



# Recent and Historical Gene Flow in Cultivars, Landraces, and a Wild Taxon of *Cucurbita pepo* in Mexico

**Carmina Martínez-González<sup>1</sup>, Gabriela Castellanos-Morales<sup>2\*</sup>, Josué Barrera-Redondo<sup>1</sup>, Guillermo Sánchez-de la Vega<sup>1</sup>, Helena S. Hernández-Rosales<sup>1</sup>, Jaime Gasca-Pineda<sup>1</sup>, Erika Aguirre-Planter<sup>1</sup>, Alejandra Moreno-Letelier<sup>3</sup>, Ana E. Escalante<sup>4</sup>, Salvador Montes-Hernández<sup>5</sup>, Rafael Lira-Saade<sup>6\*</sup> and Luis E. Eguiarte<sup>1\*</sup>**

## OPEN ACCESS

### Edited by:

Raymond L. Tremblay,  
University of Puerto Rico, Puerto Rico

### Reviewed by:

Harry Paris,  
Agricultural Research Organization  
(ARO), Israel

Antonio Gonzalez-Rodriguez,  
Universidad Nacional Autónoma  
de México, Mexico

### \*Correspondence:

Gabriela Castellanos-Morales  
gcastellanos@ecosur.mx  
Rafael Lira-Saade  
rlira@unam.mx  
Luis E. Eguiarte  
fruns@unam.mx

### Specialty section:

This article was submitted to  
Evolutionary and Population Genetics,  
a section of the journal  
Frontiers in Ecology and Evolution

**Received:** 20 January 2021

**Accepted:** 27 April 2021

**Published:** 28 May 2021

### Citation:

Martínez-González C, Castellanos-Morales G, Barrera-Redondo J, Sánchez-de la Vega G, Hernández-Rosales HS, Gasca-Pineda J, Aguirre-Planter E, Moreno-Letelier A, Escalante AE, Montes-Hernández S, Lira-Saade R and Eguiarte LE (2021) Recent and Historical Gene Flow in Cultivars, Landraces, and a Wild Taxon of *Cucurbita pepo* in Mexico. *Front. Ecol. Evol.* 9:656051. doi: 10.3389/fevo.2021.656051

<sup>1</sup> Laboratorio de Evolución Molecular y Experimental, Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad Universitaria, Mexico City, Mexico, <sup>2</sup> Departamento de Conservación de la Biodiversidad, El Colegio de la Frontera Sur, Unidad Villahermosa, Villahermosa, Mexico, <sup>3</sup> Jardín Botánico, Instituto de Biología, Universidad Nacional Autónoma de México, Ciudad Universitaria, Mexico City, Mexico, <sup>4</sup> Laboratorio Nacional de Ciencias de la Sostenibilidad (LANCIS), Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad Universitaria, Mexico City, Mexico, <sup>5</sup> Campo Experimental Bajío, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP), Celaya, Mexico, <sup>6</sup> Laboratorio de Recursos Naturales, Unidad de Biotecnología y Prototipos (UBIPRO), Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, Tlalnequantepec, Mexico

Gene flow among crops and their wild relatives is an active study area in evolutionary biology and horticulture, because genetic exchange between them may impact their evolutionary trajectories and increase the genetic variation of the cultivated lineages. Mexico is a center of diversity for the genus *Cucurbita* that includes pumpkins, squash and gourds. Gene flow between domesticated and wild species has been reported as common in *Cucurbita*; but gene flow among populations of *C. pepo* ssp. *pepo* from Mexico and its wild relative has not been studied. We used 2,061 SNPs, derived from tunable genotyping by sequencing (tGBS) to estimate gene flow among 14 Mexican traditional landraces of *C. pepo* ssp. *pepo*, also including individuals from five improved cultivars of *C. pepo* ssp. *pepo* and *C. pepo* ssp. *ovifera* var. *ovifera*, and individuals of their wild relative *C. pepo* ssp. *fraterna*. We found moderate to high levels of genetic diversity, and low to moderate genetic differentiation. In the test of introgression between lineages, we found that all possible arrangements for ancestral and derived sites between the lineages showed similar frequencies; thus, incomplete lineage sorting, but also gene flow, might be taking place in *C. pepo*. Overall, our results suggest that gene flow between these subspecies and cultigens, incomplete lineage sorting and the retention of ancestral characters shaped the evolutionary trajectory of *C. pepo* in its area of origin and diversification. In addition, we found evidence of the use of Mexican landraces as genetic material for the improvement of commercial cultivars. The landraces of Mexico are an important source of genetic diversity for *C. pepo*, which has been preserved both by management practices of small farmers and by the natural gene flow that exists between the different crop fields of the region.

**Keywords:** domestication, GBS, gene flow, landraces, SNPs, admixture, wild relatives

## INTRODUCTION

Gene flow has an important role in the evolution of populations, as for instance even a small amount of gene flow is able to counteract the effects of genetic drift and selection (Ellstrand et al., 1999), and may also introduce useful genetic variation and increase adaptation (Aguirre-Planter, 2007). The amount of gene flow in wild plants, including crop wild relatives (CWRs), can strongly vary among species, and in a given species, among populations and seasons (Ellstrand, 1992), playing an important role in defining population structure. The levels of gene flow between domesticated plants and CWRs are also expected to be highly variable depending on space-time factors, including sexual compatibility, geographic proximity, phenology and selection regimes (Ellstrand et al., 1999; Papa, 2005).

For some domesticated species and CWRs, including landraces of beans and cultivated maize in Mesoamerica, gene flow is usually greater from crops to the wild populations (Papa and Gepts, 2004; Martínez-Castillo et al., 2007; Moreno-Letelier et al., 2020). This is probably due to the higher population sizes of domesticated populations compared to wild populations, as well as different regimes of selection in cultivated and wild environments (Papa and Gepts, 2004). In addition, as alleles of the genes associated with domestication traits are usually recessive, first generation hybrids will show a phenotype similar to that of the wild parent, so they will be much better adapted to nature than to the cultivated environment. For this reason, selection in the cultivated environment will act mainly through the first hybrid generation, while in the wild it will act mainly in the following generations, allowing recombination and thus the introgression of alleles from the domesticated population (Papa, 2005).

Wild populations are a potential source of genetic diversity for crops (Hufford et al., 2013; Sawler et al., 2013). Reports from the early eighteenth-century describe the use of wild species to transfer, through gene flow, resistance to pests and diseases to crops (Prescott-Allen and Prescott-Allen, 1986). Modern plant breeders continue to exploit the genetic diversity of CWRs to detect genes and traits for crop improvement (Egan et al., 2018).

Pumpkins and squash (*Cucurbita* spp.) are among the most important horticultural crops worldwide, providing food products (Lira, 1995; Ferriol and Picó, 2008; Kates, 2019; Wehner et al., 2020). Mexico is the center of origin and center of diversity of some species of *Cucurbita*, as well as a diversification area for pumpkins CWRs (Lira et al., 2016). Most species of wild *Cucurbita* distribute in Mexico and coexist in sympatry with their domesticated counterparts (Lira et al., 2016). One of the characteristics of the genus *Cucurbita* is that there is the possibility of reproduction between domesticated plants and their CWR (Whitaker and Bemis, 1965; Decker-Walters and Walters, 1988; Wilson, 1990). Nevertheless, in Mexico few studies have focused on analyzing gene flow between CWRs (i.e., *C. argyrosperma* ssp. *sororia* and *C. pepo* ssp. *fraterna*) and domesticated pumpkins (Wilson et al., 1994; Montes-Hernández and Eguiarte, 2002; Sánchez-de la Vega et al., 2018). These authors found evidence supporting gene flow at a local level (Wilson et al., 1994; Montes-Hernández and Eguiarte, 2002),

but low levels of gene flow at a regional scale (Sánchez-de la Vega et al., 2018). However, the evaluation of gene flow in pumpkins at the genomic level for *C. pepo* ssp. *pepo* in Mexico has not been explored.

Within cultivated pumpkins, *Cucurbita pepo* is the most widely cultivated and economically important of the genus *Cucurbita* (Paris, 2016). *Cucurbita pepo* consists of five taxa, three CWRs: *C. pepo* ssp. *fraterna*, *C. pepo* ssp. *ovifera* var. *texana*, and *C. pepo* ssp. *ovifera* var. *ozarkana*; and two cultivated: *C. pepo* ssp. *pepo* and *C. pepo* ssp. *ovifera* var. *ovifera*. Two independent domestication events have been documented for *C. pepo* (Decker, 1988; Wilson et al., 1992; Decker-Walters et al., 1993; Sanjur et al., 2002; Ferriol et al., 2003; Paris et al., 2003; Gong et al., 2012; Castellanos-Morales et al., 2018). One is presumed to have occurred 10,000 years ago in Mexico, and corresponds to *C. pepo* ssp. *pepo* (Smith, 1997; Sanjur et al., 2002) possibly from *C. pepo* ssp. *fraterna* (Andres, 1987; Castellanos-Morales et al., 2019). However, there is also genetic evidence indicating that *C. pepo* ssp. *fraterna* is more closely related to *C. pepo* ssp. *ovifera* than to *C. pepo* ssp. *pepo* (Decker-Walters et al., 1993; Gong et al., 2012; Kates et al., 2017). The second domestication event could have taken place 5,000 years ago in Southeast United States for *C. pepo* ssp. *ovifera* var. *ovifera*, likely from *C. pepo* ssp. *ovifera* var. *ozarkana* (Decker-Walters et al., 1993; Sanjur et al., 2002; Smith, 2006).

Gene flow between different crops and sympatric CWRs for *C. pepo* has been previously studied (Decker and Wilson, 1987; Kirkpatrick and Wilson, 1988; Spencer and Snow, 2001; Decker-Walters et al., 2002), but not yet in Mexico, where traditional and local landraces of *C. pepo* ssp. *pepo*—including güiches, güicoy and other pumpkins—are found (Lira et al., 1995). These landraces have evolved in diverse environments, resulting in remarkable variation among them in traditional farming communities from Mexico to Central America (Lira, 1995; Castellanos-Morales et al., 2019).

Cultivation of *Cucurbita pepo* ssp. *pepo* in Mexico is carried out in both, the traditional (milpa) system—where landrace pumpkins are grown—and the intensive agricultural system, where modern cultivars, zucchini and vegetable marrow are grown (Mera Ovando et al., 2011). Therefore, in this study we estimated the ancestral and recent gene flow between the different subspecies and cultigens of both wild and domesticated *C. pepo*, and we discuss the consequences that this may represent for wild populations and for Mexican landraces. Moreover, the presence of commercial cultivars in the vicinity of traditional Mexican landraces and wild populations is a relevant factor for gene flow, and the direction of gene flow between taxa may vary according to the effective size of the populations (Ellstrand and Elam, 1993). In home gardens, where *C. pepo* Mexican landraces are grown, they usually have few individuals, so gene flow between crops and CWRs is expected to be low (Ellstrand et al., 1999), while in a horticultural crop managed with industrial farming practices, which can contain thousands of plants, the potential for such gene flow is expected to be higher (Ellstrand et al., 1999).

## MATERIALS AND METHODS

### Sampling

We initially analyzed a total of 95 plants of different subspecies and cultigens of *C. pepo*: *C. pepo* ssp. *pepo* from landraces collected in Mexico ( $n = 70$ ), *C. pepo* ssp. *fraterna* ( $n = 8$ ), *C. pepo* ssp. *ovifera* var. *ovifera* ( $n = 5$ ) and commercial cultivars ( $n = 12$ ) (Table 1). For the Mexican landraces of *C. pepo* ssp. *pepo*, we determined 13 sampling points from the distribution documented by CONABIO project KE004 (database available at <http://www.snib.mx/iptconabio/resource?r=SNIB-KE004>), covering most of its distribution range in Mexico (Figure 1). The samples come from collections carried out during the years 2014–2015 directly from markets and farmlands in different areas of the country, as well as accessions donated by the Germplasm Bank of the Bajío Experimental Field of the INIFAP located in Celaya, Guanajuato (Mexico), including *C. pepo* ssp. *fraterna* sampled in 2011 that were collected and identified by Dr. Salvador Montes-Hernández (Table 1).

### Seed Germination and DNA Extraction

In order to have a good representation of the genetic diversity found in *C. pepo*, 15–20 seeds per locality and accession were sown. In localities where there was more than one fruit, seeds were taken from different fruits. For the samples from the Germplasm Bank, we selected those samples that came from the fruits originally collected and that have not been bred, to minimize both the risk of genetic contamination and the loss of genetic diversity. Each plant was grown in a pot with a mixture of fertile soil under controlled temperature conditions (at an average 35°C). The plants were kept in a greenhouse for 30 days, at which time a leaf tissue sample of approximately 2 cm<sup>2</sup> was collected, which was used to carry out DNA extraction.

DNA extraction was performed with the DNeasy Plant Mini Kit (Qiagen), following the supplier's instructions. We ran 1% agarose gels to evaluate the integrity of the DNA. We used nanodrop (Thermo Fisher Scientific) to determine the purity of the extractions and Qubit BR dsDNA reactions (Thermo Fisher Scientific) to estimate the DNA concentration per sample. DNA extractions that had purity values 1.8–2.0 for A260/A280 and DNA concentrations >20 ng/ml were sent to DATA2BIO<sup>1</sup> for library preparation for tunable genotyping-by-sequencing (tGBS; Ott et al., 2017). Sequencing was performed with Ion Torrent platform also by DATA2BIO (see text footnote 1), resulting in a total of 124,783,053 reads (Available in the National Center of Biotechnology Information, under the BioProject accession PRJNA485527).

### SNP Calling

We checked the PHRED quality of the sequence reads with FastQC V. 0.11.1 (Andrews, 2010). Subsequently, the data were cleaned with fastxtools V 0.0.14 (Gordon and Hannon, 2010), retaining the sequences with PHRED values > 15 in 80% of their bases. The remaining sequences were mapped using Segemehl V. 0.3.4, which is a recommended program for data sequenced

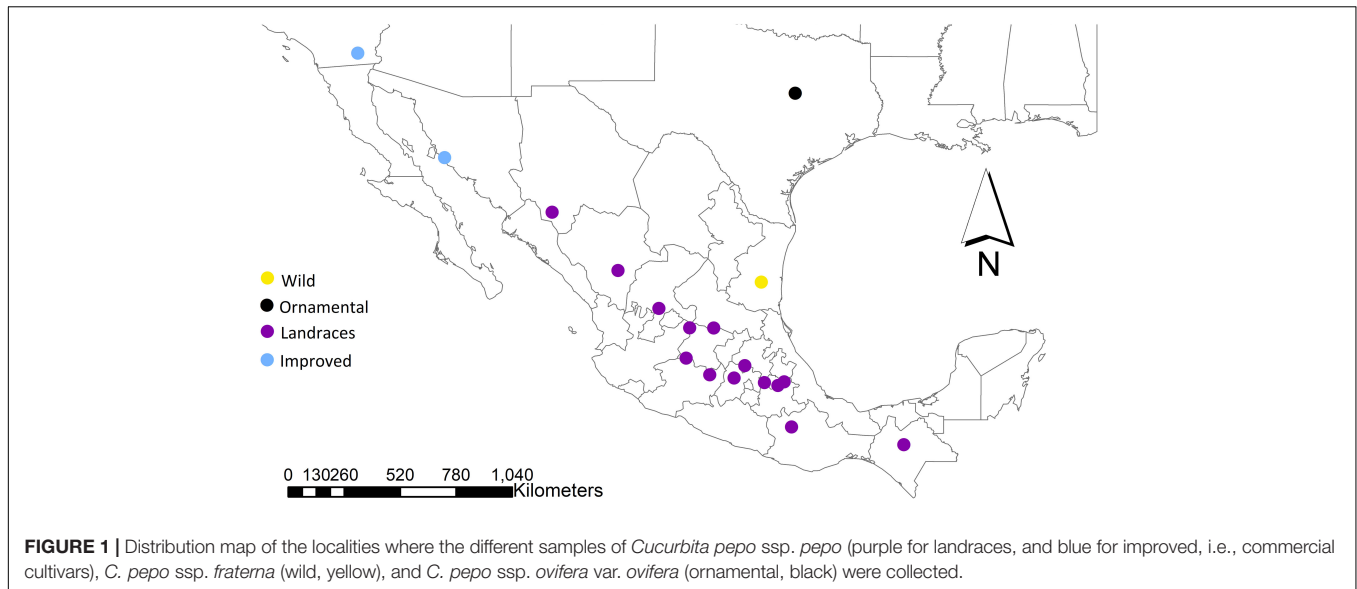
<sup>1</sup>[www.data2bio.com](http://www.data2bio.com)

**TABLE 1** | Sampling information for each subspecies with State, Origin (germplasm collection or field trip), and number of individuals grown for each locality.

Subspecies	State/Seed source	Origin	No. individuals
<i>C. pepo</i> ssp. <i>pepo</i>	Chihuahua	BG CU0879, 0880 10	5
<i>C. pepo</i> ssp. <i>pepo</i>	Durango	BG CU0654	4
<i>C. pepo</i> ssp. <i>pepo</i>	Zacatecas	Field trip Nov. 2015	3
<i>C. pepo</i> ssp. <i>pepo</i>	San Luis Potosí	Field trip Feb. 2016	2
<i>C. pepo</i> ssp. <i>pepo</i>	Jalisco	BC CU0759	3
<i>C. pepo</i> ssp. <i>pepo</i>	Guanajuato	BG CU0766	4
<i>C. pepo</i> ssp. <i>pepo</i>	Hidalgo	Field trip Nov. 2014	6
<i>C. pepo</i> ssp. <i>pepo</i>	Michoacán	BG CU0972	6
<i>C. pepo</i> ssp. <i>pepo</i>	Estado de México	BG CU1114	6
<i>C. pepo</i> ssp. <i>pepo</i>	Tlaxcala	Field trip Nov. 2014	8
<i>C. pepo</i> ssp. <i>pepo</i>	Puebla	Field trip Nov. 2014	8
<i>C. pepo</i> ssp. <i>pepo</i>	Oaxaca	BG CU1052	6
<i>C. pepo</i> ssp. <i>pepo</i>	Chiapas	Field trip Nov. 2014	8
<b><i>C. pepo</i> ssp. <i>pepo</i> improved cultivars*</b>			
<i>C. pepo</i> ssp. <i>pepo</i> Vegetable Marrow morphotype "Vegetable Spaghetti"	Sonora	Field trip Nov. 2015	4
<i>C. pepo</i> ssp. <i>pepo</i> Coccozelle morphotype "Coccozelle"	Commercial (Botanical interests Cat No. 3107; Lot #15)	Packed 2015	3
<i>C. pepo</i> ssp. <i>ovifera</i> Straightneck morphotype "Early prolific Straightneck"	Commercial (Botanical interests Cat No. 3006; Lot #1)	Packed 2015	1
<i>C. pepo</i> ssp. <i>ovifera</i> Acorn morphotype "Honey Boat"	Commercial (Botanical interests Cat No. 3076; Lot #15)	Packed 2015	2
<i>C. pepo</i> ssp. <i>pepo</i> Zucchini morphotype "Black Beauty"	Commercial (Botanical interests Cat No. 3049; Lot #1)	Packed 2015	2
<b>Wild <i>C. pepo</i></b>			
<i>C. pepo</i> ssp. <i>fraterna</i>	Tamaulipas	Fra 629, 633	8
<b>Ornamental <i>C. pepo</i></b>			
<i>C. pepo</i> ssp. <i>ovifera</i> var. <i>ovifera</i>	Ornamental mix from supermarket	Bought Oct. 2015	6

\*Morphotypes after Paris (1986).

with Ion Torrent (Caboche et al., 2014). We used the available genome assembly of the *C. pepo* ssp. *pepo* Zucchini MU-CU-16 (Montero-Pau et al., 2018) as the reference for our read mapping. Reads that had unique mapping to the reference genome were kept. We used Bcftools V. 1.9 for SNP calling, using a mapping quality and PHRED values of at least 20 for each called SNP. Subsequently, we used VCFtools V. 0.1.1 (Danecek et al., 2011) to retain the sites with a minimum sequencing depth of 5X per site per individual. The resulting SNPs were filtered with Plink V. 1.9



(Purcell et al., 2007), removing sites and individuals with more than 75% missing data, and with a minimum allele frequency (MAF)  $< 0.01$ . SNPs that did not pass the exact Hardy-Weinberg equilibrium test were eliminated with a cutoff value  $p < 0.00001$ , as well as SNPs with linkage disequilibrium greater than 0.3 that were within 1,000,000 bp and sliding windows with 100 bp jumps.

We retained a total of 2,061 SNPs, and a total of 88 individuals of the different subspecies and cultigens of *C. pepo*, which are grouped as: (1) Mexican landraces, which refer to *C. pepo* ssp. *pepo* pumpkins from landraces that were collected in Mexico ( $n = 64$ ); (2) Wild group, which consists of samples from *C. pepo* ssp. *fraterna* ( $n = 8$ ); (3) Ornamental group which is formed by samples from *C. pepo* ssp. *ovifera* var. *ovifera* ( $n = 5$ ); (4) Improved group, which consists of several commercial cultivars described in **Table 1** ( $n = 11$ ).

## Data Analysis

### Genetic Diversity

We estimated the basic measures of genetic diversity [private alleles ( $P_a$ ), observed heterozygosity ( $H_o$ ), expected heterozygosity ( $H_E$ ), and inbreeding coefficient ( $F_{IS}$ )] with the Stacks V. 2.54 population module (Catchen et al., 2013), for the subspecies and cultigens of *C. pepo*, as well as for each commercial cultivar and each locality where the landraces of *C. pepo* were collected.

As sample size for each locality was different, we performed a difference test for  $H_E$  with the R package adegenet V. 2.1.3 (Jombart, 2008; Jombart and Ahmed, 2011), with 10,000 simulations and a two-sided alternate hypothesis, to test if genetic diversity showed statistical differences among localities.

### Population Structure and Historical Gene Flow

To infer the ancestral gene flow for *C. pepo*, three different tests were carried out. The first test was a principal component analysis (PCA), calculated with Plink V. 1.9 (Purcell et al., 2007), and the data was visualized with the R package tidyverse V.

1.3.0 (Wickham et al., 2019). The second test was an analysis of individual ancestry by maximum likelihood. For this we used the Admixture V. 1.3 program (Alexander et al., 2009), where we tested  $K$ -values from 2 to 12. We performed a cross-validation test to determine the best  $K$ -value. For the analysis, a single run was performed with the values predetermined by the program, that is, a bootstrap of 200, with random seed number selection, with a criterion of stopping each run when the probability of recording each run increases by less than  $\epsilon = 10^{-4}$  between iterations. The third test was a pairwise  $F_{ST}$  analysis, which is an indirect method for estimating historical and recent gene flow (Slatkin, 1987). We used the “exact test of sample differentiation based on genotype frequencies” implemented in Arlequin V. 3.5.2.2 (Excoffier and Lischer, 2010). This estimate is based on the hypothesis of random distribution of individuals between pairs of populations as described in Raymond and Rousset (1995) and Goudet et al. (1996). We compared the differences between wild, ornamental, commercial and Mexican landraces using 100,000 MCMC iterations and a duration of 10,000 dememorization steps, with a  $P$  cutoff of 0.05.

To detect signs of introgression between the studied lineages, we obtained a second dataset. We reanalyzed the samples for 95 individuals of *C. pepo* and included 10 individuals of *C. argyrosperma* ssp. *sororia* as an outgroup. We repeated the quality filters, mapping, and SNP calling with the same parameters used before (see “SNP Calling”). We retained a total of 1,721 SNPs, and a total 91 individuals, the same 88 from *C. pepo* as before plus three samples from *C. argyrosperma* ssp. *sororia* (Barrera-Redondo et al., 2021; **Supplementary Table 1**).

To assess evidence of introgression between lineages or gene flow between *C. pepo* ssp. *fraterna* and cultigens, we conducted an ABBA-BABA test with the Dsuite V 0.4 program (Malinsky et al., 2021). We calculated the D-statistic along with the admixture fraction  $f$  with the above dataset (Durand et al., 2011). This method analyzes the correlation between allele frequencies across populations to conduct a formal test for the history of admixture

between populations, and it has been proved robust under most demographic scenarios (Malinsky et al., 2021). The test uses a four-taxon tree {Outgroup [P1, (P2, P3)]}, including an outgroup (O) that defines the ancestral and derived alleles (defined as “A” and “B,” respectively) in biallelic SNPs. The site patterns are ordered, so a “BBAA” pattern indicates that lineages P1 and P2 share the derived allele (B), while an “ABBA” pattern suggests that P2 and P3 share the derived allele, and a “BABA” pattern reflects that P1 and P3 lineages share the derived allele (Malinsky et al., 2021). To assess which pattern is most likely, the D-statistic estimates the number of times each of these patterns (“BBAA,” “ABBA” and “BABA”) conflict with the tree. If these patterns occur in equal frequencies, the statistic has a value of  $D = 0$ ,—we accept the null hypothesis of no introgression—indicating that incomplete lineage sorting is the prevailing process. Conversely, if introgression is occurring between lineages there will be an excess of either “ABBA” or “BABA” patterns, resulting in  $D > 0$  if the introgression is between P2 and P3 or  $D < 0$  if introgression occurred between P1 and P3 (Durand et al., 2011; Malinsky et al., 2021). We analyzed all possible arrangements between lineages (Mexican landraces, wild, improved and ornamental), using *C. argyrosperma* ssp. *sororia* as outgroup, and the D-statistic significance was calculated by jackknifing blocks of 20 bp.

### Recent Gene Flow

Current migration among *C. pepo* accessions was evaluated with BayesAss V. 3.0.4 (BA3) (Mussmann et al., 2019). This Bayesian method is based on Markov chain Monte Carlo (MCMC) to estimate the posterior probabilities of recent migration rates, i.e., in the last two generations (Wilson and Rannala, 2003). Parameter values were estimated after 10 million burn-in steps to allow stabilizing or convergence out of 100 million iterations sampling every 10,000 iterations. To compute suitable allele frequency ( $a$ ), inbreeding coefficient ( $f$ ) and migration rate ( $m$ ), we tested several values until we obtained the acceptance percentages recommended by BayesAss authors (between 20 and 60%). We examined the convergence for each run using Tracer v 1.6 and we determined the final values as:  $a = 0.30$  (62% accepted),  $f = 0.08$  (37% accepted), and  $m = 0.20$  (48% accepted).

## RESULTS

### Genetic Diversity

We analyzed a total of 2,061 SNPs for 88 individuals. The Mexican landraces (*C. pepo* ssp. *pepo*;  $H_E = 0.196$ ) showed the highest  $H_E$  value. While the lowest  $H_E$  value was exhibited by the ornamental mix (*C. pepo* ssp. *ovifera* var. *ovifera*;  $H_E = 0.176$ ) (Table 2). In Mexican landraces, we found highly variable values of  $H_E$ , ranging from  $H_E = 0.065$  for the state of San Luis Potosí, to  $H_E = 0.183$  for the state of Puebla (Supplementary Table 2). Nevertheless, we did not find significant differences in  $H_E$  values between taxonomic groups (Supplementary Table 3), and among populations (Supplementary Table 4), probably due to the small sample size, so these data must be taken with care.

The landraces showed the highest inbreeding coefficients ( $F_{IS}$ ), with an average  $F_{IS} = 0.116$  and the highest number of private

alleles ( $P_a = 263$ ), while the lowest inbreeding coefficients were found in an ornamental mix ( $F_{IS} = 0.037$ ) and the wild subspecies (*C. pepo* ssp. *fraterna*;  $F_{IS} = 0.055$ ) (Table 3). However, as in  $H_E$ , the inbreeding coefficient shows highly variable ranges that go from  $F_{IS} = -0.022$  in Tlaxcala, to  $F_{IS} = 0.126$  in Michoacán (Supplementary Table 2).

### Population Structure and Historical Gene Flow

In the principal component analysis (PCA), we observed three clusters (Figure 2). The first cluster is found at the center of the quadrants and contains almost all wild individuals (*C. pepo* ssp. *fraterna*). The second cluster is located in the upper right quadrant and is formed by only individuals belonging to landraces from Mexico. The last cluster is located in the lower right quadrant and consists of individuals of the four different groups of *C. pepo*, although individuals are more widely dispersed, especially landraces.

Pairwise  $F_{ST}$  values among the three subspecies of *C. pepo* and their improved cultivars were not statistically significant (Table 3). This suggests low genetic differentiation among *C. pepo* crops and CWRs. The lowest values of genetic differentiation were found among Mexican landraces and the wild subspecies ( $F_{ST} = 0.0064$ ), as well as among the Mexican landraces and the ornamental mix ( $F_{ST} = 0.0067$ ) (Table 3).

According to the Admixture analysis, the most likely values of  $K$  were 2 ( $CV = 0.386$ ) and 4 ( $CV = 0.387$ ) (Figure 3). For  $K = 2$ , the most common cluster (Figure 3A) is found in 82.7% of the samples, and is represented by purple. For  $K = 4$  (Figure 3B), the most common cluster is found in 64.8% of the samples (purple in Figure 3B); followed by the black cluster and the yellow cluster with 15.59 and 12.06% of the samples, respectively. The cluster with the least presence is the blue one, with only 7.53% of the

**TABLE 2** | Values of genetic diversity for the three subspecies of *C. pepo* and their cultigens (Improved = *C. pepo* ssp. *pepo*, Landraces = Mexican *C. pepo* ssp. *pepo*, Wild = *C. pepo* ssp. *fraterna*, and Ornamental = *C. pepo* ssp. *ovifera*).

Population	N	$P_a$	$H_O$	$H_E$	$F_{IS}$
Improved	11	6	0.173	0.189	0.068*
Landraces	64	263	0.169	0.196	0.116*
Wild	8	2	0.178	0.185	0.055*
Ornamental	5	1	0.180	0.176	0.037*

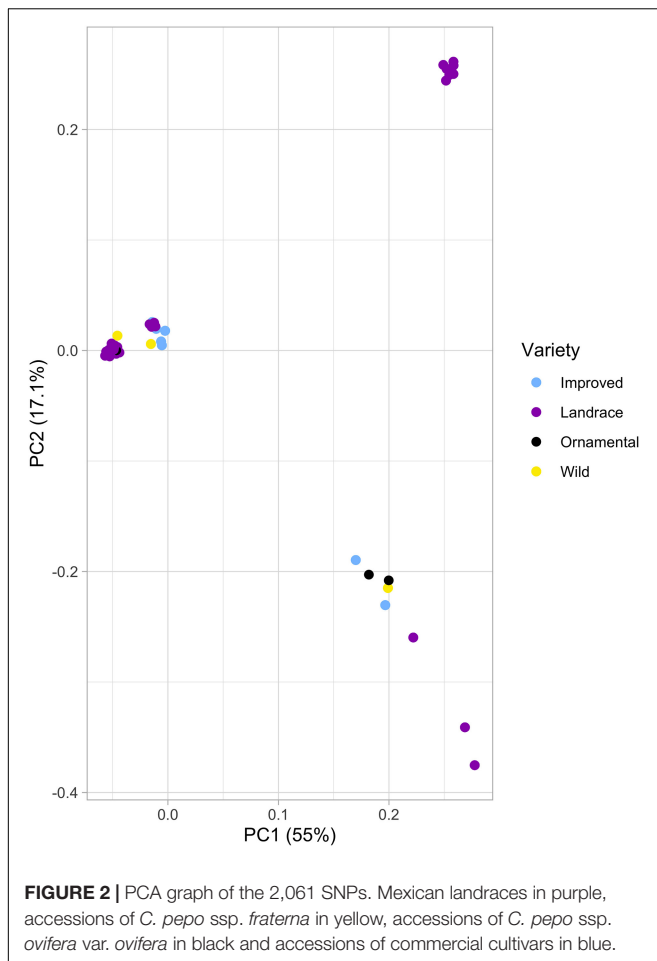
\* $P < 0.05$ .

$N$ , number of individuals;  $P_a$ , number of private alleles;  $H_O$ , observed heterozygosity;  $H_E$ , expected heterozygosity;  $F_{IS}$ , inbreeding coefficient.

**TABLE 3** | Pairwise  $F_{ST}$  values for the three subspecies and cultigens of *C. pepo*.

	Improved	Landraces	Wild	Ornamental
Improved	0.0000			
Landraces	0.0203*	0.0000		
Wild	0.0179	0.0064	0.0000	
Ornamental	0.0076	0.0067	0.0085	0.0000

\* $P < 0.05$ .



samples assigned to this genetic group. For  $K = 4$ , we see that the Mexican landraces are mainly assigned to the purple cluster, as well as individuals of the wild subspecies (i.e., CWR). Most of the individuals from commercial cultivars analyzed show a combination of genetic clusters, and there is only one individual of the ornamental mix with a small component of the yellow cluster. The blue cluster is represented mainly by the improved or commercial cultivars and the Mexican landrace accession of San Luis Potosí, as well as by a very low proportion in the rest of *C. pepo* (Figure 3B).

For the ABBA–BABA test, the counts of the sites “AABB,” “ABBA” and “BABA” are similar for all possible arrangements between the lineages, causing the D-statistics values to be low. In other words, the analysis did not show evidence of hybridization among the Mexican landraces and any other lineage (Table 4). Nevertheless, it is worth noticing that two of the tests were significant ( $P < 0.05$ ) with values of  $D > 0$ , with higher counts for “AABB” sites followed by “ABBA” sites (Table 4).

## Recent Gene Flow

BayesAss analyses (Figure 4) suggests that migration occurs mainly from the Mexican landraces of *C. pepo* ssp. *pepo* to the rest of the analyzed groups, and the wild subspecies *C. pepo* ssp. *fraterna* appears as the main receiver. In contrast, *C. pepo* ssp.

*fraterna* was the source population with the lowest fraction of migrants (ranging from 0.5 to 3.75%).

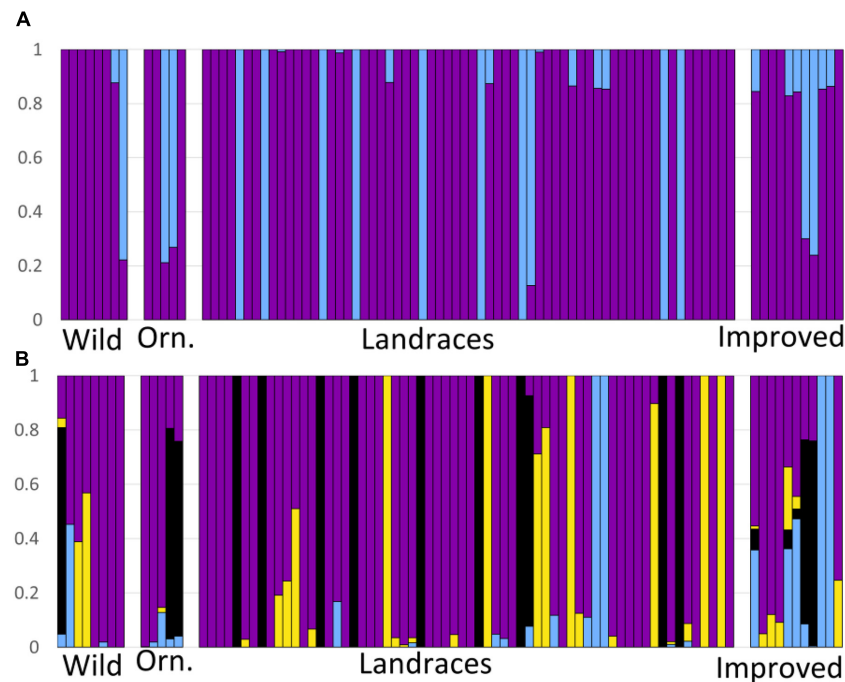
## DISCUSSION

### Genetic Diversity

In general, domesticated taxa are expected to show lower genetic diversity, when compared to CWRs, because domestication generally represents a bottleneck where domesticates are usually descendants of a few wild individuals (Doebley, 1989; Gepts and Papa, 2003; Gross and Olsen, 2010; Meyer and Purugganan, 2013; Gepts, 2014; Gaut et al., 2018). Contrary to this pattern, domesticated *C. pepo* ssp. *pepo* from Mexico ( $H_E = 0.196$ ) showed slightly higher genetic diversity than its wild relative *C. pepo* ssp. *fraterna* ( $H_E = 0.185$ ) (Table 2), but these differences were not significant (Supplementary Table 4). This pattern has been previously reported for nuclear microsatellite loci and chloroplast sequences in *C. pepo* by Castellanos-Morales et al. (2019) and may be related to a smaller effective population size in *C. pepo* ssp. *fraterna*, given its restricted distribution in the Mexican states of Tamaulipas and Nuevo León; as well as of possible genetic bottlenecks that might have occurred in these populations during the Pleistocene climatic oscillations (Castellanos-Morales et al., 2018, 2019). In contrast, the Mexican landraces of *C. pepo* ssp. *pepo* may have experienced a rapid demographic expansion after domestication (Kistler et al., 2015; Castellanos-Morales et al., 2019).

The genetic diversity in *C. pepo* is higher than that reported for *C. argyrosperma*, also estimated with SNPs (Barrera-Redondo et al., 2021). The latter may relate to *C. pepo* being the pumpkin species with higher commercial importance and higher morphological diversity (Paris, 2001). Accordingly, we expect cultivated *C. pepo* to show high effective population sizes, and high dispersal through anthropogenic movements, which is further supported by a lack of genetic structure, in contrast to *C. argyrosperma* and *C. moschata* in Mexico (Sánchez-de la Vega et al., 2018; Castellanos-Morales et al., 2019; Hernández-Rosales et al., 2020; Barrera-Redondo et al., 2021). In our analyses, the ornamental mix (*C. pepo* ssp. *ovifera* var. *ovifera*) shows the lowest value of genetic diversity ( $H_E = 0.176$ ). These cultivars have been subjected to intense agricultural management, which causes strong bottlenecks and very intense selective pressures (Formisano et al., 2010). However our results must be interpreted with caution, given that our sample size is small, and the analysis includes only a few representatives of this subspecies and none of its wild populations. We recommend in the future to analyze the wild populations of *C. pepo* ssp. *ovifera* var. *texana* and *C. pepo* ssp. *ovifera* var. *ozarkana* with large sample sizes and genome-wide data to compare the levels of genetic diversity between all *C. pepo* CWRs.

The admixture and PCA analyses indicate that the San Luis Potosi and “Vegetable Spaghetti” samples show genetic similarity (Figures 2, 3). This, together with the low genetic diversity found in San Luis Potosi, may suggest that the analyzed individuals from San Luis Potosi share ancestry with “Vegetable Spaghetti” commercial cultivar (blue genetic cluster in Figure 3B). In



**FIGURE 3 |** Admixture analysis for **(A)**  $K = 2$  and **(B)**  $K = 4$ , showing the assignment probability of individuals from *C. pepo* ssp. *fraterna* (Wild), *C. pepo* ssp. *ovifera* (Orn.), and *C. pepo* ssp. *pepo* Mexican landraces (Landraces) and improved cultivars (Improved). Each color represents a genetic cluster.

**TABLE 4 |** Results of ABBA-BABA test to detect introgression using 1,721 SNPs between all possible arrangements between lineages, considering *C. argyrosperma* ssp. *sororia* as outgroup.

P1	P2	P3	D-statistic	Z-score	P	f4-ratio	AABB	ABBA	BABA
Landrace	Improved	Ornamental	0.037	2.079	0.019	0.000	42.970	41.377	38.429
Improved	Wild	Landrace	0.019	0.928	0.177	0.000	43.688	41.768	40.220
Wild	Improved	Ornamental	0.001	0.056	0.478	0.000	44.753	39.692	39.614
Landrace	Wild	Ornamental	0.036	1.700	0.045	0.000	44.524	41.305	38.435

The numbers in "AABB" "ABBA" "BABA" indicate the number of times each pattern for the shared derived allele occurred in the analysis, standardized by the allele frequencies.

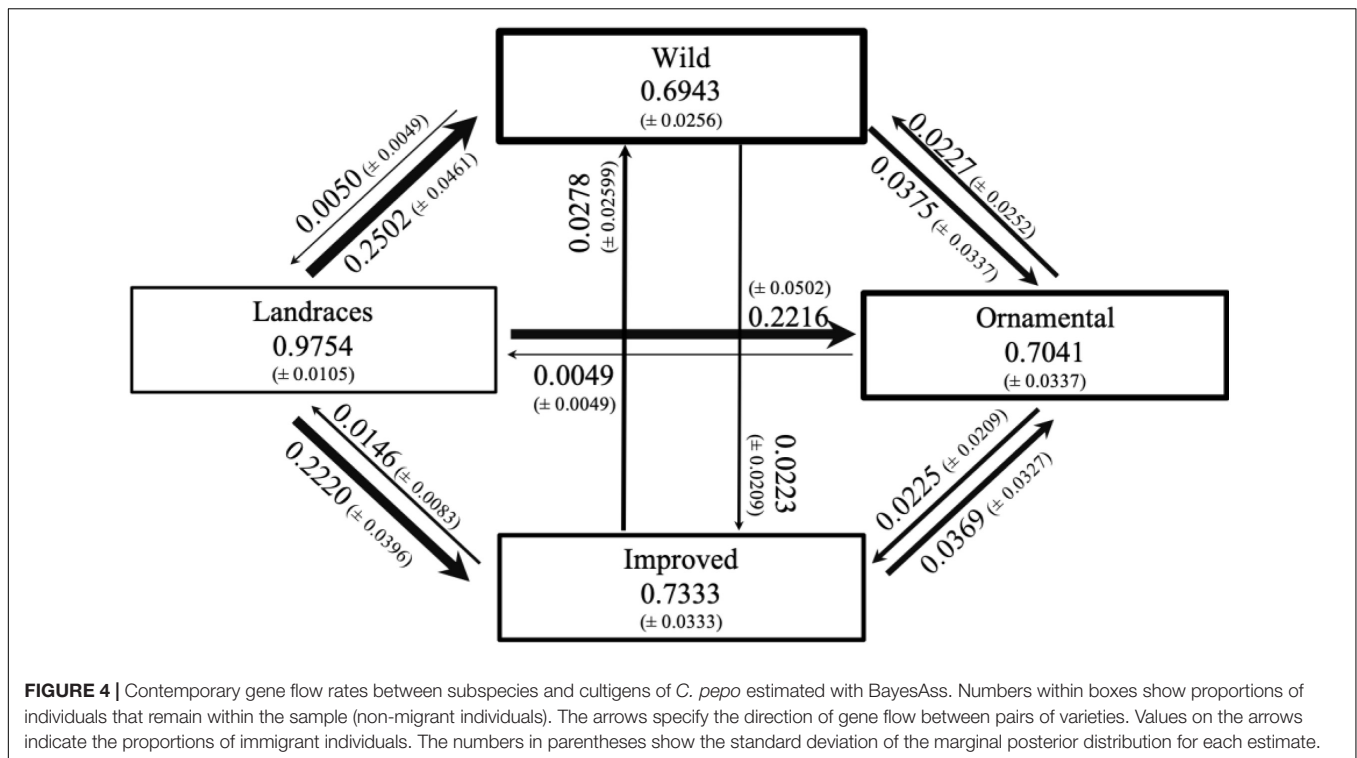
certain parts of Mexico, the traditional cultivation of milpa is still practiced, which helps maintain the genetic diversity of populations (Mastretta-Yanes et al., 2018). However, in other regions of the country –such as San Luis Potosi, which is part of the Bajío region, an important agricultural area where many crops are grown for commercial purposes– there has been a replacement by modern agricultural systems, where seeds of improved cultivars are used (Mera Ovando et al., 2011; Sánchez-de la Vega, 2017). It has been suggested that the genetic diversity stored in landraces allows crops to adapt to harsher environments and makes cultivars less susceptible to pests and diseases (Mastretta-Yanes et al., 2018). Traditional agricultural systems should be preserved to maintain the genetic diversity of *Cucurbita* crops (Montes-Hernández et al., 2005), and future analyses that explicitly test the effects of different management practices - milpa in contrast to intensive agriculture, on the genetic diversity of pumpkins need to be conducted.

We also found high inbreeding values, especially in Mexican landraces, which can be related to the use of less than 10

fruits from each harvest for next year's harvest, and also to the fact that the exchange of seeds between different farmers is uncommon (but see Montes-Hernández and Eguiarte, 2002; Montes-Hernández et al., 2005). Inbreeding can help establish favorable phenotypic characters in domesticated species, and in some cases farmers safeguard, promote, select and control the reproductive processes of a few individuals (Casas and Parra, 2016). A rapid fixation of the desired characters through artificial selection in the domestication and improvement processes can be acquired by a certain degree of inbreeding (Ellstrand and Elam, 1993). However, we cannot rule out the presence of Wahlund effect in our estimates of inbreeding, since our samples come from many fields across Mexico.

## Genetic Structure and Historical Gene Flow

The genetic differentiation among subspecies and cultigens of *C. pepo* estimated with SNPs was very low and not significant,



except between Mexican landraces and improved cultivars ( $F_{ST} = 0.0203$ ) (Table 3). This result may relate to (1) possible historical or recent gene flow, (2) incomplete lineage sorting, or (3) retention of ancestral characters.

The facility for spontaneous hybridization in *C. pepo* has been well documented since the 18th century (Paris, 2000, Paris, 2001). In particular, gene flow between wild (*C. pepo* ssp. *fraterna*) and *C. pepo* cultigens was not expected, because of their current allopatric distributions. Nevertheless, species distributions may change through time, as a response to climate and other environmental changes. In this sense, species distribution models suggest that *C. pepo* ssp. *fraterna* could have had a wider range during the mid-Holocene (Kistler et al., 2015; Castellanos-Morales et al., 2019), reaching areas near the Tehuacán and Ocampo Caves, where archeological remains have been found (Whitaker et al., 1957; Hanselka and King, 2017). Therefore, past gene flow between *C. pepo* ssp. *fraterna* and ancestral Mexican landraces cannot be ruled out.

Incomplete lineage sorting in *C. pepo* may also result in low levels of genetic differentiation and has been cited as an explanation for the differences in the relationship between taxa found in phylogenetic analyses based on different molecular markers and methods of analyses (Sanjur et al., 2002; Kistler et al., 2015; Kates et al., 2017; Castellanos-Morales et al., 2018, 2019). Accordingly, previous genetic studies, and linguistic evidence (Brown et al., 2012), have suggested a rapid human mediated dissemination after domestication or geographically widespread evolution of domesticated forms, as well increases and decreases in the effective population size of *C. pepo* ssp. *fraterna* throughout its history, and the increase in the effective population size

of Mexican local varieties of *C. pepo* ssp. *pepo* (Zizumbo and Colunga, 2008; Kistler et al., 2015; Castellanos-Morales et al., 2019). All of the above may have favored incomplete lineage sorting.

Retention of ancestral alleles is an equally likely explanation for low levels of genetic differentiation between *C. pepo* groups found in this study. Phylogenetic analyses, dates of divergence estimated based on molecular clock and the archeological records suggest that the divergence of these taxa is very recent (Whitaker et al., 1957; Smith, 1997, 2006; Kistler et al., 2015; Hanselka and King, 2017; Castellanos-Morales et al., 2018, 2019). Also, Xanthopoulou et al. (2019) reported moderate linkage disequilibrium in *C. pepo*, and a high proportion of shared SNPs between subspecies. All of the above suggest that not enough time has passed for the differential fixation of alleles in neutral nuclear markers in these taxa. Moreover, pumpkins of Mexico and Guatemala are morphologically the most similar to wild gourds, reflecting a possible earlier stage of domestication, which is further observed in the persistence of the round shape of the fruit (Paris, 2001). Deeper genomic analysis will be necessary to analyze the retention of ancestral alleles in this species.

The admixture analysis (Figure 3B) showed that the improved cultivars have the highest assignment probability to the blue clusters, as well as the presence of crossed individuals with the other genotypes. The latter may occur because different cultigens of *C. pepo* have been crossed to breed improved cultivars (Ferriol and Picó, 2008), as well as for the introduction of genes from other pumpkin species for genetic improvement (Whitaker, 1962; Whitaker and Davis, 1962; Paris et al., 1988; Holdsworth et al., 2016; Castellanos-Morales et al., 2019), especially to produce



disease-resistant crops (Robinson and Shail, 1987; Cohen et al., 2003; Formisano et al., 2010; Paris, 2018).

We did not find well-defined groups based on morphology. These results contrast with previous analyses in *C. pepo*, which found a correlation between genetic clusters and *C. pepo* morphotypes (Paris et al., 2003; Gong et al., 2012). These studies focused mainly on commercial cultivars, which have smaller effective population sizes, and are subject to strong artificial selection. As mentioned above, in Mexico the most common morphotype is the Pumpkin type, with morphological characteristics more similar to the ancient groups: large and elongated seeds, lignified shell, round shape, and marked keels (Lira et al., 1995; Paris, 2001). In fact, previous studies with maize landraces from Mexico have reported that the morphological structure of landraces has no relationship with the genetic structure (Arteaga et al., 2016; Moreno-Letelier et al., 2020), and people select phenotypes that they recognize as their local crop of interest, regardless of the genotype that these crops may have (Louette et al., 1997). In consequence, the lack of genetic differentiation within Mexican landraces of *C. pepo* might relate to cultural preference toward our local landraces. As a matter of fact, in Mexican markets or supermarkets it is common to find mature round pumpkins (in autumn) or immature zucchini (throughout the year); cultivar such as “Vegetable Spaghetti” are considered a rarity and grown mostly for export.

In the ABBA-BABA analysis, all possible arrangements between the lineages were found in a similar number of times (Table 4), causing the values of the D-statistics to be close to zero. This means that there are no strong signs of introgression between lineages in *C. pepo*, and incomplete lineage sorting may explain the patterns of genetic structure we have found in Mexico. Nevertheless, two of the values were significant with  $D > 0$ , suggesting possible introgression between improved and ornamental mix and between wild and ornamental mix (Table 4). These results further suggest that gene flow and incomplete lineage sorting may explain the distribution of genetic diversity in *C. pepo*.

For the “pipiana” squash (*C. argyrosperma* ssp. *argyrosperma*) and its wild relative (*C. argyrosperma* ssp. *sororia*), Barrera-Redondo et al. (2021) found that many derived sites are shared between taxa—that is, they found an “AABB” pattern—. The latter means that lineages in *C. argyrosperma* are well differentiated (monophyletic). Also, for *C. argyrosperma* there was genomic evidence of extant gene flow after initial domestication (Barrera-Redondo et al., 2021). The observed differences between *C. pepo* and *C. argyrosperma* may relate to differences in their domestication history—reticulate domestication followed by rapid expansion in *C. pepo* vs. single domestication followed by gene flow in *C. argyrosperma* (Andres, 1987; Castellanos-Morales et al., 2019; Barrera-Redondo et al., 2021)—. Also, trade and management patterns differ between these species, as *C. argyrosperma* ssp. *argyrosperma* is traded mainly at a local level and there are well defined genetic clusters in this taxon (Sánchez-de la Vega et al., 2018; Barrera-Redondo et al., 2021), while trade for *C. pepo* ssp. *pepo* occurs at a larger scale (Lira, 1995; Zizumbo and Colunga,

2008; Mera Ovando et al., 2011). Larger effective population size in *C. pepo* and gene flow between different geographic regions, may be causing the patterns of derived sites vs. ancestral sites to blur in this taxon. However, as previously mentioned, deeper genomic analyses that include a wider sample of the wild subspecies should be conducted in *C. pepo*.

## Recent Gene Flow

The BayesAss analysis suggests that gene flow is more common from landraces to CWRs and to commercial cultivars of pumpkins (Figure 4). As previously mentioned, these signals of high levels of gene flow could relate to the confusing effect of gene flow, incomplete lineage sorting, and retention of ancestral characters.

Our study focuses on pumpkin landraces from Mexico, which may be found in sympatry with other wild *Cucurbita* species in different areas of its distribution. *Cucurbita* flowers are pollinated by bees, which can spread the different alleles between nearby cultivated plots, maintaining or increasing the genetic diversity of the populations. In fact, it has been documented that *Peponapis* and *Xenoglossa* bees (specialized pumpkin pollinators), can promote the movement of pollen at distances of at least 700 m (Kohn and Casper, 1992), while the common bees (*Apis mellifera*), as well as the small Lepidoptera and Syrphidae, promote gene flow at distances up to 1,000 m (Ellstrand and Marshall, 1985; Ellstrand et al., 1999). Furthermore, there is a report in the United States that gene flow between *C. pepo* ssp. *pepo* and *C. pepo* ssp. *ovifera* var. *texana* pollinated by *Xenoglossa angustior* and *Peponapis pruinosa* can reach distances up to 1,300 m (Kirkpatrick and Wilson, 1988). However, although gene flow can help maintain their genetic diversity, each generation of pumpkin is still subject to inbreeding and genetic drift associated with the selection by farmers of a few parental individuals for production the following year (Montes-Hernández et al., 2005).

We estimated significant gene flow from landraces to improved and ornamental cultivars of *C. pepo* (Figure 4). This may be indicating that landraces of Mexico are still being used as a base or parental group for the generation of improved cultivars, which is consistent with our admixture result (Figure 3). Nevertheless, this may also relate to incomplete lineage sorting, the retention of ancestral characters and a hybrid origin of some improved cultivars (Kates et al., 2017; Castellanos-Morales et al., 2018, 2019).

Estimated gene flow of the wild subspecies *C. pepo* ssp. *fraterna* to landraces and commercial cultivars was low (Figure 4). This is to be expected, because the area where *C. pepo* ssp. *pepo* grows in Mexico has little or no overlap with the restricted distribution of its wild relative (Castellanos-Morales et al., 2018; Kates, 2019). However, we can observe that this pattern is very different in the opposite sense, that is, that *C. pepo* ssp. *fraterna* seems to be receiving a lot of gene flow from other crops, mainly landraces, even when they do not grow in sympatry. Nonetheless, *C. moschata* and *C. argyrosperma* ssp. *argyrosperma* grow in sympatry with *C. pepo* ssp. *fraterna* in the locality of Vado el

Moro, from where our samples were collected, and as previously reported by Wilson et al. (1994), interspecific gene flow may be taking place. The latter is something relevant to consider, because gene flow can lead to the loss of genetic identity and genetic diversity of the wild species, as well as to the detriment of its populations (Ellstrand et al., 1999; Campbell et al., 2016). Future studies should focus in estimating interspecific gene flow between *Cucurbita* crops and their wild relatives where they coexist.

In this study, we found that the Mexican landraces appear to be a genetic reservoir for the management and development of the improved and ornamental cultivars, more than the wild populations, which seem to be preserved by both, the management techniques of small farmers and gene flow between different milpas in each area. Therefore, we strongly recommend implementing government programs that promote the conservation of the milpa farming system. We also recommend monitoring introgression from modern cultivars in Mexico, to avoid the loss of identity and genetic diversity of local landraces, as well as important alleles for the adaptation of crops to inhospitable environments.

## DATA AVAILABILITY STATEMENT

The SRA records is in the NCBI database. SRA records will be accessible with the following link after the indicated release date: <https://www.ncbi.nlm.nih.gov/sra/PRJNA485527>.

## AUTHOR CONTRIBUTIONS

LE, RL-S, GC-M, and CM-G designed the research. GS-V, HH-R, and GC-M collected the biological material and extracted DNA for the tGBS data. SM-H donated material from the germplasm bank. GC-M and GS-V coordinated the sequencing procedures of the tGBS data. EA-P coordinated the administrative and lab work. JB-R, CM-G, GC-M, HH-R, JG-P, and AM-L performed the bioinformatic analyses. CM-G, GC-M, and JB-R analyzed the results. CM-G drafted the manuscript with the help of LE, GC-M, EA-P, and RL-S. RL-S and LE obtained the funding. All authors corrected and approved the final version of the manuscript.

## FUNDING

This work was funded by the CONABIO KE004 “Diversidad genética de las especies de *Cucurbita* en México e hibridación

entre plantas genéticamente modificadas y especies silvestres de *Cucurbita*”, by CONABIO PE001 “Diversidad genética de las especies de *Cucurbita* en México. Fase II. Genómica evolutiva y de poblaciones, recursos genéticos y domesticación”, by CONACYT Investigación Científica Básica 2011.167826 (clave de identificación oficial CB2011/167826) “Genómica de poblaciones: estudios en el maíz silvestre, el teosinte (*Zea mays* ssp. *parviglumis* y *Zea mays* ssp. *mexicana*)”, and by CONACYT Problemas Nacionales through the grant no. 247730, awarded to Daniel Piñero. CM-G was supported by the grant no. 742706 of the CONACYT for her Ph.D. studies.

## ACKNOWLEDGMENTS

This manuscript is presented in partial fulfillment for the requirements to obtain a Ph.D. degree by CM-G in the Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México. We thank the Posgrado en Ciencias Biológicas at the Universidad Nacional Autónoma de México (UNAM) and the Consejo Nacional de Ciencia y Tecnología (CONACYT, Mexico) for supporting Ph.D. studies and a grant for CM-G. We thank the Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México, CONABIO (projects KE004 and PE001), CONACYT (grants CB2011/167826, 24773, 742706) and Daniel Piñero for their financial support to conduct this study. We also thank the Laboratorio de Evolución Molecular y Experimental, the Instituto de Ecología, and Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, and Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Campo Experimental Bajío. We thank K. Y. Ruiz-Mondragón, G. J. Soria-Ortiz, Y. Gutiérrez-Guerrero, P. Hernández, L. M. Paredes-Torres, D. C. Hernández-Rosales, T. E. Legaspi, G. M. Rosas, S. Barrientos, E. Scheinvar, V. Souza, and L. Espinosa-Asuar for their help in fieldwork, laboratory work, logistics and/or help in conducting the analyses. We thank two reviewers for their comments which helped us improve the manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fevo.2021.656051/full#supplementary-material>

## REFERENCES

- Aguirre-Planter, E. (2007). “Flujo génico: métodos para estimarlo y marcadores moleculares,” in *Ecología Molecular*, eds L. E. Eguiarte, V. Souza, and X. Aguirre (Mexico: Semarnat-Ine-UNAM-Conabio), 40–61.
- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi: 10.1101/gr.094052.109
- Andres, T. C. (1987). *Cucurbita fraterna*, the closest wild relative and progenitor of *C. pepo*. *Cucurbit Genet. Coop. Rep.* 10, 69–71.
- Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data*.
- Arteaga, M. C., Moreno-Letelier, A., Mastretta-Yanes, A., Vázquez-Lobo, A., Breña-Ochoa, A., Moreno-Estrada, A., et al. (2016). Genomic variation in recently collected maize landraces from Mexico. *Genomics Data* 7, 38–45. doi: 10.1016/j.gdata.2015.11.002
- Barrera-Redondo, J., Sánchez-de la Vega, G., Aguirre-Liguori, J. A., Castellanos-Morales, G., Gutiérrez-Guerrero, Y. T., Aguirre-Dugua, X., et al. (2021). The domestication of *Cucurbita argyrosperma* as revealed by the genome of its wild relative. *Hortic. Res.* in press, 8:109. doi: 10.1038/s41438-021-00544-9

- Brown, C. H., Luedeling, E., Wichmann, S., and Epps, P. (2012). "The paleobiolinguistics of domesticated squash (*Cucurbita* spp.)," in *Explorations in Ethnobiology: The Legacy of Amadeo Rea*, eds M. Quinland and D. Lepofsky (Denton, TX: Society of Ethnobiology, University of North Texas), 132–161.
- Caboche, S., Audebert, C., Lemoine, Y., and Hot, D. (2014). Comparison of mapping algorithms used in high-throughput sequencing: application to Ion Torrent data. *BMC Genomics* 15:264. doi: 10.1186/1471-2164-15-264
- Campbell, L. G., Lee, D., Shukla, K., Waite, T. A., and Bartsch, D. (2016). An ecological approach to measuring the evolutionary consequences of gene flow from crops to wild or weedy relatives. *Appl. Plant Sci.* 4:1500114. doi: 10.3732/appls.1500114
- Casas, A., and Parra, F. (2016). "La domesticación como proceso evolutivo," in *Domesticación en el Continente Americano*, eds A. Casas, J. Torres-Guevara, and F. Parra-Rondinel (Mexico City: Universidad Nacional Autónoma de México), 133–158.
- Castellanos-Morales, G., Paredes-Torres, L. M., Gámez, N., Hernández-Rosales, H. S., Sánchez-de la Vega, G., Barrera-Redondo, J., et al. (2018). Historical biogeography and phylogeny of *Cucurbita*: insights from ancestral area reconstruction and niche evolution. *Mol. Phylogenet. Evol.* 128, 38–54. doi: 10.1016/j.ympev.2018.07.016
- Castellanos-Morales, G., Ruiz-Mondragón, K. Y., Hernández-Rosales, H. S., Sánchez-de la Vega, G., Gámez, N., Aguirre-Planter, E., et al. (2019). Tracing back the origin of pumpkins (*Cucurbita pepo* ssp. *pepo* L.) in Mexico. *Proc. R. Soc. B Biol. Sci.* 286:20191440. doi: 10.1098/rspb.2019.1440
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., and Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22, 3124–3140. doi: 10.1111/mec.12354
- Cohen, R., Hanan, A., and Paris, H. (2003). Single-gene resistance to powdery mildew in zucchini squash (*Cucurbita pepo*). *Euphytica* 130, 433–441. doi: 10.1023/A:1023082612420
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158.
- Decker, D. S. (1988). Origin (s), evolution, and systematics of *Cucurbita pepo* (Cucurbitaceae). *Econ. Bot.* 42, 4–15.
- Decker, D. S., and Wilson, H. D. (1987). Allozyme variation in the *Cucurbita pepo* complex: *C. pepo* var. *ovifera* vs. *C. texana*. *Syst. Bot.* 21, 263–273. doi: 10.2307/2419320
- Decker-Walters, D. S., and Walters, T. W. (1988). Genealogy and gene flow among annual domesticated species of *Cucurbita*. *Can. J. Bot.* 68, 782–789. doi: 10.1139/b90-104
- Decker-Walters, D. S., Walters, T. W., Cowan, C. W., and Smith, B. D. (1993). Isozymic characterization of wild populations of *Cucurbita pepo*. *J. Ethnobiol.* 13, 55–74.
- Decker-Walters, S. D., Staub, J. E., Chung, S.-M., Naata, E., and Quemada, H. D. (2002). Diversity in free-living populations of *Cucurbita pepo* (Cucurbitaceae) as assessed by random amplified polymorphic DNA. *Syst. Bot.* 27, 19–28. doi: 10.1043/0363-6445-27.1.19
- Doebly, J. (1989). "Isozymic evidence and the evolution of crop plants," in *Isozymes in Plant Biology*, eds D. E. Soltis, P. S. Soltis, and T. R. Dudley (Dordrecht: Springer), 165–191. doi: 10.1007/978-94-009-1840-5\_9
- Durand, E. Y., Patterson, N., Reich, D., and Slatkin, M. (2011). Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* 28, 2239–2252. doi: 10.1093/molbev/msr048
- Egan, P. A., Muola, A., and Stenberg, J. A. (2018). Capturing genetic variation in crop wild relatives: an evolutionary approach. *Evol. Appl.* 11, 1293–1304. doi: 10.1111/eva.12626
- Ellstrand, N. C. (1992). Gene flow by pollen: implications for plant conservation genetics. *Oikos* 63, 77–86. doi: 10.2307/3545517
- Ellstrand, N. C., and Elam, D. R. (1993). Population genetic consequences of small population size: implications for plant conservation. *Annu. Rev. Ecol. Syst.* 24, 217–242. doi: 10.1146/annurev.es.24.110193.001245
- Ellstrand, N. C., and Marshall, D. L. (1985). Interpopulation gene flow by pollen in wild radish, *Raphanus sativus*. *Am. Nat.* 126, 606–616. doi: 10.1086/284442
- Ellstrand, N. C., Prentice, H. C., and Hancock, J. F. (1999). Gene flow and introgression from domesticated plants into their wild relatives. *Annu. Rev. Ecol. Syst.* 30, 539–563. doi: 10.1146/annurev.ecolsys.30.1.539
- Excoffier, L., and Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567.
- Ferriol, M., and Picó, B. (2008). "Pumpkin and winter squash," in *Vegetables I. Handbook of Plant Breeding*, Vol. 1, eds F. B. Whitfield, and J. H. Last (New York, NY: Springer), 317–349. doi: 10.1007/978-0-387-30443-4\_10
- Ferriol, M., Picó, B., and Nuez, F. (2003). Genetic diversity of a germplasm collection of *Cucurbita pepo* using SRAP and AFLP markers. *Theor. Appl. Genet.* 107, 271–282. doi: 10.1007/s00122-003-1242-z
- Formisano, G., Paris, H., Frusciante, L., and Ercolano, M. R. (2010). Commercial *Cucurbita pepo* squash hybrids carrying disease resistance introgressed from *Cucurbita moschata* have high genetic similarity. *Plant Genet. Resour.* 8, 198–203. doi: 10.1017/S1479262110000183
- Gaut, B. S., Seymour, D. K., Liu, Q., and Zhou, Y. (2018). Demography and its effects on genomic variation in crop domestication. *Nat. Plants* 4, 512–520. doi: 10.1038/s41477-018-0210-1
- Gepts, P. (2014). The contribution of genetic and genomic approaches to plant domestication studies. *Curr. Opin. Plant Biol.* 18, 51–59. doi: 10.1016/j.pbi.2014.02.001
- Gepts, P., and Papa, R. (2003). "Evolution during domestication," in *Encyclopedia of Life Sciences*, ed. Nature Publishing Group (New York, NY: Macmillan Publishers), 1–7. doi: 10.1038/ngp.els.0003071
- Gong, L., Paris, H., Nee, M. H., Stift, G., Pachner, M., Vollmann, J., et al. (2012). Genetic relationships and evolution in *Cucurbita pepo* (pumpkin, squash, gourd) as revealed by simple sequence repeat polymorphisms. *Theor. Appl. Genet.* 124, 875–891. doi: 10.1007/s00122-011-1752-z
- Gordon, A., and Hannon, G. (2010). *Fastx-toolkit*. FASTQ/A Short-Reads Pre-Processing Tools. Unpublished. Available online at: [http://hannonlab.cshl.edu/fastx\\_tool](http://hannonlab.cshl.edu/fastx_tool) (accessed March, 2019).
- Goudet, J., Raymond, M., de Meeüs, T., and Rousset, F. (1996). Testing differentiation in diploid populations. *Genetics* 144, 1933–1940.
- Gross, B. L., and Olsen, K. M. (2010). Genetic perspectives on crop domestication. *Trends Plant Sci.* 15, 529–537. doi: 10.1016/j.tplants.2010.05.008
- Hanselka, J. K., and King, B. (2017). Modeling agricultural potential near the Ocampo Caves, Tamaulipas: integrating archaeological and geospatial applications. *J. Ethnobiol.* 37, 260–281.
- Hernández-Rosales, H. S., Castellanos-Morales, G., Sánchez-de la Vega, G., Aguirre-Planter, E., Montes-Hernández, S., Lira-Saade, R., et al. (2020). Phylogeographic and population genetics analyses of *Cucurbita moschata* reveal divergence of two mitochondrial lineages linked to an elevational gradient. *Am. J. Bot.* 107, 510–525. doi: 10.1002/ajb2.1424
- Holdsworth, W. L., LaPlant, K. E., Bell, D. C., Jahn, M. M., and Mazourek, M. (2016). Cultivar-based introgression mapping reveals wild species-derived Pm-0, the major powdery mildew resistance locus in squash. *PLoS One* 11:e0167715. doi: 10.1371/journal.pone.0167715
- Hufford, M. B., Lubinsky, P., Pyhäjärvi, T., Devengenzo, M. T., Ellstrand, N. C., and Ross-Ibarra, J. (2013). The genomic signature of crop-wild introgression in maize. *PLoS Genet.* 9:e1003477. doi: 10.1371/journal.pgen.1003477
- Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129
- Jombart, T., and Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27, 3070–3071. doi: 10.1093/bioinformatics/btr521
- Kates, H. R. (2019). "Pumpkins, squashes, and gourds (*Cucurbita* L.) of North America," in *North American Crop Wild Relatives*, eds S. Greene, K. Williams, C. Khoury, M. Kantar, and L. Marek (Cham: Springer), 2. doi: 10.1007/978-3-319-97121-6\_6
- Kates, H. R., Soltis, P. S., and Soltis, D. E. (2017). Evolutionary and domestication history of *Cucurbita* (pumpkin and squash) species inferred from 44 nuclear loci. *Mol. Phylogenet. Evol.* 111, 98–109. doi: 10.1016/j.ympev.2017.03.002
- Kirkpatrick, K. J., and Wilson, H. D. (1988). Interspecific gene flow in *Cucurbita*: *C. texana* vs. *C. pepo*. *Am. J. Bot.* 75, 519–527. doi: 10.1002/j.1537-2197.1988.tb13470.x
- Kistler, L., Newsom, L. A., Ryan, T. M., Clarke, A. C., Smith, B. D., and Perry, G. H. (2015). Gourds and squashes (*Cucurbita* spp.) adapted to megafaunal extinction and ecological anachronism through domestication. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15107–15112. doi: 10.1073/pnas.1516109112

- Kohn, J. R., and Casper, B. B. (1992). Pollen-mediated gene flow in *Cucurbita foetidissima* (Cucurbitaceae). *Am. J. Bot.* 79, 57–62. doi: 10.2307/2445197
- Lira, R. (1995). *Estudios Taxonómicos y Ecogeográficos de las Cucurbitaceae Latinoamericanas de Importancia Económica*. Roma: International Plant Genetic Resources Institute.
- Lira, R., Andres, T., and Nee, M. (1995). “Cucurbita,” in *Estudios Taxonómicos y Ecogeográficos de las Cucurbitaceae Latinoamericanas de Importancia Económica: Cucurbita, Sechium, Sicana y Cyclanthera Systematic and Ecogeographic Studies on Crop Gene Pools*, Vol. 9, ed. R. Lira (Rome: International Plant Genetic Resources Institute), 1–115.
- Lira, R., Eguiarte, L., Montes, S., Zizumbo-Villarreal, D., Marín, P. C.-G., and Quesada, M. (2016). “*Homo sapiens*–*Cucurbita* interaction in Mesoamerica: domestication dissemination, and diversification,” in *Ethnobotany of Mexico*, eds R. Lira, A. Casas, and J. Blancas (New York, NY: Springer), 389–401. doi: 10.1007/978-1-4614-6669-7\_15
- Louette, D., Charrier, A., and Berthaud, J. (1997). In situ conservation of maize in Mexico: genetic diversity and maize seed management in a traditional community. *Econ. Bot.* 51, 20–38. doi: 10.1007/BF02910401
- Malinsky, M., Matschiner, M., and Svardal, H. (2021). Dsuite – Fast D-statistics and related admixture evidence from VCF files. *Mol. Biol. Evol.* 21, 584–595. doi: 10.1111/1755-0998.13265
- Martínez-Castillo, J., Zizumbo-Villarreal, D., Gepts, P., and Colunga-GarcíaMarín, P. (2007). Gene flow and genetic structure in the wild–weedy–domesticated complex of *Phaseolus lunatus* L. in its Mesoamerican center of domestication and diversity. *Crop Sci.* 47, 58–66. doi: 10.2135/cropsci2006.04.0241
- Mastretta-Yanes, A., Acevedo Gasman, F., Burgeff, C., Cano Ramírez, M., Piñero, D., and Sarukhán, J. (2018). An initiative for the study and use of genetic diversity of domesticated plants and their wild relatives. *Front. Plant Sci.* 9:209. doi: 10.3389/fpls.2018.00209
- Mera Ovando, L. M., Bye Boettler, R. A., Villanueva Verdusco, C., and Luna, M. A. (2011). *Documento de Diagnóstico de las Especies Cultivadas de Cucurbita L.* México: UNAM.
- Meyer, R. S., and Purugganan, M. D. (2013). Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* 14, 840–852. doi: 10.1038/nrg3605
- Montero-Pau, J., Blanca, J., Bombarely, A., Ziarsolo, P., Esteras, C., Martí-Gómez, C., et al. (2018). De novo assembly of the zucchini genome reveals a whole-genome duplication associated with the origin of the *Cucurbita* genus. *Plant Biotechnol. J.* 16, 1161–1171. doi: 10.1111/pbi.12860
- Montes-Hernández, S., and Eguiarte, L. E. (2002). Genetic structure and indirect estimates of gene flow in three taxa of *Cucurbita* (Cucurbitaceae) in western Mexico. *Am. J. Bot.* 89, 1156–1163. doi: 10.3732/ajb.89.7.1156
- Montes-Hernández, S., Merrick, L. C., and Eguiarte, L. E. (2005). Maintenance of squash (*Cucurbita* spp.) landrace diversity by farmers’ activities in Mexico. *Genet. Resour. Crop Evol.* 52, 697–707. doi: 10.1007/s10722-003-6018-4
- Moreno-Letelier, A., Aguirre-Liguori, J. A., Piñero, D., Vázquez-Lobo, A., and Eguiarte, L. E. (2020). The relevance of gene flow with wild relatives in understanding the domestication process. *R. Soc. Open Sci.* 7:191545. doi: 10.1098/rsos.191545
- Mussmann, S. M., Douglas, M. R., Chafin, T. K., and Douglas, M. E. (2019). BA3-SNPs: contemporary migration reconfigured in BayesAss for next-generation sequence data. *Methods Ecol. Evol.* 10, 1808–1813. doi: 10.1111/2041-210X.13252
- Ott, A., Liu, S., Schnable, J. C., Yeh, C. T., Wang, K. S., and Schnable, P. S. (2017). tGBS® genotyping-by-sequencing enables reliable genotyping of heterozygous loci. *Nucleic Acids Res.* 45:e178. doi: 10.1093/nar/gkx853
- Papa, R. (2005). “Gene flow and introgression between domesticated crops and their wild relatives,” in *Proceedings of the International Workshop on the Role of Biotechnology for the Characterization and Conservation of Crop, Forestry, Animal and Fishery Genetic Resources 5–7 March, 2005*, Turin, 5–7.
- Papa, R., and Gepts, P. (2004). “Asymmetric gene flow and introgression between domesticated and wild populations,” in *Introgression from Genetically Modified Plants into Wild Relatives*, eds C. M. Hans, D. den Nijs, D. Bartsch, and J. Sweet (Wallingford: CABI), 125–138.
- Paris, H., Yonash, N., Portnoy, V., Mozes-Daube, N., Tzuri, G., and Katzir, N. (2003). Assessment of genetic relationships in *Cucurbita pepo* (Cucurbitaceae) using DNA markers. *Theor. Appl. Genet.* 106, 971–978. doi: 10.1007/s00122-002-1157-0
- Paris, H. S. (1986). A proposed subspecific classification for *Cucurbita pepo*. *Phytologia* 61, 133–138.
- Paris, H. S. (2000). Paintings (1769–1774) by A. N. Duchesne and the history of *Cucurbita pepo*. *Ann. Bot.* 85, 815–830. doi: 10.1006/anbo.2000.1147
- Paris, H. S. (2001). History of the cultivar-groups of *Cucurbita pepo*. *Hortic. Rev.* 25, 71–170.
- Paris, H. S. (2016). “Genetic resources of pumpkins and squash *Cucurbita* spp.” in *Genetics and Genomics of Cucurbitaceae*, eds R. Grumet, N. Katzir, and J. Garcia-Mas (Cham: Springer International Publishing), 111–154. doi: 10.1007/978-94-007-3216-3
- Paris, H. S. (2018). Consumer-oriented exploitation and conservation of genetic resources of pumpkins and squash, *Cucurbita*. *Isr. J. Plant Sci.* 65, 202–221. doi: 10.1163/22238980-00001036
- Paris, H. S., Cohen, S., Burger, Y., and Yoseph, R. (1988). Single-gene resistance to zucchini yellow mosaic virus in *Cucurbita moschata*. *Euphytica* 37, 27–29. doi: 10.1007/BF00037219
- Prescott-Allen, C., and Prescott-Allen, R. (1986). *The First Resource: Wild Species in the North American Economy*. New Haven, CT: Yale University, 529.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- Raymond, M., and Rousset, F. (1995). GENEPOP: a population genetics software for exact tests and ecumenicism, version 1.2. *J. Hered.* 86, 248–249.
- Robinson, R. W., and Shail, J. W. (1987). Genetic variability for compatibility of an interspecific cross. *Cucurbit Genet. Coop. Rep.* 10, 88–89.
- Sánchez-de la Vega, G. (2017). De la cueva a la mesa, y ahora al laboratorio genómico: la diversidad de calabazas de México. *Oikos* 17, 11–17.
- Sánchez-de la Vega, G., Castellanos-Morales, G., Gámez, N., Hernández-Rosales, H. S., Vázquez-Lobo, A., Aguirre-Planter, E., et al. (2018). Genetic resources in the “Calabaza Pipiana” squash (*Cucurbita argyrosperma*) in Mexico: genetic diversity, genetic differentiation and distribution models. *Front. Plant Sci.* 9:400. doi: 10.3389/fpls.2018.00400
- Sanjur, O. I., Piperno, D. R., Andres, T. C., and Wessel-Beaver, L. (2002). Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from a mitochondrial gene: implications for crop plant evolution and areas of origin. *Proc. Natl. Acad. Sci. U.S.A.* 99, 535–540. doi: 10.1073/pnas.012577299
- Sawler, J., Reisch, B., Aradhya, M., Prins, B., and Zhong, G. (2013). Genomics assisted ancestry deconvolution in grape. *PLoS One* 8:e80791. doi: 10.1371/journal.pone.0080791
- Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science* 236, 787–792. doi: 10.1126/science.3576198
- Smith, B. D. (1997). The initial domestication of *Cucurbita pepo* in the Americas 10,000 years ago. *Science* 276, 932–934. doi: 10.1126/science.276.5314.932
- Smith, B. D. (2006). Eastern North America as an independent center of plant domestication. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12223–12228. doi: 10.1073/pnas.0604335103
- Spencer, L. J., and Snow, A. A. (2001). Fecundity of transgenic wild–crop hybrids of *Cucurbita pepo* (Cucurbitaceae): implications for crop-to-wild gene flow. *Heredity* 86, 694–702. doi: 10.1046/j.1365-2540.2001.00890.x
- Wehner, T. C., Naegel, R. P., Myers, J. R., Narinder, P. S., and Crosby, K. (2020). *Cucurbits*, 2nd Edn. Wallingford: CABI.
- Whitaker, T. W. (1962). An interspecific cross in *Cucurbita*: *C. lundelliana* bailey × *C. maxima* duchesne. *Euphytica* 11, 273–281. doi: 10.1007/BF00031979
- Whitaker, T. W., and Bemis, W. P. (1965). Evolution in the genus *Cucurbita*. *Evolution* 18, 553–559.
- Whitaker, T. W., Cutler, H. C., and MacNeish, R. S. (1957). Cucurbit materials from three caves near Ocampo, Tamaulipas. *Am. Antiq.* 22, 352–358. doi: 10.2307/276133
- Whitaker, T. W., and Davis, G. N. (1962). *Cucurbits — Botany, Cultivation and Utilization*. New York, NY: Interscience Publishers Inc.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., et al. (2019). Welcome to the tidyverse. *J. Open Source Softw.* 4, 1686. doi: 10.21105/joss.01686
- Wilson, G. A., and Rannala, B. (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163, 1177–1191.

- Wilson, H. D. (1990). Gene flow in squash species. *BioScience* 40, 449–455. doi: 10.2307/1311392
- Wilson, H. D., Doebley, J., and Duvall, M. (1992). Chloroplast DNA diversity among wild and cultivated members of *Cucurbita* (Cucurbitaceae). *Theor. Appl. Genet.* 84, 859–865. doi: 10.1007/BF00227397
- Wilson, H. D., Lira, R., and Rodríguez, I. (1994). Crop/weed gene flow: *Cucurbita argyrosperma* Huber and *C. fraterna* L.H. Bailey (Cucurbitaceae). *Econ. Bot.* 48, 293–300. doi: 10.1007/BF02862330
- Xanthopoulou, A., Montero-Pau, J., Mellidou, I., Kissoudis, C., Blanca, J., Picó, B., et al. (2019). Whole-genome resequencing of *Cucurbita pepo* morphotypes to discover genomic variants associated with morphology and horticulturally valuable traits. *Hortic. Res.* 6:94. doi: 10.1038/s41438-019-0176-9
- Zizumbo, D., and Colunga, P. (2008). El origen de la agricultura, la domesticación de plantas y el establecimiento de corredores biológico-culturales en Mesoamérica. *Rev. Geogr. Agrícola* 41, 85–113.
- Conflict of Interest:** The 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.
- The reviewer AG-R declared a shared affiliation though no other collaboration with the authors to the handling editor.

Copyright © 2021 Martínez-González, Castellanos-Morales, Barrera-Redondo, Sánchez-de la Vega, Hernández-Rosales, Gasca-Pineda, Aguirre-Planter, Moreno-Letelier, Escalante, Montes-Hernández, Lira-Saade and Eguiarte. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.