighbourhood function to reduce the effect of the arbitrary location of grid cells. Afterwards, all cells not containing any community points were removed from the map. The GIS software packages DIVA-GIS 4.2 (Hijmans et al. 2004) and ArcView 3.2 were used for all spatial data processing.

Understanding and modelling diversity and genetic erosion

Regression analyses were carried out to identify those factors most closely linked to the levels of diversity and genetic erosion. A univariate regression was executed to identify survey and spatial factors showing a significant relationship with the agromorphological diversity and genetic erosion indices. Factors having a continuous data type were included directly in the regression analysis, while the categorical factors were analysed per class. The class analysis of categorical data was based on the numerical proportion of farmers falling in a given class (survey data) and by transforming a multi-state factor in separate binary classes (spatial data). To improve the regression calculations a logarithmic transformation was performed on the diversity index, but the genetic erosion index was left in its original form. All factors showing probability values lower than 0.20 in the univariate regression were included in a multivariate regression model. By removing the least significant variables through manual step-wise regression, a combination of the most significant ($p < 0.10$) variables was determined. The Variance Inflation Factor (VIF) was calculated to test all variables included in the multivariate regression model for a possible multicollinearity, the undesired relationship between independent variables. The combination of variables together with their beta estimates in the multivariate model lead to an indication of the most related diversity and genetic erosion factors in our study area.

Regression model performance

Studentized residuals, the most common way of presenting standardized model errors (Hair et al. 1998), were calculated for each community for both the genetic diversity and the genetic erosion models. By mapping these residuals in a GIS, the spatial distribution of model errors could be visualized. Here again, all values per community were displayed using different circle sizes and the residual values were rasterized using neighbouring values to give a general spatial overview. Over- and under-predictions of diversity and genetic erosion values indicate that factors are missing in the regression model, or that variables included in the model are not valid in a specific location. Negative values would indicate grid cells where a higher diversity was predicted then observed, and vice versa. This test can be useful for future site visits, to look more in depth at the specific situation and determine other possible important indicators.

Alternative diversity index

In this study, cassava characterization data were used to define the cassava diversity in Ucayali department, but obtaining these agromorphological descriptions is in general time consuming. To test the use of a faster and easier way to obtain a diversity index, a diversity figure based on farmers' knowledge of the number of varieties present was compared with the agromorphological diversity index. All surveyed farmers were asked to name the cassava varieties currently grown in their fields. The average richness in cassava varieties per farmer in a community was taken as the alternative diversity index. By taking the average of the number of different varieties planted per farmer, rather then the total number of different varieties planted in a community, problems with inconsistent naming of local crop varieties (Elias et al. 2001; Wood and Lenné 1997) among the farmers could be avoided. This alternative diversity index was mapped in a GIS to visualize its spatial pattern. The difference between the diversity indices was calculated and the Spearman rank correlation test was performed to test how strongly they were associated.

Results

Diversity and genetic erosion indices and distribution

Seven quantitative agromorphological variables characterizing the cassava samples and 16 qualitative variables were selected based on the "minimal" and "principal" descriptor list of Fukuda and Guevara (1998). See Table 2 for an overview of these traits. The Kelley-Gardner-Sutcliffe penalty function distinguished eight different groups of cassava based on the hierarchical clustering of agromorphological characteristics. Figure 2 shows a map of the morphological diversity of cassava in the study area. The number of cluster groups per community varied between one and six, as represented by the different sized circles in the map. The rasterized diversity pattern shows that the highest agromorphological cassava diversity was found in the southwestern part of the department and downstream along the Aguaytía river. Alongside the Ucayali river and upstream the Aguaytía river communities with a lower diversity, with an average of 1.5–2.5 morphological groups per community. Patterns of the genetic erosion index (Fig. 3) showed that a high

Table 2 Morphological cassava descriptors, used as basis for the diversity index calculation

Quantitative descriptor	Qualitative descriptor
Colour of the apical leaves	Number of leaf lobes
Colour of mature leaves	Length of central leaf lobe (cm)
Shape of the central leaf lobe	Width of central leaf lobe (cm)
Colour of the stem epidermis	Ratio length/width of central leaf lobe
Colour of stem cortex	Length of petiole (cm)
Colour of stem	Internode distance (cm)
Colour of the petiole	Plant height (cm)
Presence or absence of flowers	
Stem growth habit	
Number of levels of branching	
Presence of storage root peduncle	
Colour of the storage root surface	
Colour of the storage root cortex	
Colour of the root pulp	
Epidermis texture of the storage root	
Presence of storage root constrictions	

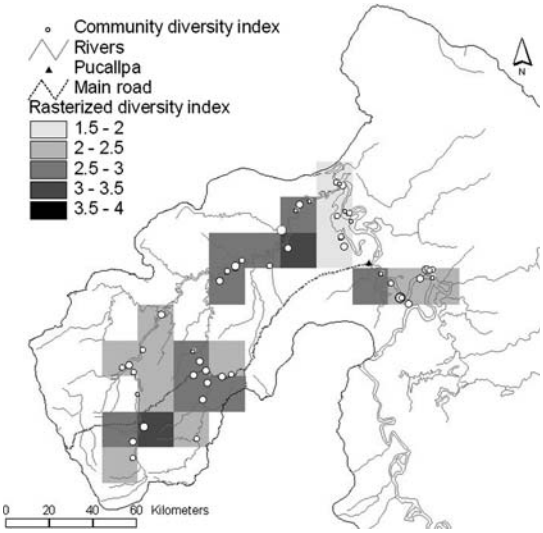

Fig. 2 Distribution of cassava diversity in the Pucallpa area. Dot size indicates the magnitude of the agromorphological diversity index, defined as the number of morphologically distinct cassava clusters per community. This index was rasterized to visualized spatial diversity patterns (grid size 0.15′, values calculated using a circular neighbourhood of 0.30′)

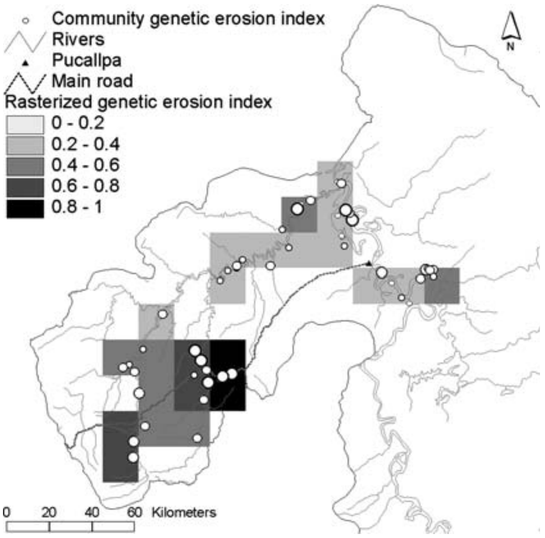

Fig. 3 Distribution of cassava genetic erosion. Dot sizes indicate the magnitude of the genetic erosion index, defined as the proportion of farmers per community having noticed a loss in cassava varieties per community. This index was rasterized to visualize spatial genetic erosion patterns (grid size 0.15′, values calculated using a circular neighbourhood of 0.30′)

proportion of farmers living near the new road joining Pucallpa with Lima noticed a loss of cassava varieties in the last decade. In communities along the two main rivers, the Aguaytía and Ucayali, less genetic erosion was reported.

Understanding and modelling diversity and genetic erosion

Univariate regression analyses were used to identify single relationships between the socio-economic and environmental data on one side and the diversity index on the other. Table 3 presents for the most related variables ($p < 0.20$) the R-square, p-value and the direction of the beta estimate. A positive beta estimate indicates higher diversity index or genetic erosion with an increasing variable value. After entering the 15 factors having a $p < 0.20$ in a multivariate model and removing the less significant variables stepwise, four variables remained significant at a $p < 0.10$ level (Table 4). Besides a high diversity estimation by farmers, the observation of an increasing trend in number of planted cassava

Table 3 Univariate regression results for cassava diversity and genetic erosion. Only the variables with a $p < 0.20$ are displayed

Factor	Diversity			Genetic erosion		
	B[a]	$R^{2,b}$	*p*-value[c]	B	R^2	*p*-value
Survey data						
Continuous						
Number of varieties known by farmer				+	0.01	0.026
Number of families members	–	0.07	0.053			
Annual cassava production (kg)				+	0.06	0.087
Farm area (ha)	+	0.07	0.072	+	0.16	0.004
Classes						
Observation of variety losses in the last decade (yes)	–	0.13	0.011			
Use of fertilizer (yes)	+	0.04	0.154	–	0.13	0.009
Possession of livestock (yes)				+	0.09	0.037
Use of contracted labour forces (yes)				+	0.08	0.042
Local origin of planted varieties (yes)				–	0.09	0.040
Farmer's intention to continue growing cassava (yes)	+	0.05	0.105			
Cultivation of commercial varieties (yes)				+	0.09	0.036
Experiencing good market possibilities (yes)				+	0.08	0.044
Experiencing problems in cultivating cassava (yes)				+	0.08	0.044
Native ethnic group (yes)				+	0.07	0.067
Cassava use (alimentation)				–	0.10	0.025
Cassava use (processing)				+	0.11	0.017
Source of planting material (relative)				–	0.07	0.069
Source of planting material (market)	–	0.03	0.194	+	0.01	0.008
Observation of trend in the under cassava cultivated area (increase)	–	0.05	0.127			
Observation of trend in number of planted cassava varieties (decrease)				+	0.09	0.035
Observation of trend in number of planted cassava varieties (stable)				–	0.09	0.039
Observation of trend in number of planted cassava varieties (increase)	+	0.06	0.087			
Estimation of present diversity (low)	–	0.11	0.018	–	0.04	0.164
Estimation of present diversity (stable)				+	0.04	0.134
Estimation of present diversity (high)	+	0.14	0.006			
Observation of trend of cassava price (increase)	+	0.05	0.116	–	0.06	0.082
Spatial data						
Continuous						
Altitude (m)				+	0.18	0.002
Mean annual temperature (°C)	–	0.04	0.146	–	0.18	0.002
Annual precipitation (mm)				+	0.13	0.011
Distance to Pucallpa (km)				+	0.11	0.018
Distance to road (km)				–	0.14	0.006
Classes						
Geomorphology (mountains)	+	0.04	0.170	+	0.15	0.005
Geomorphology (hills)				–	0.11	0.017
Geomorphology (plains)	–	0.09	0.035			

[a] The direction of the beta coefficient indicates the positive or negative contribution of each variable to the diversity or genetic erosion index estimations

[b] The *R*-square indicates the amount of variation explained by each variable

[c] The *p*-value indicates the probability that the result obtained is due to chance rather than a true relationship between measures, so the lower this value, the more likely the relationship

varieties showed a positive relationship with cassava diversity. The observation of variety losses and the number of family members had a negative relationship with the amount of agromorphological diversity present. These four variables together explained 50% of the total variation among communities' diversity index and did not show strong (VIF~1) multicollinearity.

Table 4 Multiple regression model for cassava diversity

Coefficients	Beta estimate	*p*-value
Intercept	1.215	<0.001
Estimation of present diversity (high)	1.035	<0.001
Observation of trend in number of planted cassava varieties (increase)	1.237	<0.001
Number of family members	–0.096	0.001
Observation of variety losses (yes)	–0.480	0.002
R-square	0.50	
Residual st. error	0.33	

The same tests were carried out on the genetic erosion index. Results of the univariate analyses are presented in Table 3 as well. In the multivariate regression analysis, four variables showed a significant ($p < 0.10$) relationship to reported genetic erosion (Table 5). The presence of livestock and the use of planting material acquired in markets gave positive beta estimates in the regression function, indicating a significant influence on the loss of varieties. Communities located far from the road and locations in hilly areas showed a negative relationship with farmers reporting losses of varieties. This multiple regression model covered 45% of the total variation in the genetic erosion index among the communities without showing strong multicollinearity (VIF~ 1) among the independent variables.

Regression model performance

Studentized residuals of the multivariate models, as presented in Tables 4 and 5, were calculated. To visualize spatial patterns of these errors, the residuals were plotted for each community and generalized by rasterizing them (Figs. 4 and 5).

Table 5 Multiple regression model for the presence of cassava genetic erosion

Coefficients	Beta estimate	*p*-value
Intercept	0.761	<0.001
Livestock present (yes)	1.024	0.004
Source of seeds (market)	0.503	0.044
Distance to road	–0.010	0.004
Location in hilly area	–0.301	0.002
R-square	0.45	
Residual st. error	0.24	

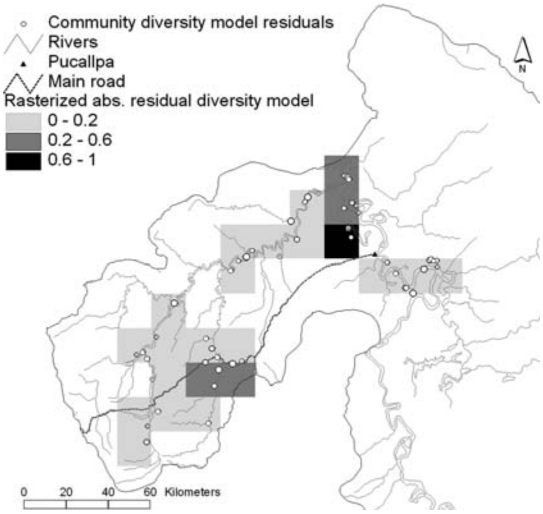

Fig. 4 Distribution of the studentized residuals of the cassava diversity multiple regression model. Dot sizes indicate the magnitude of the residuals per community, the standardized differences between the observed and predicted diversity. These values were rasterized and the absolute value was taken to visualize spatial patterns (grid size 0.15′, values calculated with a circular neighbourhood of 0.30′)

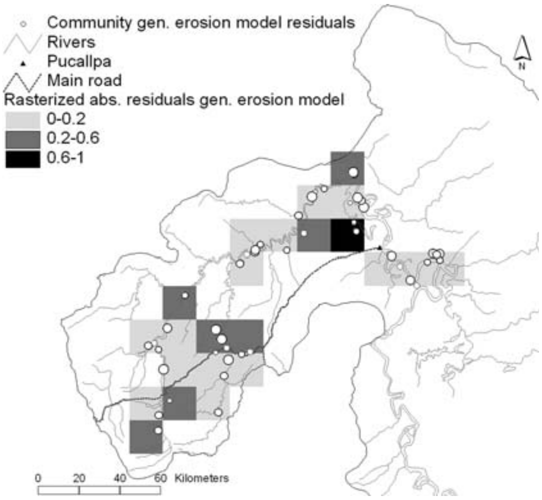

Fig. 5 Distribution of the studentized residuals of the cassava genetic erosion multiple regression model. Dot sizes indicate the magnitude of the residuals per community, the standardized differences between the observed and predicted genetic erosion. These values were rasterized and the absolute value was taken to visualize spatial patterns (grid size 0.15′, values calculated with a circular neighbourhood of 0.30′)

Significant residuals of the diversity model were found downstream along the Ucayali river and near the road in the southwest corner of the department. As can be observed from Fig. 5, residuals for the genetic erosion model showed a more scattered error pattern, but a small concentration of large model residuals was found along the road to Lima and near the city of Pucallpa.

Alternative diversity index

The spatial pattern in cassava varietal diversity as reported by farmers is presented in Fig. 6, which allows comparison with that of the agromorphological diversity index. Similar to Fig. 2, the highest diversity based on the survey data was found in communities in the southwestern part of the study area and downstream the Aguaytía river. An average difference of 0.28 was found between the diversity index based on morphological characteristics and the diversity index based on the average number of varieties per household within a community reported by the farmers. That is, on average, 0.28 more clusters are described in a community than varieties are reported by farmers. A Spearman rank correlation test showed an $r = 0.48$ with a $p < 0.001$, confirming that the indices are strongly correlated.

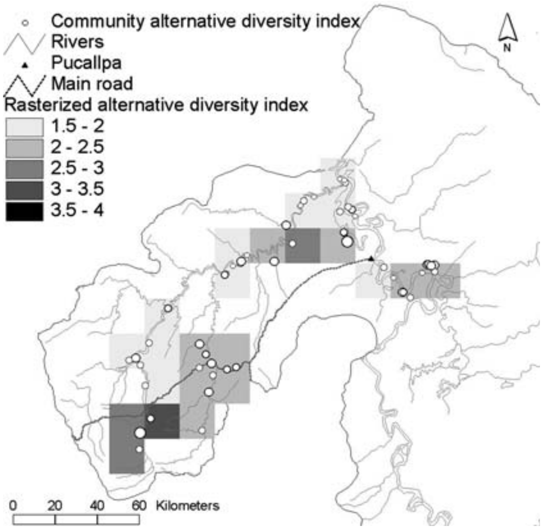

Fig. 6 Distribution of the reported cassava diversity, the average number of reported cassava varieties planted per farmer per community. Dot size indicates the magnitude of the alternative diversity index per community. This index was rasterized to visualize spatial diversity patterns (grid size 0.15′, values calculated using a circular neighbourhood of 0.30′)

Discussion and Conclusions

Cassava diversity and genetic erosion

The four variables that remained in the final diversity model all came from the survey data set. The observation of variety losses showed a negative correlation with the diversity index. Indeed, a number of communities with a low diversity index have experienced variety losses in the past. Also the number of family members remained in the model with a negative regression coefficient. This negative relation needs further exploration.

The farmers' estimation of current diversity was very strongly related to the current state of diversity, indicating, unsurprisingly, that farmers are well aware of the status of the community's diversity. Although this variable is not explaining the processes behind the diversity index, it could be used as an indicator in future diversity assessments. Likewise, the observation of an increased number of planted cassava varieties also showed a positive relationship with the number of morphological groups found in a community. So, three variables of the final model represented the farmers' observations on cassava diversity which were able to explain parts of the agromorphological diversity in the area.

No spatial factors derived from the digital maps remained in the final diversity model, suggesting that the idea of a rapid assessment of cassava diversity based on the analysis of secondary spatial data is unfortunately not viable in this case, but from the univariate regression calculations it can be seen that significantly less cassava diversity was found in the river plains. This coincides with the findings by Brush (1995) that areas showing marginal agronomic conditions, as those found in mountainous areas, and economically isolated areas, generally have a higher diversity. Plains often have more homogenous environmental conditions, which could lead to stronger competition with commercial varieties. In the Ucayali area, the reason for finding lower cassava diversity in the plains could also be that cassava typically is cultivated on slopes, as the lowlands are often used for other crops, such as groundnut.

The proportion of farmers in a community that possess livestock is correlated with higher genetic erosion. As in many other areas of the Amazon, land in the Ucayali region has been turned into pasture (Fujisaka and White 1998), mostly by new settlers. Such land use change has already been put forward as a possible cause of genetic erosion by Hutchinson and Weiss (1999). Fujisaka et al. (1998) also reported the highest losses of wild plant diversity in the Brazilian Amazon in areas where conversion to pasture has taken place. Communities where a high number of farmers reported sourcing of planting material from markets also reported a loss of varieties. This could indicate that the planting material available at local markets does not possess a high diversity, which could lead to further erosion of the present diversity. This finding corresponds with other studies that stress the importance of local planting material in relationship to diversity trends (Brush et al. 2003; Coomes and Ban 2004). Communities closer to the main road to Lima have clearly noticed more variety losses than settlements further away. In the Ucayali region, newcomers mostly settled near the new road to Lima (results not shown here), an important link to markets, and are thus subject to an increasing pressure to adopt commercial agricultural practices. As expected and stated before (Brush 1999), in communities located in hilly areas fewer farmers had noticed any variety losses.

One factor generally expected to influence diversity and genetic erosion strongly, a farmer's ethnicity, was not found to do so in this study. Adin et al. (2004) did not find convincing results that indigenous communities maintained a higher agrobiodiversity in the Peruvian Amazon than colonist communities. Guarino (1995) proposed a widely applicable empirical model containing genetic erosion risk factors. These included cultivated area extent, use of modern varieties, influence of agricultural extension services, mechanization, herbicide and fertilizer use and farming population. None of these factors showed a significant relationship to diversity losses in our specific study area.

Alternative diversity index

Although this study did not find clear socioeconomic indicators that allow the prediction of cassava diversity, it did find that farmers can play a key role in the rapid assessment of the current state of the number of varieties maintained in their community. The diversity index was based on morphological similarities, which is also the way farmers identify their varieties (Mkumbira et al. 2003), and a significant similarity between the two indices was found. Additionally, three variables in the diversity model were based on farmer's observations of cassava diversity, which also supports the conclusion that farmers are well aware of the number of cassava varieties present in their community. This opens perspectives for future rapid assessments of diversity, avoiding – at least initially – the process of morphological characterization.

General applicability of rapid assessment methodology

The methodology used here found indicators explaining 50 to 45% of the variation in respectively cassava diversity and genetic erosion among the communities studied. These indicators provide researchers and policy makers the opportunity to develop conservation strategies aimed at maximizing diversity and mitigating genetic erosion. As the study showed that farmers who buy their planting material from markets tend to experience more genetic erosion, one possible strategy could be the organization of diversity fairs, which allow farmers to acquire and exchange a wide range of local varieties from a large area. It would also allow regular monitoring of changes in diversity. This strategy has been applied successfully with tuber and grain crops in the Andean region (Gonzalez and Terrazas 2004).

The different indicators included in the models confirm that diversity and genetic erosion are two different concepts that can be explained by different processes. The assumption that simply all negative beta estimates in the diversity model could be used to assess genetic erosion risks is in fact not supported by this study, although both indices are strongly related. None of the variables in the multivariable model having a negative beta value in the diversity model showed a significant relationship with the amount of genetic erosion reported.

As 50–55% of the variation in diversity and genetic erosion among the communities remained unexplained by the models, the maps showing the over- and under-predictions of the models could help to define additional possible factors related to diversity or genetic erosion. Possible additional drivers, some mentioned in other studies, could include fragmentation and heterogeneity of land holdings per farmer (Brush 1995), the share of off-farm work (Brush et al. 2003), calamities, both natural and man-made (Maxted et al. 1997), and the number of crops present at farm level. This last could be important as farmers might be focusing on other crops rather than on cassava, which could thus explain relatively lower cassava diversity. Genetic erosion still has not been measured directly in the field in our study area. However, the baseline data collected can be used to compare with future diversity in the same area. This would allow assessing actual genetic erosion, which would enable fine-tuning the model even further.

Something to keep in mind is that the indicators presented are factors correlated with diversity and reported genetic erosion at the community scale. As in many spatial models, scale issues are important. Many communities around the city of Pucallpa showed a low cassava richness based on morphological characteristics, but when looking at a higher scale, the area around Pucallpa turned out to have one of the highest richness figures (results not presented here). In this area, a low number of different morphological groups was present per community, but different communities had different morphological groups, so in total they shared a higher diversity. It seems that a community around Pucallpa adopts only a limited number of varieties, based on local preferences, while a more isolated community in the hilly areas prefers cultivating a wider spectrum of cassava varieties. One community in this area is an exception to this and holds six different agromorphological groups. Brush (1999) mentioned the difficulties of choosing the "right" scale for measuring diversity. In another study (Brush et al. 2003), he describes maize farmers in Mexico cultivating only one variety but, because of a dynamic seed management system, the region has a diverse maize population. This implies that researchers or policy makers must have clear goals set on what level they want to measure diversity and genetic erosion.

Besides the scale issue, the indicators presented in this paper are only valid for the specific areas and crop studied and they still need to be confirmed in other crops and other regions. Carrying out the same analyses in other ecogeographic zones or regions and on other crops would therefore be an important first step in the development of a general rapid assessment system.

Ramanatha Rao and Hodgkin (2002) questioned whether crop diversity could be explained largely in terms of ecogeographic factors or if socio-economic factors have the greatest impact. In our study both types of variables made significant contributions in the final models. It is, however, difficult to fully separate environmental and socio-economic factors. Hilly areas in the Ucayali Department tend to have more cassava diversity but it is difficult to assess if this is due to differences in cropping conditions, such as soil and climate, or due to the more isolated status of these areas.

This study has shown that farmers can be consulted directly to derive diversity information, so that resource-consuming morphological characterization may not necessarily be needed as part of an initial rapid assessment of diversity. A possible strategy for a rapid assessment of diversity and its losses therefore could be to carry out a quick survey with local farmers (only targeting present diversity and knowledge of variety losses) in a representative number of study units (depending on the chosen scale) and to analyse these, using multivariate regressions, with a large set of easily accessible spatial data (both environmental and socio-economic). This would, extrapolating the regression model to a larger area, allow obtaining estimations of the diversity and its dynamics, which can be represented on maps. The results obtained would allow local policy makers to develop strategies to conserve the present diversity. As mentioned above, this strategy, although proven to be valid in this study, should first be tested, by comparing with real field data, in other regions and on other crops to get more information on its applicability.

Coming back to our introductory question, we can conclude that more research is needed before agrobiodiversity and its dynamics can be assessed purely based on indicators. Until then, caution is warranted when making general assumptions on drivers behind crop diversity.

Acknowledgements The authors wish to thank the Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) for providing the resources to carry out this study through the project "Patterns of genetic diversity and genetic erosion of traditional crops in Peru: rapid assessment and risk prediction using GIS tools" (Project 99.7860.2-001). Special thanks go to, Jorge Alcántara, Brenda Castro, Wilfredo Guillén, Jaime Mori, Llerme Ríos and all other INIEA staff members who were responsible for the project's fieldwork. The study would not have been possible without the local spatial data provided by the Instituto de Investigaciones de la Amazonía Peruana (IIAP) and the Centro Internacional de Agricultura Tropical (CIAT). Our thanks also go to Michael Hermann and Veerle Van Damme for their critical review of this paper. Finally, the authors would like to express their sincerest gratitude to all participating farmers and local authorities of the Ucayali Department, whose participation in this study was crucial to obtain the presented results.

References

Adin A, Weber JC, Sotelo Montes C, Vidaurre H, Vosman B, Smulders MJN (2004) Genetic differentiation and trade among populations of peach palm (*Bactris gasipaes* Kunth) in the Peruvian Amazon implication for genetic resource management. Theor Appl Genet 108:1564–1573

Brantestam AK, Von Bothmer R, Rashal I, Weibull J (2004) Changes in the genetic diversity of barley of Nordic and Baltic origin, studies by isozyme electrophoresis. Plant Genet Resour 1(2-3):143–149

Brush S (1995) In situ conservation of landraces in centers of crop diversity. Crop Sci 35:346–354

Brush S (1999) Genetic erosion of crop populations in centers of diversity: a revision. In: Serwinski J, Faberová I (eds) Proceedings of the technical meeting on the methodology of the FAO world information and early warning system on plant genetic resources, Prague, Czech Republic, 21–23 June 1999. Food and Agriculture Organization of the United Nations, Rome, Italy

Brush SB, Tadesse D, Van Dusen E (2003) Crop diversity in peasant and industrialized agriculture: Mexico and California. Soc Nat Resour 16:123–141

Coomes OT, Ban N (2004) Cultivated species diversity in home garden of an Amazonian peasant village in Northeastern Peru. Econ Bot 58(3):420–434

Coomes OT, Burt GJ (1997) Indigenous market-oriented agroforestry: dissecting local diversity in western Amazonia. Agrofor Syst 37:27–44

Elias M, McKey D, Panaud O, Anstett MC, Robert T (2001) Traditional management of cassava morphological and genetic diversity by the Makushi Amerindians (Guyana, South America): Perspectives for on-farm conservation of crop genetic resources. Euphytica 120:143–157

Frankel OH (1970) Genetic conservation in perspective. In: Frankel OH, Bennett E (eds) Genetic Resources in plants — their exploration and conservation. Oxford: International Biological Programme Handbook 11. Blackwell Scientific Publications, pp 469–489

FAO (1996) Global plan of action for the conservation and sustainable utilization of plant genetic resources for food and agriculture. FAO, Rome, Italy, 63 pp

FAO (1997) Report on the state of the world's plant genetic resources for food and agriculture. FAO, Rome, Italy, 510 pp

Fujisaka S, White D (1998) Pasture or permanent crops after slash-and-burn cultivation? Land-use choice in three Amazon colonies. Agroforest Syst 42:45–59

Fujisaka S, Escobar G, Veneklaas E (1998) Plant community diversity relative to human land uses in an Amazon forest colony. Biodivers Conserv 7:41–57

Fukuda WMG, Guevara CL (1998) Descritores morfologicos e agronômicos para a caracterização de mandioca (*Manihot esculenta* Crantz). Cruz das Almas: EMBRAPA-CNPMF, Brasil, 38 pp

Gonzales R, Terrazas F (2004) Ferias de semilla en el microcentro de diversidad de tubérculos andinos (TAs) Candelaria, Componente de la estrategia de monitoreo y apoyo a los flujos locales de mantenimiento de la diversidad de TAs. Proceedings of the XI International Congress on Andean Crops. Cochabamba, Bolivia, 3–6 February 2004. PROINPA, Cochabamba, Bolivia

Guarino L (1995) Assessing the threat of genetic erosion. In: Guarino L., Ramanatha Rao V, Reid R (eds), Collecting plant genetic diversity: Technical guidelines. CAB International, Wallingford, pp. 67–74

Goa LZ (2003) The conservation of Chinese rice biodiversity: genetic erosion, ethnobotany and prospects, Genet Resour Crop Evol 50:17–32

Hutchinson CF, Weiss E (1999) Remote sensing contribution to an early warning system for genetic erosion of agricultural crops. In: Serwinski J, Faberová I (eds) Proceedings of the technical meeting on the methodology of the FAO world information and early warning system on plant genetic resources, Prague, Czech Republic, 21–23 June 1999. Food and Agriculture Organization of the United Nations, Rome, Italy

Hair JF, Anderson RE, Tatham RL, Black WC (1998) Multivariate data analysis, 5th edn. Prentice-Hall Inc., Upper Saddle River, New Jersey, USA

Hijmans RJ, Guarino L, Bussink C, Barrantes I, Rojas E (2004) DIVA-GIS, Version 4. A geographic information system for the analysis of biodiversity data. Manual. (Available at http://www.diva-gis.org)

Hijmans RJ, Cameron SE, Parra JL, Jones PG and Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25, 1965–1978 (Data available at http://www.biogeo.berkeley.edu/)

INEI (1993) IX Censo Nacional de Poblacíon y IV de Vivienda, http://www.inei.gob.pe

IPGRI (2004) Diversity for Well-being. Making the most of agricultural biodiversity. International Plant Genetic Resources Institute. Rome, Italy. http://www.ipgri.cgiar.org/publications/pubfile+.asp?ID_PUB=996

Kaufman L, Rousseeuw PJ (1990) Finding groups in data: an introduction to cluster analysis. Wiley, New York

Kelley LA, Gardner SP, Sutcliffe MJ (1996) An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally-related subfamilies. Protein Eng 9:1063–1065

Maxted N, Hawkes JG, Guarino L, Sawkins M (1997) Towards the selection of taxa for plant genetic conservation. Genetic Resour Crop Evol 44:337–348

Millennium Ecosystem Assessment (2005) Ecosystems and Human Well-being: Biodiversity Synthesis. World Resources Institute, Washington, DC. 86 pp http://www.millenniumassessment.org/proxy/document.354.aspx

Mkumbira J, Chiwona-Karltun L, Lagercrantz U, Mahungu NM, Saka J, Mhone A, Bokanga M, Brimer L, Gullberg U, Rosling H (2003) Classification of cassava into ‘bitter’ and ‘cool’ in Malawi: From farmers’ perception to characterisation by molecular marker. Euphytica 132:7–22

Morin S, Calibo M, Garcia-Belen M, Pham J-L, Palis F (2002) Natural hazards and genetic diversity in rice. Agric Human Values 19:133–149

Oliveira de LO, Martins ER (2002) A quantitative assessment of genetic erosion in ipecac (*Psychotria ipecacuanha*). Genet Resour Crop Evol 49:607–617

R Development Core Team (2004) R: A language and environment for statistical computing, R Foundation for Statistical Computing Vienna, Austria http://www.R-project.org

Ramanatha Rao V, Hodgkin T (2002) Genetic diversity and conservation and utilization of plant genetic resources. Plant Cell Tiss Org Cult 68:1–19

Secretariat of the Convention on Biological Diversity (2001) Global biodiversity outlook. Secretariat of the Convention on Biological Diversity. Montreal, Canada, http://www.biodiv.org/gbo/gbo-pdf.asp

Secretariat of the Convention on Biological Diversity (2002) Strategic Plan for the Convention on Biological Diversity. Montreal, Canada, http://www.biodiv.org/sp/

Synnevåg G, Huvio T, Sidibé Y, Kanoutë A (1999) Farmer’s indicators for decline and loss of local varieties from traditional farming systems. A case study from Northern Mali. In: Serwinski J and Faberová I (eds) Proceedings of the technical meeting on the methodology of the FAO world information and early warning system on plant genetic resources, Prague, Czech Republic, 21–23 June 1999. Food and Agriculture Organization of the United Nations, Rome, Italy

Vadez V, Reyes-García V, Godoy RA, Apaza VL, Byron E, Huanca T, Leonard WR, Pérez E, Wilkie D (2004) Brief communication: does integration to the market threaten agricultural diversity? panel and cross-sectional data from a horticultural-foraging society in the Bolivian Amazon. Human Ecol 32(5):635–646

Vavilov NI (1951) The origin, variation, immunity and breeding of cultivated plants. Selected writings of N.I. Vavilov, translated from the Russian by K. Starr Chester. Chronica Bot 13:1–366

Wood D, Lenné JM (1997) The conservation of agrobiodiversity on-farm: questioning the emerging paradigm. Biodivers Conserv 6:109–129

WWF (2004) Living Planet Report 2004. World Wilde Fund for Nature. Gland, Switzerland. 40 pp http://www.panda.org/downloads/general/lpr2004.pdf