Contemporary evolution of maize landraces and their wild relatives influenced by gene flow with modern maize varieties

Idalia C. Rojas-Barrera[a,b,1], Ana Wegier[c], José de Jesús Sánchez González[d], Gregory L. Owens[e,f], Loren H. Rieseberg[e], and Daniel Piñero[b]

[a]Posgrado en Ciencias Bioquímicas, Universidad Nacional Autónoma de México (UNAM), C.P. 04510 Ciudad de México, México; [b]Departamento de Ecología Evolutiva, Instituto de Ecología, UNAM, 04510 Ciudad de México, México; [c]Laboratorio de Genética de la Conservación, Jardín Botánico, Instituto de Biología, UNAM, 04510 Ciudad de México, México; [d]Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara, Nextipac Zapopan, 44600 Jalisco, México; [e]Department of Botany and Biodiversity Research Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; and [f]Department of Integrative Biology, University of California, Berkeley, CA 94720

Edited by Edward S. Buckler, US Department of Agriculture-Agricultural Research Service/Cornell University, Ithaca, NY, and approved September 6, 2019 (received for review October 12, 2018)

Mexico is recognized as the center of origin and domestication of maize. Introduction of modern maize varieties (MVs) into Mexico raised concerns regarding the possible effects of gene flow from MVs into maize landraces (LRs) and their wild relatives (WRs), teosintes. However, after more than 60 y from the release of the first MVs, the impact of the sympatry with LRs and their WRs has not been explored with genetic data. In this work, we assessed changes in the genomes of 7 maize LRs and 2 WR subspecies from collections spanning over 70 y. We compared the genotypes obtained by genotyping by sequencing (GBS) for LRs and WRs before and after the adoption of MVs, and observed introgression from sympatric MVs into LRs and into the WR *Zea mays* ssp. *mexicana* sampled after the year 2000. We also found a decrease in the paired divergence index (F_{ST}) between MV-LR and MV-WR over the same time frame. Moreover, we determined that LR genetic diversity increased after 2000, probably as a result of gene flow from MVs introduced in the 1990s. Our findings allowed us to identify ongoing changes in the domesticated and wild maize genetic pools, and concur with previous works that have evaluated short-term gene flow from MVs into LRs in other crops. Our approach represents a useful tool for tracking evolutionary change in wild and domesticated genetic resources, as well as for developing strategies for their conservation.

modern varieties | gene flow | landraces | crop wild relatives | population genomics

The Mesoamerican centers of crop origins are characterized by crop landraces (LRs) living in sympatry with their wild relatives (WRs). Both are important genetic resources; thus, conservation of their unique genetics and minimizing their genetic erosion are priorities for the Convention on Biological Diversity and to ensure food security (1–3).

Maize is a genetically diverse crop (4–6), probably because of the broad range of environments in which it grows, as well as both ancient and contemporary introgression with sympatric WRs (5, 7–9). Contemporary gene flow with modern varieties (MVs) has been assessed for genetically modified (GM) cultivars only (10–12), without considering that in recent history, non-GM MVs have been cultivated in sympatry with both LRs and WRs. This raises the possibility that MV alleles may be spreading into LR or WR gene pools, potentially replacing indigenous alleles. Nevertheless, the impact of sympatry of Mexican LRs and WRs with MVs has not yet been explored using a wide range of temporal sampling to show the complete time line of effects.

There is evidence that adaptive introgression from the WR *Zea mays* ssp. *mexicana* (hereafter, *mexicana*) into maize LRs promoted the latter's adaptation to the highlands of Mexico (9). Likewise, adaptive introgression by artificial selection has been documented for maize LRs in Italy, where changes in population structure, as well as new alleles acquired from MVs, were identified in local LRs (13). The use of MVs also has provoked concerns about the displacement of LRs by MVs and the loss of alleles present in LRs adapted to local and heterogeneous environments (1, 14–16). This occurred in fewer than 5 y in the midwestern United States, where open-pollinated maize cultivars were replaced by maize MVs (17). In contrast, it has been reported for both potato and rice that after MV introductions, genetic diversity initially increases (18–20), although this increase is followed by a decline in diversity (21, 22) that stops after the partial substitution of LRs by MVs (15).

In Mexico, the first maize MVs were open-pollinated varieties under mass selection released in 1946. These were produced by the Mexican Department of Agriculture in agreement with The Rockefeller Foundation. A few years later, in 1950, the first hybrid maize lines were released, which represented crosses between open-pollinated varieties that had been selfed for 1 generation prior to hybrid production. At this stage, the efforts of improvement programs focused on the central Mexican Plateau and the central northwest and southwest of Mexico (23). The next

Significance

Crop diversity may be essential for adaptation to diverse future climates, and its conservation depends on human practices and preferences. Besides these, we show here that regulations promoting the adoption of modern cultivars can promote rapid changes in the genetic pools of indigenous landraces (LRs) and crop wild relatives (WRs). We compared a wide range of temporal samples of maize LRs and WRs. Modern varieties (MVs), LRs, and their WRs have been in sympatry for over 60 y. We provide genomic evidence of ongoing evolution of *Zea mays* L. due to introgression from MVs. These findings should foster monitoring strategies and policies that use and safeguard the genetic diversity of maize and its WRs at their center of origin.

Author contributions: I.C.R.-B., A.W., and D.P. designed research; I.C.R.-B. performed research; J.d.J.S.-G., G.L.O., L.H.R., and D.P. contributed new reagents/analytic tools; I.C.R.-B. analyzed data; and I.C.R.-B., L.H.R., and D.P. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Open Science Framework (OSF), https://osf.io/pqvt4/, and GitHub, https://github.com/yetzehev/Ongoing_Evol_Landraces.

[1]To whom correspondence may be addressed. Email: icrojasb@iecologia.unam.mx.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1817664116/-/DCSupplemental.

First published September 30, 2019.

decade, the 1960s, has been recognized as the beginning of the green revolution that led to a worldwide production increase of different cereal grains. This revolution relied on multiple strategies, including mechanization, advances in plant genetics and breeding, synthesis and application of fertilizers and pesticides, and adoption of high-yielding modern crop varieties (24, 25). Mexico is historically linked to the green revolution because of the successful wheat breeding program headed by Norman Borlaug. In the 1960s, 2 important institutions were created: the National Research Institute in Agriculture (INIA, its acronym in Spanish) and the International Maize and Wheat Improvement Center (CIMMYT, its acronym in Spanish). However, the acceptance of maize MVs was limited and occurred gradually. By 1976, maize MVs covered less than 15% of land dedicated to maize production (26). In the southeast of Mexico, MVs were adopted and replanted as open-pollinated varieties; this process, known as creolization, led to a decrease in land dedicated to LRs (14, 16, 20, 22, 27, 28). During this period, the trade of maize MVs was controlled by the government, whereas it is currently controlled by private companies (29, 30).

The first Mexican Seed Act of 1961 established the authority of the National Seed Producer (PRONASE, its acronym in Spanish) over the trade of genetically improved seeds produced by INIA. However, the private sector found this to be restrictive and pressed for authorization to perform seed research. Permission was granted in 1976, and its first MVs were released in 1983 (30, 31). In 1991, the seed act was modified to allow unrestricted private sector participation in seed research, production, and trade, and the Mexican Seed Act of 1996 gave the private seed companies access to INIA MVs. The next decade was crucial; the seed act was modified again in 2007 and PRONASE disappeared, and by 2010, private companies had almost total control of the MV seed markets (30, 32).

Current estimates of the proportion of land used for maize production that is devoted to MVs range from 42.5 (29) to 58% (33). However, reports on the penetrance of MVs into the maize market (34) imply that the latter proportion could be overestimated (35).

Maize MVs are mainly produced and consumed in the northwest and central west of Mexico, in contrast to the rest of the country, where LRs are predominant (29). The partial adoption of MVs can be explained by the added value of maize LRs, such as sensory attributes, higher quality for specialty foods (36–38), and adaptation to marginal environments and poor soil conditions (22, 37). Consequently, LRs are the base of traditional Mexican cuisine (39, 40), while MVs are used for tortillas, for the corn flour agroindustrial market, or to feed cattle (29, 40).

The partial adoption of MVs has led to frequent sympatry with local LRs and their WRs, which are grown in neighboring fields or in the same locality. Sympatry and creolization appear to have facilitated gene flow from MVs into LRs, which is promoted by farmers in some cases. This genetic exchange has only been documented through phenotypic evidence; so far, there is no genetic evidence of its extent or consequences (20, 39, 41).

Here, we tested for gene flow from maize MVs into local LRs, and their WRs, teosintes: *Zea mays* ssp. *parviglumis* (hereafter, *parviglumis*) and *mexicana*. These wild subspecies currently are sympatric with MVs mainly in the central Mexican Plateau and the Balsas River Basin, located in the central-southwestern part of Mexico, the region recognized as the center of domestication of maize (42–44). The teosintes *mexicana* and *parviglumis* exhibit different altitudinal distributions: *parviglumis* inhabits subhumid tropical environments between 143 and 1,960 meters above sea level (masl), whereas *mexicana* is found in subtropical to temperate conditions in highlands from 1,500 to 2,990 masl (45). We hypothesized that gene flow has continuously occurred from MVs into LRs and WRs in the Mexican Plateau and the Balsas River Basin regions.

To assess levels of introgression from maize MVs into native LRs and their WRs, we compared the genotypes of individuals currently in sympatry with MVs, including 7 LRs and the 2 teosinte subspecies mentioned above. The LRs included were classified according to Anderson and Cutler (46), who defined a LR as a group of related individuals with enough characteristics in common to permit their recognition as a group. The 7 LRs were sorted by sampling period (before 1960, from 1960 to 1980, and after 2000); *mexicana* and *parviglumis* individuals were sampled during the 1980s and after 2000. We also included MVs found in sympatry with LRs and WRs and, finally, allopatric populations of WRs to assess the role of sympatry. Genotypes obtained via genotyping by sequencing (GBS) (47) were compared according to sampling time period. Comparisons of LRs and WRs sampled at different time periods allowed us to confirm gene flow from maize MVs into LRs and into their WRs, detect rapid changes in the population structure of maize LRs, and document higher nucleotide diversity in the most recent LRs and WRs.

Materials and Methods

Samples and Accessions. We sampled LRs currently cultivated in sympatry with their WRs and commercial MVs, and then compared them with the oldest samples and accessions sorted into different periods in order to evaluate genetic exchange between these 3 groups.

To accomplish this aim, we examined the Native Maize Project record from the National Commission for the Knowledge and Use of Biodiversity. The purpose of this project was to update the distribution information for maize LRs and their WRs and to determine diversity centers in Mexico (48). We looked for LRs sampled in Chalco, the central Mexican Plateau, the Balsas River Basin, and the Oaxacan regions where WRs are distributed (49). We set 3 temporal periods: (1) The first was used as a negative control and included samples collected before 1960, prior to the adoption of MVs; (2) the second period was set from 1960 to 1980;and (3) the third period included samples collected after 2000. Since we were interested in potential LR-WR introgression, we selected both LR and WR samples from regions where their distributions overlap and are currently in close geographic proximity (*SI Appendix*, Fig. S1).

Seven maize LRs (Chalqueño, Zamorano, Mushito, Pepitilla, Tabloncillo, Bolita, and Vandeño) met our criteria for inclusion. We considered each sympatric group as a single analysis unit (i.e., 1 LR, 1 WR, 1 or more MVs). To analyze global patterns per sampling period, we grouped the genotypes as shown in Table 1.

Once we selected the groups of LRs and WRs, we visited the localities where these LRs were already reported as sympatric with *mexicana* or *parviglumis*. Four of the 7 sympatric LRs-WRs were growing in proximity to 1 or more commercial MVs. We included the 3 localities where we did not find MVs to test if sympatry at the time sampled was required to detect gene flow. All of the MVs included are the first generation (F1) of commercial varieties produced through controlled mating, which were collected between 2014 and 2015. As a control to test for gene flow from sympatric MVs into WRs, allopatric populations of *mexicana* and *parviglumis* were included.

For LRs and MVs sampled in 2014 to 2015, we collected at least 5 complete ears from randomly selected plants, with the exception of Vandeño LR (Huetamo, Michoacán), for which farmers could only provide us with seeds. The identity of LRs sampled in 2014 was established by Juan Manuel Hernández Casillas from the National Research Institute in Forestry, Agriculture and Livestock (INIFAP, its acronym in Spanish), and the same localities were revisited in 2015. For *mexicana* and *parviglumis*, we sampled seeds from at least 5 plants at each locality. Seeds from each locality were pooled, and 5 randomly selected seeds were germinated and genotyped. LR and WR accessions collected before or during the 1980s were obtained from CIMMYT and INIFAP. Samples from 2002 to 2004 were obtained from the University Center of Biological and Agricultural Sciences of the Universidad de Guadalajara. The passport information for all samples and accessions included in this work is available in *SI Appendix*, Table S1.

DNA Extraction and Genotyping. Maize and WR seeds were germinated in a CONVIRON environmental chamber at 25 °C with a period of 12-h/12-h darkness and light, respectively. We used fresh leaf tissue from each plant to extract DNA using the DNeasy Plant Mini Kit (QIAGEN, Inc.). The quality of the DNA was assessed on an 0.8% agarose gel, and the DNA samples were quantified by fluorometry using Qubit 3.0 and the Quantification Starter Kit (Thermo Fisher Scientific, Inc.).

Table 1. Time periods of sample collections: Maize LRs, MVs, and WRs

Population	Population code	*N*
LR sampled between 1940 and 1959	LR < 1960	49
LR samples between 1960 and 1980	LR 1960 to 1980	35
LR sampled between 2000 and 2015	LR > 2000	106
MV sampled between 2014 and 2015	MVs	47
Z. mays ssp. *mexicana* sampled in 1984	Zmx1980	12
Z. mays ssp. *mexicana* sampled between 2000 and 2015	Zmx2000	48
Allopatric *Z. mays* ssp. *mexicana* sampled between 2002 and 2015	Allopatric *mexicana*	11
Z. mays spp. *parviglumis* sampled between 1978 and 1984	Zpr1980	11
Z. mays ssp. *parviglumis* sampled between 2000 and 2015	Zpr2000	44
Allopatric *Z. mays* ssp. *parviglumis* sampled in 2015	Allopatric *parviglumis*	10
Z. diploperennis	*Z. diploperennis*	12
Total		385

Zmx, *Z. mays* ssp. *mexicana*; Zpr, Z. *mays* ssp. *parviglumis*.

One DNA sample for each plant was submitted to the Biotechnology Resource Center of Cornell University for GBS on a 96-plex plate. Samples were digested with ApeKI enzyme, and libraries were developed according to the protocol standardized for maize by Elshire et al. (47). Four samples were submitted twice and used as controls to assess the reproducibility of sequencing.

Samples included 7 native maize LRs, collected from 1943 to 2015, and their WRs, collected from 1978 to 2015; all of the MVs were collected between 2014 and 2015. We included the species *Zea diploperennis* as an outgroup. Altogether, 385 samples were genotyped in this work.

Bioinformatics Workflow.

Variants discovery. Fastq files were demultiplexed with GBSx (50), and reads were trimmed with Trimmomatic 0.36 (51). Alignments were performed with Nextgenmap 0.5.3 (52) using the B73 genome (AGPv4) (https://www.maizegdb.org/assembly) as a reference, and all alignments were converted to binary files with Samtools 1.5 (53). Variants were discovered for each sample using the HaplotypeCaller, and genotypes were merged with GenotypeGVCFs; both tools are from the Genome Analysis Toolkit (GATK 3.8.0) (54). We filtered single-nucleotide polymorphisms (SNPs) with a minor allele frequency (MAF) of 0.01, maximum missing data of 20% per SNP, and minimum mean depth of 2X; this step was done with vcftools 0.1.15 (55), and scripts are available at Open Science Framework (OSF) (https://osf.io/pqvt4/) (56) and GitHub (https://github.com/yetzehev/Ongoing_Evol_Landraces). After filtering, we kept 316,294 SNPs distributed across the 10 chromosomes and contigs in the reference genome.

Ancestry analysis. Admixture 1.3.0 (57) was used to infer the ancestry assignment for WRs, LRs, and MVs. We ran models from 2 to 8 genetic groups (K) and selected the K with the lowest cross-validation (CV) error. The Admixture results were used as criteria to remove accessions TbTp2014 and TbTp2015, which were collected as *parviglumis* (*SI Appendix*, Table S1) but assigned to *mexicana* ancestry, as well as 2 individuals collected as Vandeño LR that clustered with MVs. We ran a second ancestry assignment for 463 genotypes, including the 2 clusters that grouped the MVs collected for this work and the tropical breeding pool genotypes from the 2013 dataset of Romay et al. (65). We ran models from 1 to 20 groups and again selected the K with the lowest CV error.

Gene flow test (ABBA-BABA). To test for gene flow from MVs into native LRs and WRs, we calculated Patterson's *D*-score, also known as the ABBA-BABA test. The test is based on a resolved phylogeny among 4 taxa ([(H1,H2), H3], H4) and determines if the proportion of derived states is influenced by gene flow.

To compute this test, we used the BAM files from each individual and ran the analysis with the multipop ABBA-BABA module from the ANGSD package (Analysis of Next Generation Sequencing Data; https://github.com/ANGSD/angsd) (58); the parameters can be accessed at https://github.com/yetzehev/Ongoing_Evol_Landraces. The statistical threshold employed was $P < 0.05$ with a Bonferroni correction.

In our model, we placed LR samples from different time periods in positions H1 and H2, and then asked if 1 time period shared more derived alleles with any of the MVs in position H3, with *Z. diploperennis* as the H4. If the tree has an excess of ABBA or BABA patterns, it indicates gene flow between H3 and H2 (ABBA) or between H3 and H1 (BABA) (59–61).

We also ran models for each analysis unit represented by LRs-WRs sympatric or allopatric with MVs. Then, we tested for gene flow with MVs that were sympatric at the sampling moment, and, again, we asked whether allele sharing between LRs and MVs differed between time periods. For those LRs that were in sympatry with only one of the 2 genetic clusters identified for MVs with the ancestry analysis, or for those that were allopatric for MVs, we used the accession ZmH12015 (*SI Appendix*, Table S1) to survey gene flow from MV cluster 1 (MV1) and the accession SJH12015 (*SI Appendix*, Table S1) to survey gene flow from MV cluster 2 (MV2).

Divergence* index (F_{ST}) *across the genome. To survey population divergence between the MV clusters, LRs and WRs sampled at different periods, we calculated the divergence index (F_{ST}) with custom scripts using biallelic SNPs (https://github.com/owensgl/reformat/blob/master/vcf2vertical_bi_basic.pl and https://github.com/owensgl/pop_gen/blob/master/SNPtable2Fst.pl). We then averaged F_{ST} values in 5-kb windows considering linkage disequilibrium decay (62) and performed a nonparametric Kruskal–Wallis test, followed by a pairwise Mann–Whitney *U* test, to survey for changes in F_{ST} over time. In addition, we calculated F_{ST} values in 20-Mb windows per chromosome to plot the distributions with ggplot (63).

To assess temporal population divergence at the level of each analysis unit (*Samples and Accessions*), we calculated F_{ST} between the MVs and each of the 7 LRs and their sympatric WRs (*SI Appendix*, Fig. S3). For the LRs Chalqueño, Pepitilla, and Vandeño, which were not found in sympatry with MVs, we calculated F_{ST} using the accession ZmH12015 for MV1 and the accession SJH12015 for MV2.

Clustering analysis. To assess the germplasm source of MVs sampled in sympatry with LR-WR populations, we performed SNP calling for both LRs and MVs, using Tassel 5 (64) and the SNP database Tags On Physical Map v2.7 (http://cbsusrv04.tc.cornell.edu/users/panzea/download.aspx?filegroupid=4), which uses B73 version 2 as reference genome (AGPv2). Then, the genotypes generated for this work were merged and compared with a comprehensive public dataset of 2,578 genotypes from the US national maize inbred seed bank (65) available for AGPv2. Data were filtered with a MAF of 0.01 and maximum missing data of 20% per SNP, for a final total of 13,953 SNPs.

A first principal component analysis (PCA) was performed for the LRs and the MVs with our set of 319,294 SNPs described in *Variants discovery*, using the R package SNPRelate v1.12.0 (66). Subsequently, to explore the genetic relationships between the MVs sampled for this work and other breeding programs, 3 additional PCAs were run for MVs and the genotypes from the US national maize inbred seed bank, using the set of 13,953 SNPs described in the previous paragraph.

In addition, we constructed 2 maximum-likelihood (ML) trees, based on the generalized time-reversible (GTR) model, for 1,002 samples and the set of 13,953 SNPs using FastTree software (67) and plotted with FigTree. The first tree (*SI Appendix*, Fig. S6) included MV1 and MV2, as well as US inbred seed bank genotypes having a breeding program declaration (65). The second tree was run for 463 genotypes; this dataset included MV1; MV2; and the tropical breeding pools from Mexico, Nigeria, and Cameroon (*SI Appendix*, Fig. S6). Both trees were rooted with *parviglumis*.

Nucleotide diversity. We calculated the nucleotide diversity (π) in 20-, 50-, and 100-kb windows (*SI Appendix*, Table S2); the Watterson estimator (θ_W); and Tajima's *D* (*SI Appendix*, Table S3) with ANGSD (*SI Appendix*, Table S3) (58) for the LRs and WRs sorted into different periods. The π and θ_W distributions were compared to assess changes in genetic diversity over time; statistical significance was evaluated with a nonparametric Kruskal–Wallis test, followed by pairwise Mann–Whitney *U* tests (*SI Appendix*, Fig. S6). Significance was established at $P < 0.05$ with Bonferroni correction. The frequency distribution was plotted with ggplot2 (63).

Results

Ancestry Analysis (Admixture). Our ancestry analysis for LRs, MVs, and WRs found that K = 7 exhibited the lowest CV error, so we used 7 clusters for further analyses (Fig. 1). The MVs were grouped into 2 clusters; MV1 (Fig. 1, *Center*, red) was more abundant and included 6 of the 9 MVs analyzed. MV1 shares ancestry with subtropical and lowland LRs and is present in all of the temporal periods. MV2 (Fig. 1, *Upper*, yellow) included 3 MVs sampled at 2 different localities. MV2 ancestry is missing in the first time period and is only detected at low levels (0.06) in samples of Bolita LR in the second time period. However, after 2000, MV2 ancestry is detected in a low proportion in 6 of the 7 LRs, except Chalqueño, and it became more abundant in lowland LRs. Furthermore, 2 individuals collected in 2015 as Vandeño LR have ancestry that is entirely MV2 (Fig. 1, *Upper*, yellow).

We also found that LR ancestry is significantly correlated with altitude (Fig. 1, *Lower Right*), and that all LRs, except Chalqueño, include some ancestry from an MV cluster. The ancestry group that included MV1 and middle and lowland LRs (Fig. 1, *Lower Right*, red) increases its abundance over time (*SI Appendix*, Fig. S2). Meanwhile, Chalqueño LR appears as a single homogeneous group, and its ancestry significantly decreases with altitude and time (Fig. 1 and *SI Appendix*, Fig. S2).

With respect to the WRs, we observed 2 genetic clusters for *mexicana* (Fig. 1, *Upper*, orange and brown) and 2 for *parviglumis* (Fig. 1, *Upper*, dark green and light green). As with the LRs, their ancestries are related to the altitude gradient; however, we did not find a difference in ancestry proportions of individuals collected prior to 1980 versus those sampled after 2000 (*SI Appendix*, Fig. S2). Surprisingly, we collected individuals assigned to *mexicana* ancestry in 2014 and 2015 from croplands near San Lorenzo, Jalisco. San Lorenzo is located in the western portion of central Mexico, a low-altitude region where *parviglumis* is typically found.

Gene Flow Tests (ABBA-BABA). The ancestry assignment of samples into different time periods suggests gene flow from MVs into LRs. Based on this, we ran several ABBA-BABA tests to evaluate whether the gene flow from MVs was associated with the change in population structure of LRs during the study period.

We tested 12 gene flow models to survey for gene flow from MV1 and MV2 into LR > 2000 and WR > 2000, and we set old LRs and old WRs as the sister groups (H1), respectively (Fig. 2). In effect, this test is asking if there are more MV alleles in more recent samples than in older collections. We found gene flow from MV1/MV2 into LR > 2000 ($P < 0.0001$). This result was independent of the H1 group used, which was LR < 1960 or LR 1960 to 1980. We also explored the particular cases of each analysis unit. The results revealed gene flow from both MV clusters into 6 of the 7 LRs collected after 2000. The exception was Chalqueño, the highland LR sampled from 2,200 to 2,700 masl (*SI Appendix*, Table S4).

We observed gene flow from MVs into sympatric *mexicana* in both time periods, when allopatric *mexicana* was used as a negative control. This signal is stronger in the later sympatric population (>2000), as seen in *D*-score values with the allopatric comparison and when explicitly testing earlier and later *mexicana* populations ($P < 0.05$). This was seen when either MV1 or MV2 was used as the H3 sample. To assess whether the gene flow signal from MVs to *mexicana* was associated with ancient and/or current exchange between *mexicana* and LRs, we set LR > 2000 in the H3 position as the donor for *mexicana* individuals collected in 1980 or after 2000. We observed gene flow from LR > 2000 into *mexicana* > 2000 (*SI Appendix*, Table S4). The *D*-score is higher when the donor group is LR > 2000 in comparison to MVs, which indicates that the gene flow signal for MVs-*mexicana* > 2000 is picking up gene flow between LRs-*mexicana*. Nevertheless, this signal is heterogeneous in time and space; when we assessed gene flow for each analysis unit, there was no gene flow from MVs into *mexicana* from the Central Plateau (ZmTm2015, *SI Appendix*, Table S4).

In addition, we detected gene flow from MV1 and MV2 into *parviglumis* sampled in 1980 ($P < 0.05$); this result indicates that gene flow was higher in the past, but it does not rule out the possibility of gene flow for *parviglumis* sampled after 2000. We did not run a gene flow model for allopatric *parviglumis* populations because the classification of the Amatlán population remains unclear (9, 68), and the ancestry assignment for *parviglumis* sampled at Malinalco showed admixture with *mexicana* (Fig. 1).

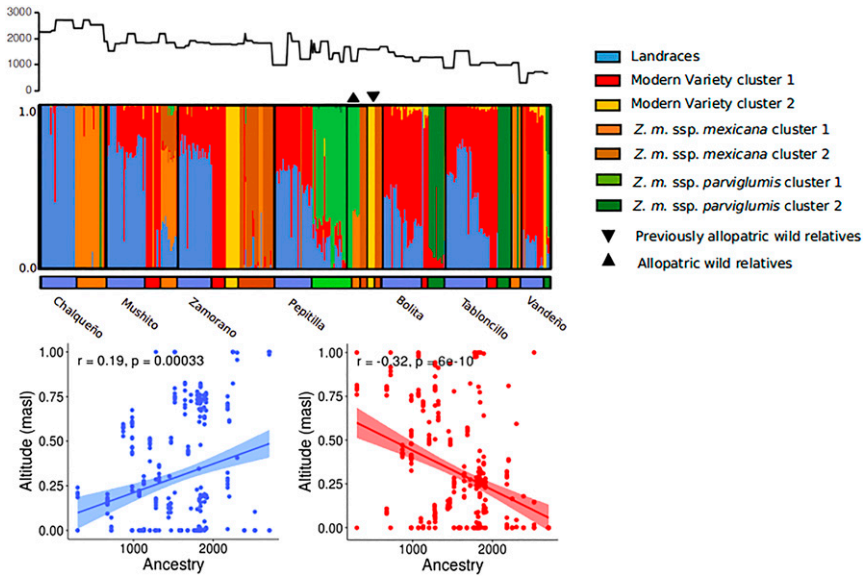

Fig. 1. Ancestry analysis. (*Upper*) Altitudinal distributions of the samples. (*Center*) From left to right, samples are ordered from the oldest to the most recently collected, and from higher to lower altitude sampling unit. Each unit is separated by a black line, and these were named after the LRs found in sympatry with *mexicana* or *parviglumis* and MVs, ancestry for K = 7 (lowest CV error). (*Lower*) Spearman rank correlation test between ancestry assignment and altitudinal distribution of samples.

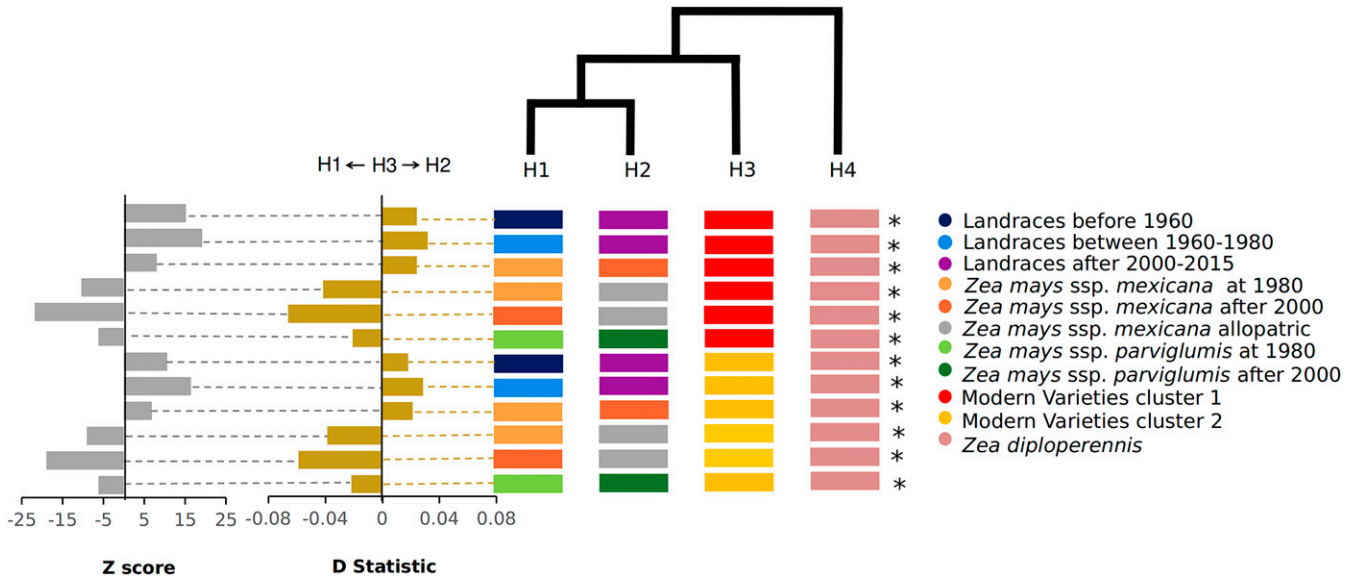

Fig. 2. Gene flow models (ABBA-BABA). (*Right*) Tree shows the phylogenetic relations assumed for the gene flow scenarios explored. Each model is represented by the color codes assigned to the different groups. *$P < 0.05$ (statistical significance). (*Left*) Plots show the Z-score and Patterson's *D*-statistic: When *D* is positive, there is gene flow between H3 and H2; when *D* is negative, there is gene flow between H3 and H1.

Differentiation Across the Genome (F_{ST}). We used the F_{ST} index to estimate the temporal genetic divergence between LR-MV and also between WR-MV. We observed that F_{ST} decreased significantly throughout the period studied, which means that allelic frequencies of LRs, *parviglumis* and *mexicana*, sampled after 2000 are more similar to MVs (Fig. 3).

In addition, F_{ST} was significantly lower between LR-MV1 than between LR-MV2 in the 3 periods evaluated (Fig. 3 *A* and *B*). The same trend was observed when we compared F_{ST} between the 2: WR-MV1 and WR-MV2. Again, both WRs, *mexicana* and *parviglumis*, are significantly more similar genetically to MV1 than to MV2 (Fig. 3 *C–F*). When we calculated the F_{ST} for each of the 7 analysis units (*SI Appendix*, Fig. S3), we observed that for 6 of the 7 LRs and their sympatric WRs, there were lower differentiation estimates for MV1 compared with those found for MV2. This trend was independent of sympatry.

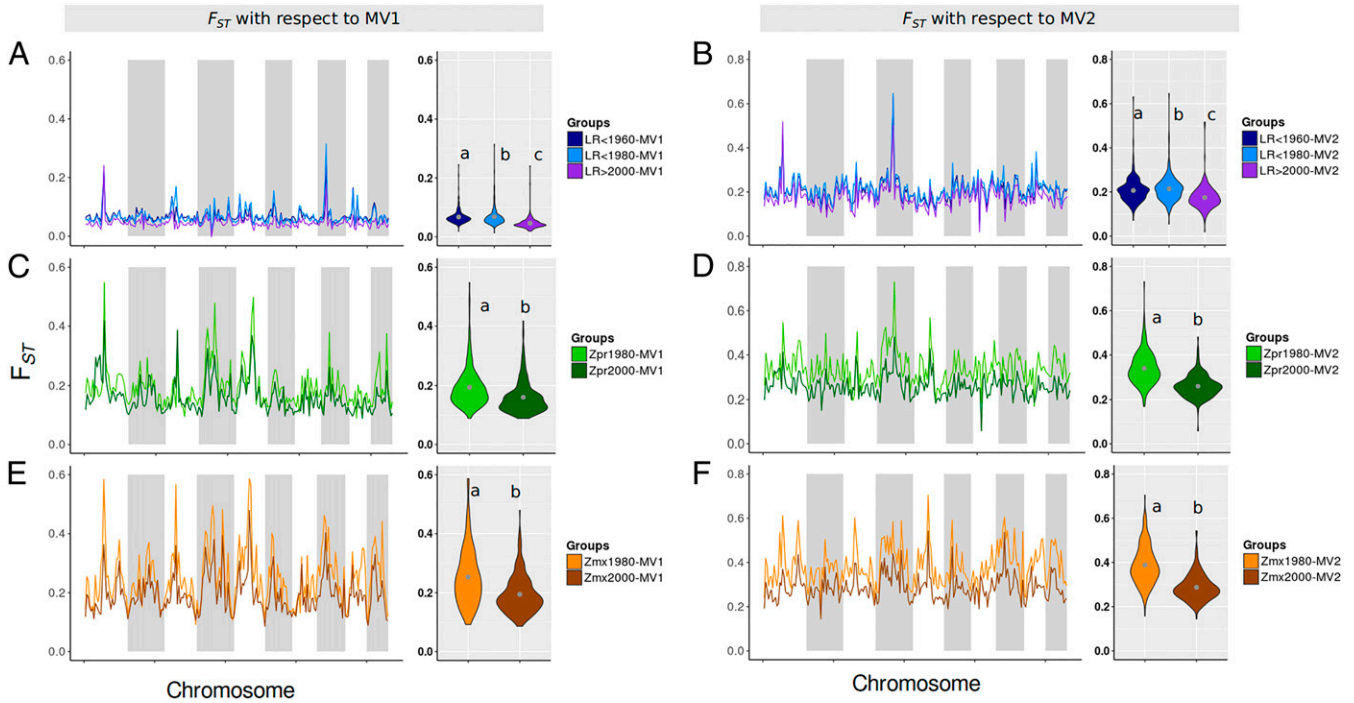

Fig. 3. Pairwise F_{ST} across the 10 chromosomes of maize and their WRs, *mexicana* and *parviglumis*, with respect to MVs. F_{ST} with respect to MV1 (*Left*) and MV2 (*Right*) is shown. The alternating shaded areas represent the 10 chromosomes of maize and their WRs; letters above the violin plots show statistical significance ($P < 0.01$). F_{ST} of LRs was sampled at different periods, with respect to MV1 (*A*) and MV2 (*B*). F_{ST} of *Z. mays* ssp. *parviglumis* (Zpr) was sampled at different periods with respect to MV1 (*C*) and MV2 (*D*). F_{ST} of *Z. mays* ssp. *mexicana* (Zmx) was sampled at different periods with respect to MV1 (*E*) and MV2 (*F*).

Genotypic Clustering. When we ran the PCA, we observed that the distance between MV1 and MV2 was larger than the distance between MV1 and the LRs (Fig. 4*A*). To better understand this unexpected distribution, we used a dataset from the US seed bank, which included CIMMYT maize line (CML) genotypes from the Mexican breeding program and genotypes belonging to other breeding programs around the world (65). Both MV clusters fell within the distribution of the tropical breeding programs from Mexico, Nigeria, and Cameroon (Fig. 4 *B* and *C*).

Ancestry analysis further showed that the tropical Mexican breeding program is the main germplasm source for MV1 and MV2 (*SI Appendix*, Fig. S4), although 2 MVs belonging to MV1 shared ancestry with genotypes from the Cameroon breeding program. These results agreed with the maximum likelihood (ML) tree for MV1, MV2, and the tropical lines, which showed that a group of CMLs is basal to both clusters (*SI Appendix*, Fig. S5).

Nucleotide Diversity. Finally, to test if changes in ancestry proportion and gene flow have had an impact on π, we compared LR < 1960, LR 1960 to 1980, and LR > 2000 (*SI Appendix*, Fig. S5). We found that π for LR > 2000 (2.03e-04) is significantly higher than π for the oldest samples, sorted into LR < 1960 (1.87e-04) and LR 1960 to 1980 (1.84e-04). Interestingly, π for the intermediate period, LR 1960 to 1980, was significantly lower than for the other 2 periods. In addition, π for MV2 was significantly lower than π for MV1 (*SI Appendix*, Table S2). Because π is influenced by intermediate frequency alleles, we decided to estimate θ_W to survey rare allele dynamics. In contrast to π, θ_W increased over the period evaluated (*SI Appendix*, Fig. S6). This trend matched Tajima's *D*, which is positive for LR < 1960 (0.105 ± 0.505), and negative for LR 1960 to 1980 (−0.032 ± 0.453) and LR > 2000 (−0.077 ± 0.546).

Discussion

Ancestry Changes with Altitude and Sampling Time. There was a significant correlation between the LR and WR ancestries with altitude (Fig. 1, *Lower Right*). This correlation concurs with previous works (4, 69), and it is consistent with the successful adoption of MVs in medium- and low-altitude regions (22, 36, 70). We observed 1 predominant ancestry group in middle and lowland LRs and

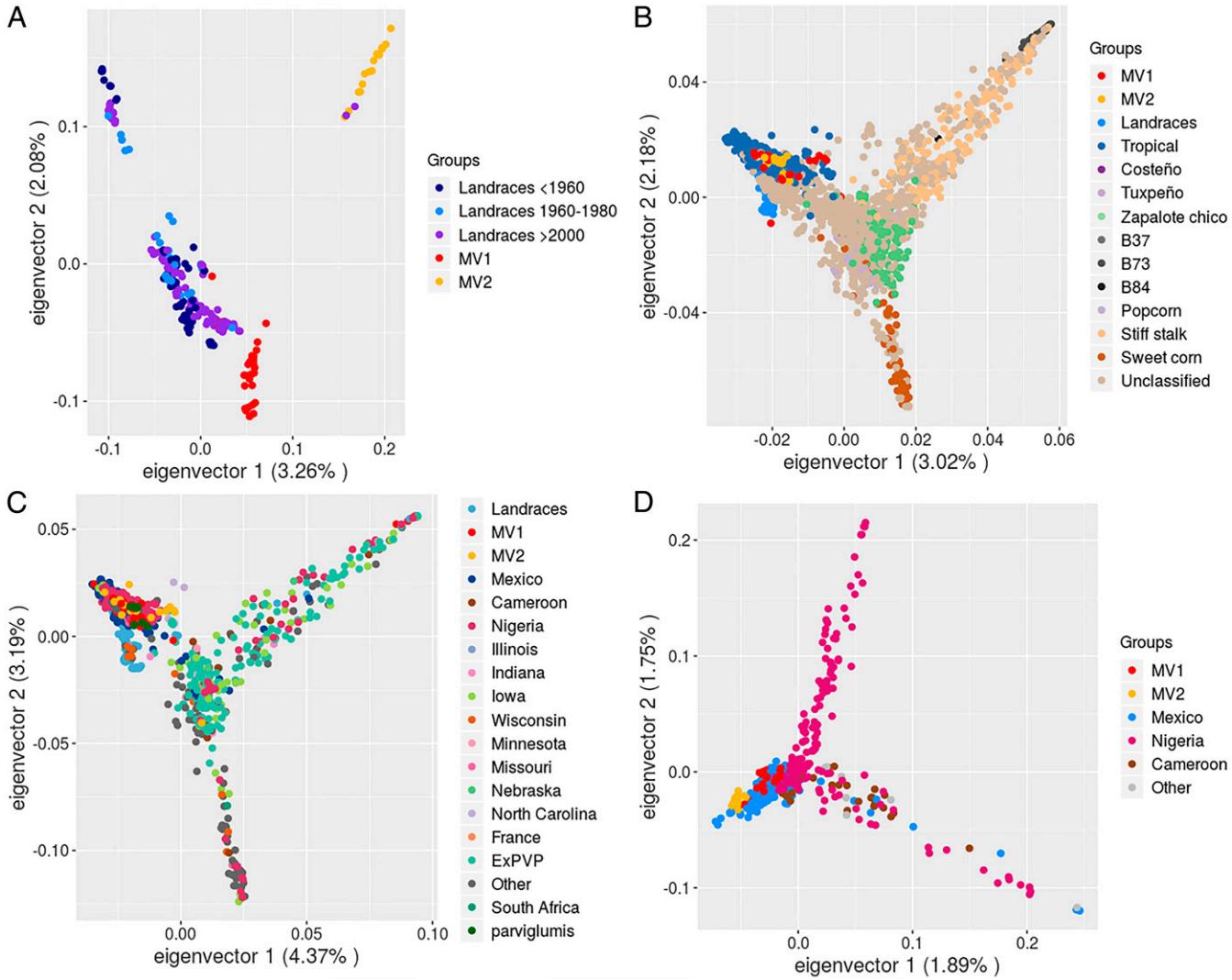

Fig. 4. MVs clustering with germplasm from different breeding programs. (*A*) Genotype distribution of MV1 and MV2 sampled in sympatry with LRs. LRs are colored by sampling period: earlier than 1960 (LR < 1960), dark blue; between 1960 and 1980 (LR 1960 to 1980); light blue; and later than 2000 (LR > 2000), purple. (*B*) Distribution for MV1, MV2, and the LRs collected for this work with a sample subset from the US national maize inbred seed bank (2,578 genotypes and 13,953 SNPs). (*C*) Clustering of genotypes colored by the breeding program (1,002 genotypes, 13,953 SNPs). ExPVP, expired plant variety protection. (*D*) Tropical breeding pools from Mexico, Nigeria, Cameroon, MV1, and MV2 (463 genotypes and 13,953 SNPs).

MV1. This ancestry increases its proportion in LRs over time (*SI Appendix*, Fig. S2), probably due to the continuous increase in the adoption of MVs. We also expected a broader geographic distribution for MV1 ancestry, because lowland alleles are more broadly adapted, as measured by grain yield, than highland alleles (71, 72).

The system to develop MVs changed drastically after 1990 when the Mexican Seed Act was modified (30, 32), which allowed the private sector to capture 95% of the seed market (30). The LR ancestry associated with MV2 was less prevalent than MV1 ancestry. According to the ancestry assignment with the US inbred seed bank genotypes (*SI Appendix*, Fig. S4), the MV2 genetic pool shares ancestry with the tropical breeding programs from Mexico. MV2 ancestry has increased over time, and if it is a recently introduced genetic pool, then it likely will follow the same trend as MV1. In fact, 2 individuals that were referred to as LRs by a farmer in 2015 clustered with MV2. It is likely that these individuals are the offspring of MVs that were replanted, which is a common practice known as creolization (39, 41).

Despite efforts to develop MVs genetic pools adapted to the highlands (23, 73, 74), their presence in the highlands is still scarce (34, 36). We did not find any MVs in sympatry with the Chalqueño highland LR. It has been reported that highland LRs lack tolerance to heat and inbreeding (70), features that are associated with strong local adaptation and probably have limited the introduction of MVs in the highlands. It can further explain why small farmers have adopted new technology, but they have not adopted MVs (36).

The pattern that we observed for WR ancestries (Fig. 1, *Upper*, orange and green clusters) matched the known altitudinal distribution of *mexicana* and *parviglumis* (45), except for individuals collected in San Lorenzo, Jalisco, located at 1,071 masl, a region where *parviglumis* is typically distributed. We collected teosintes at this locality in 2014 and 2015 in different croplands; these samples are clustered with subspecies *mexicana*. This unusual distribution can be attributed to the mobility of seeds fostered by human practices; it has been suggested that maize mimetic teosinte traveled from the Chalco region to Puebla and Toluca in trucks that transport manure (14). The nonintentional human transfer of teosintes also has been documented in Spain, where a nonclassified variety became a weed that damaged maize production in Western Europe and the north of Spain (75). The changes in the geographic distribution of teosintes require further evidence and a formal analysis of the conditions that promote their mobility.

Gene Flow from MVs into LRs and WRs. Gene flow from both MV clusters into LR > 2000 appears to be a consequence of the gradual adoption of MVs. Even though the first MVs were released in 1960, ancestry analysis does not show major changes during the 1960 to 1980 period (*SI Appendix*, Fig. S2), perhaps because the first MVs were based on high-yield LRs, such as Tuxpeño and Celaya; thus, gene flow from MV to LR would transfer alleles already present in LRs (22, 23). In 1970, the private sector began developing its own MVs, but these were not released until 1980. During this period, private companies did not have access to the seeds produced by INIA (30); therefore, MVs from this period are probably based on improved CIMMYT lines or foreign germplasm. Beyond the sociopolitical causes, the resowing of locally grown MVs led to deliberate and/or accidental crosses between LRs and MVs (22, 39, 41). Gene flow from MVs into LRs has been previously documented with phenotypic evidence (14, 39). However, this work provides genomic evidence of the occurrence and extent of gene flow from MVs into LRs.

Specifically, we detected gene flow from MVs into 6 LRs distributed at medium and low altitudes, but not for Chalqueño, a highland LR sampled from 2,200 to 2,700 masl. The lack of gene flow into the only highland LR included can be explained by the rarity of MVs in highland regions (34, 36). However, it is also possibly associated with the strength of local adaptation due to a higher number of maize environments per land unit in the highlands than in the midlands and lowlands (70). The tropical and subtropical MVs are adapted to more homogeneous environmental conditions, which has facilitated large-scale adoption of MVs. Under this hypothesis, MV introgressions into highland LRs are likely to be removed by selection against the generalist MV alleles versus the locally adapted LR alleles (72, 76).

However, the strength of local adaptation is a trade-off that may limit the capacity of highland LRs to respond to climate change and might make small farmers located in these regions more vulnerable (77). Conversely, the scarcity of MVs in the highlands also has been attributed to the failure of institutional breeding programs. In 2003, it was projected that in a fairly short time frame (10 to 20 y), highland LRs could be displaced by MVs (36). However, 12 y after the publication of this prediction, MVs are still rare in the highlands, and the highland genetic group had minimal changes, although we would expect that the highlands would follow the same trend for midland and lowland crop areas, if MVs are adopted at a higher frequency in the highlands.

Z. mays has a pollen/pistil incompatibility system that is heritable and nonreciprocal. Two main loci have been described: Gametophyte factor (Ga1) and Teosinte cross-barrier 1 (Tcb1). While we considered gametophytic incompatibility between the MVs and the Chalqueño highland LR as a possible explanation for the lack of gene flow, the codominant allele *Ga1-m* has been detected in Chalqueño and in a representative sample of MVs previously evaluated (78, 79). Thus, cross-fertilization incompatibility does not appear to account for the absence of gene flow from MVs into Chalqueño.

We selected LRs growing in sympatry with MVs when collected to increase the probability of detecting gene flow. However, our results indicate that sympatry at time of sampling was not needed to detect gene flow. For example, we observed gene flow from MVs into Vandeño and Pepitilla in the absence of MVs at sampled locations, which implies that gene flow occurred in previous generations or is a consequence of sympatry with creolized MVs, such as the 2 individuals that were collected as Vandeño but assigned genetically to MV2.

We found evidence of gene flow from MVs into *mexicana*. Our results are supported by experimental evidence: Crosses of teosintes from the Central Plateau and Chalco regions as the female parent with the MV P36D14 as male parent showed that 5 to 10% of the seeds produced were viable (80). Even when the proportion of viable seeds is low, such hybrids are sufficiently fertile to produce the signal of gene flow we detected. Furthermore, gene flow is bidirectional between *mexicana* and maize in natural populations (9), although it occurs at an asymmetrical rate that favors introgression from WRs into maize (80).

However, the gene flow test between LR > 2000-*mexicana* revealed that the gene flow signal between MVs-*mexicana* is picking up current gene flow between LRs-*mexicana*. Although this signal is heterogeneous in space, we detected gene flow from LRs into the Chalco region teosintes, but not for Central Plateau teosintes. In order to distinguish between these gene flow events, it is necessary to identify introgressed genomic regions to look for markers associated with domestication or improvement (81).

We found that the direction of gene flow signal from both MV groups goes toward *parviglumis* individuals sampled in 1980, a result suggesting that gene flow was stronger in the past. However, it does not preclude the possibility of gene flow into *parviglumis* collected after 2000, although to test this, we would require an allopatric population. Even beyond the objectives of the present paper, direct estimates of fitness and local adaptation will help to establish the consequences of introgression. Additionally, analyses of local adaptation by assessing the abundance of deleterious alleles and signatures of selection in candidate

genes would clarify if the ongoing gene flow is being counteracted by artificial and natural selection, and would help to establish the possible consequences for conservation.

F_{ST} between MVs with LRs and WRs Decays over Time. We found that F_{ST} between MVs and LRs, and MVs with *mexicana*, decays over time (Fig. 3) and that this decay can be partially explained by modification of LR and WR genomes by recent or continuous gene flow from MVs (82). Also, F_{ST} between *parviglumis* and MVs decreased through time despite the fact that according to Patterson's *D*-statistics, gene flow was stronger for older *parviglumis* samples (Fig. 3 *C* and *D*). However, the effect size is small, consistent with previous observations that *parviglumis* rarely hybridizes with maize (83). Moreover, since the ABBA-BABA test is polarized with *Z. diploperennis*, it is based on a smaller number of markers than F_{ST}. Thus, the results from the ABBA-BABA test in this instance should be viewed with caution. Nonetheless, it would be of interest to identify introgressed regions in LR and WR, and to compare their divergence levels with genomic regions that have not been affected by gene flow, in order to determine whether other processes are affecting genome-wide F_{ST} values.

F_{ST} between MV1 with LRs, *mexicana* and *parviglumis*, was significantly lower in all 3 periods than the divergence with MV2. This result is consistent with the early presence of MV1 and the abundance of the cultivated lines from MV1 versus MV2.

MV1 and MV2 Appear to Be Derived from the Tropical Mexican Breeding Program. More than 80% of MVs produced by the private sector in Mexico contain CIMMYT germplasm (84, 85), and we found that both MV1 and MV2 fell into the germplasm of the tropical breeding programs (Fig. 4). When we performed an ancestry analysis with the tropical breeding programs only (*SI Appendix*, Fig. S4), we observed that MV1 was more heterogeneous and appears to be mostly derived from the CIMMYT Mexican breeding program. However, a group of individuals from MV1 grouped with accessions from the Cameroon breeding program started in the 1980s. Germplasm for this program was based on inbred maize populations resistant to the streak virus, as well as other diverse sources (86). Thus some MVs collected for this work might be derived from foreign tropical breeding programs, although the Mexican breeding program appears to be the main source, as shown by the ML phylogenetic tree (*SI Appendix*, Fig. S5).

Finally, according to the results obtained with the PCA, ancestry assignment, and F_{ST} analysis, MV2 is a clearly differentiated group that likely derives from a few genotypes from the Mexican breeding program. Such a bottleneck and associated genetic drift may account for its reduced nucleotide diversity (*SI Appendix*, Table S2) and high divergence (87) (*SI Appendix*, Fig. S3).

Nucleotide Diversity Changes with Time. The nucleotide diversity for LRs decreased for the intermediate period, LR 1960 to 1980, but π is significantly higher for LR > 2000. The decrease of π in the 1960 to 1980 period was unexpected. However, it could be associated with the adoption and creolization of MVs released in 1960 that were based on a few outstanding Mexican LRs (23, 88). A study in the state of Chiapas in southeast Mexico showed that creolized MVs were planted in a higher proportion than local LRs and were resown up to 10 y after their introduction (22, 28).

On the other hand, although the participation of the private sector in trade started in 1980, the adoption of MVs was not instantaneous. Thus, it is improbable that MVs produced by the private sector were responsible for the significant decrease in π for LR 1960 to 1980. However, the gradual adoption of MVs over time (33) and the drastic changes in MV production after 1990 could explain the significant increase in π in LR > 2000. By the 1990s, 45 companies controlled 55% of the maize seed market. There was further consolidation over the next decade; in 2009, 2 companies produced 95% of the maize MVs used (30, 31). We expected that π would increase after gene flow from MVs, as has been reported for potato and rice (13, 18, 89), particularly if the MVs were foreign. However, it has been reported that the trend for increased diversity is followed by a decline when LRs are displaced by MVs (1, 15).

In contrast to π, the estimate θ_W tends to increase over the period evaluated (*SI Appendix*, Fig. S6). It matches with Tajima's *D* (*SI Appendix*, Table S3), which is positive for LR < 1960 and negative for LRs sorted into later periods. These parameters suggest a demographic expansion that is probably related to the sustained increase in demand for maize.

Although our approach was not designed to evaluate genetic erosion, our results are relevant to the controversy over genetic erosion at the center of origin of maize (90–92). Genetic erosion can be measured at crop, variety, and allele levels (15). We resampled the 7 LRs reported in the maize native record before 1960, and π was significantly higher for the most recent LRs (22, 39, 93). Also, note that the approach used in this work cannot detect genetic erosion at the level of specific alleles. However, the increase over time in the proportion of the predominant ancestry of MV1 in LRs, as well as the gene flow from MVs into LRs reported here, indicates that the gene pools present in LR < 1960 have been substantially modified.

Mexican agriculture is quite heterogeneous because farmers keep growing local LRs due to their sensory, aesthetic, ritualistic, and cultural values, in comparison to MVs (22, 29, 38, 39). Nevertheless, our results support a gradual replacement of traditional management systems with improved agricultural technologies, based on both MVs and their technology packages (94). This transition from community organization to crops seen only as exchange value can be socially, economically, and ecologically disadvantageous for small farmers (38), and also may imply a modernization bottleneck (15), which should be monitored in view of the nutritional and evolutionary services of small-scale farmers (35, 95).

Keeping LRs in a state of stasis is not possible and perhaps not desirable (39). However, it is possible to promote efforts to preserve wild and domesticated gene pools and allow for evolution. These efforts should include improving local LRs, encouraging seed savers networks, and developing markets for LRs that recognize their added value and the evolutionary services of small-scale farmers (1, 35).

In conclusion, our results show that the widespread adoption of MVs in Mexico has had significant impacts on the genetic composition of LRs and their WRs. We have demonstrated that there is substantial gene flow from MVs into LRs and into *mexicana*. As a result, genetic ancestry and allelic frequencies have changed across the time period evaluated. The changes detected provide short-term evidence of the ongoing evolution of maize LRs and their WRs at the center of origin of maize, and support a transition from traditional toward commercial agricultural systems, which suggests the early stages of a modernization bottleneck (1, 15). These findings can be used to design monitoring strategies and agricultural policies to reach biodiversity targets (3) for the conservation of gene reservoirs at the centers of diversity.

ACKNOWLEDGMENTS. We thank Azalea Guerra, Leopoldo Vázquez, Alejandro Martínez, Diana Rivera, Alicia Mastretta, Myriam Campos, Verónica Reyes, and Alfredo Villarruel for fieldwork assistance; Tania Garrido for technical laboratory assistance; Ernesto Campos Murillo for bioinformatic assistance in executing analyses in a cluster environment; Juan Manuel Hernandez Casillas for LR identification; and Mauricio Bellon for reading and suggestions made to this work. We also thank the reviewers for their constructive comments and feedback on earlier versions of the manuscript. Computational analyses were performed in the computing cluster of the National Commission for the Knowledge and Use of Biodiversity (CONABIO, its acronym in Spanish), which was partially funded by the Secretaría de Medio Ambiente y Recursos Naturales through the grant "Contribución de

la Biodiversidad para el Cambio Climático" to CONABIO. We thank the CIMMYT, INIFAP, and University Center of Biological and Agricultural Sciences germplasm banks for providing old seeds included in this work, and especially the Mexican farmers who shared their seeds and knowledge. This work was supported by the Consejo Nacional de Ciencia Tecnología through PhD Scholarship 384904 (to I.C.R.-B.) and by the Secretaría de Medio Ambiente y Recursos Naturales and CONABIO through the grant "Monitoreo de las razas criollas y de los linajes geográficos del maíz en México usando un enfoque genómico: una estrategia para su uso sustentable y su conservación" (to D.P.). This work is part of the PhD thesis of I.C.R.-B. from the Programa de Doctorado en Ciencias Bioquímicas, Universidad Nacional Autónoma de México.

1. S. B. Brush, In situ conservation of landraces in centers of crop diversity. *Crop Sci.* **35**, 346–354 (1995).
2. P. Gepts, Plant genetic resources conservation and utilization. *Crop Sci.* **46**, 2278–2292 (2006).
3. Convention on Biological Diversity, Aichi Biodiversity Targets (Convention on Biological Diversity, 2010). https://www.cbd.int/sp/targets/. Accessed 26 September 2018.
4. M. C. Arteaga *et al*., Genomic variation in recently collected maize landraces from Mexico. *Genom. Data* **7**, 38–45 (2015).
5. J. F. Doebley, M. M. Goodman, C. W. Stuber, Isoenzymatic variation in Zea (Gramineae). *Syst. Bot.* **9**, 203–218 (1984).
6. Y. Vigouroux *et al*., An analysis of genetic diversity across the maize genome using microsatellites. *Genetics* **169**, 1617–1630 (2005).
7. A. Eyre-Walker, R. L. Gaut, H. Hilton, D. L. Feldman, B. S. Gaut, Investigation of the bottleneck leading to the domestication of maize. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4441–4446 (1998).
8. L. Wang *et al*., The interplay of demography and selection during maize domestication and expansion. *Genome Biol.* **18**, 215 (2017).
9. M. B. Hufford *et al*., The genomic signature of crop-wild introgression in maize. *PLoS Genet.* **9**, e1003477 (2013).
10. A. Piñeyro-Nelson *et al*., Transgenes in Mexican maize: Molecular evidence and methodological considerations for GMO detection in landrace populations. *Mol. Ecol.* **18**, 750–761 (2009).
11. S. Ortiz-García *et al*., Absence of detectable transgenes in local landraces of maize in Oaxaca, Mexico (2003-2004). *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12338–12343 (2005).
12. D. Quist, I. H. Chapela, Transgenic DNA introgressed into traditional maize landraces in Oaxaca, Mexico. *Nature* **414**, 541–543 (2001).
13. E. Bitocchi *et al*., European flint landraces grown in situ reveal adaptive introgression from modern maize. *PLoS One* **10**, e0121381 (2015).
14. J. A. Serratos, *Gene Flow Among Maize Landraces, Improved Maize Varieties, and Teosinte: Implications for Transgenic Maize*, J. A. Serratos, M. A. Willcox, F. Castillo, Eds. (CIMMYT, 1997).
15. M. van de Wouw, C. Kik, T. van Hintum, R. van Treuren, B. Visser, Genetic erosion in crops: Concept, research results and challenges. *Plant Genet. Resour.* **8**, 1–15 (2010).
16. J. van Heerwaarden, J. Hellin, R. F. Visser, F. A. van Eeuwijk, Estimating maize genetic erosion in modernized smallholder agriculture. *Theor. Appl. Genet.* **119**, 875–888 (2009).
17. J. F. Crow, 90 years ago: The beginning of hybrid maize. *Genetics* **148**, 923–928 (1998).
18. K. A. Steele *et al*., Has the introduction of modern rice varieties changed rice genetic diversity in a high-altitude region of Nepal? *Field Crops Res.* **113**, 24–30 (2009).
19. E. Bitocchi *et al*., Introgression from modern hybrid varieties into landrace populations of maize (Zea mays ssp. mays L.) in central Italy. *Mol. Ecol.* **18**, 603–621 (2009).
20. R. A. Ortega Paczka, Variación en maíz y cambios socioeconómicos en Chiapas, Méx. 1946-1971. M.Sc. thesis, Colegio de Postgraduados, Chapingo, Mexico (1973).
21. S. B. Brush, J. E. Taylor, M. R. Bellon, Technology adoption and biological diversity in Andean potato agriculture. *J. Dev. Econ.* **39**, 365–387 (1992).
22. M. R. Bellon, S. B. Brush, Keepers of maize in Chiapas, Mexico. *Econ. Bot.* **48**, 196–209 (1994).
23. E. J. Wellhausen, El mejoramiento del maíz en México: Avances actuales y proyección hacia el futuro. *Revista de la Sociedad Mexicana de Historia Natural* **21**, 435–462 (1961).
24. R. E. Evenson, D. Gollin, Assessing the impact of the green revolution, 1960 to 2000. *Science* **300**, 758–762 (2003).
25. G. S. Khush, Green revolution: The way forward. *Nat. Rev. Genet.* **2**, 815–822 (2001).
26. CIMMYT, CIMMYT review 1976 (CIMMYT, 1976). https://repository.cimmyt.org/xmlui/handle/10883/3599. Accessed 27 May 2018.
27. D. M. Jones, *The Green Revolution in America: Success or Failure?* (University of Texas Press, 1977), pp. 55–63.
28. M. R. Bellon, J. Risopoulos, Small-scale farmers expand the benefits of improved maize germplasm: A case study from Chiapas, Mexico. *World Dev.* **29**, 799–811 (2001).
29. J. A. García-Salazar, R. Ramírez-Jaspeado, El mercado de la semilla mejorada de maíz (Zea mays L.) en México: Análisis del saldo comercial por entidad federativa. *Rev. Fitotec. Mex.* **37**, 69–77 (2014).
30. B. M. Luna Mena, M. A. Hinojosa Rodríguez, Ó. J. Ayala Garay, F. Castillo González, J. A. Mejía Contreras, Perspectivas de desarrollo de la industria semillera de maíz en México. *Rev. Fitotec. Mex.* **35**, 1–7 (2012).
31. G. Aboites, F. Martínez, G. Torres, El negocio de la producción de semillas mejoradas y su rol en el proceso de privatización de la agricultura mexicana. *Espiral Estudios sobre Estado y Sociedad* **6**, 151–185 (1999).
32. A. Espinosa-Calderón *et al*., Ley de Semillas y Ley Federal de Variedades Vegetales y transgénicos de maíz en México. *Rev. Mex. Cienc. Agric.* **5**, 293–308 (2014).
33. SIAP, Uso de tecnología y de servicios en el campo Cuadros tabulares 2016 (Servicio de Información Agroalimentaria y Pesquera, 2016). https://www.gob.mx/siap/documentos/tecnificacion. Accessed 5 June 2018.
34. M. L. Donnet *et al*., *El potencial de mercado de semillas mejoradas de maíz en México* (CIMMYT, 2012).
35. M. R. Bellon *et al*., Evolutionary and food supply implications of ongoing maize domestication by Mexican campesinos. *Proc. Biol. Sci.* **285**, 20181049 (2018).
36. R. H. Perales, S. B. Brush, C. O. Qualset, Landraces of Maize in Central Mexico: An altitudinal transect. *Econ. Bot.* **57**, 7–20 (2003).
37. M. R. Bellon, J. E. Taylor, "Folk" soil taxonomy and the partial adoption of new seed varieties. *Econ. Dev. Cult. Change* **41**, 763–786 (1993).
38. J. A. Zarazua-Escobar, G. Almaguer-Vargas, J. G. Ocampo-Ledesma, The direct field support program (PROCAMPO) from and its impact on agricultural productive and commercial knowledge management in Estado de México. *Agric. Soc. Desarro.* **8**, 89–105 (2011).
39. R. H. Perales, S. B. Brush, C. O. Qualset, Dynamic management of maize landraces in Central Mexico. *Econ. Bot.* **57**, 21–34 (2003).
40. C. U. Miramontes-Piña, *Situación actual y perspectivas del maíz en México 1996-2010* (SIAP Distrito Federal, México, 2006).
41. M. R. Bellon, J. Berthaud, Transgenic maize and the evolution of landrace diversity in Mexico. The importance of farmers' behavior. *Plant Physiol.* **134**, 883–888 (2004).
42. Y. Matsuoka *et al*., A single domestication for maize shown by multilocus microsatellite genotyping. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6080–6084 (2002).
43. D. R. Piperno, A. J. Ranere, I. Holst, J. Iriarte, R. Dickau, Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5019–5024 (2009).
44. A. J. Ranere, D. R. Piperno, I. Holst, R. Dickau, J. Iriarte, The cultural and chronological context of early Holocene maize and squash domestication in the Central Balsas River Valley, Mexico. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5014–5018 (2009).
45. J. J. Sánchez González *et al*., Ecogeography of teosinte. *PLoS One* **13**, e0192676 (2018).
46. E. Anderson, H. C. Cutler, Races of Zea Mays: I. Their recognition and classification. *Ann. Mo. Bot. Gard.* **29**, 69–88 (1942).
47. R. J. Elshire *et al*., A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* **6**, e19379 (2011).
48. CONABIO, Proyecto global de maíces nativos (Comisión Nacional Para el Conocimiento y Uso de la Biodiversidad, 2011). https://www.biodiversidad.gob.mx/genes/proyectoMaices.html. Accessed 23 April 2018.
49. K. Fukunaga *et al*., Genetic diversity and population structure of teosinte. *Genetics* **169**, 2241–2254 (2005).
50. K. Herten, M. S. Hestand, J. R. Vermeesch, J. K. J. Van Houdt, GBSX: A toolkit for experimental design and demultiplexing genotyping by sequencing experiments. *BMC Bioinformatics* **16**, 73 (2015).
51. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
52. F. J. Sedlazeck, P. Rescheneder, A. von Haeseler, NextGenMap: Fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics* **29**, 2790–2791 (2013).
53. H. Li *et al*.; 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
54. A. McKenna *et al*., The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
55. P. Danecek *et al*.; 1000 Genomes Project Analysis Group, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
56. I. C. Rojas-Barrera *et al*., Contemporary evolution of maize landraces and their wild relatives influenced by gene flow with modern maize varieties. Open Science Framework (OSF). https://osf.io/pqvt4/. Deposited 9 September 2019.
57. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
58. T. S. Korneliussen, A. Albrechtsen, R. Nielsen, ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).
59. E. Y. Durand, N. Patterson, D. Reich, M. Slatkin, Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* **28**, 2239–2252 (2011).
60. R. J. Kulathinal, L. S. Stevison, M. A. F. Noor, The genomics of speciation in Drosophila: Diversity, divergence, and introgression estimated using low-coverage genome sequencing. *PLoS Genet.* **5**, e1000550 (2009).
61. R. E. Green *et al*., A draft sequence of the Neandertal genome. *Science* **328**, 710–722 (2010).
62. J. Yan *et al*., Genetic characterization and linkage disequilibrium estimation of a global maize collection using SNP markers. *PLoS One* **4**, e8451 (2009).
63. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
64. P. J. Bradbury *et al*., TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635 (2007).
65. M. C. Romay *et al*., Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biol.* **14**, R55 (2013).
66. X. Zheng *et al*., A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–3328 (2012).
67. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **26**, 1641–1650 (2009).
68. J. d. J. Sánchez González *et al*., Distribución geográfica del teocintle (Zea spp.) en México y situación actual de las poblaciones (CONABIO, 2008). https://www.biodiversidad.gob.mx/genes/pdf/proyecto/Anexo8_ResultadosProyectos/FZ002/Informe%20final/Teocintle/Informe-Fina%20Teocintle_FZ002.pdf. Accessed 27 June 2018.

69. J. van Heerwaarden *et al*., Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1088–1092 (2011).
70. M. Bjarnason, Ed., *The Subtropical, Midaltitude, and Highland Maize Subprogram* (CIMMYT, 1994).
71. C. Jiang *et al*., Genetic analysis of adaptation differences between highland and lowland tropical maize using molecular markers. *Theor. Appl. Genet.* **99**, 1106–1119 (1999).
72. K. Mercer, Á. Martínez-Vásquez, H. R. Perales, Asymmetrical local adaptation of maize landraces along an altitudinal gradient. *Evol. Appl.* **1**, 489–500 (2008).
73. M. G. Vázquez Carrillo *et al*., Interacción genotipo-ambiente del rendimiento y calidad de grano y tortilla de híbridos de maíz en Valles Altos de Tlaxcala, México. *Rev. Fitotec. Mex.* **35**, 229–237 (2012).
74. M. Tadeo Robledo *et al*., Productividad de variedades precoces de maíz de grano amarillo para Valles Altos. *Rev. Mex. Cienc. Agric.* **3**, 1417–1423 (2012).
75. M. Trtikova *et al*., Teosinte in Europe–Searching for the origin of a novel weed. *Sci. Rep.* **7**, 1560 (2017).
76. CIMMYT, Maize production regions in developing countries (CIMMYT, 1988). https://repository.cimmyt.org/xmlui/handle/10883/705. Accessed 5 June 2018.
77. M. R. Bellon, D. Hodson, J. Hellin, Assessing the vulnerability of traditional maize seed systems in Mexico to climate change. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13432–13437 (2011).
78. L. de La Cruz Larios, "Sistemas de incompatibilidad genética en maíz y teocintle (Zea spp.) en México," PhD thesis, Universidad de Guadalajara, Guadalajara, Mexico (2007). http://repositorio.cucba.udg.mx:8080/xmlui/bitstream/handle/123456789/5149/Gonzalez_Ledesma_J_Santos.pdf?sequence=1. Accessed 8 June 2018.
79. J. M. Padilla García *et al*., Incompatibilidad gametofítica en las razas mexicanas de maíz. *Rev. Mex. Cienc. Agric.* **3**, 525–537 (2012).
80. B. M. Baltazar, J. de Jesús Sánchez-Gonzalez, L. de la Cruz-Larios, J. B. Schoper, Pollination between maize and teosinte: An important determinant of gene flow in Mexico. *Theor. Appl. Genet.* **110**, 519–526 (2005).
81. M. B. Hufford *et al*., Comparative population genomics of maize domestication and improvement. *Nat. Genet.* **44**, 808–811 (2012).
82. K. E. Holsinger, B. S. Weir, Genetics in geographically structured populations: Defining, estimating and interpreting F(ST). *Nat. Rev. Genet.* **10**, 639–650 (2009).
83. J. Doebley, Molecular evidence for gene flow among Zea species. *Bioscience* **40**, 443–448 (1990).
84. P. R. Aquino Mercado, R. J. Peña Bautista, I. Ortiz-Monasterio, México y el CIMMYT (CIMMYT, 2008). https://repository.cimmyt.org/handle/10883/657. Accessed 23 April 2018.
85. M. L. Morris, *Impactos del mejoramiento de maíz en América Latina, 1966-1997* (CIMMYT, 2000).
86. A. Menkir, A. Melake-Berhan, C. The, I. Ingelbrecht, A. Adepoju, Grouping of tropical mid-altitude maize inbred lines on the basis of yield data and molecular markers. *Theor. Appl. Genet.* **108**, 1582–1590 (2004).
87. L. Excoffier, N. Ray, Surfing during population expansions promotes genetic revolutions and structuration. *Trends Ecol. Evol.* **23**, 347–351 (2008).
88. A. J. Trueba Carranza, Semillas mexicanas mejoradas de maíz: Su potencial productivo (Estado de México: Colegio de Postgraduados, 2012).
89. E. Bellucci *et al*., Population structure of barley landrace populations and gene-flow with modern varieties. *PLoS One* **8**, e83891 (2013).
90. G. A. Dyer, A. López-Feldman, A. Yúnez-Naude, J. E. Taylor, Genetic erosion in maize's center of origin. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 14094–14099 (2014).
91. S. B. Brush *et al*., Assessing maize genetic erosion. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E1 (2015).
92. G. A. Dyer, A. López-Feldman, A. Yúnez-Naude, J. E. Taylor, J. Ross-Ibarra, Reply to Brush et al.: Wake-up call for crop conservation science. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E2 (2015).
93. D. Louette, A. Charrier, J. Berthaud, In situ conservation of maize in Mexico: Genetic diversity and maize seed management in a traditional community. *Econ. Bot.* **51**, 20–38 (1997).
94. D. Byerlee, E. H. de Polanco, Farmers' stepwise adoption of technological packages: Evidence from the Mexican Altiplano. *Am. J. Agric. Econ.* **68**, 519–527 (1986).
95. D. P. Faith *et al*., Evosystem services: An evolutionary perspective on the links between biodiversity and human well-being. *Curr. Opin. Environ. Sustain.* **2**, 66–74 (2010).