



## Genomic variation in recently collected maize landraces from Mexico



María Clara Arteaga<sup>a,1</sup>, Alejandra Moreno-Letelier<sup>b,1</sup>, Alicia Mastretta-Yanes<sup>c,\*</sup>, Alejandra Vázquez-Lobo<sup>e</sup>, Alejandra Breña-Ochoa<sup>e</sup>, Andrés Moreno-Estrada<sup>d,2</sup>, Luis E. Eguiarte<sup>e</sup>, Daniel Piñero<sup>e</sup>

<sup>a</sup> Departamento de Biología de la Conservación, División de Biología Experimental y Aplicada, Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California, Carretera Ensenada-Tijuana 3918, 22860 Ensenada, B.C., Mexico

<sup>b</sup> Jardín Botánico, Instituto de Biología, Universidad Nacional Autónoma de México, Mexico, DF, 04510, Mexico

<sup>c</sup> CONACYT Research Fellow - Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, Mexico, Liga Periférico – Insurgentes Sur, No. 4903, 14010, México, DF, Mexico

<sup>d</sup> Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA

<sup>e</sup> Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico, DF, 04510, Mexico

### ARTICLE INFO

#### Article history:

Received 29 July 2015

Received in revised form 3 November 2015

Accepted 6 November 2015

Available online 7 November 2015

#### Keywords:

Maize

Teosinte

Maize SNP50K BeadChip

Mexican landraces

Proyecto Global de Maíces Nativos

### ABSTRACT

The present dataset comprises 36,931 SNPs genotyped in 46 maize landraces native to Mexico as well as the teosinte subspecies *Zea mays* ssp. *parviglumis* and ssp. *mexicana*. These landraces were collected directly from farmers mostly between 2006 and 2010. We accompany these data with a short description of the variation within each landrace, as well as maps, principal component analyses and neighbor joining trees showing the distribution of the genetic diversity relative to landrace, geographical features and maize biogeography. High levels of genetic variation were detected for the maize landraces ( $H_E = 0.234$  to  $0.318$  (mean  $0.311$ )), while slightly lower levels were detected in *Zea m. mexicana* and *Zea m. parviglumis* ( $H_E = 0.262$  and  $0.234$ , respectively). The distribution of genetic variation was better explained by environmental variables given by the interaction of altitude and latitude than by landrace identity. This dataset is a follow up product of the Global Native Maize Project, an initiative to update the data on Mexican maize landraces and their wild relatives, and to generate information that is necessary for implementing the Mexican Biosafety Law.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### 1. Introduction

The astonishing phenotypic diversity within maize (*Zea mays* ssp. *mays*) may only be paralleled by the range of variation in dogs, with the difference that maize was domesticated only ca. 9000 years ago [22,33] and dogs up to ca. 35,000 years ago [32]. The morphological and physiological variation of maize is evidenced by its hundreds of landraces [12]. Landraces are dynamic populations with a historical origin, distinct identity, often genetically diverse and locally adapted, and associated with a set of farmers' practices of seed selection and field management as well as with traditional knowledge [7]. In Mexico 59 maize landraces are currently grown [30], including Wellhausen's [34] landraces, although more may remain undiscovered [19]. These landraces are particularly important from the genetic perspective because

they are grown under contrasting environmental conditions; and because Mexico is where maize was domesticated most likely from Balsas teosinte (*Z. m. ssp. parviglumis*; [18,33]) and where admixture with another teosinte (*Z. m. ssp. mexicana*) widely occurred ever since [33].

The Mexican landraces thus represent an important element to explore the evolution of the maize genome and could act as a genetic reservoir for further adapting crops to new conditions and pathogens [12]. However, in order to conserve, monitor and better use this variation it is necessary to understand it at its molecular, geographic and biocultural levels.

The high diversity of Mexican maize landraces is related to the biocultural processes by which they emerged: landraces are a product of indigenous selection to satisfy quality and variety requirements of the indigenous diet and traditions including religious ideas related to the color and shape of the cob [19]. This has been carried out by more than 60 indigenous groups and mestizo farmers of Mexico for over 9000 years [17] and it is still an on-going process. For instance, some landraces traits (e.g. color, shape) are associated to specific products of the Mexican cuisine and are consumed both the rural producers and the population of urban areas [16], thus driving preferences and selection over landraces and traits. In this way, both for food security and cultural preferences, the Mexican maize landraces are still grown in the country. This is done mostly by smallholders, typically in <5 ha,

\* Corresponding author.

E-mail addresses: [mariaclaraarteaga@yahoo.com](mailto:mariaclaraarteaga@yahoo.com) (M.C. Arteaga), [almolet@gmail.com](mailto:almolet@gmail.com) (A. Moreno-Letelier), [amastretta@conabio.gob.mx](mailto:amastretta@conabio.gob.mx), [a.mstt.yanes@gmail.com](mailto:a.mstt.yanes@gmail.com) (A. Mastretta-Yanes), [ale.lobolobo@gmail.com](mailto:ale.lobolobo@gmail.com) (A. Vázquez-Lobo), [alesaurus\\_bo@yahoo.com](mailto:alesaurus_bo@yahoo.com) (A. Breña-Ochoa), [amoreno@langebio.cinvestav.mx](mailto:amoreno@langebio.cinvestav.mx) (A. Moreno-Estrada), [fruns@unam.mx](mailto:fruns@unam.mx) (L.E. Eguiarte), [pinero@unam.mx](mailto:pinero@unam.mx) (D. Piñero).

<sup>1</sup> These authors contributed equally to this work.

<sup>2</sup> Current address: Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO), Centro de Investigación y de Estudios Avanzados del IPN, Irapuato, Guanajuato, Mexico.

but together accounting for 85% of the productive land of Mexico [23, 31]. These smallholders perform rainfed and traditional agriculture, often growing more than one variety per cycle because a single one does not contain all the desired characteristics and because growing landraces with different vulnerabilities allows for a yield even under adverse conditions ([4,16,28]. Such management generates diverse opportunities for gene flow, thus promoting a complex genetic mosaic among these landraces. To such gene flow scenario we must add that some of the Mexican landraces are grown in sympatry with wild teosinte subspecies, which in Central Mexico occur as weeds, but that are also considered a valuable genetic resource, especially regarding resistance to adverse conditions and diseases [26].

On top of the diversity driven by cultural preferences, there is the diversity of environments provided by the Mexican topography to which maize has been exposed here over historical time. This means that the Mexican maize landraces have been grown over thousands of years from sea level to more than 2900 m of altitude, from 12.0 to 29.1 °C growing season mean temperature and from 400 to 3555 mm

growing season rainfall [29]. In other words, these landraces can be grown in a wide range of environments including arid and cold conditions, where commercial hybrids perform poorly [4,16].

Despite their importance, several Mexican maize landraces are threatened due to the socio-economic problems that Mexican agriculture is currently facing [1,5,16]. Under such scenario, several instances of the Mexican government started the “Global Native Maize Project” (<http://www.biodiversidad.gob.mx/genes/proyectoMaices.html>), an initiative to update the data on Mexican maize landraces and their wild relatives, and to generate information that is necessary for implementing the Mexican Biosafety Law [10]. This legislation requires the scientific and detailed description of the areas of origin for crops native to Mexico and analyses of their genetic diversity, and thus needs molecular tools and data that could help monitoring and managing the Mexican germplasm [5].

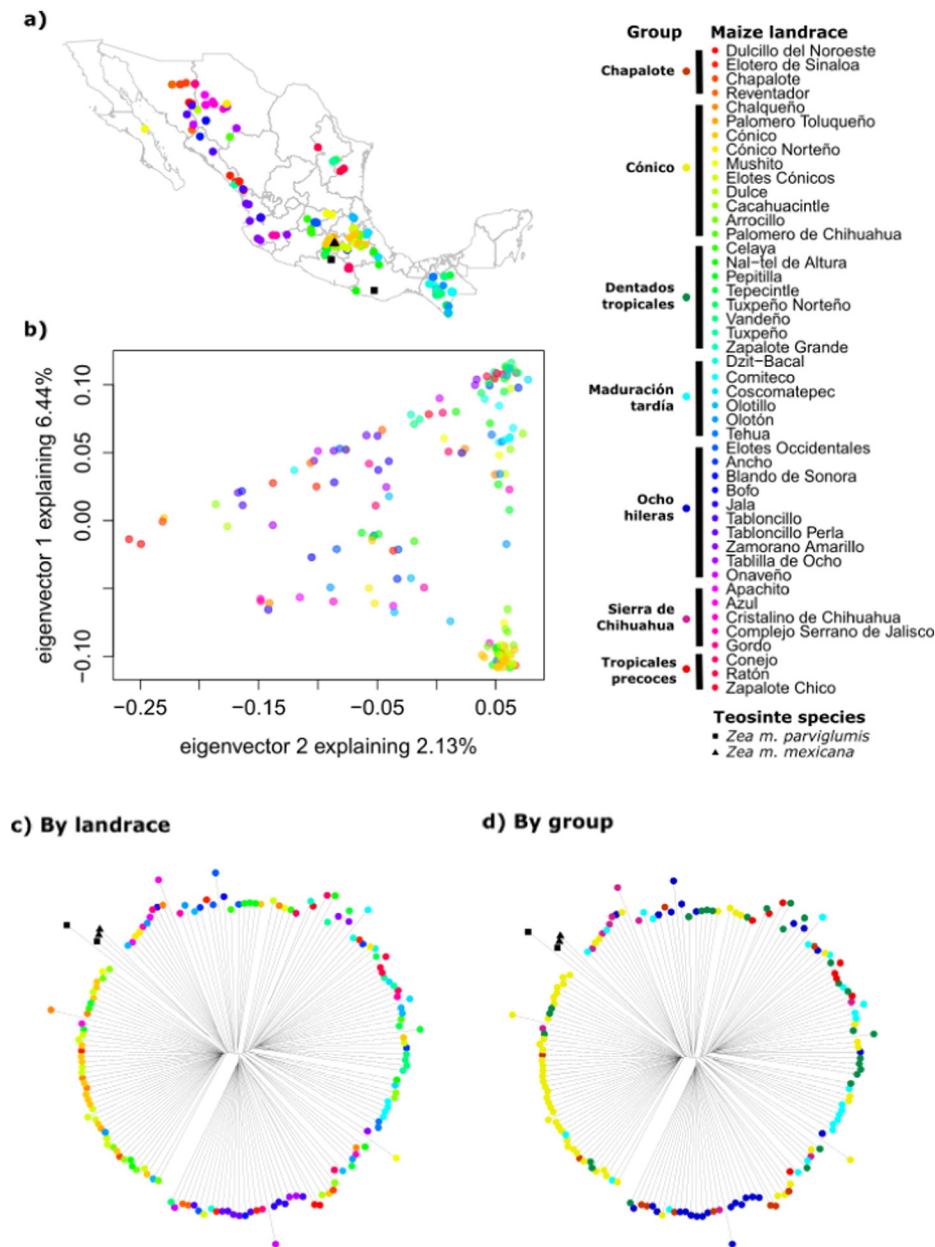
Here we provide 36,931 SNPs genotyped using Illumina MaizeSNP50 BeadChip in 46 maize landraces and two teosinte subspecies aiming to: (1) aid programs for the conservation, monitoring and better use of this

**Table 1**

Population genetics statistics for the genotyped Mexican landraces and teosinte species using the Illumina MaizeSNP50 BeadChip.

Landrace or sp.	n	% Missing data	Total number of alleles	$H_E$	$H_O$	$F_{IS}$
Ancho	3	0.90	61,314	0.308	0.341	−0.106
Apachito	2	0.95	57,153	0.307	0.379	−0.234
Arrocillo	4	0.93	64,856	0.326	0.309	0.051
Azul	2	0.88	56,711	0.300	0.374	−0.246
Blando de Sonora	1	0.79	46,896	0.294	0.560	−0.907
Bofo	1	0.90	46,656	0.290	0.550	−0.893
Cacahuacintle	5	1.06	65,784	0.319	0.271	0.149
Celaya	3	1.03	59,772	0.288	0.332	−0.154
Chalqueño	7	1.00	68,346	0.329	0.282	0.142
Chapalote	2	1.11	56,672	0.302	0.322	−0.067
Comiteco	5	1.10	66,071	0.324	0.272	0.161
Complejo Serrano de Jalisco	2	0.91	57,257	0.309	0.376	−0.218
Conejo	4	0.84	64,725	0.323	0.336	−0.040
Coscomatepec	3	1.15	61,334	0.308	0.334	−0.085
Cristalino de Chihuahua	2	0.87	57,319	0.309	0.378	−0.223
Cónico	16	0.89	70,749	0.328	0.289	0.120
Cónico Norteño	3	1.10	61,352	0.310	0.281	0.093
Dulce	1	0.87	47,433	0.311	0.591	−0.904
Dulcillo del Noreste	2	0.93	55,987	0.289	0.374	−0.293
Dzit-Bacal	3	1.04	61,848	0.317	0.326	−0.031
Elotero de Sinaloa	5	0.83	65,725	0.317	0.314	0.012
Elotes Cónicos	14	0.86	70,559	0.327	0.292	0.107
Elotes Occidentales	4	0.85	64,872	0.326	0.342	−0.049
Gordo	2	0.78	58,426	0.327	0.384	−0.175
Jala	4	0.78	64,820	0.324	0.335	−0.034
Mushito	3	0.96	62,086	0.321	0.332	−0.036
Nal-tel. de Altura	5	1.30	65,075	0.312	0.242	0.225
Olotillo	6	0.86	67,455	0.329	0.309	0.059
Olotón	4	0.87	64,614	0.322	0.310	0.038
Onaveño	2	1.01	56,516	0.297	0.328	−0.104
Palomero Toluqueño	1	0.68	48,012	0.320	0.618	−0.928
Palomero de Chihuahua	1	0.91	46,579	0.289	0.546	−0.891
Pepitilla	4	0.83	64,520	0.321	0.318	0.008
Ratón	3	1.10	61,788	0.315	0.348	−0.102
Reventador	2	0.90	58,103	0.322	0.403	−0.251
Tablilla de ocho	2	0.97	57,625	0.314	0.384	−0.223
Tabloncillo	4	0.95	63,947	0.312	0.310	0.008
Tabloncillo	3	1.18	61,823	0.319	0.315	0.013
Tehua	2	1.07	56,262	0.296	0.339	−0.146
Tepecintle	4	0.83	64,540	0.322	0.318	0.012
Tuxpeño	4	0.92	63,965	0.314	0.311	0.008
Tuxpeño Norteño	2	1.09	55,851	0.289	0.363	−0.259
Vandeno	4	0.79	65,028	0.328	0.333	−0.015
Zamorano Amarillo	3	0.86	61,726	0.314	0.348	−0.110
Zapalote Chico	1	1.31	44,134	0.234	0.422	−0.800
Zapalote Grande	1	0.83	47,346	0.307	0.586	−0.908
<i>Zea m. mexicana</i>	2	2.69	53,687	0.262	0.306	−0.171
<i>Zea m. parviglumis</i>	2	2.88	51,530	0.234	0.189	0.191

n: number of individuals used per landrace or species,  $H_E$ : expected heterozygosity correcting for sampling size,  $H_O$ : observed heterozygosity correcting for sampling size,  $F_{IS}$ : inbreeding coefficient.

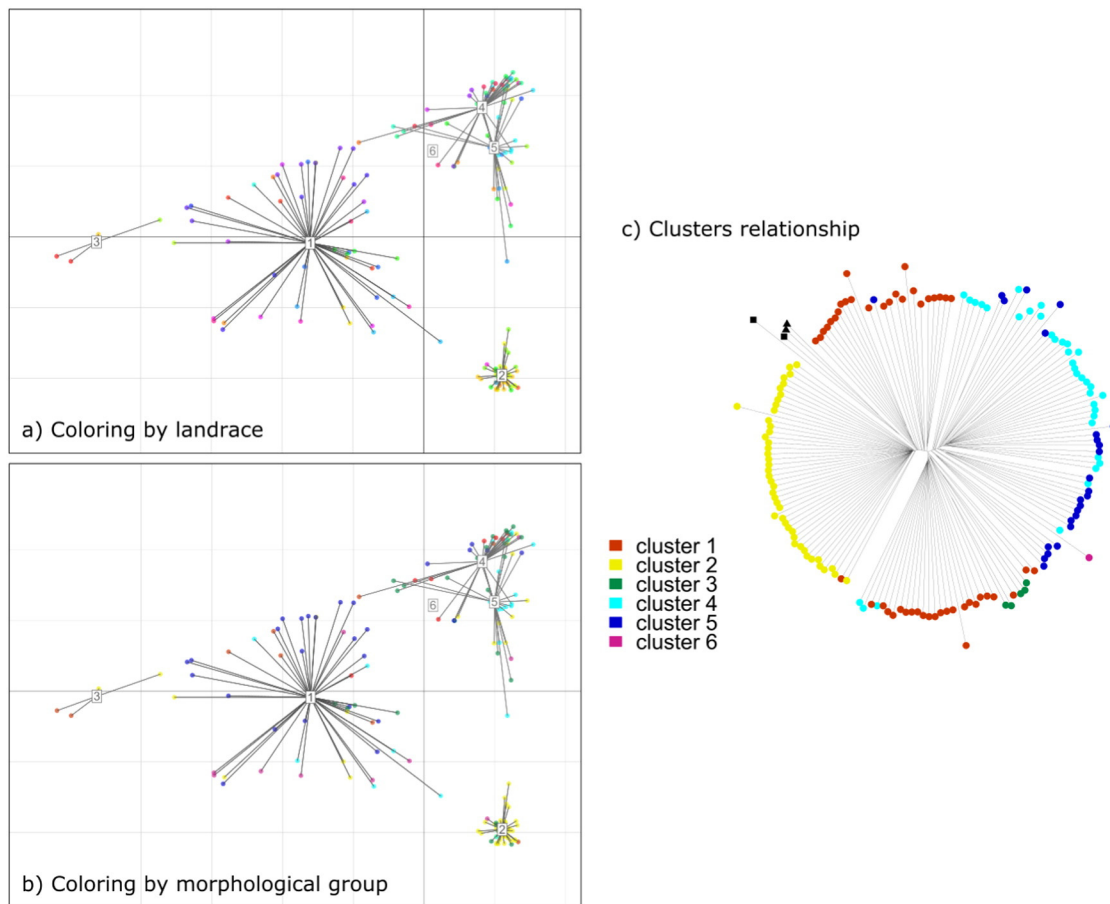


**Fig. 1.** Genetic variation by maize landrace and morphological groups in Mexico. a) Geographic origin of the samples; b) PCA showing the first two components, each point represents a sample colored by landrace (teosinte spp. are not shown); c) neighbor joining trees colored by landrace; and d) by morphological group according to CONABIO [10]. Maize landrace color code is shown to the right of the black bar and group code to the left.

variation, by providing a summary of the genetic diversity present in the Mexican landraces sampled during the Global Native Maize Project; and (2) to facilitate further research on the genetic variation of maize by adding to the existing data this publically available dataset of recently collected samples. We accompany these data with a short description of the variation in each landrace (Table 1), as well as maps, principal component analyses and neighbor joining trees showing the distribution of the genetic diversity relative to landrace (Figs. 1 and 2), geographical features (Figs. 3 and 4a) and maize biogeography (Fig. 5b). The results presented here can also be explored with interactive figures available at <https://conabio.shinyapps.io/MaicesMx50kSNP-english>. The present dataset is a following up product of the abovementioned Global Native Maize Project, and represents a contribution to the molecular data needed to monitor and manage the Mexican genetic diversity of maize.

## 2. Results and discussion

One to 16 (mean 3.4) accessions (one individual each) per landrace and two per teosinte subspecies were included (Table 1). For the landraces, the expected heterozygosity ranged from  $H_E = 0.234$  to  $0.318$  (mean  $0.311$ ), and for *Zea m. mexicana* and *Zea m. parviglumis* it was  $H_E = 0.262$  and  $0.234$ , respectively (Table 1). Up to 70,749 total alleles were found in a single landrace (Cónico,  $n = 16$ ) and a minimum of 44,134 (Zapalote Chico,  $n = 1$ ), with a mean of 59,799. Most  $F_{IS}$  values were negative (65%) or close to zero, ranging from  $F_{IS} = -0.928$  (Palomero Toluqueño,  $n = 1$ ) to  $0.225$  (Nal-tel. de Altura,  $n = 5$ , Table 1). These results show that there is high genetic variation among the Mexican maize landraces, and that most of the sampled landraces form out-crossing groups.



**Fig. 2.** Clustering analyses for the maize landraces from Mexico. a–b) PCA showing the first two components and the six groups found. Samples are colored by (a) landrace; and (b) morphological group according to CONABIO [10]. Color codes as in Fig. 1. c) Clusters identity is shown in the neighbor joining tree. Maize landraces are represented with circles and teosinte with other characters as in Fig. 1.

The first three components of the PCA accounted for 6.44, 2.13 (Fig. 1b) and 1.42% of the variation, respectively, and 50% of the variation was explained by 60 components. Mexican maize landraces do not form distinct clusters (Figs. 1b, c and 2). The samples form clearer groups when examining the morphological groups (Fig. 1d), but these do not correspond to the six clusters identified by the clustering analysis (Fig. 2). In fact, only the group Cónico corresponds to most samples from Cluster 2 (Fig. 2) and the rest of the clusters include a mixture of landraces and morphological groups, which shows that most landraces are not monophyletic units.

Altitude and latitude have a visible effect on the structure seen in the PCA (Figs. 3 and 4). The samples can be divided according to the low (<1000 m) or high (>2000 m) altitude range where they are grown (Fig. 3c, top), with some samples falling in an intermediate range. This variation accounts for most of the structure seen in the first component of the PCA. The admixture analysis (Fig. 4) is similar to previously published results [33] and also shows that altitude is an important variable driving genetic structuring. However, latitude also structures the variation by creating a North and South split around 19.8° N, which accounts for the gradient seen in the second component of the PCA. Noteworthy, in the admixture analysis some of the individuals of low altitudes showed assignment values of the group mostly present at high altitudes (Fig. 4, red bars), a possible explanation could be gene flow given the small distances separating maize populations despite altitudinal differences. However a larger sampling size would be necessary to test this further.

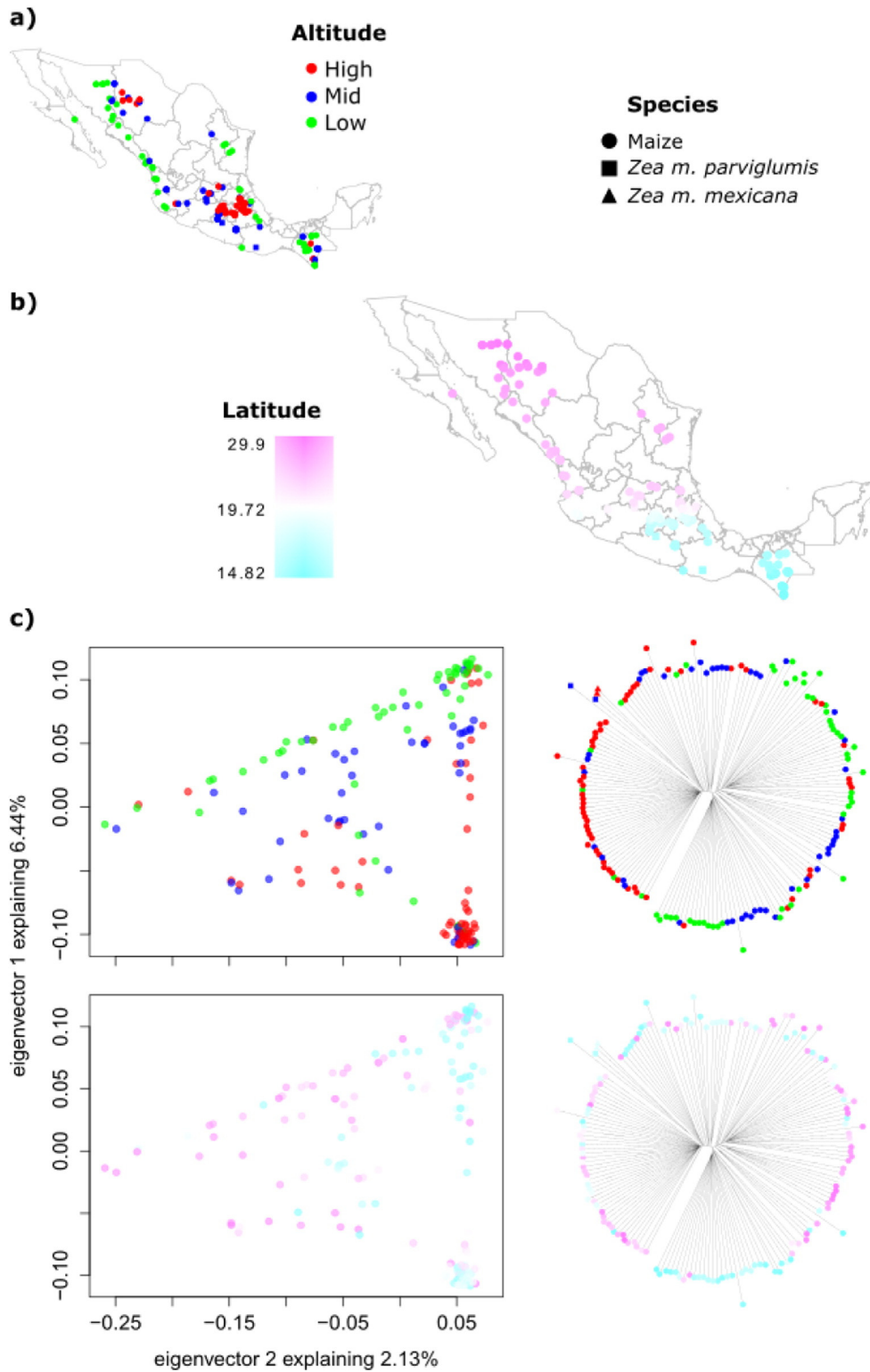
Given the long shape and complex topography of Mexico, the interaction of altitude and latitude derives in a wide range of environmental conditions where maize landraces are grown. This range of conditions is

also reflected in the PCA and the neighbor joining trees, as the ecological groups according to [29]; (Figs. 4 and 5c, top) explain the clustering of samples with more detail than altitude alone. Overlapping this information with maize biogeographic regions of ([21]; Fig. 5c, bottom) and landrace distribution (Fig. 1) shows that the biogeographical regionalization in general terms describes both geography and environmental conditions. For instance, the samples from the Central Plateau (i.e. mostly the Cónico group) are grown at high elevations of temperate climate, while the Northwest Sierras region groups several landraces of low to mid elevations (bottom clade in the neighbor joining trees) that are grown in semi-hot environments forming a northwards genetic cline (gold in Fig. 5c, bottom).

Given the rapid decay of linkage disequilibrium reported for maize, it is expected that selection at a given locus will have little effect on its neighboring loci [3,27]. Because farmers perform directional selection based on a limited number of traits [28], it is possible for landraces to keep their distinguishing morphological characteristics despite extensive gene flow (both human mediated seed exchange and natural pollen dispersal). Thus, contrasting local adaptation to different environmental conditions (Figs. 3 and 4) and farming practices (e.g. deep plating; [9]) can widely occur between and within landraces.

### 3. Conclusions

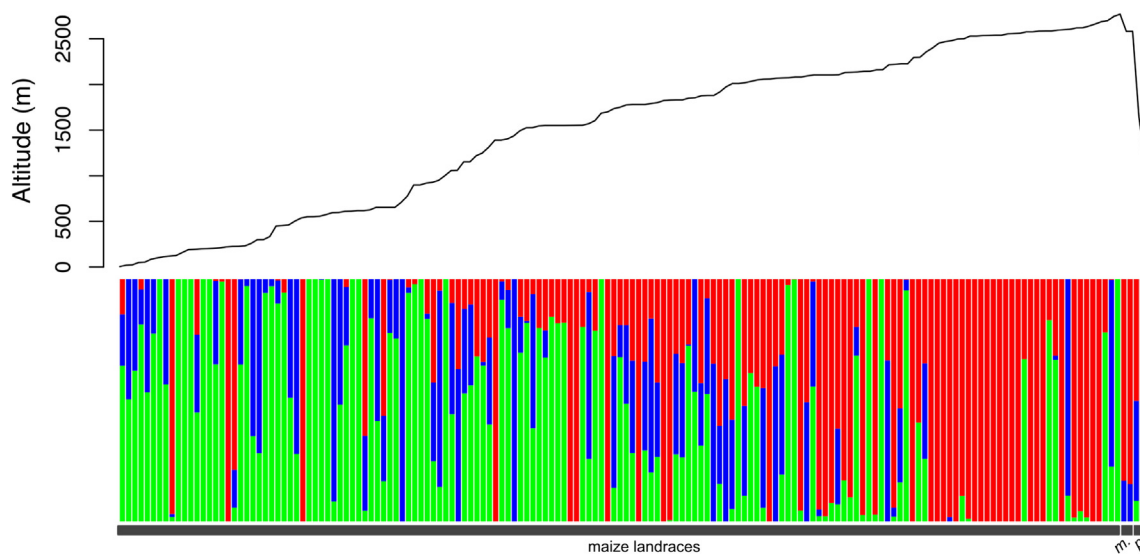
The Mexican maize landraces have a high genetic diversity that is structured mostly according to the interaction of latitude and altitude that the Mexican landscape offers. At the same time, this genetic diversity is exposed to management practices that promote not only directional selection, but also gene flow with other landraces or *Z. mays*



**Fig. 3.** Genetic variation with respect to altitude and latitude in maize landraces from Mexico. a) Altitude category of sampled accessions: high (2000 to >2750 m), mid (1000 to 2000 m) and low (<1000 m); b) latitudinal range of sampled accessions; c) PCA showing the first two components (left) and neighbor joining trees (right) colored by altitude category (top) and latitude (bottom). The shape of the symbol indicates maize or teosinte subspecies. Teosintes are not shown in the PCA.

subspecies. In this way, a maize landrace can better be considered an open and evolving genetic system rather than a fixed category [24], in which the conservation of its genetic diversity likely deepens on preserving the evolutionary processes and traditional management

practices that have generated such diversity. The present dataset adds-up to the already published material, so that such evolutionary processes could be studied further and the conservation of the Mexican landraces could be monitored with the inclusion of genetic information.



**Fig. 4.** Top: altitude at which individuals were sampled. Bottom: bar plot of assignment values of the admixture analysis. Each bar corresponds to an individual. The grey line below indicates if the individual is a maize landrace (left) or teosinte (m: *mexicana*, and p: *parviglumis*).

## 4. Materials and methods

### 4.1. Sampling

Maize samples include 46 landraces grown in Mexico (Table 1), they are represented by 161 accessions provided by the Maize Germplasm Bank of the Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP), which were collected directly from farmers mostly between 2006 and 2010 as part of the Global Maize Project. Seed collection used herein is a good representation of the geographic occurrence of maize in Mexico, and covers the entire country (Fig. 1a). Teosinte species were sampled during 2012 from wild populations or were obtained from botanical gardens collections. Metadata of samples is available at: <http://dx.doi.org/10.5061/dryad.4t20n>.

### 4.2. DNA extraction

One seed per accession was germinated in the greenhouse at the Instituto de Ecología-UNAM and leaf tips were collected from three-week old seedlings. DNA was extracted with a modified CTAB protocol [36] and purified with the Qiagen MinElute PCR purification kit. DNA quality was determined with a Nanodrop spectrophotometer. Samples with a 260/280 ratio below 1.6 were re-extracted.

### 4.3. Maize SNP50K BeadChip genotyping and quality control

Single nucleotide polymorphisms were genotyped at the Instituto Nacional de Medicina Genómica (INMEGEN) with the Illumina MaizeSNP50 BeadChip on an Infinium HD assay (Illumina, San Diego, CA, USA). Automated allele calling was implemented in GenomeStudio 2010.1 (Genotyping module 1.7.4; Illumina) after excluding loci with GenTrain score < 0.3, manually checking loci with scores 0.3–0.45 and excluding loci with more than 30% missing data. Data was exported to PLINK [25], where individuals with more than 10% of missing data were removed. Final dataset comprises 36,931 SNPs, out of which 862 were monomorphic.

### 4.4. Genetic descriptive analyses and geographic data

To visualize the distribution of genetic variation a neighbor joining tree was constructed using the *adegenet* and *ape* packages of R [14,20]. A PCA of all polymorphic SNP genotypes (i.e. not considering possible

linkage disequilibrium effects) was also performed using the package *SNPRelate* [35] and a clustering analysis was performed using *adegenet* and *ade4* [11]. The first two components were used for plotting. Only maize samples were plotted because of the known issue of unequal sampling in PCA of genetic data [6]. Samples were colored by landrace, landrace morphological group [10,30; Fig. 1], altitude (Fig. 3), latitude (Fig. 3), by landrace ecological group according to Ruiz Corral et al. [29] (Fig. 5a) and by the biogeographic regions of maize according to Perales & Golicher [21] biocultural analyses (Fig. 5b). Interactive plots were also constructed using the package *shiny* [8] and are available at CONABIO's website <https://conabio.shinyapps.io/MaicesMx50kSNP-english>.

Observed ( $H_O$ ) and expected ( $H_E$ ) heterozygosity were calculated separately for each landrace using the *adegenet* package in R [14] and corrected for small sampling size using  $H(2n/(2n-1))$  following Hedrick [13]. The inbreeding coefficient  $F_{IS}$  was calculated as  $(H_E - H_O) / H_E$  (Table 1).

### 4.5. Admixture analysis

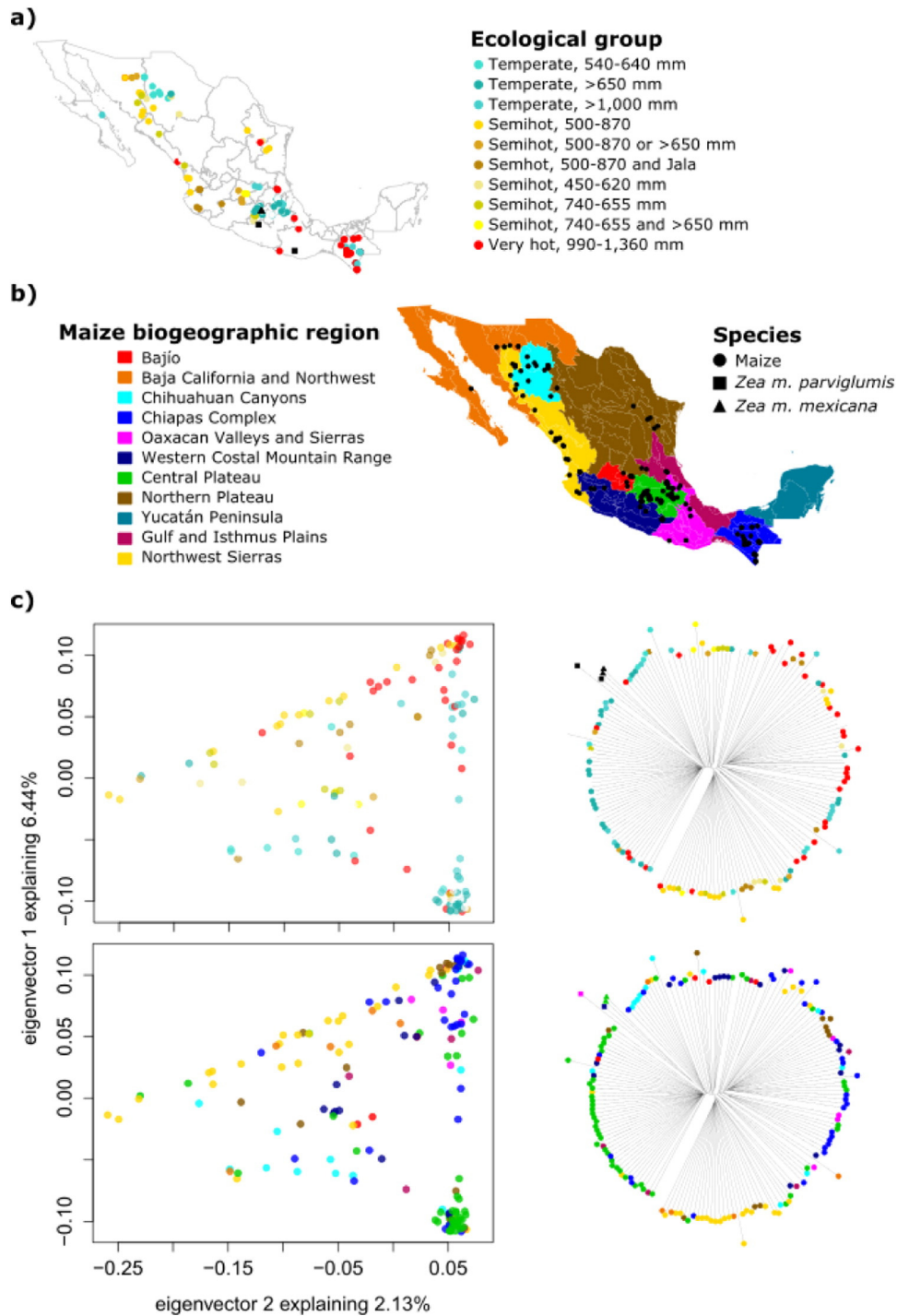
To further verify if the present dataset shows the altitude-driven structure found by previous studies [33] an admixture analysis was performed using the software ADMIXTURE v. 1.23 [2] with  $K = 3$ . This value of  $K$  was chosen to examine if our samples form the three altitudinal groups previously reported (Meso-America Lowland, West Mexico and Mexican Highlands; [33]).

### 4.6. Data availability

SNP data, maps, metadata, admixture files and scripts used here are available at the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.4t20n> (available upon acceptance). Interactive figures are available at CONABIO's website as part of the Global Native Maize Project results: <https://conabio.shinyapps.io/MaicesMx50kSNP-english>.

## Conflict of interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.



**Fig. 5.** Distribution of genetic variation in ecological and maize biogeographic regions in Mexico. a) Ecological group of sampled accessions as in Ruiz Corral et al. [29], numbers refer to mean seasonal rainfall (mm); b) Sampling over the maize biogeographic regions of Perales & Golicher [21]; and, c) PCA showing the first two components (left) and neighbor joining trees (right) colored by ecological group (top) and maize biogeographic region (bottom). Teosintes are not shown in the PCA. The ecological group for some samples is missing, those samples are not colored but their tree branches are shown.

## Acknowledgments

Funding was provided by the Dirección General del Sector Primario y Recursos Naturales Renovables (DGSPRR) of the Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT) through a grant (2009-1) to the Comisión Nacional para el Uso y Conocimiento de la Biodiversidad (CONABIO). Teosinte sampling and AML and AVL posdoc salaries were paid by CONACYT grant CB2011/167826, awarded to LEE. We thank Instituto Nacional de Investigaciones Forestales, Agrícolas y

Pecuarias (INIFAP) and Dr. Juan Manuel Hernández Casillas for seed samples. Genotyping was performed with the help of Fabiola Morales Mandujano Instituto Nacional de Medicina Genómica (INMEGEN).

## References

- [1] F. Acevedo, E. Huerta, C. Burgeff, P. Koleff, J. Sarukhán, Is transgenic maize what Mexico really needs? *Nat. Biotechnol.* 29 (2011) 23–24.
- [2] D.H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19 (2009) 1655–1664.

- [3] N. Barton, B.O. Bengtsson, The barrier to genetic exchange between hybridising populations. *Heredity* 57 (3) (1986) 357–376.
- [4] M.R. Bellon, J. Berthaud, M. Smale, et al., Participatory landrace selection for on-farm conservation: an example from the central valleys of Oaxaca, Mexico. *Genet. Resour. Crop. Evol.* 50 (2003) 401–416.
- [5] C. Burgeff, E. Huerta, F. Acevedo, J. Sarukhán, How much can GMO and non-GMO cultivars coexist in a megadiverse country? *AgBioforum* 17 (2014) 90–101.
- [6] W. Burgos-Paz, S.E. Ramos-Onsins, M. Pérez-Enciso, L. Ferretti, Correcting for unequal sampling in principal component analysis of genetic data. *Proceedings of the 10th World Congress of Genetics Applied to Livestock Production*, 2014.
- [7] T.C. Camacho-Villa, N. Maxted, M. Scholten, B. Ford-Lloyd, Defining and identifying crop landraces. *Plant Genet. Resour.* 3 (2005) 373–384.
- [8] W. Chang, J. Cheng, J.J. Allaire, et al., Shiny: Web Application Framework for R. 2015.
- [9] G.N. Collins, Pueblo Indian maize breeding varieties specially adapted to arid regions developed by Hopis and Navajos—their work not sufficiently appreciated—probably much yet to be learned from them. *J. Hered.* 5 (1914) 255–268.
- [10] CONABIO, Proyecto Global de Maíces Nativos: Recopilación, Generación, Actualización y Análisis de Información Acerca de la Diversidad Genética de Maíces y sus Parientes Silvestres en México. 2011 (México).
- [11] S. Dray, A.-B. Dufour, The ade4 package: implementing the duality diagram for ecologists. *J. Stat. Softw.* 22 (2007) 1–20.
- [12] M.M. Goodman, W.L. Brown, G. Sprague, J. Dudley, *Races of corn. Corn and Corn Improvement*, American Society of Agronomy, Madison, Wisconsin, USA 1988, pp. 33–79.
- [13] P.W. Hedrick, *Genetics of Populations*. Jones & Bartlett Learning, 2010.
- [14] T. Jombart, adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24 (2008) 1403–1405.
- [16] E. Lazos-Chavero, Consideraciones socioeconómicas y culturales en la controvertida introducción del maíz transgénico: el caso de Tlaxcala. *Sociológica* 29 (2014) 201–240.
- [17] C. Mapes Sánchez, L.M. Mera Ovando, Manejo de la diversidad. in: T.A. Kato, C. Mapes, L.M. Mera, J.A. Serratos, R.A. Bye (Eds.), *Origen y diversificación del maíz: una revisión analítica*, UNAM, CONABIO, México, D.F., 2009.
- [18] Y. Matsuoka, Y. Vigouroux, M.M. Goodman, et al., A single domestication for maize shown by multilocus microsatellite genotyping. *Proc. Natl. Acad. Sci. U. S. A.* 99 (2002) 6080–6084.
- [19] R. Ortega-Paczka, La Diversidad del Maíz en México. in: G. Esteva, C. Marielle (Eds.), *Sin maíz no hay país*, CONACULTA/Museo Nacional de Culturas Populares, México 2003, pp. 123–154.
- [20] E. Paradis, J. Claude, K. Strimmer, APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20 (2004) 289–290.
- [21] H. Perales, D. Golicher, Mapping the diversity of maize races in Mexico. *PLoS One* 9 (2014), e114657.
- [22] 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.* (2009) (pnas.0812525106).
- [23] J. Polanco, T. Flores, Bases Para una política de I&D e innovación Para la cadena de valor del maíz. *Foro Consultivo Científico y Tecnológico*, 2008 12–94.
- [24] G. Pressoir, J. Berthaud, Population structure and strong divergent selection shape phenotypic diversification in maize landraces. *Heredity* 92 (2003) 95–101.
- [25] S. Purcell, B. Neale, K. Todd-Brown, et al., PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81 (2007) 559–575.
- [26] R.G. Reeves, The use of teosinte in the improvement of corn inbreds. *Agron. J.* 42 (1950) 248–251.
- [27] D.L. Remington, J.M. Thornsberry, Y. Matsuoka, et al., Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc. Natl. Acad. Sci.* 98 (2001) 11479–11484.
- [28] E. Rice, M. Smale, J.-L. Blanco, Farmers' use of improved seed selection practices in Mexican maize: evidence and issues from the Sierra de Santa Marta. *World Dev.* 26 (1998) 1625–1640.
- [29] J.A. Ruiz Corral, N. Durán Puga, J.d.J. Sánchez González, et al., Climatic adaptation and ecological descriptors of 42 Mexican maize races. *Crop Sci.* 48 (2008) 1502.
- [30] J. Sanchez, M.M. Goodman, C.W. Stuber, Isozymatic and morphological diversity in the races of maize of Mexico. *Econ. Bot.* 54 (2000) 43–59.
- [31] SIAP, Situación actual y perspectivas del maíz en México 1996–2012. 2008.
- [32] P. Skoglund, E. Ersmark, E. Palkopoulou, L. Dalén, Ancient wolf genome reveals an early divergence of domestic dog ancestors and admixture into high-latitude breeds. *Curr. Biol.* 25 (2015) 1515–1519.
- [33] J. Van Heerwaarden, J. Doebley, W.H. Briggs, et al., Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proc. Natl. Acad. Sci.* 108 (2011) 1088–1092.
- [34] E.J. Wellhausen, Razas de maíz en México, su Origen, características y distribución. *Oficina de Estudios Especiales, Secretaría de Agricultura y Ganadería México*, 1951 (**Folleto técnico no. 5**).
- [35] X. Zheng, D. Levine, J. Shen, et al., A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28 (2012) 3326–3328.
- [36] J.J. Doyle, J.L. Doyle, Genomic plant DNA preparation from fresh tissue—the CTAB method. *Phytochemical Bulletin* 19 (1987) 11–15.