Genomics Data 7 (2016) 38–45

Contents lists available at ScienceDirect

Genomics Data

journal homepage: www.elsevier.com/locate/gdata

Genomic variation in recently collected maize landraces from Mexico

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

[a] 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
[b] Jardín Botánico, Instituto de Biología, Universidad Nacional Autónoma de México, Mexico, DF, 04510, Mexico
[c] CONACYT Research Fellow - Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, Mexico, Liga Periférico — Insurgentes Sur, No. 4903, 14010, México, DF, Mexico
[d] Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA
[e] 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 maiz* 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.

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 (M.C. Arteaga), almolet@gmail.com (A. Moreno-Letelier), amastretta@conabio.gob.mx, a.mstt.yanes@gmail.com (A. Mastretta-Yanes), ale.lobo17@gmail.com (A. Vázquez-Lobo), alesaurus_bo@yahoo.com (A. Breña-Ochoa), amoreno@langebio.cinvestav.mx (A. Moreno-Estrada), fruns@unam.mx (L.E. Eguiarte), pinero@unam.mx (D. Piñero).
[1] These authors contributed equally to this work.
[2] Current address: Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO), Centro de Investigación y de Estudios Avanzados del IPN, Irapuato, Guanajuato, Mexico.

http://dx.doi.org/10.1016/j.gdata.2015.11.002

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 Ñorteño	2	1.09	55,851	0.289	0.363	−0.259
Vandeño	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
Zea m. mexicana	2	2.69	53,687	0.262	0.306	−0.171
Zea m. parviglumis	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.