ORIGINAL PAPER

Estimating maize genetic erosion in modernized smallholder agriculture

Joost van Heerwaarden · J. Hellin ·
R. F. Visser · F. A. van Eeuwijk

Abstract Replacement of crop landraces by modern varieties is thought to cause diversity loss. We studied genetic erosion in maize within a model system; modernized smallholder agriculture in southern Mexico. The local seed supply was described through interviews and in situ seed collection. In spite of the dominance of commercial seed, the informal seed system was found to persist. True landraces were rare and most informal seed was derived from modern varieties (creolized). Seed lots were characterized for agronomical traits and molecular markers. We avoided the problem of non-consistent nomenclature by taking individual seed lots as the basis for diversity inference. We defined diversity as the weighted average distance between seed lots. Diversity was calculated for subsets of the seed supply to assess the impact of replacing traditional landraces with any of these subsets. Results were different for molecular markers, ear- and vegetative/ flowering traits. Nonetheless, creolized varieties showed low diversity for all traits. These varieties were distinct from traditional landraces and little differentiated from their ancestral stocks. Although adoption of creolized maize into the informal seed system has lowered diversity as compared to traditional landraces, genetic erosion was moderated by the distinct features offered by modern varieties.

Introduction

Since the advent of modern plant breeding, there has been concern that the substitution of improved seed for traditional crop varieties poses a threat to biological diversity (Harlan and Martini 1936; Harlan 1975). Such decrease in diversity is commonly referred to as genetic erosion (Frankel and Bennett 1970): "the loss of genetic diversity, in a particular location and over a particular period of time, including the loss of individual genes, and the loss of particular combinations of genes, such as those manifested in landraces or varieties" (FAO/IPGRI 2002). This broad description hides a complex phenomenon that is hard to measure in practice (Brush 1999). The traditional perception of genetic erosion has been that of the loss of a stable and diverse set of locally adapted landraces resulting from the adoption of a small number of modern varieties (Hawkes 1983; Brush 1999). Under this view, genetic erosion may be observed as the disappearance of named varieties in regions, where they used to be present (e.g. Hammer et al. 1996).

In recent years, a different picture of genetic erosion has emerged. Particularly for maize, it has been shown that landraces often persist after the introduction of improved seed (Bellon 1996; Perales et al. 2003a). The commercial, formal seed system, thus, coexists with the traditional, informal seed system based on seed recycling and exchange (Almekinders et al. 1994). Moreover, the varietal composition of the informal seed system has been found to

J. van Heerwaarden (✉) · F. A. van Eeuwijk
Biometris–Applied Statistics, Wageningen University,
P.O. Box 100, 6700 AC Wageningen, The Netherlands
e-mail: jvheerwaarden@gmail.com

J. Hellin
Impacts Targeting and Assessment Unit, International Maize
and Wheat Improvement Center (CIMMYT), Apdo Postal 6-641,
06600 Mexico D.F., Mexico

R. F. Visser
Laboratory of Plant Breeding, Wageningen University,
P.O. Box 386, 6700 AJ Wageningen, The Netherlands

be dynamic. Over time, varieties are lost and new ones are introduced from elsewhere (Louette et al. 1997). Commonly, improved varieties are incorporated into the informal system (Almekinders et al. 1994), a process that is known as creolization (Bellon and Risopoulos 2001). These creolized varieties are often given local names, becoming part of what farmers consider to be landrace seed.

The effect of modern varieties on local diversity is far from obvious. Diversity could even increase if improved germplasm is genetically more heterogeneous than local seed or if it offers traits that are not present in traditional landraces (Wood and Lenne 1997; Louette and Smale 2000). Also, creolized varieties may diverge from their parental stock by local gene flow and selection (Pressoir and Berthaud 2004b; Perales et al. 2005), leading to the generation of new diversity. The complex, dynamic composition of the traditional seed system, thus, poses challenges to the study of genetic erosion. The disappearance of named landrace varieties is not sufficient proof for diversity loss. It is necessary to quantify the diversity consequences of changes in the composition of the seed supply.

This raises the issue of defining and measuring diversity in the field. The following quantities are often suggested: (1) numbers of different entities or richness (2) evenness of distribution of these entities and (3) the extent of the difference between entities (FAO/IPGRI 2002). All three measures rely on the definition of discrete taxonomic units. This is relatively straightforward for modern varieties, as the formal system supplies certified seed of known identity. The nature of the informal seed system makes the designation of discrete entities rather difficult, however (Cromwell 1990; Almekinders et al. 1994; Louette et al. 1997). Local names do not necessarily reflect the genetic history of crops. Different names may be given to identical seeds while, conversely, a single name may apply to heterogeneous material (Jarvis et al. 2008).

Measuring diversity within the informal seed system, thus, requires the definition of a unit of seed identity that does not depend on nomenclature. One such unit is the seed lot. A seed lot can be defined as "the grain used by a single farmer to plant a single variety in a single season" (Louette and Smale 2000). It may be considered as the basic entity of seed management. Seed lots are frequently replaced by the farmers who grow them (e.g. Rice et al. 1998), so the value of individual lots to conservation is limited. However, seed lots that are genetically different are of potential value to farmers and represent the diversity that is accessible through seed exchange. In this study we, therefore, define diversity as the average distance, measured using molecular markers or phenotypic traits, between seed lots. We may divide seed lots collected from the field into several groups, such as informal-, formal-, creolized- and landrace seeds and calculate diversity within each of these groups. The effect of changes in seed stock composition can then be estimated by calculating how diversity changes when one group of seed is replaced by another.

This paper presents a case study on genetic erosion in maize agriculture in Mexico, a country that is considered the center of origin of maize. Production in most of the country is dominated by smallholder agriculture that relies mainly on traditional landraces. Our study area, La Frailesca in Southern Chiapas, has seen a strong increase in the use of formal seed (Bellon and Brush 1994). The informal seed sector does persist in this region but replacement of landraces by creolized maize varieties has occurred (Bellon and Risopoulos 2001). This provides an opportunity to determine if replacement of the remaining landraces by modern germplasm is leading to a decrease in diversity, particularly in the informal seed system.

Although work on genetic diversity in landraces and modern varieties exists (e.g. Reif et al. 2005; Huang et al. 2007), our study is unique in that it compares diversity in the field, including creolized varieties. We present an approach that combines biological data with information on seed history and local abundance to give diversity estimates for different seed groups within the seed system. Information obtained from local farmers and seed companies was used to estimate the current composition of the seed supply. Distance between seed lots was calculated based on SSR markers and ear-, vegetative- and flowering traits and converted to a measure of weighted average within-group diversity. We investigate whether replacement of traditional maize by modern seed has a negative impact on biological diversity. We address the hypothesis that commercial varieties are less diverse than traditional landraces currently present in the area under study. We thereby evaluate the consequences of increased adoption of improved maize varieties, both directly and through creolization, on local levels of biological diversity.

Materials and methods

General approach and sampling

Our aim was to perform a case study on genetic erosion in a region, where the formal and informal seed systems coexist. We chose smallholder maize agriculture in the La Frailesca region in southern Chiapas, Mexico as a model system. The region lies at an average altitude of 600 m and comprises several municipalities south of the state capital Tuxtla Guttierrez (Fig. 1). Maize production in La Frailesca is much more commercialized than in other parts of Chiapas. Landholdings are small, however, and farmers
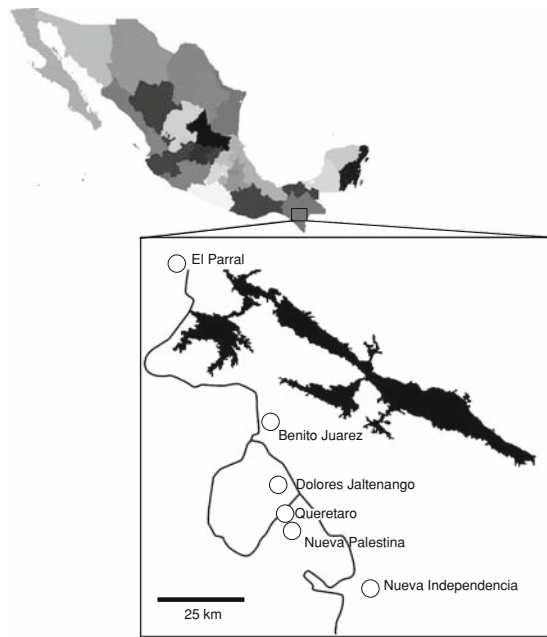
Fig. 1 Sampling area in La Frailesca Chiapas. Names of the sampled villages are marked in the magnified area

generally consume part of their own production for subsistence. The formal sector provides both commercial hybrids and open-pollinated varieties (OPVs). In contrast to hybrids, OPVs may be replanted without any yield penalty and are hence frequently introduced into the informal seed supply as creolized seed. The informal seed sector provides both traditional landraces and creolized varieties.

A 2005 survey on local planting materials (D. Flores, CIMMYT, unpublished), together with vendor information (data from 2003), was used to determine the relative abundance of the following seed types: landrace seed, creolized seed and individual commercial varieties. Sampling was stratified according to seed type to ensure sufficient sample size within each type but was otherwise random. Informal seed was collected from six villages in the region (Fig. 1). Formal seed was obtained from commercial seed vendors. In all, 30 seed lots were collected (Table 1). For commercial seed types, which are sold as certified seed, a single seed lot was collected per type/variety. Seed lots were considered creolized if they were bought as certified seed and subsequently replanted for at least 7 years. Traditional landraces were identified as such if they had a long history in the area and were not reported by any farmers as originating from certified seed.

Seed lots were characterized for phenotypic traits and SSR allele frequencies. Principal component- and cluster

analysis were performed on phenotypic traits to describe the differences between seed lots and to identify the traits contributing most to these differences.

Average between-seed lot distances were calculated within groups that represent alternative subdivisions of the seed supply: formal seed, subdivided into OPVs and hybrids; informal seed, subdivided into creolized varieties and landraces; landrace group, which was subdivided into named varieties, where appropriate. As sampling was stratified according to seed type, seed lots from different seed types were not always represented in accordance to their field frequencies within each group. Within-group distances between seed lots were therefore weighted by the probability of observing each distance in a random sample (details provided below).

We addressed the consequences of changes in the seed supply by comparing diversity estimates between the different groups of seed. First, we compared diversity within landraces sharing the same name to that found in creolized seed, OPVs and hybrids to measure diversity within traditional varieties as compared to the three main groups of non-landrace seed. Second, we compared diversity within the group of landraces to that observed in the creolized- and formal seed groups. This was done to determine if replacement of traditional landraces with either formal or creolized seed affects diversity. Finally, we compared informal- and formal seed diversity. This served the dual purpose of estimating the diversity effect of abandoning the informal seed system and determining if creolization has led to a reduction of diversity in the informal seed sector.

Field experiment

To evaluate phenotypic diversity, a field experiment was planted at CIMMYT's Tlaltizapan field station in May 2006. The experiment consisted of a split plot type design, with five replicate blocks containing all 30 seed lots. To minimize effects of competition between different kinds of varieties, seed lots were divided into four classes (landrace, creolized, hybrid and OPV) that were expected to be more or less homogeneous for vegetative and growth characteristics. These seed lot classes were randomly assigned to main plots within blocks and seed lots were subsequently randomized as sub-plots within seed lot class main plots. Each sub-plot consisted of 50 plants of a single seed lot, planted at 20-cm intervals in two 5 m rows. The first two plants in each row were discarded. The field was irrigated and fertilized throughout the experiment according to CIMMYT standard protocols. Vegetative traits were measured after flowering. Tassels were harvested and stored in a cold room before measurement. Measured traits are given in Table 2.

Table 1 Sampled seed lots

	Seed-lot	Local name	Farmer class	Race	Origin	W. I	W. II	W. III
Formal	OV1	V424	Improved	Tuxpeño	Buena Vista A.C.	**0.140**	**0.033**	0.033
	OV2	V424	Improved	Tuxpeño	CIMMYT	–	–	–
	OV3	V524	Improved	Tuxpeño	CIMMYT	–	–	–
	OV4	V526	Improved	Tuxpeño	PROASE	**0.400**	**0.095**	0.095
	OV5	V534	Improved	Tuxpeño	PROASE	**0.460**	**0.110**	0.110
	HB1	Nutria	Improved	Tuxpeño	ASGROW	0.267	**0.203**	0.203
	HB2	S-3G	Improved	Tuxpeño	Cristiani Burkard	0.004	**0.002**	0.002
	HB3	S-5G	Improved	Tuxpeño	Cristiani Burkard	0.069	**0.053**	0.053
	HB4	Z-30	Improved	Tuxpeño	Hartz	0.083	**0.063**	0.063
	HB5	Z-31	Improved	Tuxpeño	Hartz	0.021	**0.016**	0.016
	HB6	3086	Improved	Tuxpeño	Pioneer	0.140	**0.106**	0.106
	HB7	30F94	Improved	Tuxpeño	Pioneer	0.420	**0.318**	0.318
Informal	CC1	Conejo	Landrace	Zapalote chico	El Parral	–	1	**0.27**
	CO1	Olotillo	Landrace	Olotillo	Dolores Jaltenango	1		
	CO2	Olotillo	Landrace	Olotillo	El Parral			
	CO3	Olotillo	Landrace	Olotillo	Guadalupe Victoria			
	CO4	Olotillo	Landrace	Olotillo	Nueva Palestina			
	CO5	Olotillo	Landrace	Olotillo	Nueva Palestina			
	CT1	Jarocho	Landrace	Tuxpeño	El Parral	1		
	CT2	Jarocho	Landrace	Tuxpeño	N. Independencia			
	CT3	Jarocho	Landrace	Tuxpeño	Nueva Palestina			
	CT4	Jarocho	Landrace	Tuxpeño	Queretaro	–	–	–
	RV1	V424	Landrace	Tuxpeño	Benito Juarez	1	1	**0.73**
	RV2	Precoz	Landrace	Tuxpeño	Dolores Jaltenango			
	RV3	Tuxpeño precoz	Landrace	Tuxpeño	Dolores Jaltenango			
	RV4	San Gregoreo	Landrace	Tuxpeño	Dolores Jaltenango			
	RV5	Pronase	Landrace	Tuxpeño	N. Independencia			
	RV6	Pronase	Landrace	Tuxpeño	Queretaro			
	RV7	Tuxpeño	Landrace	Tuxpeño	Queretaro			
	RV8	Sardina	Landrace	Tuxpeño	Queretaro			

OV OPVs, *HB* hybrids, *CC, CO, CT* landraces, *RV* creolized varieties. W. I, W. II and W. III, are the relative within-group frequencies used to calculate weighted average between-seed lot distance for the comparisons I, II and III. For OPVs and hybrids, W. I. represents relative abundance calculated from sales data. Values in bold indicate the groups within each comparison

SSR genotyping

For each seed lot, 30–35 individuals were germinated under greenhouse conditions. Hybrid seed was considered to be of a single genotype and only a single seed was planted for each seed lot. DNA of a total of 838 plants was extracted using CIMMYT's standard CTAB protocol from ground, lyophilized leaf tissue. Ten easily scorable SSR loci were selected from previous studies (Matsuoka et al. 2002; Warburton et al. 2002) based on bin number and product size, to achieve the highest possible genome coverage while allowing for multiplexing of individual PCR products. Fluorescently labeled primers (Applied Biosystems, Sigma-Aldrich) were ordered for the following markers: phi034, phi093, phi061, phi014, umc1061, phi227562, phi96100, bnlg1784, phi029 and bnlg2047 (Maize GDB, http://www.agron.missouri.edu/ssr.html). PCR reactions were performed in a 10 µl reaction volume, containing 1–2 µl of 2 µM primer, 1.2 µl of 10 mM dNTP, 0.4 µl of 50 mM $MgCl_2$ and 1 µl of 10× PCR buffer. After PCR, 1.5 µl of pooled product was denatured in 9 µl of HiDi formamide containing 1 µl of ROX500 (Applied Biosystems) size standard. Samples were analyzed on an ABI 3100 capillary sequencer (Applied Biosystems). Fragment sizes were scored and converted to allele tables using Genotyper 2.1 (Perkin Elmer/Applied Biosystems) software.

Table 2 Traits measured in the field experiment and their units of measurement

	Trait	Unit	Code
Plant/tassel	Days to anthesis	Days	DA
	Days to silking	Days	DS
	Anthesis silking interval[a]	DS-DA	ASI
	Plant height	5 cm	PH
	Ear height	5 cm	EH
	Leaf number above the ear	#	LN
	Width of ear leaf	mm	LW
	Length of ear leaf	mm	LL
	Stem diameter	mm	SD
	Tassel branch number	#	TN
	Tassel length	#	TL
	Primary tassel branch length	cm	PL
	Secondary tassel branch length	cm	SL
	Tassel fresh weight	0.1 g	TW
Ear	Ear diameter	mm	ED
	Kernel row number	#	KN
	Kernel thickness	mm	KT
	Ear weight	g	EW
	Kernel weight	0.01 g	KW
	Cob length	mm	CL
	Cob diameter	mm	CD
	Cob weight	0.01 g	CW
	Total grain weight[a]	EW-CW	TGW
	Kernel length[a]	ED-CD	EKL
	Relative ear diameter[a]	ED/CL	RED
	Estimated kernel width[a]	$\pi(CD + ED)/2KN$	EKI

[a] Derived measures

Phenotypic description of collected seed lots

Phenotypic trait data were aggregated at the seed lot level by fitting a mixed model and calculating Best Linear Unbiased Estimates (BLUEs) using the LMER function in R (R Development Core Team 2005). Information for each trait was, thus, summarized in a vector of seed lot means. The mixed model for individual traits included a fixed seed lot effect besides random terms for block, row and column. Row and column random effects were included to correct for fertility trends in the field. The phenotypic data were first used for an exploratory analysis of seed lot differences and to identify discriminatory traits. To this end, data were scaled to correct for different measurement units. Scaling of seed lot BLUEs for each trait, was done by subtracting the trait mean, and dividing the result by the average standard error of the mean to give $x_{standardized} = \frac{x-\bar{x}}{SE}$. We applied this scaling by the mean experimental standard error (rather than the usual standard deviation of mean values), to emphasize traits with higher discriminatory value. We generated a biplot with trait loadings projected on the plane of the first two principal components of the scaled data (functions prcomp and biplot, R statistical software). Because of the scaled values, more discriminatory variables have a longer representation in the biplot. We tested for trait differences between specific seed lots using a mixed model equivalent of the two sample t test.

Distance definitions and cluster analyses

We used Gower's distance (Gower 1971) as a measure of pairwise phenotypic differentiation between seed lots. Phenotypic distances were calculated separately for vegetative/flowering- and ear/kernel traits.

Genotypic data were analyzed using the MSA program (Microsatellite Analyser, Dieringer and Schlotterer 2003). A modification of Reynolds' coancestry-based distance was used to measure molecular differentiation between seed lots (Reynolds et al. 1983). Coancestry, θ_{ij}, was calculated according to Weir and Cockerham (1984) based on the variance components of allelic frequencies due to populations, individuals within populations and alleles within individuals (Weir and Cockerham 1984). The latter component accounts for within-population variance due to non-random mating. This provides an advantage over θ_{ij} as proposed by Reynolds et al., especially since homozygote excess is common in maize (Pressoir and Berthaud 2004a; Reif et al. 2006). Unrooted neighbor-joining trees of pairwise differences (Saitou and Nei 1987) were constructed using the program SplitsTree (Huson and Bryant 2006).

Analysis of within-group diversity

Distances between individual seed lots were converted to weighted average within-group distance (See Cox et al. 1986, for a comparable approach using pedigree data to estimate wheat diversity in the United States). An estimate for average diversity within group g ($g = 1,...,G$), where group g contains n_g seed types (strata) is given by:

$$D_g = \sum_{i,j}^{n_g} p_i^g p_j^g d_{ij}^g$$

where for group g, d_{ij}^g is the average between-seed lot distance between seed types i and j, and d_{ii}^g and d_{jj}^g are the average between-seed lot distance within seed types i and j. p_i^g and p_j^g are the relative frequencies of seed types i and j within this group, with $\sum_{i,j}^{n_g} p_i^g p_j^g = 1$. We distinguished the following seed types: traditional landraces (1 type), creolized varieties (1 type) and individual commercial cultivars (10 types). Sampling was random within these seed types. Relative within-group frequencies for different seed types were obtained from survey and vendor interviews.

For each group, the distribution of D_g was evaluated by a bootstrap procedure. Data were re-sampled 1,000 times using a specially written R script (R Development Core Team 2005, code available on request). For phenotypic data, five experimental replicate blocks and 48 plants for each seed lot within a replicate were sampled with replacement. For the sake of computational efficiency, arithmetic seed lot mean instead of BLUEs were calculated for each bootstrap sample. For molecular data, re-sampling was done over all 10 loci. D_g was calculated for each iteration to generate an estimated sampling distribution. For formal seed, identical seed types (i.e. $i = j$) that are planted repeatedly, have a non-zero expected phenotypic difference because of experimental error. We therefore adjusted D_g for formal seed by setting d_{ij}, $i = j$ equal to the mean difference between bootstrap iterations for that seed lot. As θ takes account of sampling error (Weir and Cockerham 1984), no such correction was required for molecular distance.

Results

Description of the seed supply

The composition of the seed supply in La Frailesca (Fig. 2) can be summarized as follows: 48% of all planted seed lots

were commercial hybrids; 15% were open-pollinated varieties (OPVs); 27% were creolized seed; and 10% were traditional landraces. The formal sector (hybrids and OPVs), thus, provides 63% of all seed and the informal sector 37%. Within the informal sector, 73% of seed lots are creolized varieties while only 27% are traditional landraces. These numbers, together with the sales volumes of individual commercial varieties, were used to calculate group specific weights for each seed type. All 30 collected seed lots, together with within-group frequencies for each seed type, are listed in Table 1.

Within the formal seed system, we collected seven hybrid seed lots. Hybrids have been introduced only recently in La Frailesca, the first varieties dating back to the late 1990s. Open-pollinated varieties have a much longer history in the region. They mainly originate from public breeding programs and have been released in central Chiapas since the late 1970s by the National Agricultural Research Service (INIFAP). Five seed lots were obtained: V-524 (OV3), a variety created by CIMMYT and released by INIFAP in 1975. It was very popular until its removal from the market in 2001. Previous studies suggest that many creolized seed lots are derived from this variety (Bellon and Risopoulos 2001). V-424 (OV1-2), a variety selected for earliness by CIMMYT from the same population and released in 1981 by INIFAP. Both, sample of commercially sold seed (OV1) as well as seed from

Fig. 2 Schematic representation of the composition of the seed supply in La Frailesca
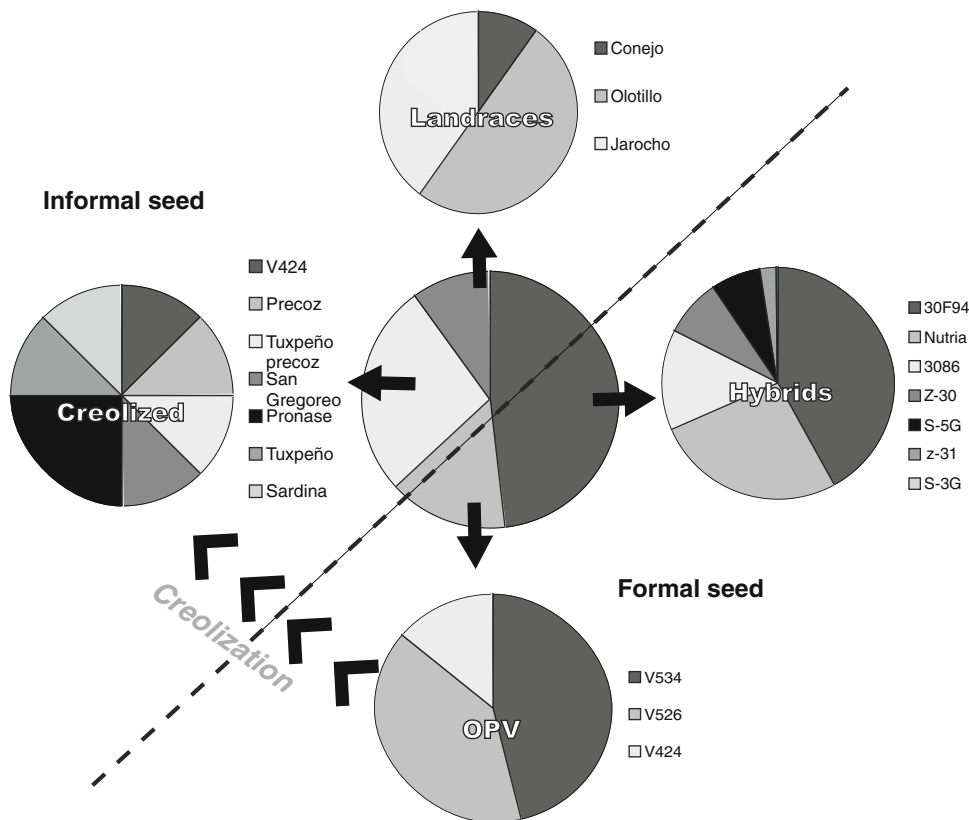
CIMMYT's gene bank (OV2) were included. V-526 (OV4) and V-534 (OV5), these varieties were released by INIFAP in 1982 and 1989, respectively.

Within the informal seed system, three named landrace varieties were collected. The two main varieties Olotillo (CO) and Jarocho (CT) belong to two distinct races, Olotillo and Tuxpeño, respectively (Wellhausen et al. 1952). The name Jarocho suggests that it is an introduced variety, since the same word is commonly used to indicate the inhabitants of the neighboring state of Veracruz. It has a long history in the area, however, and there was no evidence of it being a creolized variety.

The name Olotillo, meaning thin cob, refers to the most obvious trait that distinguishes this race from most other races that have been described (Wellhausen et al. 1952). A single seed lot called Conejo (CC) was collected: based on its ear and plant traits, it probably belongs to the zapalote chico race.

Creolized varieties had very diverse nomenclature, reflecting their introduction history or population of origin. The varieties named V424 (RV1), Precoz (RV2) and Tuxpeño precoz (RV3), were most probably derived from V-424 which was sold under the popular name Tuxpeño Precoz. Seed lots called Tuxpeño (RV7) are likely to originate from V-524 (OV3) and to a lesser extent from V-526 (OV4) both of which went under this name. Pronase (RV5) refers to the former state-owned seed company PRONASE. This company sold different OPVs, so the name sheds little light on the seed's identity. Similarly, the name San Gregoreo (RV4) reflects the label on seed bags containing unknown varieties that were distributed by the government around 1989. The variety Sardina (RV8), allegedly owes its name to the man who introduced it in the early 1980s and promoted it because its ears produced grain like "sardines in a can". This variety was said to derive from V-424 (OV1-2), but this could not be verified. We found no evidence of creolized seed lots that were derived from hybrid maize. Most seed lots were planted in quantities of around 30 kg (equivalent to about 1.5 ha) regardless of whether they were formal or informal varieties. Olotillo (CO) and Conejo (CC) were planted in much smaller volumes (1–20 kg, 6 kg on average) and were apparently used for special purposes.

Phenotypic description of collected seed lots

Figure 3 presents the biplot based on the principal component analysis of the scaled trait values for each seed lot. Traits with the highest loadings on the first and second principal components are shown as gray arrows. Five clusters of seed lots could be distinguished: Three landrace clusters (CC, CO and CT), a cluster of improved varieties including all hybrids, creolized varieties and most OPVs

(M) and an outlier pair formed by the two seed lots of V-424 (OV1-2). Conejo (CC) and V-424 are both early maturing varieties with a short plant height and small, sturdy ears. Conejo can be distinguished from V424 by its somewhat taller plants, narrower leaves and stems, shorter ears with fewer kernel rows and lower total grain weight. The two main landraces, Olotillo (CO) and Jarocho (CT), were separated from the other seed lots by being tall and late, presenting longer tassel branches as well as a lower number of kernel rows with slightly wider kernels and longer slimmer ears. Compared to hybrids and OPVs these two landraces had relatively narrow ear-leaves and lower ear and grain weight. CO and CT differed mainly in cob diameter and weight and relative ear diameter. Total grain weight differed quite substantially between the different seed lots (Fig. 4). Hybrids had the highest total grain weight as expected. With the exception of the early maturing V-424 (OV1-2) and the creolized variety RV6, all improved and creolized material had a higher individual grain yield than the traditional landrace seed lots. Their grain weight was generally equivalent to that of the classic V-524 variety (OV3). There was differentiation within the CT, CO and M clusters as well as within the set of creolized varieties for most traits. In the biplot, the creolized varieties RV3 and RV6 fell in between OV1-2 and the taller, later flowering OPVs (Fig. 3). When tested against the most probable ancestors, OV3 and OV1-2, significant differences were found for several traits. Surprisingly, one seed lot of the Jarocho landrace (CT4) grouped together with the improved and creolized varieties. It is relatively early flowering and short and has 14 kernel rows instead of the typical 10–12. This seed lot is likely to be a creolized variety in spite of its name.

Cluster analysis

The dendrogram based on pairwise distances for ear traits (Fig. 5a) confirmed the clear separation between landraces and improved and creolized varieties that was observed in the biplot. Hybrids (HB), formed a single cluster that included OV4, probably due to higher grain yield. The two V-424 seed lots (OV1-2) formed a separate cluster reflecting shorter cobs and lower grain weight. CT4 grouped closely with the latter cluster, just as RV3, RV6 and RV8. This again suggests a close relation of CT4 to improved maize. Among the landraces, all Olotillos (CO) grouped together, in accordance to their racial classification. Distances within this cluster were quite large, however, reflecting considerable variability in ear traits. Similar heterogeneity was present among the Jarocho (CT) seed lots.

Vegetative and flowering distances also produced separate clusters for improved (HB, OV and RV) and

Fig. 3 Biplot for seed lots and plant traits (first two axes). Length of gray arrows are proportional to the discriminative value of the different traits (presented in *black*, abbreviations given in Table 2). *Ellipses* enclose the different types of germplasm. M: modern and creolized varieties, OV1: V424, CT: Jarocho landrace, CO: Olotillo landrace, CC: Conejo landrace
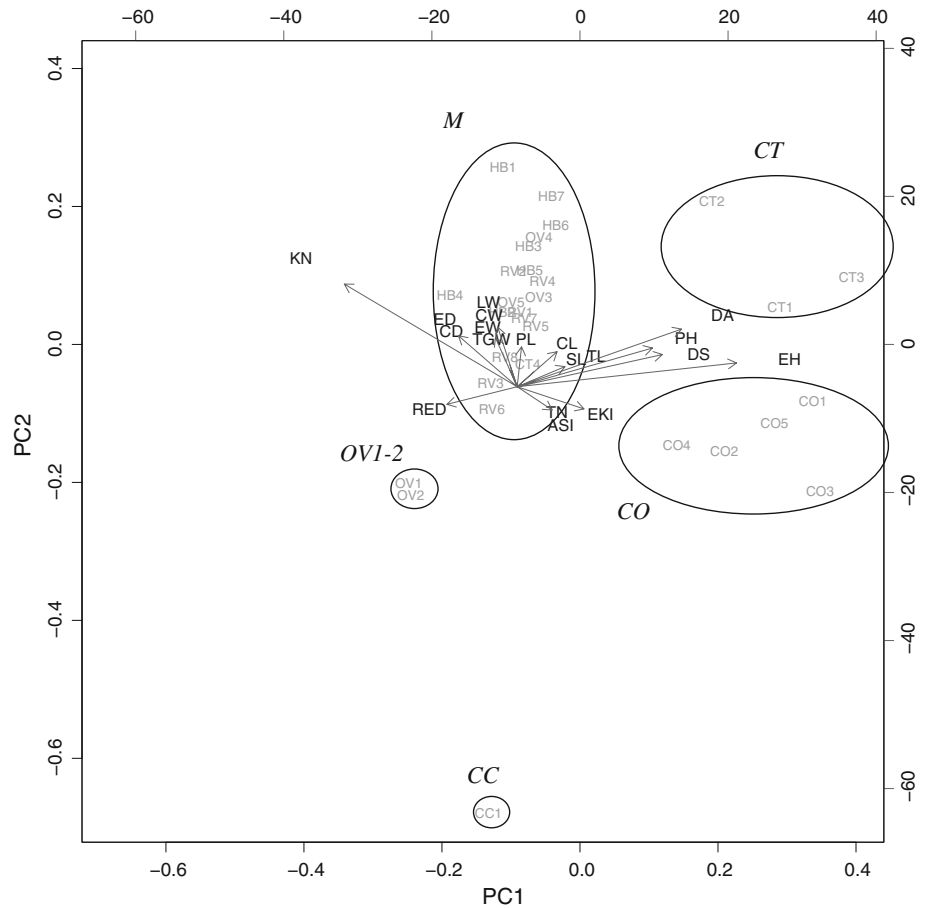

Fig. 4 Total grain weight (g) for all seed lots. Landraces are given in *white*, Creolized varieties in *light gray*, OPVs in *dark gray* and hybrids in *black*. CO: Olotillo, CT: Jarocho, RV: Creolized, OV: OPV, HB: Hybrid
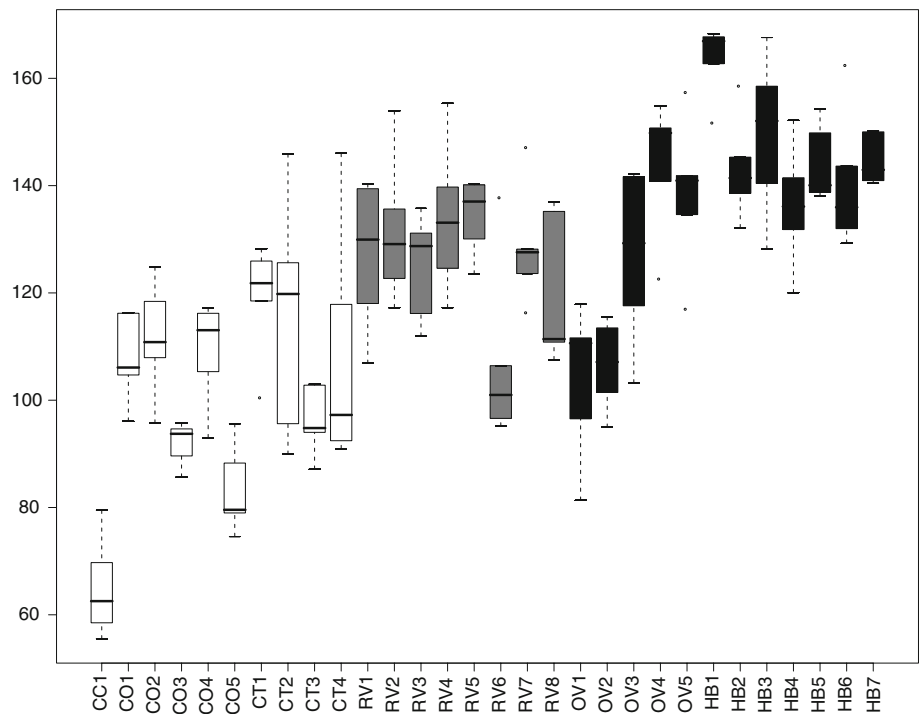
Fig. 5 a Dendrogram based on ear traits (Gower distance). **b** Dendrogram based on vegetative and flowering traits (Gower distance). **c** Dendrogram based on molecular distance (Reynolds), hybrids excluded. CO: Olotillo, CT: Jarocho, RV: Creolized, OV: OPV, HB: Hybrid
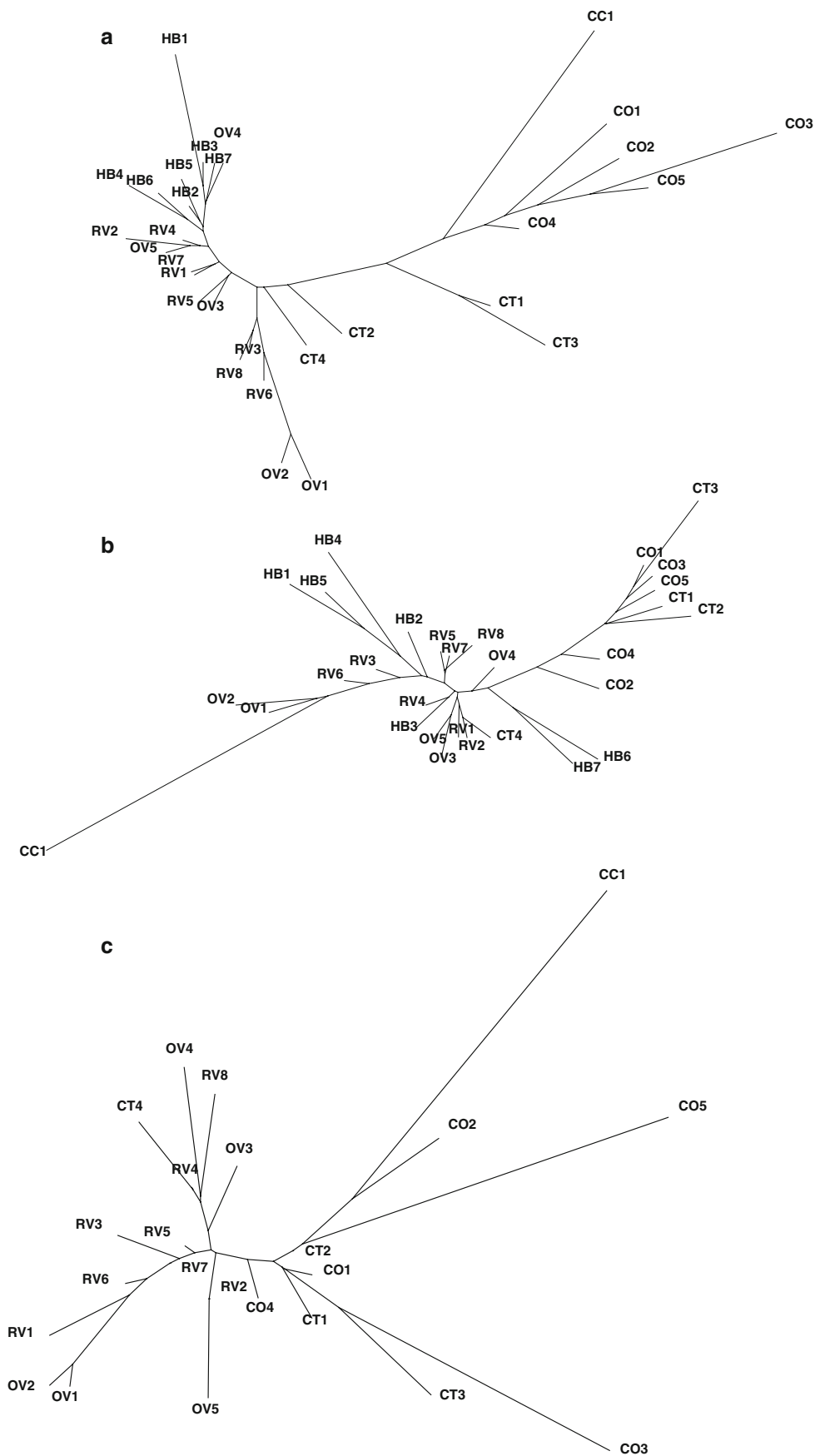
traditional varieties (CO and CT), with the exception of Conejo (CC) (Fig. 5b). Contrary to the results based on ear traits, Olotillo (CO) and Jarocho (CT) did not form separate clusters, suggesting that there is little difference in vegetative and flowering traits between the two races. Branch-lengths within the CT/CO cluster were short, revealing that these landraces are relatively homogeneous for these traits. In contrast, hybrids showed high diversity, with long branches separating the different varieties. Most of the OV and RV populations formed a poorly differentiated cluster that again included CT4. Only V-424 (OV1-2) was clearly different and clustered together with Conejo (CC) due to its earliness and short plant stature.

Clustering based on molecular distances showed extremely long branch-lengths between the different hybrids compared to those observed between landrace, creolized and OPV seed lots (not shown). This was not unexpected, given the fact that hybrids are produced by crossing two inbred lines. Each inbred line is fixed for a single allele at each locus, so the variance in allele frequency between different hybrids is expected to be large. We therefore excluded hybrids for a better appreciation of the relationships between the remaining seed lots (Fig. 5c). Modern varieties (OV and RV) were again clustered separately from traditional landraces (CO, CT and CC). The identity of CT4 as a creolized variety was confirmed by it falling within the cluster of creolized seed lots and OPVs.

Genetic differences between clusters were small. Pairwise θ between the OV and the CO seed lots was only 0.027 for example. The two main landraces Olotillo (CO) and Jarocho (CT) did not form separate clusters. Pairwise θ between the two races was only 0.015. This in contrast to average θ between seed lots within races which was 0.050 and 0.092 for Jarocho and Olotillo, respectively. The single seed lot of Conejo (CC) had a pairwise θ of 0.090 with respect to the other landraces. This relatively high differentiation of the CO and CC seed lots was probably caused by stronger drift due to small population sizes.

Analysis of within-group diversity

Average weighted distance between seed lots was calculated for different groups, representing subsets of the seed supply. The following comparisons were made: (I) The two main traditional landraces Jarocho (CT) and Olotillo (CO), creolized seed lots (RV), open-pollinated varieties (OV) and hybrids (HB). (II) Landraces (C), creolized varieties (RV) and seed lots from the formal system (F). (III) Formal seed (F), versus informal seed (I), containing all landraces and creolized varieties.

Comparison I showed that ear traits were more diverse within the Olotillo (CO) and Jarocho (CT) landraces compared to the other groups (Fig. 6); Hybrids (HB),

OPVs (OV) and creolized (RV) seed lots were both more homogeneous than the two landraces ($P < 0.025$, estimated from the bootstrap distributions), in spite of the fact that they contained distinct varieties. Results were different for vegetative/flowering traits and molecular markers. As was suggested by the cluster analysis, hybrids were the most diverse group for these traits (although the difference in vegetative/flowering traits was not significant when compared to the CT group). OPVs and creolized varieties contained less vegetative/flowering diversity than both the Olotillo, Jarocho and Hybrid groups ($P < 0.025$). Molecular distance was similar within the OPV-, creolized- and Jarocho groups. Olotillo showed higher molecular differentiation than the other non-hybrid groups ($P < 0.01$). This probably reflects drift due to the smaller field sizes of Olotillo compared to CT, RV, OV and HB seed lots.

The traditional landraces as a whole were phenotypically more diverse than both creolized- and formal sector seed (comparison II, Fig. 6). The difference between landraces and formal seed being less pronounced for vegetative/flowering traits. The high diversity of the landrace group was in part due to the inclusion of the Conejo (CC) seed lot. Molecular differentiation was highest for formal seed, owing to the high differentiation of Hybrids. Creolized seed was the least diverse group both for phenotypic traits and molecular markers.

Formal seed was significantly less diverse than informal seed for phenotypic traits but not for molecular differentiation (comparison III, Fig. 6). Contrasting the results of comparisons II and III showed that the inclusion of creolized seed into the informal seed sector decreases diversity for all traits as compared to that found within the traditional landraces proper. The reduction, although significant for all traits, was minor relative to the diversity loss that would result from a complete replacement of landraces with creolized varieties (comparison II, Fig. 6). Moreover, the level of diversity observed for informal seed was higher than expected based on within-type distances only. For ear traits for example, we expected an average weighted distance of 0.17 instead of the observed value of 0.23. This higher diversity resulted from the phenotypic distance between creolized and landrace seed lots, which increased the distance between seed lots within the informal seed group.

Discussion

The aim of the present study was to estimate the diversity impacts of modern variety adoption, using maize in La Frailesca as a model system. The complexity of the traditional seed supply makes analysis of genetic erosion far from trivial. Changes in varietal composition, poorly
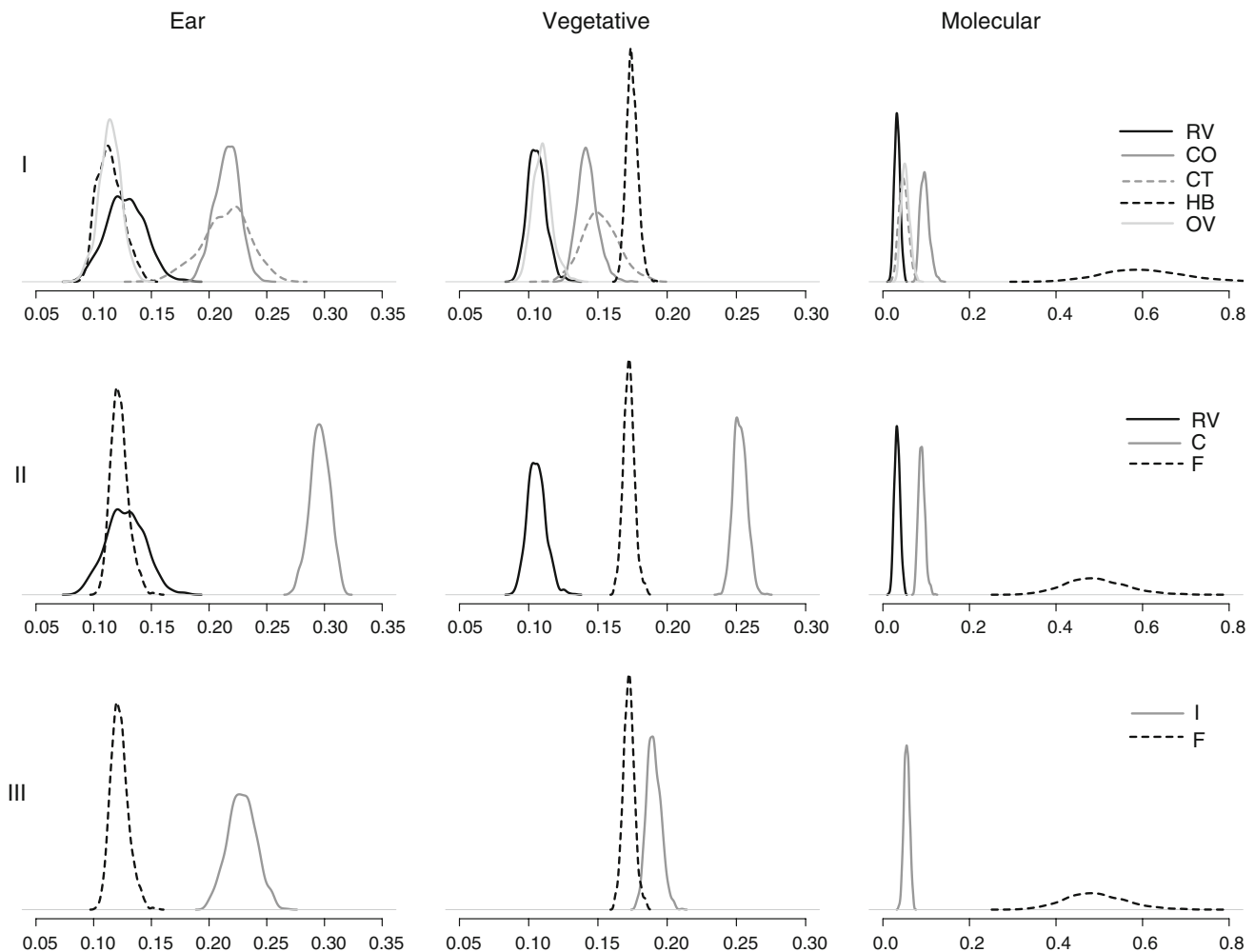
Fig. 6 Bootstrap distribution (1,000 bootstraps) of within-group differentiation between seed lots for three different comparisons (I–III). Weighted average distances are shown on the horizontal axis. Results are shown for ear traits, vegetative/flowering traits (Gower's distance) and molecular markers (Reynolds' distance). Diversity is shown on the horizontal axis. CO: Olotillo, CT: Jarocho, RV: Creolized, OV: OPV, HB: Hybrid, C: Landrace, F: Formal seed, I: Informal seed

defined local taxonomy and the process of creolization limit the use of counts based on nomenclature as a tool to study diversity loss. The alternative we have presented here is to measure biological variation within different groups of formal and informal seed and to test if changes in preponderance of these groups of germplasm affect diversity. Our study represents a new approach to the evaluation of genetic erosion in traditional agricultural systems. We combined different sources of data to generate a detailed description of the current composition of local seed stocks.

The main premise of our study was that individual seed lots represent the basic unit of seed movement and selection in traditional agriculture (Louette et al. 1997). We hence defined diversity as the weighted average distance between seed lots for either molecular marker allele frequencies or phenotypic traits. This definition provides a measure of biological diversity without making assumptions about the identity of named informal varieties. Our method allows for stratified sampling by weighting within-group distances using data on the abundance of specific seed types. This allows researchers to focus sampling on rarer types without biasing diversity estimates for specific seed groups. Comparison of different groups within the seed stock can be made quite easily by adjusting seed type weights for each group. These groups can be chosen to reflect different compositions of the seed supply so that we may evaluate if changes in composition affect diversity.

Commercial seed dominates maize agriculture in La Frailesca. This shows that the formal sector is providing seed that is suited to local production demands. Dominance of improved germplasm may affect diversity in two ways: by displacement of the informal seed supply or by modification of the composition of informal seed through

creolization. In the former case, farmers will rely solely on commercial seed. We have shown that formal seed, and particularly hybrids, may offer considerable diversity in terms of flowering/vegetative traits and molecular differentiation. Hence, adoption of commercial varieties in La Frailesca does not necessarily represent a significant loss of diversity in these traits.

From a conservation perspective, however, the composition of the formal seed supply is largely irrelevant in that it simply reflects what is offered by the market (Louette et al. 1997). This diversity is sensitive to changes in seed supply and cannot be regenerated by local evolution that acts on varieties within the informal seed sector.

In this respect, diversity within the informal seed sector is of greater interest. The informal sector is characterized by what has been described as "dynamic management" of genetic resources (Perales et al. 2003b). Diversity in the informal seed supply is not only determined by the kind of seed that farmers choose to plant but also by the extent to which seed lots become differentiated through genetic drift, gene flow and natural/human selection. We may argue that in such a system it is the balance between creation and loss of diversity that is of concern, rather than the maintenance of particular varieties. Under this view, the adoption of modern varieties into the traditional seed supply is not different from adopting a new landrace. The high diversity typically found in traditional landraces can be attributed to long divergence times separating seed lots. Creolized modern varieties may similarly add to local diversity if they represent a distinct genetic origin and history from local landraces. Displacement of traditional landraces with creolized seed of homogeneous origin and characteristics may obviously reduce local diversity. However, loss of diversity will only occur if evolutionary divergence of creolized seed lots is slow compared to their rate of introduction. Both the initial diversity represented by traditional and creolized varieties as well as estimates of their subsequent divergence are thus relevant to evaluating the diversity impact of creolization.

At present, the informal seed system persists in La Frailesca and mostly consists of creolized seed together with a small proportion of traditional landraces. Traditional landraces showed higher molecular and phenotypic differentiation than creolized varieties, particularly for ear traits. This diversity is partly explained by the presence of three separate races. Interestingly, however, the Olotillo and Jarocho seed lots were distinguishable only by ear characteristics. Although they are considered to belong to two different races (Wellhausen et al. 1952), they did not cluster separately for plant traits and molecular markers. This suggests that the two races represent the outcome of differential selection on ear shape by farmers (Louette and Smale 2000) rather than forming two separate genetic

entities. It is relevant to point out the low genetic differentiation between the two races, compared to the differentiation between individual seed lots. Genetic differences between seed lots as measured by molecular markers are likely to be strongly affected by drift due to limited population size. The relatively high differentiation shown by Olotillo and Conejo seed lots seem to confirm this notion. This may explain why a recent marker study, based on single seed lots, showed Olotillo and Tuxpeño to be clearly differentiated (Reif et al. 2006). Our results suggest that racial studies based on single accessions may not always be appropriate.

Creolized seed was diverse in terms of local nomenclature, with seven different names found among eight seed lots. These names reflect the introduction history of different seed lots and are not predictive of actual seed identity. Farmers generally consider these varieties as local and researchers should be aware of the risk of misclassifying creolized seed as local landraces. This point is made evident by our observation of the seed lot CT4 that revealed striking similarity in both morphology and marker frequencies to modern varieties, in spite of it being classified as a traditional landrace by farmers. The observation of creolized seed lots that were distinct from their most probable parental stocks confirmed that local seed management may indeed cause seed lot differentiation. Compared to landraces, creolized maize was relatively homogeneous for phenotypic traits, however, and most seed lots were very similar to their probable parental varieties. Apparently, local evolution has not yet generated much diversity. This result contradicts earlier suggestions that creolized seed lots represent new varieties combining traits from both modern and traditional materials (Wood and Lenne 1997; Bellon and Risopoulos 2001). Creolized seed lots were clearly distinct from the two main traditional landraces for plant height, leaf width, grain yield and row number. Total grain weight was consistently higher compared to traditional landraces. This may explain why the former dominate the commercialized agriculture that is currently being practiced. Given the relatively low differentiation between creolized seed lots, complete replacement of traditional landraces by creolized seed would constitute a loss of diversity.

However, creolized varieties offer traits that are distinct from those found in local landraces, thereby contributing to diversity in the informal seed supply. The coexistence of the two classes of seed at their present frequencies was hence found to cause only limited reduction of diversity.

In conclusion, this study confirms that testing the hypothesis of genetic erosion in smallholder agriculture is complex. Different classes of germplasm may be more or less diverse depending on the kind of traits considered. The question as to whether changes in seed stock composition

affect genetic diversity in La Frailesca can thus not be answered unequivocally. The fact that the formal system has become the primary source of seed does not in itself prove that genetic erosion has occurred. As we have seen, commercial seed may be diverse for certain traits. On the other hand, the persistence of the informal seed system does not mean that local genetic resources have been conserved, since the adoption of creolized seed can still reduce local diversity. The present work has allowed for actual levels of diversity loss to be assessed for different traits. Admittedly, we have analyzed a relatively arbitrary set of traits and markers. To address the issue of genetic erosion in a way that is meaningful to farmers, diversity needs to be described for traits and trait combinations that are considered valuable by the farmers themselves (Bellon 1996; Bellon et al. 2003). We hope, however, that the methodology presented here will contribute to a more quantitative approach to the problem of genetic erosion in dynamic seed systems that are typical of smallholder agriculture.

Conflict of interest The authors declare that they have no conflict of interest.

References

Almekinders C-J-M, Louwaars N-P, Bruijn G-H-D (1994) Local seed systems and their importance for an improved seed supply in developing countries. Euphytica 78:207–216

Bellon MR (1996) The dynamics of crop infraspecific diversity: a conceptual framework at the farmer level. Econ Bot 50:26–39

Bellon MR, Brush SB (1994) Keepers of maize in Chiapas, Mexico. Econ Bot 48:196–209

Bellon MR, Risopoulos J (2001) Small-scale farmers expand the benefits of improved maize germplasm: a case study from Chiapas, Mexico. World Dev 29:799–811

Bellon MR, Berthaud J et al (2003) Participatory landrace selection for on-farm conservation: an example from the Central Valleys of Oaxaca, Mexico. Genet Resour Crop Evol 50:401–416

Brush SB (1999) Genetic erosion of crop populations in centers of diversity: a revision. In: Proceedings of the technical meeting on the methodology of the FAO world information and early warning system on plant genetic resources held at the Research Institute of Crop Production, Prague, Czech Republic, 21–23 June 1999. http://apps3.fao.org/wiews/Prague/Paper5.jsp

Cromwell EA (1990) Seed diffusion mechanisms in small farmer communities. Lessons from Asia, Africa and Latin America. Network Paper Agricultural Administration Research and Extension Network (21) Overseas Development Administration, London, UK

Cox TS, Murphy JP, Rodgers DM (1986) Changes in genetic diversity in the red winter wheat regions of the United States. Proc Natl Acad Sci USA 83(15):5583–5586

Dieringer D, Schlotterer C (2003) Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Mol Ecol Notes 3:167–169

FAO/IPGRI (2002) Review and development of indicators for genetic diversity, genetic erosion and genetic vulnerability (GDEV): summary report of a joint FAO/IPGRI workshop (Rome, 11–14 September, 2002)

Frankel OH, Bennett E (1970) Genetic resources in plants—their exploration and conservation. Blackwell Scientific Publications, Oxford

Gower JC (1971) A general coefficient of similarity and some of its properties. Biometrics 27:857–874

Hammer K, Knupffer H, Xhuveli L, Perrino P (1996) Estimating genetic erosion in landraces—two case studies. Genet Resour Crop Evol 43:329–336

Harlan JR (1975) Our vanishing genetic resources. Science 188:618–621

Harlan H-R, Martini M-L (1936) Problems and results of barley breeding. USDA yearbook of agriculture: 303–346. US Government Printing Office, Washington DC

Hawkes J-G (1983) The diversity of crop plants. Harvard University Press, Cambridge, MA

Huang X, Wolf M et al (2007) Did modern plant breeding lead to genetic erosion in European winter wheat varieties? Crop Sci 47:343–350

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23:254–267

Jarvis DI, Brown AHD, Cuong PH, Collado-Panduro L, Latournerie-Moreno L, Gyawali S, Tanto T, Sawadogo M, Mar I, Sadiki M, Hue NT-N, Arias-Reyes L, Balma D, Bajracharya J, Castillo F, Rijal D, Belqadi L, Rana R, Saidi S, Ouedraogo J, Zangre R, Rhrib K, Chavez JL, Schoen D, Sthapit B, De Santis P, Fadda C, Hodgkin T (2008) A global perspective of the richness and evenness of traditional crop-variety diversity maintained by farming communities. Proc Natl Acad Sci USA 105(14):5326–5331

Louette D, Smale M (2000) Farmers' seed selection practices and traditional maize varieties in Cuzalapa, Mexico. Euphytica 113:25–41

Louette D, Charrier A, Berthaud J (1997) In situ conservation of maize in Mexico: genetic diversity and maize seed management in a traditional community. Econ Bot 51:20–38

Matsuoka Y, Vigouroux Y et al (2002) A single domestication for maize shown by multilocus microsatellite genotyping. Proc Natl Acad Sci USA 99:6080–6084

Perales H, Brush SB, Qualset CO (2003a) Landraces of maize in Central Mexico: an altitudinal transect. Econ Bot 57:7–20

Perales H, Brush SB, Qualset CO (2003b) Dynamic management of maize landraces in Central Mexico. Econ Bot 57:21–34

Perales HR, Benz BF, Brush SB (2005) Maize diversity and ethnolinguistic diversity in Chiapas, Mexico. Proc Natl Acad Sci USA 102:949–954

Pressoir G, Berthaud J (2004a) Patterns of population structure in maize landraces from the Central Valleys of Oaxaca in Mexico. Heredity 92:88–94

Pressoir G, Berthaud J (2004b) Population structure and strong divergent selection shape phenotypic diversification in maize landraces. Heredity 92:95–101

R Development Core Team (2005) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3, URL http://www.R-project.org

Reif J, Hamrit S et al (2005) Trends in genetic diversity among European maize cultivars and their parental components during the past 50 years. Theor Appl Genet 111:838–845

Reif J, Warburton M et al (2006) Grouping of accessions of Mexican races of maize revisited with SSR markers. Theor Appl Genet 113:177–185

Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the co ancestry coefficient basis for a short-term genetic distance. Genetics 105:767–779

Rice E, Smale M, Blanco J (1998) Farmers' use of improved seed selection practices in Mexican maize: evidence and issues from the Sierra de Santa Marta. World Dev 26:1625–1640

Saitou N, Nei M (1987) The neighbor-joining method a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Warburton ML, Xia XC et al (2002) Genetic characterization of CIMMYT inbred maize lines and open pollinated populations using large scale fingerprinting methods. Crop Sci 42:1832–1840

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the analysis of population structure. Evolution 38:1358–1370

Wellhausen E-J, Roberts L-M et al (1952) Races of maize in Mexico. Their origin, characteristics and distribution. Bussey Institution, Harvard University Press, Cambridge, MA

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