



OPEN ACCESS

EDITED BY

Wenwu Wu,
Zhejiang Agriculture and Forestry University,
China

REVIEWED BY

Sakina Elshibli,
University of Helsinki, Finland
Haiyun Yang,
Zhejiang Agriculture and Forestry University,
China
Julie Sardos,
Alliance Bioversity International and CIAT,
France

*CORRESPONDENCE

June Simpson
✉ june.simpson@cinvestav.mx

PRESENT ADDRESS

Adriana Ceniceros Ojeda,
Grupo Solena SAPI,
de CV. Departamento de Agrogenómica. Av.
Olimpica 3020-D, Col. Villas de San Juan,
Leon, Guanajuato, Mexico

RECEIVED 06 January 2023

ACCEPTED 30 May 2023

PUBLISHED 22 June 2023

CITATION

Martínez O, Ojeda AC, Hayano-Kanashiro C,
Reyes Valdés MH, Hernández JLP and
Simpson J (2023) Criteria for prioritizing
selection of Mexican maize landrace
accessions for conservation *in situ* or *ex situ*
based on phylogenetic analysis.
Front. Ecol. Evol. 11:1139377.
doi: 10.3389/fevo.2023.1139377

COPYRIGHT

© 2023 Martínez, Ojeda, Hayano-Kanashiro,
Reyes Valdés, Hernández and Simpson. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Criteria for prioritizing selection of Mexican maize landrace accessions for conservation *in situ* or *ex situ* based on phylogenetic analysis

Octavio Martínez¹, Adriana Ceniceros Ojeda^{2†},
Corina Hayano-Kanashiro³, M. Humberto Reyes Valdés⁴,
José Luis Pons Hernández⁵ and June Simpson^{2*}

¹Unidad de Genómica Avanzada (UGA/LANGEBIO), CINVESTAV Irapuato, Irapuato, Guanajuato, Mexico,

²Department of Genetic Engineering, CINVESTAV, Unidad Irapuato, Irapuato, Guanajuato, Mexico,

³DICTUS, Universidad de Sonora, Hermosillo, Sonora, Mexico, ⁴Universidad Autónoma Agraria Antonio
Narro, Calzada Antonio Narro, Saltillo, Coahuila, Mexico, ⁵INIFAP Campo Experimental del Bajío, Celaya,
Guanajuato, Mexico

Introduction: A combination of factors including migration, climate change and the encroachment of growing cities are associated with the extinction of species and a reduction in biodiversity. In Mexico maize landraces are an important source of germplasm that could be invaluable for future breeding programs in response a changing environment. Efforts to develop conservation strategies are hampered by the costs of housing and maintaining large germplasm collections. Effective criteria are therefore needed to identify rare genotypes that are in danger of being lost and to define minimal cost-effective core collections.

Methods: A large-scale genotyping analysis of Mexican maize landraces was carried out using 14 SSR markers spread across the 10 maize chromosomes. DNA was extracted from 30 individual plants of each accession and 3 bulks containing 220ng of DNA from each of 10 individual plants were prepared. Amplification reactions were carried out individually on each bulk sample and SSRs were detected using an automatic sequencer. Binary scores were used to produce phylogenies and a rareness index Ri was developed to identify a minimal collection of accessions.

Results: A total of 40,140 individual plants from accessions obtained in 14 different Mexican states and covering 47 maize races were analyzed. Accessions from Northern and Central-Southern states could be distinguished based on genotype and accessions from Michoacán state were found to be distinct to all other accessions. A minimal collection of 56 accessions that covered all identified alleles and included rare allele combinations was determined.

Discussion: The strategy employed successfully distinguished accessions from different regions of Mexico and identified rare germplasm and offers a cost effective method to determine minimal germplasm collections for conservation *in situ* or *ex situ*. The strategy could easily be adapted for use in other native crop species such as common bean, chili pepper and tomato among others.

KEYWORDS

maize landraces, conservation, rareness index, phylogeny, core collection

Introduction

Mexico although dependent on maize as a basic staple food crop, is not self-sufficient in maize production and imports corn mainly from the United States. As the origin and center of domestication of maize (Lyu, 2017), Mexico houses an incomparable wealth of diverse maize germplasm that has traditionally been maintained by small farmers throughout the country (Dyer and Lopez-Feldman, 2013; Acosta-Díaz et al., 2014; Perales and Golicher, 2014; Donovan et al., 2022). These landrace populations are adapted to local environmental and climatic conditions and are often grown for specific applications. As such they represent an invaluable genetic resource for maize breeding programs. A recent report (Ramirez-Villegas et al., 2022), estimates that around 70% of the global diversity of maize germplasm is currently conserved in *ex situ* genebanks. Several maize germplasm collections exist in institutions throughout Mexico (see Supplementary Table S1) and at a global level,¹ some collections harbor materials collected up to 60–80 years previously and may not accurately represent the landraces currently grown in Mexico. In this respect, a report by McLean-Rodríguez et al. (2021), comparing *ex situ* germplasm stored for long periods and *in situ* materials currently grown has shown based on GBS analysis that at least in the specific case studied, the overall genome polymorphism was very similar between both groups. Currently grown *in situ* germplasm however showed some selection for alleles at specific loci associated with grain size. Effective *ex situ* preservation of germplasm demands specially designed infrastructure and the capacity to catalog, administer and periodically regenerate the accessions to avoid loss of seed viability due to long term storage (Guzzon et al., 2021). The need to reduce costs by ensuring the optimized management of *ex situ* collections by defining “core collections” has been discussed previously (Crossa et al., 1994) and more recently Singh et al. (2019) reported a strategy based on GBS genotyping to identify duplicated or closely related materials within *Aegilops tauschii*.

Landrace populations of maize in Mexico even without any official status, contribute enormously to the *in situ* conservation of maize germplasm diversity. The introduction of hybrid corn varieties to Mexico under the free trade agreement (NAFTA, 1994) and subsequent uncertainty as to the presence and effects of genetically modified maize, led to the suggestion that hybrid varieties could rapidly replace traditional landraces with a consequent loss of germplasm diversity (Dyer et al., 2014). A study conducted among small farmers in Chiapas (Mexico), addressed this hypothesis with the conclusion that hybrids are planted by small farmers mainly for commercial reasons, whereas landraces are preserved to ensure food security for individual farmers and their families and to satisfy local preferences and uses in the preparation of specific dishes (Bellon and Hellin, 2011).

The essential role played by smallholder farmers in conserving and promoting maize germplasm diversity and their participation in providing food security by maintaining specific maize landraces has been documented in detail (Bellon et al., 2018). However these traditional farming systems are under threat and multiple factors contribute to the loss of landrace diversity. These include: adoption of hybrid varieties, a transition to other crops such as avocado or tomato,

the migration of younger generations from rural areas to cities or abroad, leaving the responsibility to maintain landrace populations to an aging population and changes in the commercialization and processing of maize among others (McLean-Rodríguez et al., 2019). In particular, before the moratorium on cultivation of transgenic maize varieties in Mexico was introduced it was feared that the cultivation of transgenic hybrids would lead to irreparable losses of valuable germplasm resources (Bellon and Berthaud, 2004; Bellon and Berthaud, 2006). In order to alleviate this situation and provide incentives to remain in rural areas and maintain traditional varieties of maize and other endogenous species, the idea to offer more directed support for *in situ* conservation has been receiving attention.² In common with *ex situ* initiatives, the main problem for conservation programs *in situ* is to balance costs and effectiveness in order to identify the germplasm which should be given priority. Several criteria for germplasm selection can be applied simultaneously based on phenotype, geographical location and applications of different varieties, however the presence of rare genetic combinations may be easily overlooked in such schemes. Molecular markers such as SSR, ISSR, and GBS offer an advantage for the analysis of collections where diversity is expected within individual accessions or where morphological classification is not clear. Several reports employing these markers have been applied previously to analyze collections of maize landraces either specific to Mexico, Latin America or outside this region (Matsuoka et al., 2002; McLean-Rodríguez et al., 2021; Soliman et al., 2021).

With the aim to develop a molecular-genetic based criteria to identify rare alleles or combinations of alleles within maize germplasm, a cost effective, SSR based method for genotyping maize landraces on a large scale and a novel algorithm based on phylogeny analysis was developed (Hayano-Kanashiro et al., 2017). The algorithm can be applied to determine the minimal number of samples from a given collection that would include the widest range of genetic diversity. The method was originally tested successfully as a proof of concept, initially by analyzing only maize landraces from Puebla state (Hayano-Kanashiro et al., 2017). In this report we describe the analysis of maize landraces from 14 Mexican states involving 40,140 individual plants and representing 47 maize races, based on the same methodology. We show that the method is efficient in determining a minimal number of maize landrace accessions that could be prioritized for inclusion in conservation strategies and provides a useful criteria for the process of selecting germplasm for *in situ* and/or *ex situ* conservation which could be easily extended to landrace collections of other crops such as chili pepper or common bean.

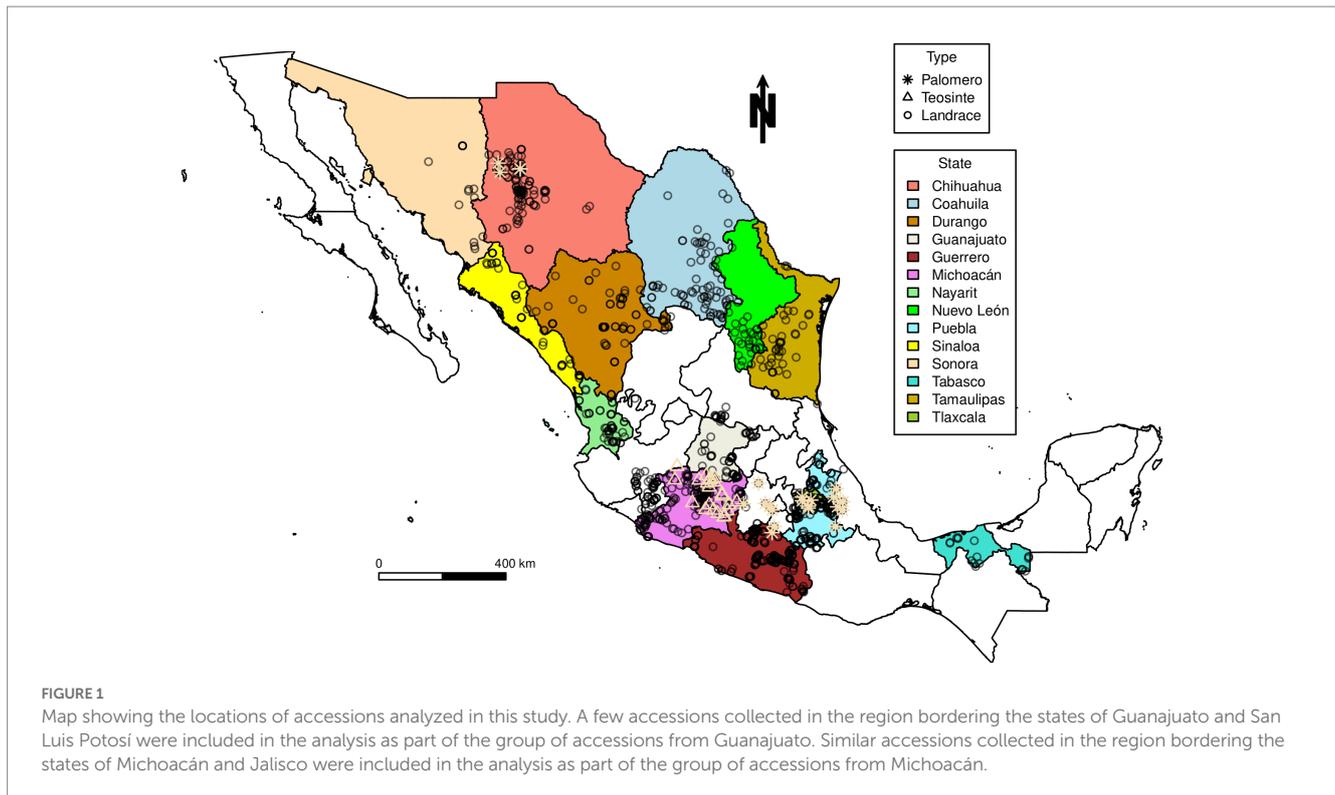
Materials and methods

Plant material

The aim of the study was to exploit recently collected samples/accessions of a wide range of maize landrace germplasm. Other criteria used to select samples were the availability of GPS localization

¹ <https://www.genebanks.org/resources/crops/maize/>

² https://www.gob.mx/cms/uploads/attachment/file/325810/biodiversidad_1_uno-32-47_eckart_boege.pdf



of collection sites and a basic morphological characterization in terms of race and/or color. Samples from regions important in terms of maize diversity and where landrace cultivation is significant such as Guerrero, Puebla and Michoacán, were essential to the study. Regions with lower levels of diversity from the northern states were also included for comparison (Perales and Golicher, 2014; Bellon et al., 2018). Although not all geographical regions and all maize races could be included, the sample represents one of the most extensive analyses in terms of combining molecular, geographical and morphological data carried out to date on maize landraces.

Maize landrace accessions were generously provided from local collections by institutions and universities throughout Mexico (see Annex 1 in Supplementary material). Landrace accessions were obtained from the states of Chihuahua, Coahuila, Durango, Guanajuato, Guerrero, Michoacán, Nuevo Leon, Nayarit, Puebla, Tabasco, Tamaulipas, Tlaxcala, Sinaloa, and Sonora and additionally for comparison, some teosinte and Palomero Toluqueño accessions were also included. The distribution and location of accessions analyzed is shown in Figure 1.

Forty seven of the 59 previously described maize races were included in the analysis namely: Pepitilla, Tuxpeño, Elotes Occidentales, Conejo, Elotes Cónicos, Ancho, Olotillo, Perla, Celaya, Vandeño, Arrocillo, Tepecintle, Tuxpeño Norteño, Zapalote Grande, Ratón, Onaveño, Bolita, Chalqueño, Tehuacanero, Maíz Dulce, Mushito, Cónico Norteño, Tabloncillo Zamorano, Mil granos, Tablilla de Ocho, Amarillo de Montaña, Complejo Serrano de Jalisco, Elotero de Sinaloa, Palomero Toluqueño, Reventador, Comitico, Zapalote Chico, Cacahuacintle, Apachito, Cristalino de Chihuahua, Palomero de Chihuahua, Gordo, Azul, Bofo, Jala, Dulce de Jalisco, Elotero de Sinaloa, Tabloncillo Perla, Dulcillo del Noroeste, Blando de Sonora, Chapalote, Morado Oaxaca. In addition, 3 races of Teosinte:

Chapala-*Zea mays*-subspecies Mexicana, Mesa central-*Zea mays*-subspecies Mexicana and Balsas-*Zea mays*-subspecies Parviglumis were also analyzed. The race of each accession is indicated in Supplementary Table S1. Some accessions could not be accurately classified as belonging to a particular race and this information has been left blank.

DNA extraction and molecular marker analysis

Preparation of DNA was as described in Hayano-Kanashiro et al. (2017) using the ZR-96 Plant/Seed DNA kit (ZYMO Research). The total number of accessions analyzed was 1,338 based on a bulk sampling strategy where DNA was extracted from 30 individual plants from each accession. Three bulked samples for each accession were obtained by pooling 220 ng of DNA from 10 individual plants. Therefore in total 40,140 individual plants were sampled. A selection of 14 microsatellite markers (SSR) distributed throughout the maize genome, determined based on data obtained from MAIZE GDB (Maize genetics and Genomics Database)³ was used for the analysis (Reyes-Valdes et al., 2013). PCR amplification conditions for SSR analysis were also described previously in Hayano-Kanashiro et al. (2017). Briefly, one primer of each pair was labeled with one of the fluorescent dyes: 6-FAM, PET, NED or VIC. Individual PCR reactions for each primer pair were carried out in 30 μ l final volume, using AmpliTaq Gold[®] PCR Master Mix (Applied Biosystems) and 100

³ <http://www.maizegdb.org/>

nanograms of template genomic DNA from each bulk. PCR conditions were as follows: 95°C initial denaturation for 5 min, 35 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 40 s, and a final extension at 72°C for 10 min. Amplification products were pooled in 3 groups of 4 different colored fluorescent dyes and one group of 2 colored dyes to avoid overlap of fragment sizes and were separated and visualized on an Applied Biosystems ABI 3730XL sequencer at the Genomic sequencing facility at LANGEBIO, CINVESTAV-Irapuato and analyzed using GENEMAPPER V. 4.0 and Peak scanner V. 1.0 software programs (Applied Biosystems).

Data analysis

Binary scores (1 presence, 2 absence) for each allele were compiled for all accessions and captured in a relational database (MySQL, Oracle© 2013). Analyses were carried out using R (R Development Core Team, 2011). Euclidean distance was measured within and between accessions to determine the genetic structure and a *t* test was also applied to determine the average differences between and within accessions. UPGMA (Unweighted Pair Group Method with Arithmetic Mean) was employed for cluster analysis and Jackknife resampling was used to evaluate the robustness of the dendrograms obtained. Collapsed clusters were obtained by defining the states of origin of the accessions as individual classes, then testing the statistical support for this classification. It was determined that the null hypothesis of equality of mean distances within and between states was rejected with a highly significant *p* value, $p < 2.2 \times 10^{-16}$ in both the Welch Two Sample *t*-test and the Wilcoxon rank sum test. A “collapsed” dendrogram could therefore be constructed grouping accessions by state. To test whether closely related collapsed clusters are strongly supported, a test of “robustness” of the collapsed clusters was carried out by determining the statistical support for the segregation of the accessions into two groups. This is done by testing the null hypothesis $H_0: w = b$ versus the alternative $H_a: w < b$, where *w* and *b* are the true means of the distances within and between the groups, respectively and applying the Welch Two Sample *t*-test and the Wilcoxon rank sum test.

A coefficient of rareness, R_i was determined for all accessions as described in detail in Hayano-Kanashiro et al. (2017). A brief outline of the method is as follows: For each accession the average of the square differences between the score of each marker/allele combination in relation to the mean score in the whole collection is determined. Accessions with higher uncommon marker/allele combinations have higher R_i values. A looping algorithm was used to determine a set of accessions that contain all marker/allele combinations by first selecting the accession with the highest R_i . Then for the remaining accessions, the gain in number of marker/allele combinations is determined until all marker/allele combinations are included in the set of accessions. Supplementary Tables S2, S3 present the set of all marker alleles and the matrix of marker/allele combinations per accession, respectively. Data can be accessed at: <https://zenodo.org/record/7798787#.ZCwvluzMKa0> and the algorithm employed to obtain the minimum collection of genotypes can be accessed at: <https://zenodo.org/record/7812240#.ZDM5Q-zMjN>.

TABLE 1 Comparison of numbers and size ranges of alleles.

Marker	No. alleles identified	Size range (bp) of alleles	Rarest alleles (found in a single bulk)
phi015	24	63–140	114, 138
phi031	24	173–241	237
phi033	23	224–302	284
phi034	26	95–166	166
phi051	11	127–152	152
phi053	30	127–210	152, 210
phi064	41	65–151	0
phi072	27	121–170	0
phi093	20	249–303	266, 303
phi109188	22	112–182	0
phi127	20	93–154	151, 154
phi427913	21	104–164	104, 158
phi96100	24	226–305	257, 300
phi96342	20	205–259	242*

Total number of alleles: 333, * most common allele.

Results and discussion

A simple method to compare diversity between studies is to determine the overall numbers and size ranges of the molecular alleles identified. Hayano-Kanashiro et al. (2017) showed that the analysis of samples from Puebla State alone using the methodology developed, uncovered a greater number of alleles and a wider range of allele sizes than any previous report. As would be expected the global analysis presented here, again uncovered new alleles and expanded the size range of the markers as shown in Table 1. Supplementary Table S2 indicates the sizes of all alleles identified and Supplementary Table S3 shows the incidence of each allele within each accession where 1,2 or 3 indicate presence in 1,2 or all 3 bulks of each accession. The most common allele was for SSR Phi96342, allele 242, found in 3798 of the bulked accession samples. Sixteen alleles (listed in Table 1) were found in only a single bulked sample. These results confirm the high levels of diversity of Mexican maize germplasm and emphasize the need to develop cost effective criteria for the conservation of these materials.

In order to determine the overall diversity and relationship between samples, genetic distances based on molecular data were calculated and the results are presented in the dendrogram in Figure 2. Strikingly and perhaps unexpectedly, the accessions are found to cluster strongly in relation to geographic region (colors represent different states) rather than morphological characteristics or race. On a larger scale samples obtained from northern states also cluster separately from those in central and southern states (indicated by A and B, respectively, in Figure 2). The clusters show that in the central and southern region genotypes are clearly regional, however in the northern states, a mixture of colors within some clusters can be observed, indicating that related genotypes are not so restricted to discrete geographical areas. Selection of landraces for adaptation to a

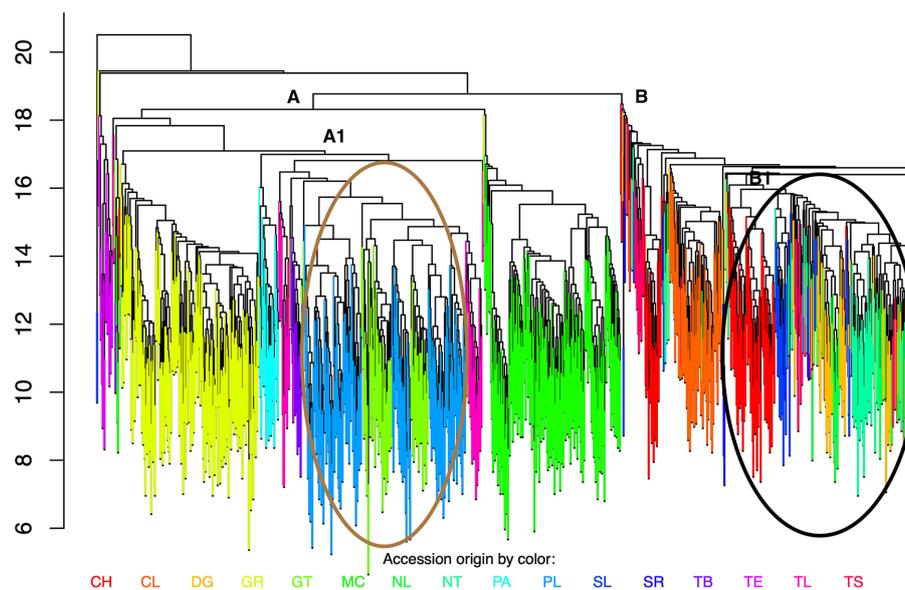


FIGURE 2

Global analysis of diversity. Dendrogram showing relationship between all accessions. Accessions from different regions, Teosinte and Palomero are shown in different colors. CH-Chihuahua, DG-Durango, CL-Coahuila, GR-Guerrero, GT-Guanajuato, MC-Michoacán, NL-Nuevo León, NT-Nayarit, PA-Palomero, PL-Puebla, SL-Sinaloa, SR-Sonora, TE-Teosinte, TB-Tabasco, TL-Tlaxcala, TS-Tamaulipas. (A) Cluster of accessions from central and southern regions, (B) Cluster of accessions from the northern region, Black circle indicates undefined clusters, Brown circle-defined clusters.

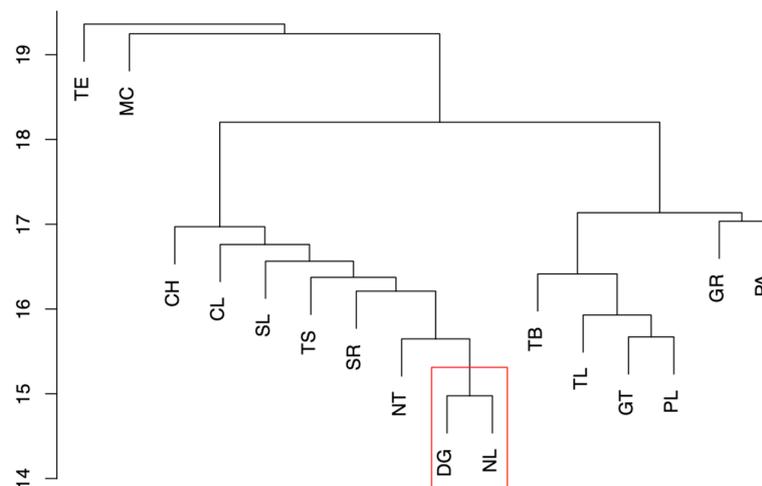


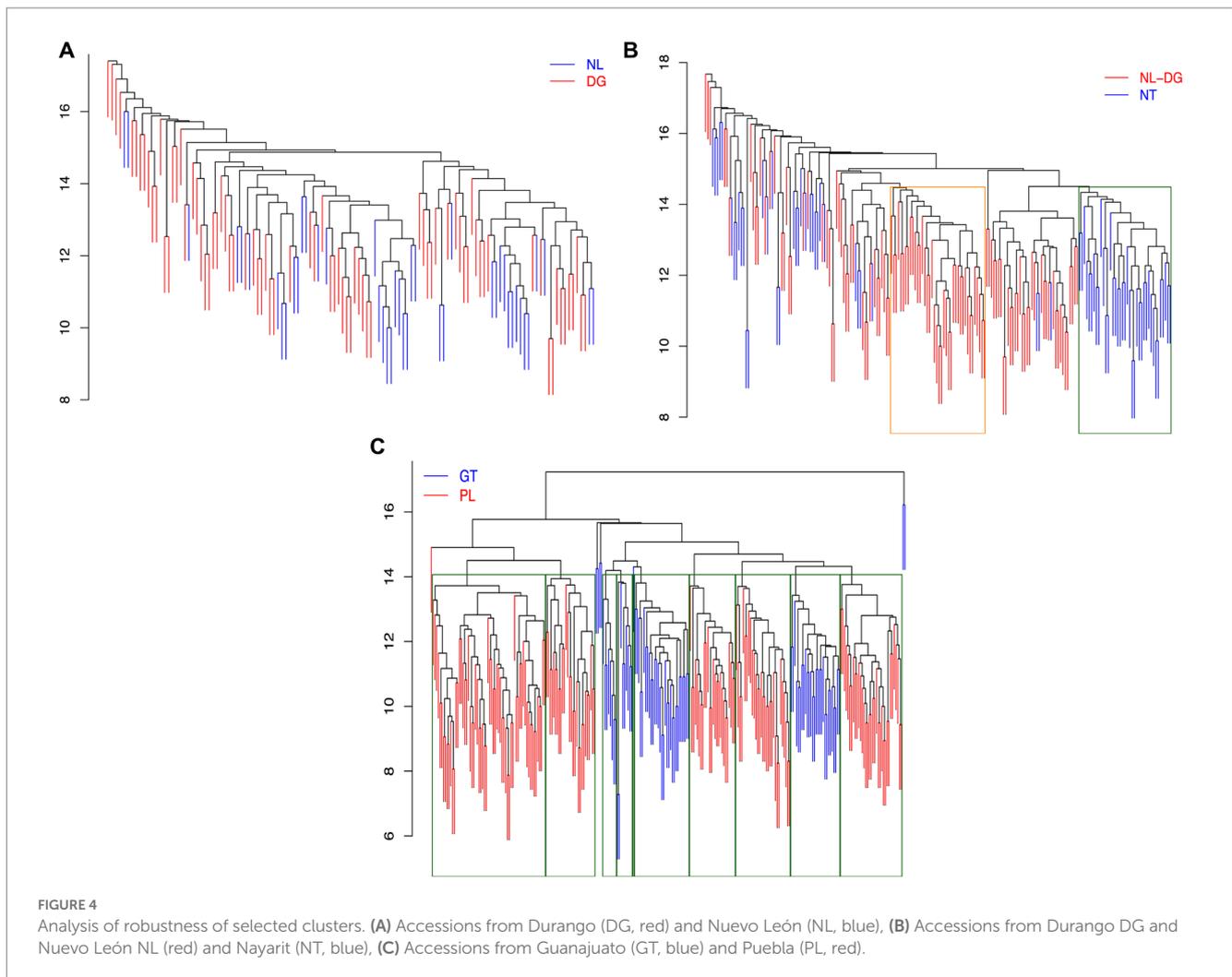
FIGURE 3

Initial representation of collapsed clusters. CH-Chihuahua, DG-Durango, CL-Coahuila, GR-Guerrero, GT-Guanajuato, MC-Michoacán, NL-Nuevo León, NT-Nayarit, PA-Palomero, PL-Puebla, SL-Sinaloa, SR-Sonora, TE-Teosinte, TB-Tabasco, TL-Tlaxcala, TS-Tamaulipas. Red box indicates the group tested for robustness.

specific environment may account for the association with geographical regions. Hayano-Kanashiro et al. (2017) reported no association between seed color and genotype and a low association with race and genotype. However a significant association with the altitude at which the landraces were grown and genotype was observed. Similar observations have been reported for maize landraces in Italy and France and for wheat in Afghanistan (Hartings et al., 2008; Diaw et al., 2021; Tehseen et al., 2021). In this case there are significant differences in the races/genotypes cultivated in the northern region in comparison to those grown in the central and southern region as documented in

Supplementary Table S1. These observations are consistent with a selection for adaptation to a specific environment and therefore could explain the separation between samples from the northern states (lower altitudes and arid or semi-arid terrain) and those from the central/southern region (higher altitudes and temperate climate).

In order to simplify the analysis a collapsed cluster dendrogram was obtained (Figure 3). Although this dendrogram reflects the topology in the dendrogram in Figure 2, the samples from Nuevo León (NL) and Durango (DG) are shown as separate but closely related branches. In Figure 2 however, a clear separation by state of



origin within group B1 cannot be determined. In order to resolve these differences, the robustness of the topology of group B1 was determined and compared with a group of clearly defined clusters (Figure 2A1). Figure 4A shows that when this analysis is carried out for the clusters containing samples from Nayarit, Durango, and Nuevo León (northern states) within cluster B1 in Figure 2, the topology was not robust in regard to the samples from Durango and Nuevo León and no discrete separation was obtained in relation to the geographical origin of the samples. In comparison, the samples from Nayarit although also included in cluster B1, did meet the criteria to accept the formation of a separate cluster (Figure 4B). When the same analysis was carried out for cluster A1 in Figure 2, containing samples from Guanajuato and Puebla in the central region, a robust separation of clusters was confirmed for individual genotype groups even though the samples share many genotypic components (Figure 4C). These results in general suggest that although similar landrace genotypes are grown in Guanajuato and Puebla they remain distinct and that no or very low levels of exchange of germplasm occurs between these regions. In contrast the lack of robustness between the clusters from Durango and Nuevo León indicates that similar germplasm is grown throughout these states and to a lesser extent is shared with areas in Nayarit. The reason behind the overlap in genotypes between the accessions from Nuevo León and Durango is unclear. Bellon et al. (2018) however have

discussed the importance of sharing and distribution of landrace materials between farmers. This will most commonly occur within local populations and from the map in Figure 1, it can be observed that many accessions from Nuevo León and Durango were obtained in areas close to or spanning the borders between the 2 states. The fact that this region is also served by major highways within the Mexican transport network may also facilitate the sharing of germplasm in this particular region.

Based on the analysis of robustness a corrected “collapsed” cluster dendrogram was produced which allows the topology and geographical associations of the genotypes to be clearly observed as shown in Figure 5 and reflects more accurately the topology of the dendrogram in Figure 2.

The teosinte samples in Figure 5 form a completely separate cluster as would be expected and the dendrogram shows clearly the separation between the two main geographical regions. Whereas in general, accessions from northern states are more closely related and less diverse, the central and southern regions show much higher diversity. Samples from Guerrero group closely with those of the Palomero landrace and are clearly distinct in relation to the other states of the central region. The samples from Michoacán however are shown to be extremely diverse and completely separated from all other genotypes. These results indicate that germplasm from Guerrero and especially Michoacán harbors

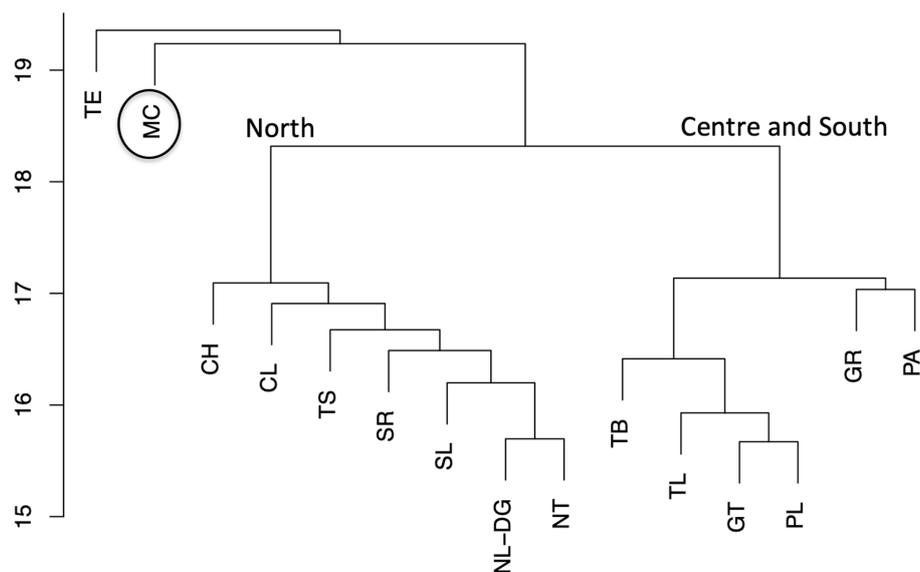


FIGURE 5

Global analysis summarized as robust collapsed clusters. Dendrogram summarizing the relationship between all accessions from different regions. CH-Chihuahua, DG-Durango, CL-Coahuila, GR-Guerrero, GT-Guanajuato, MC-Michoacán, NL-Nuevo León, NT-Nayarit, PA-Palmero, PL-Puebla, SL-Sinaloa, SR-Sonora, TE-Teosinte, TB-Tabasco, TL-Tlaxcala, TS-Tamaulipas. Two main clusters representing the North and Center/South are indicated, black circle indicates cluster of accessions from Michoacán.

diverse genotypes that could be at the level of specific alleles or combinations of different alleles across the genome. This suggests that materials from these regions should be a priority when criteria for germplasm conservation are determined. A prevalence of 5 of the 7 defined teosinte races including *Z. mays* spp. *Parviglumis* and *Z. mays* spp. *Mexicana* has been reported in Michoacán (Carrera-Valtierra et al., 2012). Ellstrand et al. (2007) have also shown that spontaneous hybridization between *Z. mays* spp. *Parviglumis* and *Z. mays* spp. *Mexicana* and maize (*Zea mays* spp. *mays*) can occur under field conditions and it is therefore probable that hybridization between maize landraces and wild teosinte genotypes could occur and may be a contributing factor in the diversity of this germplasm.

Although attention for conservation should be paid to germplasm from Guerrero and Michoacán this does not mean that all accessions from these regions should be given priority or that all germplasm from other regions should be excluded. The application of the Rarity index R_i (Hayano-Kanashiro et al., 2017) is therefore an extremely useful tool for the identification of rare materials and the determination of a core collection that should be given priority for conservation.

Application of the R_i analysis to the whole collection of samples produced a set of selected accessions where the dendrogram in Figure 6 shows a very distinct topology to the overall diversity analysis in Figure 5. As expected not all samples from either Michoacán or Guerrero were included in the minimal set of genotypes and some rare genotype combinations were identified in accessions from other regions as summarized in Table 2. Consistent with the observations from Figure 5 Michoacán harbors more diverse samples and the highest numbers of rare genotypes were identified in this region. In contrast, no rare genotypes were identified in Durango, Nuevo León, Nayarit, Sonora, Tabasco, and

TABLE 2 Numbers of rare accessions identified in each state/genotype.

State/Genotype	No. rare accessions
Chihuahua state	5
Coahuila state	1
Tamaulipas state	2
Michoacán state	16
Guerrero state	8
Puebla state	8
Guanajuato state	1
Tlaxcala state	4
Teosinte genotype	5
Palmero genotype	6

No rare accessions were identified in the states of Sinaloa, Nayarit, Nuevo León, Durango, Sonora or Tabasco.

Sinaloa. A surprising result is the observation that although the accessions from Puebla State as a whole did not form a separate cluster, 8 accessions were identified as rare. This is the same number as determined for Guerrero even though Guerrero germplasm in general is more diverse and forms a distinct cluster in the global diversity analysis. Interestingly, two main groups denominated A and B in Figure 6 that are distinct from those determined in the global diversity analysis were identified by the R_i analysis. On careful observation it can be seen that cluster B shows a similar topology and distribution of samples as Figure 2 where the teosinte samples (C) are most distant to all other samples. The Michoacán samples (D) form a discrete cluster and within the second main cluster the Guerrero accessions also form discrete cluster (E) whereas the palmero accessions and accessions from Puebla,

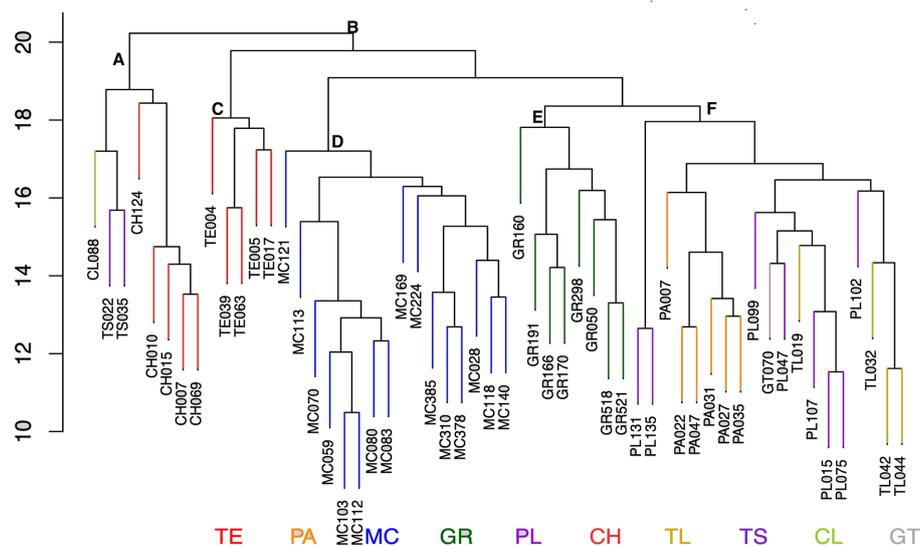


FIGURE 6

Core subset of accessions. Dendrogram showing the relationship between the accessions that form a core subset which includes all 333 alleles detected in the analysis. CH-Chihuahua, CL-Coahuila, GR-Guerrero, GT-Guanajuato, MC-Michoacán, PA-Palomero, PL-Puebla, TE-Teosinte, TL-Tlaxcala, TS-Tamaulipas, (A) Northern region, (B) All other regions, teosinte, and Palomero, (C) Teosinte, (D) Michoacán, (E) Guerrero, (F) Central region/Palomero.

Guanajuato and Tlaxcala are distributed throughout the remaining cluster (F). Perhaps the most unexpected result is the identification of rare genotypes in the northern states of Chihuahua, Coahuila, and Tamaulipas that group in a separate discrete cluster (A) distant from the other samples. On first sight this result is surprising, however a parallel analysis carried out to determine the presence of transgenes in the same samples used in this study (Under revision) revealed significant levels of transgenes in accessions obtained from regions close to the border with the United States. It is tempting to speculate therefore that the separation of cluster A may be due to introgression of genetic material from hybrid commercial maize into local landraces in these regions leading to new/rare allele combinations which were detected in the present analysis. The R_i identified 56 accessions as a minimal core collection that contains all 333 distinct SSR alleles identified in this study. Therefore the R_i can be used as a simple but effective tool to provide criteria for selection of germplasm that should be prioritized for conservation either *in situ* and/or *ex situ*. The R_i index could also be applied to monitor changes in landrace populations over time and in different regions. Although the methodology was applied to maize in this study, it could easily be adapted for landrace populations of other cultivated species such as chili pepper or common bean in Mexico or other important landrace species throughout the world.

Conclusion

The large-scale SSR analysis of Mexican landraces uncovered new alleles in comparison to previous studies (Gethi et al., 2002; Matsuoka et al., 2002), confirming the wealth of diversity represented by this local germplasm. Landrace genotypes are strongly associated with local geographic regions where farmers probably share and distribute seeds

saved from their harvest. The only region where this did not occur was between the states of Durango, Nuevo León and to a lesser extent Nayarit suggesting that seeds may be shared and distributed over this much more extensive area. Analysis of the robustness of the dendrogram topology also supports these conclusions. The global diversity analysis identified Michoacán and Guerrero as regions with comparatively higher diversity and these findings were confirmed by the R_i analysis that additionally identified rare genotypes within accessions from central region states. The lack of rare genotypes in samples from northern states is consistent with previously reported lower levels of diversity in this region (Perales and Golicher, 2014). The separate cluster of northern states in the R_i analysis was unexpected but may be due to introgression of commercial hybrid germplasm into local landraces.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://zenodo.org/record/7798787#.ZCwvluzMKa0>, 7798787 and <https://zenodo.org/record/7812240#.ZDM5Q-zMjJN>, 7812240.

Author contributions

JS conceived and coordinated the project. OM and MHRV developed and applied statistical analytical methods. CH-K, ACO, and JLPH coordinated accession management, sample preparation, and carried out molecular marker analysis. All authors participated in interpretation of results, contributed to the article, and approved the submitted version.

Funding

This work was funded by CIBIOGEM, Mexico grants 143919 and 266518.

Acknowledgments

The authors would like to thank Katia Gil Vega for administrative support. The authors also indebted to the many individuals and institutions that provided samples for this analysis as listed in the [Supplementary Annex](#). Qiagen and Labsengen (UGA-Cinvestav) provided equipment and services. CHK and ACO acknowledge postdoctoral fellowships provided under CIBIOGEM Mexico grants 143919 and 266518 respectively.

Conflict of interest

ACO is employed by Grupo Solena SAPI de CV.

References

- Acosta-Díaz, E., Zavala-García, F., Valadez-Gutiérrez, J., Hernández-Torres, I., Amador-Ramírez, M. D., and Padilla-Ramírez, J. S. (2014). Exploración de germoplasma nativo de maíz en Nuevo León, México. *Revista Mexicana de Ciencias Agrícolas* 8, 1477–1485. ISSN: 2007-0934
- Bellon, M. R., and Berthaud, J. (2004). Transgenic maize and the evolution of landrace diversity in Mexico. The importance of farmers' behavior. *Plant Physiol.* 134, 883–888. doi: 10.1104/pp.103.038331
- Bellon, M. R., and Berthaud, J. (2006). Traditional Mexican agricultural systems and the potential impacts of transgenic varieties on maize diversity. *Agric. Hum. Values* 23, 3–14. doi: 10.1007/s10460-004-5861-z
- Bellon, M. R., and Hellin, J. (2011). Planting hybrids, keeping landraces: agricultural modernization and tradition among small-scale maize farmers in Chiapas, Mexico. *World Development* 39, 1434–1443. doi: 10.1016/j.worlddev.2010.12.010
- Bellon, M. R., Mastretta-Yanes, A., Ponce-Mendoza, A., Ortiz-Santamaría, D., Oliveros-Galindo, O., Perales, H., et al. (2018). Evolutionary and food supply implications of ongoing maize domestication by Mexican campesinos. *Proc. R. Soc. B Biol. Sci.* 285:20181049. doi: 10.1098/rspb.2018.1049
- Carrera-Valtierra, J. A., Ron Parra, J., Jiménez Cordero, A. A., Morales Rivera, M. M., Iranda Medrano, R., Sahagún Castellanos, L., et al. (2012). *Razas de teocintle en Michoacán, Su origen, distribución y caracterización morfológica*. Texcoco: Universidad Autónoma de Chapingo.
- Crossa, J., Taba, S., Eberhart, S. A., Bretting, P., and Vencovsky, R. (1994). Practical considerations for maintaining germplasm in maize. *Theor. Appl. Genet.* 89, 89–95. doi: 10.1007/BF00226988
- Diaw, Y., Tollon-Cordet, C., Charcosset, A., Nicolas, S. D., Madur, D., Ronfort, J., et al. (2021). Genetic diversity of maize landraces from the south-west of France. *PLoS One* 16:e0238334. doi: 10.1371/journal.pone.0238334
- Donovan, J., Rutsaert, P., Domínguez, C., and Peña, M. (2022). Capacities of local maize seed enterprises in Mexico: implications for seed systems development. *Food Sec.* 14, 509–529. doi: 10.1007/s12571-021-01247-8
- Dyer, G. A., and Lopez-Feldman, A. (2013). Inexplicable or simply unexplained? The management of maize seed in Mexico. *Plos One* 8:e68320. doi: 10.1371/journal.pone.0068320
- Dyer, G. A., Lopez-Feldman, A., Yunez-Naude, A., and Taylor, J. E. (2014). Genetic erosion in maize's center of origin. *Proc. Natl. Acad. Sci. U. S. A.* 111, 14094–14099. doi: 10.1073/pnas.1407033111
- Ellstrand, N. C., Garner, L. C., Hegde, S., Guadagnuolo, R., and Blancas, L. (2007). Spontaneous hybridization between maize and teosinte. *J. Hered.* 98, 183–187. doi: 10.1093/jhered/esm002
- Gethi, J. G., Labate, J. A., Lamkey, K. R., Smith, M. E., and Kresovich, S. (2002). SSR variation in important U.S. maize inbred lines. *Crop Sci.* 42, 951–957. doi: 10.2135/cropsci2002.9510
- Guzzon, F., Gianella, M., Velazquez Juarez, J. A., Sanchez Cano, C., and Costich, D. E. (2021). Seed longevity of maize conserved under germplasm bank conditions for up to 60 years. *Ann. Bot.* 127, 775–785. doi: 10.1093/aob/mcab009
- Hartings, H., Berardo, N., Mazzinelli, G. F., Valoti, P., Verderio, A., and Motto, M. (2008). Assessment of genetic diversity and relationships among maize (*Zea mays* L.) Italian landraces by morphological traits and AFLP profiling. *Theor. Appl. Genet.* 117, 831–842. doi: 10.1007/s00122-008-0823-2
- Hayano-Kanashiro, C., de la Vega Martínez, O., Reyes-Valdés, M. H., Pons-Hernández, J.-L., Hernández-Godínez, F., Alfaro-Laguna, E., et al. (2017). An SSR-based approach incorporating a novel algorithm for identification of rare maize genotypes facilitates criteria for landrace conservation in Mexico. *Ecol. Evol.* 7, 1680–1690. doi: 10.1002/ece3.2754
- Lyu, J. (2017). Maize domestication: an ancient genome speaks. *Nat. Plants* 3:16215. doi: 10.1038/nplants.2016.215
- Matsuoka, Y., Vigouroux, Y., Goodman, M. M., Sanchez, G. J., Buckler, E., and Doebley, J. (2002). A single domestication for maize shown by multilocus microsatellite genotyping. *Proc. Natl. Acad. Sci. U. S. A.* 99, 6080–6084. doi: 10.1073/pnas.052125199
- McLean-Rodríguez, F. D., Camacho-Villa, T. C., Almekinders, C. J. M., Pè, M. E., Dell'Acqua, M., and Costich, D. E. (2019). The abandonment of maize landraces over the last 50 years in Morelos, Mexico: a tracing study using a multi-level perspective. *Agric. Hum. Values* 36, 651–668. doi: 10.1007/s10460-019-09932-3
- McLean-Rodríguez, F. D., Costich, D. E., Camacho-Villa, T. C., Pè, M. E., and Dell'Acqua, M. (2021). Genetic diversity and selection signatures in maize landraces compared across 50 years of *in situ* and *ex situ* conservation. *Heredity* 126, 913–928. doi: 10.1038/s41437-021-00423-y
- NAFTA. (1994). Available at: [https://www.trade.gov/north-american-free-trade-agreement-nafta#:~:text=TheNorthAmericanFreeTradeAgreement\(NAFTA\)whichwas,U.S.Mexicobilateralcommercialrelationship](https://www.trade.gov/north-american-free-trade-agreement-nafta#:~:text=TheNorthAmericanFreeTradeAgreement(NAFTA)whichwas,U.S.Mexicobilateralcommercialrelationship) (Accessed June 07, 2023).
- Perales, H., and Golicher, D. (2014). Mapping the diversity of maize races in Mexico. *PLoS One* 9:e114657. doi: 10.1371/journal.pone.0114657
- Ramirez-Villegas, J., Khoury, C. K., Achicanoy, H. A., Diaz, M. V., Mendez, A. C., Sosa, C. C., et al. (2022). State of *ex situ* conservation of landrace groups of 25 major crops. *Nat. Plants* 8, 491–499. doi: 10.1038/s41477-022-01144-8
- Reyes-Valdes, M. H., Santacruz-Varela, A., Martinez, O., Simpson, J., Hayano-Kanashiro, C., and Cortes-Romero, C. (2013). Analysis and optimization of bulk DNA sampling with binary scoring for germplasm characterization. *PLoS One* 8:e79936. doi: 10.1371/journal.pone.0079936
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: The R Foundation for Statistical Computing. Available at: <http://www.R-project.org/>
- Singh, N., Wu, S., Raupp, W. J., Sehgal, S., Arora, S., Tiwari, V., et al. (2019). Efficient curation of genebanks using next generation sequencing reveals substantial duplication of germplasm accessions. *Sci. Rep.* 9:650. doi: 10.1038/s41598-018-37269-0

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fevo.2023.1139377/full#supplementary-material>

Soliman, E. R. S., El-Shazly, H. H., Börner, A., and Badr, A. (2021). Genetic diversity of a global collection of maize genetic resources in relation to their subspecies assignments, geographic origin, and drought tolerance. *Breed. Sci.* 71, 313–325. doi: 10.1270/jsbbs.20142

Tehseen, M. M., Istiqlil, D., Kehel, Z., Sansaloni, C. P., da Silva Lopes, M., Kurtulus, E., et al. (2021). Genetic diversity and population structure analysis of *Triticum aestivum* L landrace panel from Afghanistan. *Genes* 12:340. doi: 10.3390/genes12030340