Received: 12 February 2022 | Accepted: 13 June 2022

DOI: 10.1002/agj2.21153

Agronomy Journal

ARTICLE

Agronomic Application of Genetic Resources

Genetic diversity of subtropical double-haploid maize lines selected for high oil content

Grethel P. Gaytán-Pinzón[1,†] | **Eduardo Sandoval-Castro**[1,†] | **Luis A. Peinado-Fuentes**[2] | **Juan P. Valenzuela-Apodaca**[1] | **Abraham Cruz-Mendívil**[3] | **Carlos L. Calderón-Vázquez**[1]

[1]Centro Interdisciplinario de Investigación para el Desarrollo Integral Regional (CIIDIR) Unidad Sinaloa Instituto Politécnico Nacional, Departamento de Biotecnología agrícola. Guasave, Sinaloa, México

[2]Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP), Campo Experimental Valle del Fuerte, Guasave, Sinaloa, México

[3]CONACyT-Instituto Politécnico Nacional, CIIDIR Unidad Sinaloa, Laboratorio de Genómica Funcional, Guasave, Sinaloa, México

Correspondence
Carlos L. Calderón-Vázquez, Centro Interdisciplinario de Investigación para el Desarrollo Integral Regional (CIIDIR) Unidad Sinaloa, Instituto Politécnico Nacional, Departamento de Biotecnología agrícola. Guasave, Sinaloa, México.
Email: ccalderon@ipn.mx

[†]These authors contributed equally to this work.

Assigned to Associate Editor Paulo Teodoro.

Funding information
Instituto Nacional de Investigaciones Forestales, Agrtícolas y Pecuarias , Grant/Award Number: 10561934518; Instituto de Apoyo a la Investigación e Innovación Sinaloa, Grant/Award Number: PIVISE 017/2017

Abstract

The inclusion of high-oil maize germplasm into breeding programs may be an excellent alternative for increasing grain nutritional quality. Knowing the germplasm genetic diversity is crucial for assisting breeding programs. Here, the genetic diversity and population structure of four high-oil-content maize (*Zea mays* L.) populations (Bajio yellow population [BYP], northwestern yellow population [NYP], Bajio white population [BWP], and northwestern white population [NWP]) were analyzed by Diversity Arrays Technology sequencing. Three-hundred ten double-haploid (DH) lines were genotyped, and 19,078 single-nucleotide polymorphism (SNP) markers were uniformly detected among populations after filtering by missing data >20% and minor allele frequency ≥0.05. Genetic diversity indexes showed polymorphic information content (PIC) values of 0.346, 0.352, 0.353, and 0.353; observed heterozygosity values of 0.221, 0.194, 0.284, and 0.177; and expected heterozygosity values of 0.188, 0.165, 0.219, and 0.152, for BYP, NYP, BWP, and NWP, respectively. Genetic structure results showed variations in pairwise genetic distance comparisons among the 310 DH lines, ranging from 0.119 to 0.385. Multidimensional scaling analysis and discriminant analysis of principal components grouped the DH lines into three different and five clusters, respectively, based on their origin region and grain color. On the other hand, STRUCTURE analysis revealed the presence of two different groups unrelated to grain color or origin region. The wide genetic variability among the analyzed DH lines highlights their potential to contribute new beneficial alleles into subtropical maize breeding programs and will facilitate the selection of parental lines and the identification of heterotic groups to generate high-oil maize hybrids.

Abbreviations: BWP, Bajio white population; BYP, Bajio yellow population; CML, CIMMYT maize inbred line; DAPC, discriminant analysis of principal components; DArT, Diversity Arrays Technology; DArT-seq, DArT sequencing; DH, double haploid; H_e, expected heterozygosity; H_o, observed heterozygosity; HOC, high oil content; MDS, multidimensional scaling; NWP, northwestern white population; NYP, northwestern yellow population; PCR, polymerase chain reaction; PIC, polymorphic information content; UPGMA, unweighted pair group method with arithmetic mean.

1 | INTRODUCTION

Maize (*Zea mays* L.) is the most important cereal grain in terms of worldwide production (García-Lara & Serna-Saldívar, 2019) and it is cultivated in most of temperate and subtropical regions of the planet. Maize production has high nutritional value and economic importance worldwide not only as source of human food but also as animal feeding and as a feedstock for industrial products (Food & Agricultural Organization, 2019). The demand for maize in the developing world is expected to double by 2050 (Singh et al., 2020), thus breeding programs have mainly focused on increasing maize production. However, to combat global malnutrition, maize breeding programs must include nutritional quality aspects (Ortiz-Islas et al., 2019).

Maize with high oil content (HOC) is considered specialty maize and has a grain oil concentration >6%, while in commercial maize hybrids grain oil concentration is ~4.5% on a dry weight basis (Yang & Li, 2018). High-oil-content maize hybrids possess characteristics that might have a great effect on the agricultural, livestock, and industrial sectors in Mexico because of their high energetic value. High-oil-content maize varieties have been bred for temperate environments using conventional recurrent selection (Lambert et al., 2004) and the doubled-haploid (DH) technology (Battistelli et al., 2013). By using DH technology, homozygosity is reached in a shorter time and thus breeding programs become more efficient (Prigge & Melchinger, 2012). Doubled-haploid technology also offers the opportunity to carry out association studies between molecular markers and phenotypic traits for its implementation in marker-assisted selection, allowing higher efficiency and precision in selection of parental lines (Prasanna et al., 2013).

A maize breeding program that prioritized HOC was started in Mexico with four subtropical populations of white and yellow grains adapted to the northwestern and Bajio regions of Mexico (Ortega-Corona et al., 2015; Preciado-Ortiz et al., 2013). After eight cycles of recurrent selection by half siblings for HOC and agronomic traits, the oil content in grain was increased by 33–60% (7.5–8.1% of oil content) compared with original populations (4.3–5.7% of oil content), which had a significant positive correlation with the germ size increase. In addition, a small but significant increase in grain yield was observed (100–110 Kg ha^{-1} per cycle), suggesting that it is feasible to develop high-yielding subtropical HOC maize with acceptable agronomic performance and grain physical properties (Preciado-Ortiz et al., 2013). These lines showed values of oleic acid that comprised 34–43% and 33–44% of the total fatty acids in yellow and white hybrids, respectively, whereas linoleic acid in yellow and white hybrids ranged from 37 to 52% and from 34 to 51%, respectively, indicating that these DH lines have potential for direct use in food and feed industries (Ortiz-Islas et al., 2019) and demonstrating that the development of HOC hybrids adapted to subtropical environments is viable.

Core Ideas

- Genetic diversity was used as a tool for high-oil-content maize breeding.
- High genetic diversity was identified in four subtropical maize breeding populations.
- This study Identified heterotic groups based on SNP genotyping.

Recently, these families were selected to generate homozygous DH lines. After 11 cycles of recurrent selection, DH lines were obtained from the four improved populations (Bajio yellow population [BYP], northwestern yellow population [NYP], Bajio white population [BWP], and northwestern white population [NWP]) to generate inbred lines that could be used as parentals in simple crosses. However, the genetic characterization is helpful to make the breeding process more efficient and could also increase the probability of maximizing the heterosis in the new hybrids (Singode et al., 2017). Molecular markers in maize are used for population and genome evolution studies, germplasm characterization, trait mapping, and breeding (Romay, 2018). This information, in combination with phenotypic data, is used by breeders to identify sources of possible favorable alleles to be included into their breeding programs (Wang et al., 2017). Therefore, the evaluation of genetic diversity and population structure among germplasm through molecular markers is mandatory to fully exploit the potential of such breeding material by increasing the heterosis.

Molecular characterization and population genetic structure of maize germplasm have been successfully applied for selecting parental lines and assigning heterotic groups (Wu et al., 2016). Ogugo et al. (2015) performed a molecular characterization by genotyping-by-sequencing of 417 maize DH lines using 97,190 single-nucleotide polymorphism (SNP) markers, reporting that 97% of the DH lines were genetically pure (<2% heterogeneity), and the genetic distance between pairwise comparison ranged from 0.055 to 0.457. Similar analyses revealed high genetic variability of 157 maize inbred lines genotyped with 4,976 polymorphic SNPs using genotyping-by-sequencing; 91.1% of the inbred lines were considered pure with <5% heterogeneity, and the remaining lines had a heterogeneity ranging from 5.5 to 40%. Also, cluster and model-based population structure analyses clustered the 157 lines into four groups (Leng et al., 2019).

Recently (Osuman et al., 2020), 162 early maturing white and yellow tropical maize inbred lines were genotyped with 9,684 SNP markers, showing a mean genetic diversity of 0.30, a mean polymorphic information content (PIC) value of 0.25, and a mean genetic distance of 0.42. Their population structure analysis revealed six different groups based on the history of selection, grain color, and pedigree. Thus, studying the genetic diversity and population structure of maize lines will allow breeders to identify more efficiently the best parents in order to generate maize hybrids with improved agronomic characteristics previously selected by recurrent selection. Here, 310 DH lines from four subtropical maize yellow and white populations with HOC obtained after 11 cycles of half-sib recurrent selection were selected aiming to evaluate the genetic diversity and population structure by DArT-seq SNP discovery technology (Sansaloni et al., 2011).

2 | MATERIALS AND METHODS

2.1 | Plant material

Four HOC subtropical maize populations with high yield potential were developed by integrating different germplasm adapted to the most important agricultural regions of Mexico, Bajio and northwest, including white and yellow grain for each region: BWP, NWP, BYP, and NYP. The four original populations belonged to opposite heterotic groups of white and yellow populations and involved several landraces as well as improved germplasm; the pedigree details have been summarized previously (Ortega-Corona et al., 2015). The half-sib recurrent selection scheme used to develop the four HOC subtropical populations was previously documented by Preciado-Ortiz et al. (2013). Briefly, for each cycle of selection, 50 half-sib HOC families from each population were planted in a 5-m row as a female parent (detasseled) in an isolated plot. The male pollinators were a bulk composite integrated with HOC seeds from each family and were planted every other row (two females and one male). At harvest, 200 ears of the population were selected from the best four plants in each family (special emphasis was done to select healthy plants). Then, a bulk seed of each of the 200 ears was used to identifying the best 50 HOC families using a near-infrared spectrometer (Infratec 1241, food and feed analyzer, Tecator AB). Finally, 100 seeds were selected and used to initiate a new cycle in each of the four populations. After 11 cycles of recurrent selection for HOC in the four populations, families were selected to generate homozygous DH lines at International Maize and Wheat Improvement Center (CIMMYT). A total of 310 subtropical DH lines (26 BYP, 51 NYP, 73 BWP, and 160 NWP) were used in this study.

2.2 | DNA extraction and genotyping

Total genomic DNA was isolated from seeds by using a modified CTAB method (Doyle & Doyle, 1987) with sarcosyl. Briefly, digestion buffer containing 2% (w/v) CTAB and 1% (w/v) sarcosyl were added into 1.5-ml microcentrifuge tube containing 50 mg of powdered maize seed and was shaken for 1.5 h at room temperature. Then, phenol/chloroform (1:1) was added and then shaken for 20 min and the top aqueous layer were removed. The DNA integrity and quality were evaluated by electrophoresis in 2% agarose gel and by spectrophotometry (Nanodrop2000, Thermo Scientific) using the ratios of optical densities 260/280 (1.80–1.90). DNA was stored at −20 ˚C until library preparation.

Genotyping of DH lines was performed at Genetic Analysis Service for Agriculture (SAGA) laboratory in CIMMYT Texcoco, Mexico, following the DArT-seq protocol (Sansaloni et al., 2011). Briefly, total DNA was digested by two restriction enzymes *PstI* (recognition site 5′-CTGCAG-3′) and *HpaII* (recognition site 5′-CCGG-3′) to reduce genomic complexity. Barcoded adapters were ligated to identify each sample (digestion–ligation reaction); only mixed fragments (*PstI–HpaII*) were effectively amplified by polymerase chain reaction (PCR). After PCR, equimolar amounts of amplification products of each sample of the 96-well microtiter plate were bulked and applied to c-Bot (Illumina) bridge PCR, followed by fragment sequencing on Illumina HiSeq 2500 System (https://www.illumina.com). Single-end sequencing was run for 69 cycles. Sequences generated from each lane were processed by proprietary DArT analytical pipelines (http://www.diversityarrays.com). First, samples were barcode separated and a single file with all reads (length of 68–69 pb) was generated, then the quality control of fastq files analysis was carried out (Phred > 30). The filtered reads were aligned to a data set of tropical and subtropical maize lines available in CIMMYT to accomplish single-nucleotide polymorphism (SNP) calling, leading a presence–absence matrix of SNP markers. Also, to obtain the physical location of the genotyped markers, sequences were aligned with BLAST to the B73.v4 maize genome reference (Jiao et al., 2017).

2.3 | Genetic diversity

The SNP markers with >20% of missing data and minor allele frequency <0.05 were eliminated using TASSEL (v5.2.74) software (Bradbury et al., 2007). Observed (H_o) and expected heterozygosity (H_e) were calculated for each DH line by using Bio-R software v2.0 (Pacheco et al., 2016). Polymorphic information content for each SNP was calculated using PowerMarker v3.25 (Liu & Muse, 2005).

2.4 | Population structure analysis

Genetic distances among lines and populations were calculated according to Rogers (1972) using Bio-R software. Based on the genetic distance, dendrograms were constructed using unweighted pair group method with arithmetic mean (UPGMA) in Mega software version X (Kumar et al., 2018). Additionally, a multidimensional scaling (MDS) analysis was conducted in CurlyWhirly software v1.19.09.04 (https://ics.hutton.ac.uk/curlywhirly) for exploring the population structure in a three-dimensional plot.

With the purpose to find the grouping of DH lines, the genetic structure was determined by Bayesian clustering analysis in STRUCTURE software v2.3.4 (Pritchard et al., 2000). Runs were performed using an admixture model, varying the number of groups (K) from 1 to 5, each K repeated 10 times with a burn-in period of 10,000 and 10,000 Markov chain Monte Carlo replications after burn-in. The best K was identified by inputting the STRUCTURE results into the STRUCTURE HARVESTER software using the Evanno method. Based on the output log likelihood of data [LnP(D)], the ad hoc statistic delta K (ΔK) was used to determine the optimal number of groups (Evanno et al., 2005). Results of 10 replicate files were integrated using the CLUMPP software v1.1.2b (Jakobsson & Rosenberg, 2007).

Finally, to complement the Bayesian analysis to identify the population structure, a discriminant analysis of principal components (DAPC) was carried out using the 'adegenet' package (Jombart et al., 2010).

3 | RESULTS

The DArT-seq data of 310 subtropical maize DH lines generated 45,868 raw SNP markers. After filtering (missing data >20% and minor allele frequency ≤0.05 were removed), 19,078 high-quality SNP markers were selected for further statistical analysis.

To get the physical position across the 10 chromosomes, the 19,078 SNPs were mapped in silico to the maize genome reference B73.v4. The chromosome coverage ranged from 752 SNPs on chromosome 10 to 1,756 SNPs on chromosome 1. 8,214 SNPs (43.05% of selected SNP) did not map to the maize genome reference B73, which might indicate exclusive alleles for tropical maize because B73 is from temperate climate origin (Supplemental Table S1).

3.1 | Genetic diversity

To measure the genetic diversity present in the four analyzed populations, genetic diversity indexes (PIC, H_o, and H_e) for the four maize populations were determined (Table 1). The NWP had the lowest average values of both H_o and H_e, while the BWP had the highest average values.

3.2 | Population structure

3.2.1 | Genetic distance

Genetic distances among populations ranged from 0.243 (BYP–NWP) to 0.268 (NYP–BWP) (Table 2). Pairwise genetic distance between the 310 DH lines ranged from 0.119 to 0.385; most of them showed values between 0.201 and 0.250 (Figure 1) with a percentage of 38.1 (Figure 1a) and 46.6% (Figure 1b) for yellow and white DH lines, respectively. The higher genetic distance, the higher heterosis when crossing maize lines. Only a few pairwise comparisons of the DH lines showed genetic distance higher than 0.351 for both yellow and white maize DH lines. In yellow lines, the highest genetic distance (0.365) was observed between NYP-164 and NYP-247. In white maize lines, the highest genetic distance (0.385) was observed between NWP-430 and NWP-603.

3.2.2 | Clustering analysis

Radial dendrograms were constructed based on the genetic distances to show the genetic relationships of DH maize lines adapted to subtropical environments. The UPGMA dendrogram assigned the 310 DH lines into two main clusters followed by several subclusters (Supplemental Figure S1). Because single crosses will be made between yellow and white maize (yellow DH lines will not cross with white DH lines), the dendrograms were constructed by separating yellow and white kernel color. Figure 2 shows a dendrogram for the 77 yellow maize lines from the northwest and Bajio regions of Mexico, evidencing two main clusters; in Cluster 1, only the NYP-206 DH line was grouped, and the rest 76 DH lines were placed in Cluster 2, showing a mixture among the DH yellow lines. Based on the Figure 3, the 233 white maize lines were grouped into two main clusters. A total of 12 DH lines were placed in Cluster 1, all derived from the NWP; the second cluster included 221 DH lines a separated according to their origin (Bajio or northwest).

In addition, a MDS analysis clustered the 310 DH lines into three groups according to the region or environmental adaptation (Bajio or northwest) and grain color (Figure 4). The first group was a mixture composed of yellow maize lines from both populations (BYP and NYP), the second group was only composed of BWP lines, and the third group was only composed of NWP lines.

Clustering analysis by STRUCTURE (Figure 5) suggests that the 310 DH maize lines fit into two groups ($K = 2$). The

TABLE 1 Genetic diversity of the 310 doubled-haploid maize lines based on 19,078 single-nucleotide polymorphism markers

Population	No. of lines	Polymorphic information content	Observed heterozygosity	Expected heterozygosity
BYP	26	0.346 ± 0.03	0.221 ± 0.002	0.188 ± 0.001
NYP	51	0.352 ± 0.02	0.194 ± 0.002	0.165 ± 0.001
BWP	73	0.353 ± 0.02	0.284 ± 0.006	0.219 ± 0.004
NWP	160	0.353 ± 0.04	0.177 ± 0.002	0.152 ± 0.001
Total	310	0.354 ± 0.02	0.186 ± 0.003	0.159 ± 0.002

Note. BYP, Bajio yellow population; NYP, northwest yellow population; BWP, Bajio white population; NWP, northwest white population.

TABLE 2 Rogers genetic distance between four doubled-haploid maize populations based on 19,078 single-nucleotide polymorphism markers

Population	BYP	NYP	BWP	NWP
BYP	0	0.247 ± 0.069	0.249 ± 0.062	0.243 ± 0.069
NYP	–	0	0.268 ± 0.045	0.245 ± 0.064
BWP	–	–	0	0.253 ± 0.057
NWP	–	–	–	0

Note. BYP, Bajio yellow population; NYP, northwest yellow population; BWP, Bajio white population; NWP, northwest white population.

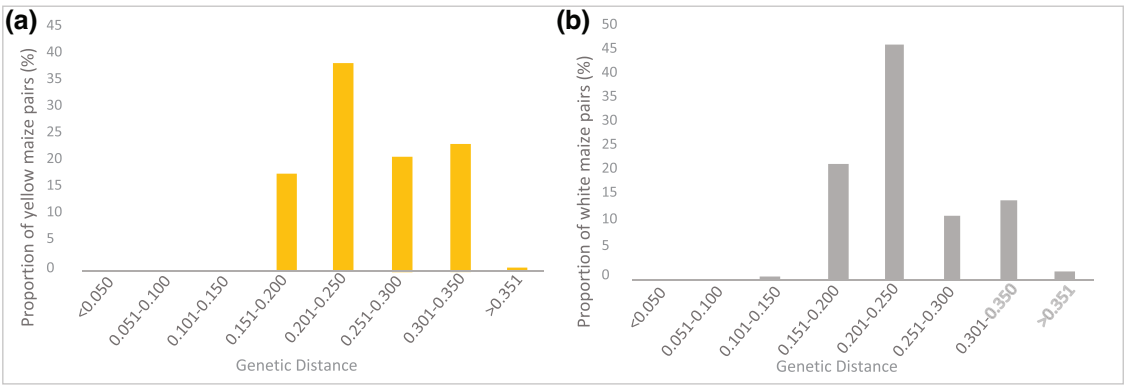

FIGURE 1 Frequency distribution categories of pairwise genetic distance based on 19,078 SNP markers of (a) 77 double-haploid yellow maize lines and (b) 233 double-haploid white maize lines

Group 1 consisted of 49 DH lines (15.8%), while Group 2 clustered 261 DH lines (84.2%) (Figure 5c). Interestingly, the organization of both groups showed no relation to grain origin or color, observing admixture of maize lines from the four populations of origin.

On the other hand, the results of DAPC suggest the presents of five groups (Figure 6). The *k*-means clustering method implemented by the find.clusters function of the Adegenet package found that the lowest value of BIC coincided with the value of $K = 5$ (Figure 6a). The DAPC was performed retaining 50 principal components and all discriminant functions (four eigenvalues) (Figure 6b). The size of the five groups found to K1, K2, K3, K4, and K5 were 127, 31, 65, 17, and 70, respectively. The 310 DH lines are separated according to grain color through the *x* axis, closely grouping the lines of the BYP (K4, orange) and NYP (K3, yellow) populations; while the lines from NWP formed two groups, K1 and K2 (blue and gray) in the negative quadrant of the *x* axis. The fifth group (red) was forming for the DH lines from BWP (Figure 6b).

Both, UPGMA and STRUCTURE clustering analyses showed the presence of two groups (Supplemental Figure S1; Figure 5, respectively). Two subgroups were observed in the UPGMA analysis, then, an additional analysis was performed separating lines by grain color (Figures 2 and 3). The results from DAPC analysis were consistent with the results from MDS analysis (Figure 4).

4 | DISCUSSION

High-oil maize varieties have been selected for breeding in temperate environments using conventional recurrent

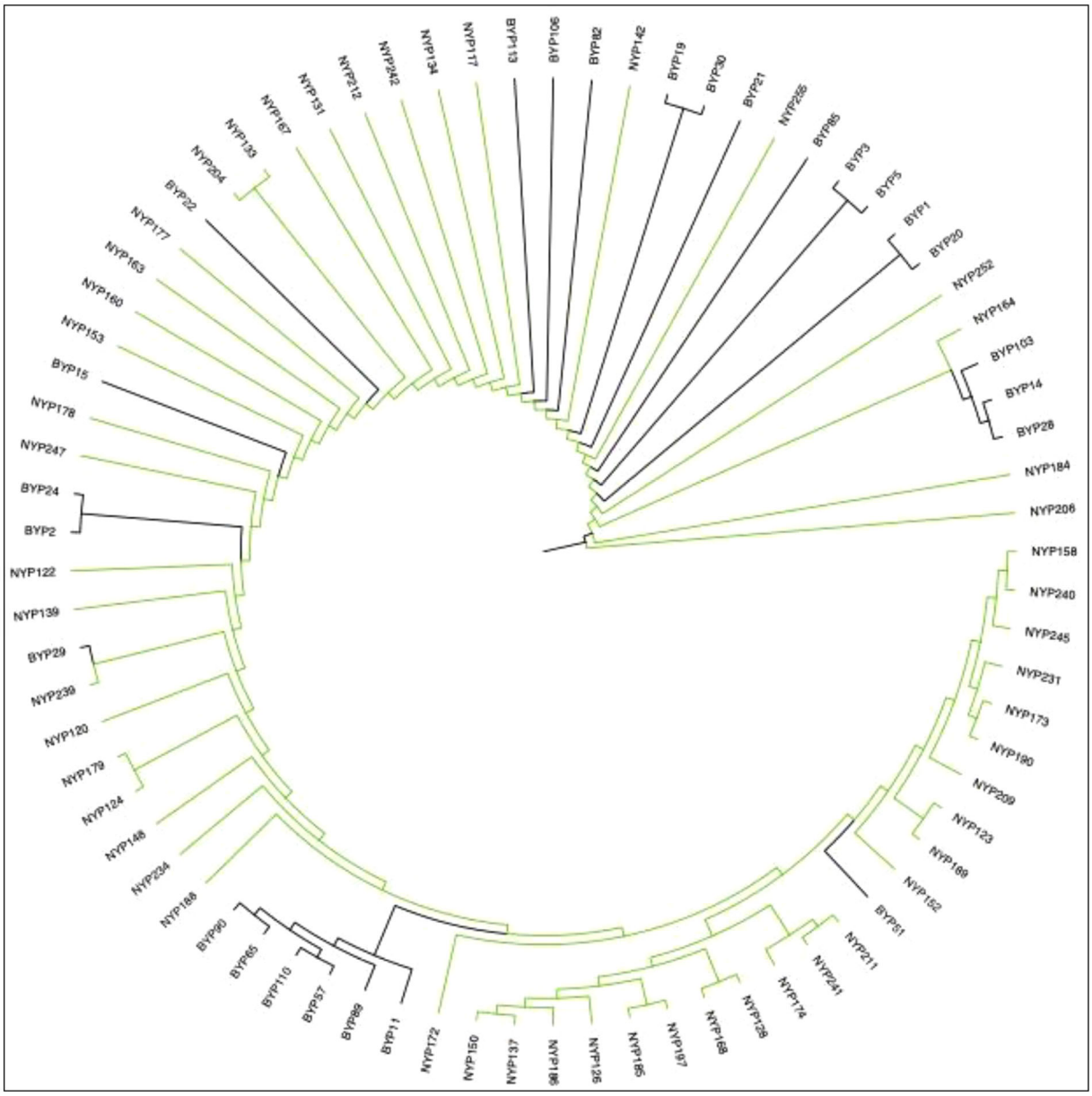

FIGURE 2 Unweighted pair group method with arithmetic mean dendrogram based on Roger's genetic distance calculated from 19,078 SNPs of 77 double-haploid yellow maize lines from the northwest and Bajio of Mexico. NYP, northwest yellow population (black); BYP, Bajio yellow population (green)

selection (Ortiz-Islas et al., 2019). On the other hand, subtropical maize genotypes own higher genetic diversity compared with those from temperate regions (Yan et al., 2009). Exploring this abundant diversity could be useful within genetic improvement programs around the world to generate competitive HOC maize hybrids. Modern breeding programs focus on selecting inbred lines to form single crosses that exhibit the highest heterosis for grain yield (Tomkowiak et al., 2019). However, facing human malnutrition, it is also necessary to identify genotypes with high nutritional value (Ortiz-Islas et al., 2019). In this sense, the enhancement of nutritional quality has become a popular objective in maize breeding, HOC being one of the main targets (Kahriman et al., 2020).

Maize grain should contain >6% oil to be considered as HOC maize (Lambert, 2001). High-oil-content maize varieties have shown positive nutritional effects for humans (Zhang et al., 2012) because of the high content of both oleic and linoleic acids (Serna-Saldivar, 2010), and for animals, the higher metabolizable energy in HOC compared with regular maize hybrids especially in terms of feed efficiency, growth, and overall productivity (Weiss & Wyat, 2000). In subtropical regions, an increasing interest exists to generate high-yield

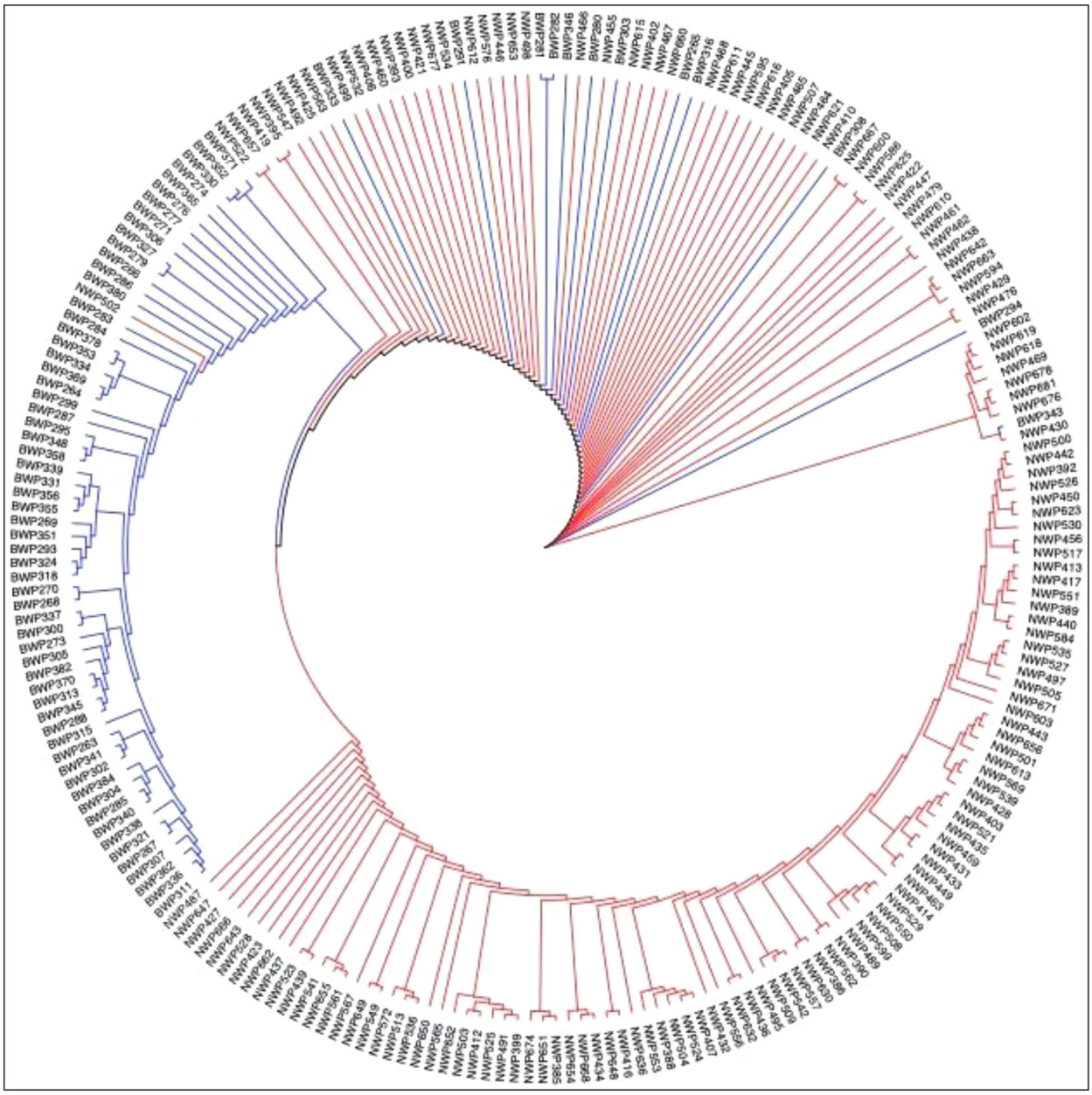

FIGURE 3 Unweighted pair group method with arithmetic mean dendrogram based on Roger's genetic distance calculated from 19,078 SNPs of 233 double-haploid white maize lines from the northwest and Bajio of Mexico. BWP, Bajio white population (blue); NWP, northwest white population (red)

maize hybrids with value-added properties such as HOC (Silva-Venancio et al., 2019). Currently, maize breeding programs in Mexico have joined efforts in order to develop HOC white and yellow maize varieties without affecting the grain yield (Ortega-Corona et al., 2015; Preciado-Ortiz et al., 2013) that are suitable for the elaboration of flours and tortillas (Vázquez-Carrillo et al., 2015) with an improved fatty acid profile (Ortíz-Islas et al., 2019) and good performance maize hybrids that would affect the livestock and food industry in Mexico in short term (Picón-Rico et al., 2018; Silva-Venancio et al., 2019). In this work, molecular and bioinformatic techniques were incorporated to analyze the potential of white and yellow HOC maize DH lines (oil content of 3.7–8.1%; Supplemental Table S2) derived from subtropical germplasm as parental lines.

The genetic diversity and population structure of 310 DH lines were analyzed through 19,078 polymorphic SNP markers. Several studies have reported the genetic characterization of maize inbred lines using different numbers of markers; Boakyewaa et al. (2019) efficiently characterized 94 maize inbred lines using 15,047 SNP markers, which allowed a clearer inference of the population structure, whereas

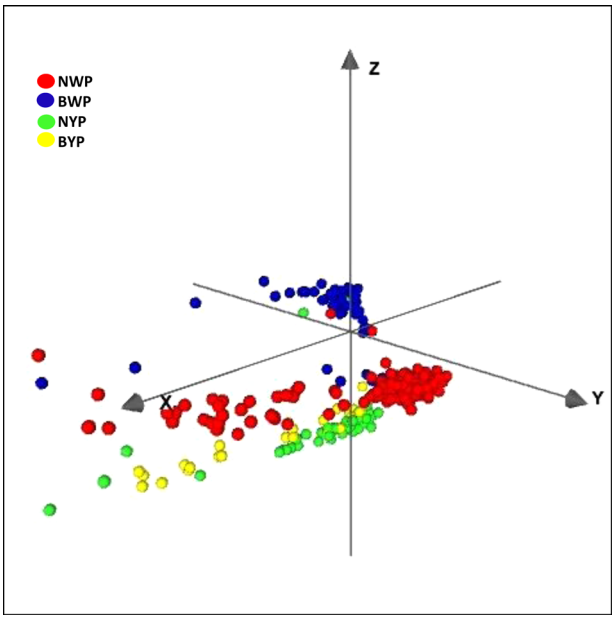

FIGURE 4 Multidimensional scaling (MDS) graph of double-haploid lines of white and yellow corn from Bajío and northwest Mexico. BWP, Bajio white population (blue); NWP, northwest white population (red); BYP, Bajio yellow population (yellow); NYP, northwest yellow population (green)

Leng et al. (2019) used a lower number of markers (4,976 SNPs) for the genetic characterization of 157 elite maize inbred lines. Thus, the 19,078 SNP markers can be considered as a suitable number to determine the genetic diversity and population structure of DH lines. In this study, SNP maps of tropical and subtropical maize were used for SNP calling from the studied DH lines. Interestingly, 8,214 SNP markers (43.05% of total) did not map to any of the maize B73 chromosomes, which could mean that these SNP loci are present only in subtropical maize. With the aim to investigate genetic variation at a genome-wide scale, Jiao et al. (2017) reported a vastly improved de novo assembly and annotation of B73 maize reference genome, and optical maps were generated for two additional tropical maize inbred lines (Ki11 and W22). Only 32% of the assembled map of Ki11 and 39% of the W22 map were mapped to the new B73 reference via common restriction patterns, thus confirming the tremendous genetic contrast between B73 (a temperate genotype) and tropical genotypes such as those described in this report. These findings highlight the importance of including tropical and subtropical maize genomes for the SNP calling in further studies.

The 19,078 SNPs had an average PIC value of 0.354, which is higher than the PIC values reported previously by Osuman et al. (2020) and Silva et al. (2019) with averages of 0.25 and 0.17, respectively. A higher mean PIC value indicates the effectiveness of the markers and implies that majority of the selected SNPs were informative and polymorphic enough to highlight the differences among the 310 DH lines.

The mean H_o value (0.186) was higher than H_e (0.159), which is normal in breeding populations because of the recurrent selection. In contrast with our results, Kasoma et al. (2021) reported a mean H_o of 0.350, lower than the H_e of 0.530 for 59 maize genotypes. A population with moderate heterozygosity is one in which most of the genetic loci are not fixed in a homozygous ($H_o = 0$) or heterozygous ($H_o = 1$) state (Elston, 2005).

Additionally, the heterozygosity is highly related to the genetic purity of inbred lines and is an important quality control criterion for maize breeding (Ertiro et al., 2017), likewise, it allows high level of heterosis in hybrid formation (Mengesha et al., 2017). Inbred lines are considered pure or fixed when the proportion of heterozygous loci does not exceed 5% (Semagn et al., 2012). Our results indicate that all DH lines had an average H_o >18%. Inbred lines or DH lines with >5% heterogeneous SNP loci are likely to have some changes in allele frequencies that may have been caused by pollen contamination during seed multiplication (Warburton et al., 2010) or seed admixtures during harvesting and handling. Also, residual heterozygosity could be expected at 5 to 10% rates even in advanced selfing generations (Nepolean et al., 2013).

Our results contrast the results of other similar studies, in which the H_o is lower. For example, Ogugo et al. (2015) reported heterozygosity ≤1% in 417 DH lines and only two DH lines presented heterozygosity >5%. Wu et al. (2016) reported that the 538 CIMMYT maize inbred lines (CMLs) showed a mean H_o of 0.01. On the other hand, our results are concordant with Ertiro et al. (2017) who reported that only 22% of 265 inbred lines were considered pure with <5% heterogeneity, while the remaining 78% of the inbred lines had a heterogeneity ranging from 5.1 to 31.5%. Finally, Leng et al. (2019) reported that 91.1% of the 157 inbred lines were considered pure with <5% heterogeneity, while the remaining 8.9% of the inbred lines had a heterogeneity ranging from 5.5 to 40.0%. When the heterozygosity exceeds 5%, it is suggested to perform additional generations of selfing in order to fix the homozygous alleles and be able to obtain the advantages of using pure lines such as easy maintenance of parental lines, high heterosis in hybrids, and easy quality control during hybrid seed production (Ertiro et al., 2017). Therefore, we suggest one or two additional cycles of selfing in the DH maize lines in order to increase their purity to be used as parental lines.

4.1 | Population structure

The population structure of 310 maize DH lines was assessed by using four different approaches: (a) an UPGMA dendrogram based on genetic distance, (b) a MDS analysis, (c) a Bayesian clustering analysis in STRUCTURE,

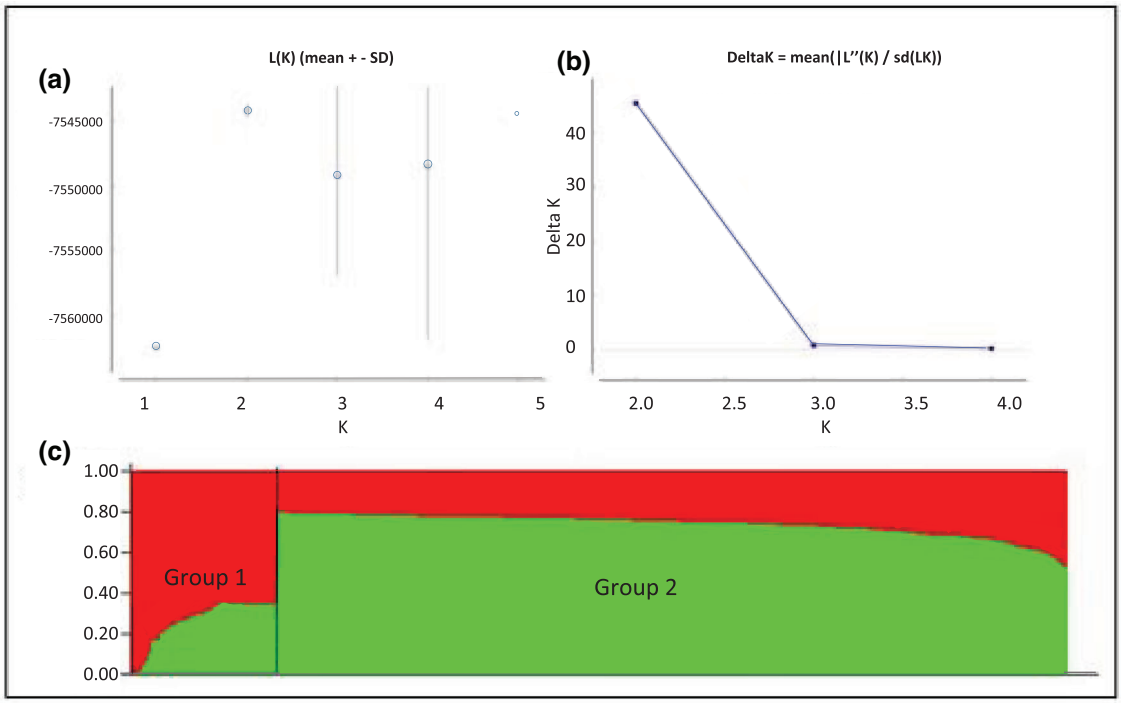

FIGURE 5 Population structure of 310 double-haploid (DH) maize lines estimated with 19,078 SNPs. (a) Plots of LnP(D) for each *K* calculated by STRUCTURE. The LnP(D) data is shown as mean ± SD (standard deviation). (b) Delta *K* (ΔK) values for *K* ranging from 2 to 4 according to Evanno et al, 2005. (c) Population structure of the panel when $K = 2$. Each of the 310 DH corn lines is represented by a thin vertical bar, which is partitioned into *K* colored segments on the *x* axis with lengths proportional to the estimated probability membership in each of the *K*-inferred clusters (*y* axis)

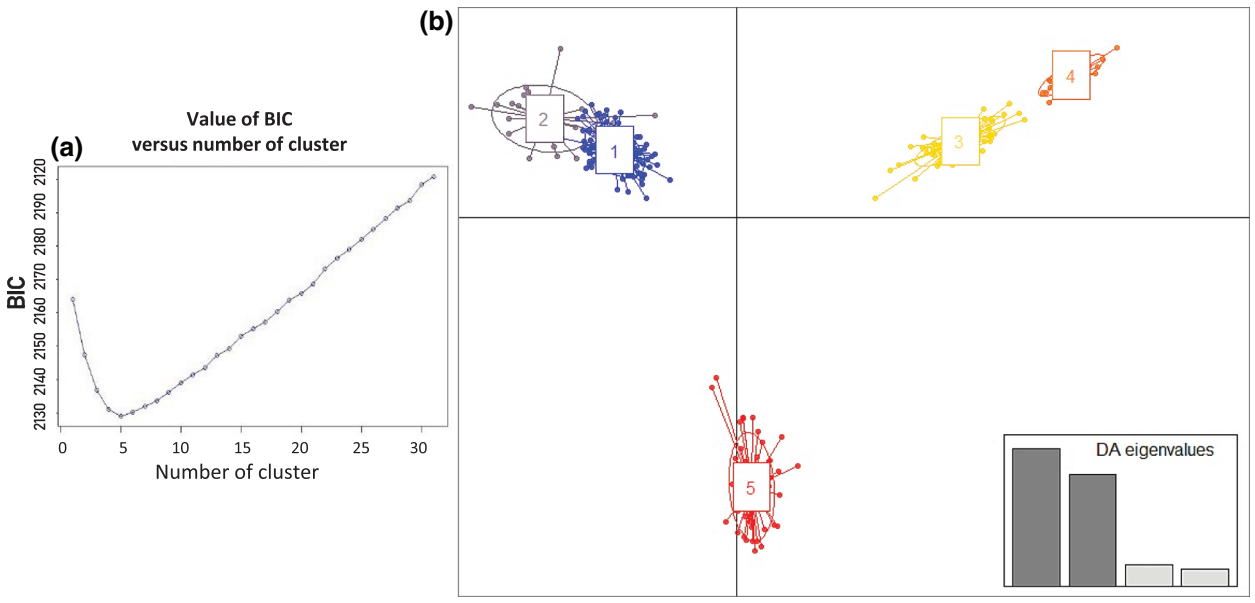

FIGURE 6 Genetic clustering of 310 high-oil content (HOC) maize lines by discriminant analysis of principal components (DAPC). (a) Bayesian information criterion (BIC) is provided for different numbers of clusters (from 1 to 31); a *K* value of 5 (the lowest value of the BIC) represents the most optimal clustering of the data. (b) Scatterplot from the DAPC analysis; HOC maize lines are indicated by dots. Number and colors mention the five genetic clusters retained from BIC values: Cluster 1 in blue, Cluster 2 in gray, Cluster 3 in yellow, Cluster 4 in orange, and Cluster 5 in red

and (d) DAPC analysis, which provided complementary information. Genetic distance is a measure of the genetic differentiation between pairs of lines or populations (Leng et al., 2019). Pairwise genetic distances showed a clear separation between the white and yellow maize lines. The high genetic diversity present in maize is a result of the high variety of agroecological environments in which they were grown, which contributed to the development of divergent populations adapted to different edaphic and climatic conditions and biological factors (Mikić et al., 2017). For example, higher genetic divergence has been observed among tropical maize than among temperate maize because tropical and subtropical maize are more diverse and thus contain rarer alleles than temperate maize (Yan et al., 2009). Analyzing other studies, genetic divergence measured through pairwise genetic distance among tropical maize germplasm have shown larger differences. For instance, pairwise genetic distances among 417 DH maize lines genotyped with 97,190 SNP markers ranged between 0.055 and 0.457 (Ogugo et al., 2015). Other study involving 265 inbred lines reported a larger range between pairwise genetic distances, but lower genetic distances, ranging from 0.011 to 0.345 (Ertiro et al., 2017). The genetic distance among the four subtropical maize populations had a smaller range from 0.243 (BYP and NWP) to 0.268 (NYP and BWP) compared with to those among the DH lines. The HOC maize populations were developed in the same genetic breeding program, which could explain the small range of genetic distance between the four populations. Wu et al. (2016) studied 538 CMLs and six temperate inbred lines and found that genetic distances between different subgroups varied from 0.590 to 0.710. This high genetic divergence could be due to contrasting germplasm origins (temperate, tropical, and subtropical environments). The SNP-based genetic distances may provide important insights for effective parental selection, avoiding crosses between genetically similar subtropical maize lines (Silva et al., 2019). The higher the genetic distance, the higher heterosis has been reported in previous studies (Tomkowiak et al., 2020). Our results show that approximately 25 and 18% of the pairwise distance among yellow and white DH maize, respectively, had a genetic distance >0.301. These results suggest that within the four HOC maize populations, there are a large number of DH lines to generate maize hybrids with high heterosis and yield potential. Improved maize hybrids with DH lines as parentals have been developed by several commercial maize breeding programs in Europe, North America, China (Molenaar & Melchinger, 2019), and Africa (Beyene et al., 2017; Chaikam et al., 2018), and their use, in combination with molecular markers, has offered efficient alternatives for increasing selection gain (Chaikam et al., 2019).

The construction of two dendrograms allowed the visualization of genetic divergence among the same grain color lines (yellow and white). Lines showing lower genetic distance were situated more closely to each other and showed a closer branch on the dendrogram, and those with higher genetic distance were less unrelated and showed a more distant branch. Yellow maize DH lines did not clearly cluster according to their origin region (Bajio or northwest), showing a mixture among the 77 lines (Figure 2). On the other hand, the 233 white DH lines showed a separation related to the origin (Figure 3). These results suggest that yellow DH lines share a greater number of markers among each other compared with the white DH lines, which might contain specific markers in relation to the region of origin. In a similar way, Wu et al. (2016) detected a separation of 538 CMLs into three groups according to the geographical adaptation of the genotypes (lowland tropical, middle tropical, and lowland tropical highlands). Also, Leng et al. (2019) reported 157 elite inbred lines that were clustered into four groups according to genetic relationship (i.e., the common ancestral parents for breeding).

Multidimensional scaling analysis suggested the presence of three groups: the first group containing DH lines from NWP, the second group from BWP, and the third group with admixture between BYP and NYP. Interestingly, besides MDS analysis included the whole set of lines (both grain color) the results are concordant and supports the UPGMA results, showing a separation by origin region only for white DH lines. The MDS analysis realized by Chen et al. (2016) using 18,028 SNPs allowed the separation of the 561 CMLs into three main groups based on adaptation environment (maize adapted to highland, subtropical, and lowland environments).

To establish the genetic groups within the 310 maize DH lines, STRUCTURE analysis based on a Bayesian model and DAPC without being based on a model were performed. The Bayesian analysis of STRUCTURE grouped the DH lines into only two groups ($K = 2$); however, this grouping was not defined either by grain color or by region of origin because both groups (Figure 5c) included DH lines from the four populations. In contrast, Ertiro et al. (2017) reported 265 inbred lines that were grouped by the STRUCTURE software into three groups ($K = 3$), which was consistent with the pedigree information. On the other hand, Leng et al. (2019) analyzed 157 lines and grouped them into four different groups ($K = 4$), in agreement to the genetic relationships but not according to their institute or company of origin.

In contrast, the DAPC analysis grouped the maize DH lines in a better way, establishing five clusters (Figure 6b) according to grain color and the origin region (white grain maize), which is in complete concordance with the MDS analysis results (Supplemental Figure S2). The differences between the STRUCTURE and DAPC results may be due to the fact that the STRUCTURE program works on the based assumptions such as linkage and Hardy–Weinberg equilibrium, while DAPC does not (Jombart et al., 2010).

Also, in our case, the 310 DH lines were developed within the same breeding program but also the DH lines of the

four populations were established in both northwest and Bajio environments, which could have caused large number of shared genetic markers among them, as observed in the Figure 5c, and therefore, not achieving a separation based on source population. Also, the low genetic divergence observed among most tropical germplasm (Guo et al., 2021) might support these results.

Molecular markers provide an alternative approach to characterizing genetic diversity and population structure on a large scale within a given germplasm collection (Leng et al., 2019). Our analysis with molecular markers in HOC DH maize lines allowed us to not only identify high levels of genetic diversity but also the genetic distance among lines, which is valuable information for breeders trying to identify potential lines for increasing the heterosis when hybrid conformation. Information of the lines, including the grain color, origin, the presence or absence of genetic markers, genetic distance, and other parameters, are also considered in breeding programs (Osuman et al., 2020). In the present study, the clustering analysis was not fully consistent through the different methods possibly not only because of the fundamentals of each analysis and the genetic breeding scheme developed but also to the low genetic divergence among lines. Besides the inconsistences in clustering analysis, the results reported here are valuable for selecting parental DH lines with HOC for the consolidation of a HOC hybrid breeding program for subtropical environments.

5 | CONCLUSION

Maize breeding program for HOC should be integrative, combining information of not only genetic diversity, genetic distance, and clustering analysis but also information about heterotic groups defined by field assays in order to identify specific and general combining ability. This research not only will help breeders to better understand how to select parental DH lines and assign heterotic groups to generate new hybrids but also provides plenty of information for genetic research such as genome-wide association study and marker-assisted selection to produce maize hybrids with HOC for subtropical regions in the future.

ACKNOWLEDGMENTS

This work was funded by grants of the INIFAP's project 10561934518 and INAPI-Sinaloa. The authors gratefully acknowledge Dr. Ricardo E. Preciado-Ortíz from INIFAP Campo Experimental Bajio, Guanajuato, Mexico, for providing the doubled-haploid lines with high oil content and to the Dr. Raúl J. Delgado-Macuil from CIBA-Tlaxcala for helping to perform DAPC analysis. Also, we gratefully acknowledge the technical staff at SAGA-CIMMYT, especially to Dr. César Petroli, Guadalupe Valdez, and Araceli Balderas for their contributions to this study. The first author acknowledges financial support from Consejo Nacional de Ciencia y Tecnología (CONACYT) through a PhD scholarship.

AUTHOR CONTRIBUTIONS

Grethel P. Gaytán-Pinzón: Formal analysis; Writing – original draft; Writing – review & editing. Eduardo Sandoval-Castro: Formal analysis; Writing – original draft; Writing – review & editing. Luis A. Peinado-Fuentes: Conceptualization; Funding acquisition. Juan P. Valenzuela-Apodaca: Methodology. Abraham Cruz-Mendívil: Data curation; Methodology. Carlos L. Calderón-Vázquez: Conceptualization; Funding acquisition; Project administration.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

ORCID

Carlos L. Calderón-Vázquez https://orcid.org/0000-0002-6674-2504

REFERENCES

Battistelli, R. G., Von Pinho, V., Justus, A., Couto, E. G. O., & Balestre, M. (2013). Production and identification of doubled haploids in tropical maize. *Genetics and Molecular Research*, *12*, 4230–4242. https://doi.org/10.4238/2013.October.7.9

Beyene, Y., Mugo, S. N., Oikeh, S. O., Juma, C., Olsen, M., & Prasanna, M. B. (2017). Hybrids performance of doubled haploid lines derived from 10 tropical bi-parental maize populations evaluated in contrasting environments in Kenya. *African Journal of Biotechnology*, *16*, 371–379. https://doi.org/10.5897/AJB2016.15697

Boakyewaa Adu, G., Badu-Apraku, B., Akromah, R., Garcia-Oliveira, A. L., Awuku, F. J., & Gedil, M. (2019). Genetic diversity and population structure of early-maturing tropical maize inbred lines using SNP markers. *PLOS ONE*, *14*, e0214810. https://doi.org/10.1371/journal.pone.0214810

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*, 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Chaikam, V., Molenaar, W., Melchinger, A. E., & Prasanna, B. M. (2019). Doubled haploid technology for line development in maize: Technical advances and prospects. *Theoretical and Applies Genetics*, *132*, 3227–3243. https://doi.org/10.1007/s00122-019-03433-x

Chaikam, V., Nair, S. K., Martinez, L., Lopez, L. A., Utz, H. F., Melchinger, A. E., & Boddupalli, P. M. (2018). Marker-assisted breeding of improved maternal haploid inducers in maize for the tropical/subtropical regions. *Frontiers in Plant Science*, *9*, 1527. https://doi.org/10.3389/fpls.2018.01527

Chen, J., Zavala, C., Ortega, N., Petroli, C., Franco, J., Burgueño, J., Costich, D. E., & Hearne, S. J. (2016). The development of quality control genotyping approaches: A case study using elite maize lines. *PLOS ONE*, *11*, e0157236. https://doi.org/10.1371/journal.pone.0157236

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, *19*, 11–15.

Elston, R. C. (2005). Genetic markers. Encyclopedia of biostatistics. John Wiley and Sons Ltd.

Ertiro, B. T., Semagn, K., Das, B., Olsen, M., Labuschagne, M., Worku, M., Wegary, D., Azmach, G., Ogugo, V., Keno, T., Abebe, B., Chibsa, T., & Menkir, A. (2017). Genetic variation and population structure of maize inbred lines adapted to the mid-altitude sub-humid maize agro-ecology of Ethiopia using single nucleotide polymorphic (SNP) markers. *BMC Genomics*, *18*, 777. https://doi.org/10.1186/s12864-017-4173-9

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, *14*, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Food and Agricultural Organization (2019). *FAOSTAT, FAO statistical databases*. http://www.fao.org/faostat/en/#data/QC/visualize

García-Lara, S., & Serna-Saldívar, S. O. (2019). Corn history and culture. In S. O. Serna-Saldivar (Ed.), *Corn* (3rd ed., pp 1–18). AACC International Press. https://doi.org/10.1016/B978-0-12-811971-6.00001-2

Guo, R., Chen, J., Petroli, C. D., Pacheco, A., Zhang, X., San Vicente, F., Hearne, S. J., & Dhliwayo, T. (2021). The genetic structure of CIMMYT and U.S. inbreds and its implications for tropical maize breeding. *Crop Science*, *61*, 1666–1681. https://doi.org/10.1002/csc2.20394

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*, 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M. C., Wang, B., Campbell, M. S., Stein, J. C., Wei, X., Chin, C. S., Guill, K., Regulski, M., Kumari, S., Olson, A., Gent, J., Schneider, K. L., Wolfgruber, T. K., May, M. R., Springer, N. M., … Ware, D. (2017). Improved maize reference genome with single-molecule technologies. *Nature*, *546*, 524–527. https://doi.org/10.1038/nature22971

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genomics*, *11*, 94. https://doi.org/10.1186/1471-2156-11-94

Kahriman, F., Aktaş, F., Songur, U., Şerment, M., & Egesel, C. (2020). Screening Turkish maize landraces for kernel oil content and oil quality traits. *Plant Genetic Resources: Characterization and Utilization*, *18*, 278–286. https://doi.org/10.1017/S1479262120000258

Kasoma, C., Shimelis, H., Laing, M. D., Shayanowako, A. I. T., & Mathew, I. (2021). Revealing the genetic diversity of maize (*Zea mays* L.) populations by phenotypic traits and DArTseq markers for variable resistance to fall armyworm. *Genetic Resources and Crop Evolution*, *68*, 243–259. https://doi.org/10.1007/s10722-020-00982-9

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*, 1547–1549.

Lambert, R. J. (2001). High-oil corn hybrids. In A. R. Hallauer (Ed.), *Specialty corns* (2nd ed., pp. 131–150). CRC Press.

Lambert, R. J., Alexander, D. E., & Mejaya, I. J. (2004). Single kernel selection for increased grain oil in maize synthetics and high-oil hybrid development. *Plant Breeding Reviews*, *24*, 153–175.

Leng, Y., Lv, C., Li, L., Xiang, Y., Xia, C., Wei, R., Rong, T., & Lan, H. (2019). Heterotic grouping based on genetic variation and population structure of maize inbred lines from current breeding program in Sichuan province, Southwest China using genotyping by sequencing (GBS). *Molecular Breeding*, *39*, 38. https://doi.org/10.1007/s11032-019-0946-y

Liu, K., & Muse, S. V. (2005). PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics*, *21*, 2128–9. https://doi.org/10.1093/bioinformatics/bti282

Mengesha, W. A., Menkir, A., Unakchukwu, N., Meseka, S., Farinola, A., Girma, G., & Gedil, M. (2017). Genetic diversity of tropical maize inbred lines combining resistance to *Striga hermonthica* with drought tolerance using SNP markers. *Plant Breeding*, *136*, 338–343. https://doi.org/10.1111/pbr.12479

Mikić, S., Kondić-špika, A., Brbaklić, L., Stanisavljević, D., Ćeran, M., Trkulja, D., & Mitrović, B. (2017). Molecular and phenotypic characterisation of diverse temperate maize inbred lines in South-east Europe. *Zemdirbyste-Agriculture*, *104*, 31–40. https://doi.org/10.13080/z-a.2017.104.005

Molenaar, W. S., & Melchinger, A. E. (2019). Production of doubled haploid lines for hybrid breeding in maize. In O. Frank & F. Wolfgang (Eds.) *Advances in breeding techniques for cereal crops* (pp. 143–163). Burlegh Dodds Science Publishing Company.

Nepolean, T., Singh, I., Hossain, F., Pandey, N., & Gupta, H. S. (2013). Molecular characterization and assessment of genetic diversity of inbred lines showing variability for drought tolerance in maize. *Journal of Plant Biochemistry and Biotechnology*, *22*, 71–79. https://doi.org/10.4025/actasciagron.v43i1.53540

Ogugo, V., Semagn, K., Beyene, Y., Runo, S., Olsen, M., & Warburton, M. (2015). Parental genome contribution in maize DH lines derived from six backcross populations using genotyping by sequencing. *Euphytica*, *202*(1), 129–39.

Ortega-Corona, A., Picón-Rico, R., Preciado-Ortíz, E., Terrón-Ibarra, A. D., Guerrero-Herrera, M., García-Lara, S., & Serna-Saldivar, S. (2015). Selection response for oil content and agronomic performance in four subtropical maize populations. *Maydica*, *603*, 1–8.

Ortiz-Islas, S., García-Lara, S., Preciado-Ortiz, R. E., & Serna-Saldívar, S. O. (2019). Fatty acid composition and proximate analysis of improved high oil corn double haploid hybrids adapted to subtropical areas. *Cereal Chemistry*, *96*, 182–192. https://doi.org/10.1002/cche.10109

Osuman, A. S., Badu-Apraku, B., Ifie, B. E., Tongoona, P., Obeng-Bio, E., & Garcia-Oliveira, A. L. (2020). Genetic diversity, population structure and inter-trait relationships of combined heat and drought tolerant early-maturing maize inbred lines from west and central Africa. *Agronomy*, *10*, 1324. https://doi.org/10.3390/agronomy10091324

Pacheco, A., Alvarado, G., Rodríguez, F., & Burgueño, J. (2016). *BIO-R (Biodiversity analysis with R for Windows) Version 2.0*. CIMMYT Research Data & Software Repository Network.

Prasanna, B. M., Chaikam, V., & Mahuku, G. (2013). *Tecnología de dobles haploides en el mejoramiento de maíz: Teoría y práctica*. CIMMYT.

Preciado-Ortiz, R. E., García-Lara, S., Ortiz-Islas, S., Ortega-Corona, A., & Serna-Saldivar, S. O. (2013). Response of recurrent selection on yield kernel oil content and fatty acid composition of subtropical maize populations. *Field Crops Research*, *142*, 27–35. https://doi.org/10.1016/j.fcr.2012.11.019

Picón-Rico, R., Preciado-Ortiz, R. E., Cervantes-Ortiz, F., Covarrubias-Prieto, J., & Terrón-Ibarra, A. (2018). Efectos heteróticos en líneas doble haploides de maíz de grano blanco y alto contenido de aceite. *Revista Fitotecnia Mexicana*, *41*, 177–186. https://doi.org/10.35196/rfm.2018.2.177-186

Prigge, V., & Melchinger, A. E. (2012). Production of haploids and doubled haploids in maize. In V. Loyola-Vargas, & N. Flint-Garcia (Eds.), *Plant cell culture protocols* (pp. 161–172). Humana Press.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *1552*, 945–959. https://doi.org/10.1093/genetics/155.2.945

Rogers, J. S. (1972). Measures of genetic similarity and genetic distance. In M. R. Wheeler (Ed.), *Studies in genetics VII* (pp. 145–154). University of Texas Publication.

Romay, M. C. (2018). Rapid, affordable, and scalable genotyping for germplasm exploration in maize. In J. Bennetzen, S. Flint-Garcia, C. Hirsch, & R. Tuberosa (Eds.), *The maize genome. Compendium of plant genomes* (pp. 31–46). Springer. https://doi.org/10.1007/978-3-319-97427-9_3

Sansaloni, C., Petroli, C., Jaccoud, D., Carling, J., Detering, F., Grattapaglia, D., & Kilian, A. (2011). Diversity Arrays Technology (DArT) and next-generation sequencing combined: Genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. *BMC Proceedings*, *5*, P54. https://doi.org/10.1186/1753-6561-5-S7-P5

Semagn, K., Beyene, Y., Makumbi, D., Mugo, S., Prasanna, B. M., Magorokosho, C., & Atlin, G. (2012). Quality control genotyping for assessment of genetic identity and purity in diverse tropical maize inbred lines. *Theoretical and Applied Genetics*, *125*, 1487–1501. https://doi.org/10.1007/s00122-012-1928-1

Serna-Saldívar, S. O. (2010). *Cereal grains: Properties, processing, and nutritional attributes* (1st ed.). CRC Press.

Silva, K. J., Guimarães, C. T., Guilhen, J. H. S., Guimarães, P. E. D. O., Parentoni, S. N., Trindade, R. D. S., & Bernardes, C. D. O. (2019). High-density SNP-based genetic diversity and heterotic patterns of tropical maize breeding lines. *Crop Science*, *60*, 779–787. https://doi.org/10.1002/csc2.20018

Silva-Venancio, S., Preciado-Ortiz, R. E., Covarrubias-Prieto, J., Ortíz-Islas, S., Serna-Saldivar, S. O., García-Lara, S., Terrón-Ibarra, A. D., & Palacios Rojas, N. (2019). Identification of superior doubled haploid maize (*Zea mays*) inbred lines derived from high oil content subtropical populations. *Maydica*, *64*.

Singh, B. P., Lenka, D., Lenka, D., & Tripathy, S. K. (2020). Physiological characterization of maize inbred lines under moisture deficit condition. *Journal of Pharmacognosy and Phytochemistry*, *9*, 112–114.

Singode, A., Manivannan, A., Ahmad, B., Srivastava, E., & Mahajan, V. (2017). Heterotic grouping in early maturing Indian maize lines. *International Journal of Agricultural. Research, Innovation and Technology*, *6*, 57–62.

Tomkowiak, A., Bocianowski, J., Kwiatek, M., & Kowalczewski, P. Ł. (2020). Dependence of the heterosis effect on genetic distance, determined using various molecular markers. *Open Life Sciences*, *15*. https://doi.org/10.1515/biol-2020-0001

Tomkowiak, A., Bocianowski, J., Radzikowska, D., & Kowalczewski, P. Ł. (2019). Selection of parental material to maximize heterosis using SNP and SilicoDarT markers in maize. *Plants*, *8*, 349. https://doi.org/10.3390/plants8090349

Vázquez-Carrillo, M. G., Santiago-Ramos, D., Gaytán-Martínez, M., Morales-Sánchez, E., & Guerrero-Herrera, M. (2015). High oil content maize: Physical, thermal and rheological properties of grain, masa, and tortillas. *LWT-Food Science and Technology*, *60*, 156–161. https://doi.org/10.1016/j.lwt.2014.07.043

Wang, C., Hu, S., Gardner, C., & Lübberstedt, T. (2017). Emerging avenues for utilization of exotic germplasm. *Trends in Plant Science*, *22*, 624–637. https://doi.org/10.1016/j.tplants.2017.04.002

Warburton, M. L., Setimela, P., Franco, J., Cordova, H., Pixley, K., Bänziger, M., Dreisigacker, S., Bedoya, C., & MacRobert, J. (2010). Toward a cost-effective fingerprinting methodology to distinguish maize open-pollinated varieties. *Crop Science*, *50*, 467–477. https://doi.org/10.2135/cropsci2009.02.0089

Weiss, W. P., & Wyatt, D. J. (2000). Effect of oil content and kernel processing of corn silage on digestibility and milk production by dairy cows. *Journal of Dairy Science*, *83*, 351–358. https://doi.org/10.3168/jds.S0022-0302(00)74886-7

Wu, Y., San Vicente, F., Huang, K., Dhliwayo, T., Costich, D. E., Semagn, K., Sudha, N., Olsen, M., Prasanna, B. M., Zhang, X., & Babu, R. (2016). Molecular characterization of CIMMYT maize inbred lines with genotyping-by-sequencing SNPs. *Theoretical and Applied Genetics*, *129*, 753–765. https://doi.org/10.1007/s00122-016-2664-8

Yan, J., Shah, T., Warburton, M. L., Buckler, E. S., McMullen, M. D., & Crouch, J. (2009). Genetic characterization and linkage disequilibrium estimation of a global maize collection using SNP markers. *PLOS ONE*, *4*, e8451. https://doi.org/10.1371/journal.pone.000845

Yang, X., & Li, J. (2018). High-oil maize genomics. In J. Bennetzen, S. Flint-Garcia, C. Hirsch, & R. Tuberosa (Eds.), *The maize genome. Compendium of plant genomes* (pp. 305–307). Springer. https://doi.org/10.1007/978-3-319-97427-9_18

Zhang, Y., Xue, R., Zhang, Z., Yang, X., & Shi, H. (2012). Palmitic and linoleic acids induce ER stress and apoptosis in hepatoma cells. *Lipids in Health and Disease*, *11*, 1. https://doi.org/10.1186/1476-511X-11-1

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Gaytán-Pinzón, G. P., Sandoval-Castro, E., Peinado-Fuentes, L. A., Valenzuela-Apodaca, J. P., Cruz-Mendívil, A., & Calderón-Vázquez, C. L. (2022). Genetic diversity of subtropical Double-Haploid maize lines selected for high oil content. *Agronomy Journal*, 1–13. https://doi.org/10.1002/agj2.21153