Revealing core heterotic germplasm and characterizing of maize inbred lines using SSRs

Chunhuai Hu1*, Long Zhang1*, Yongbin Dong1, Qiang Zhou1, Xinyu Li1, Yangliu Ren1, Sailan Zhu1, Lan Zhang1, Zhenzhen Liu1, Yuling Li1**

1College of Agronomy, Henan Agricultural University, Collaborative Innovation Center of Henan Grain Crops, National Key Laboratory of Wheat and Maize Crop Science, 63 Nongye Rd, Zhengzhou, China
*These two authors contributed equally to this manuscript
**Corresponding author: E-mail: yuling_li@126.com

Abstract
Clear knowledge about germplasm characteristics of inbred lines is very important for breeding and genetic research in maize. In this study, a total of 100 SSR markers uniformly distributed on all chromosomes were used to classify 758 lines with broad germplasm types and to reveal the germplasm relationship of 452 introduced lines with unknown pedigree and 91 lines from different sources. Great genetic diversity was observed among these lines, with total alleles 384, alleles per locus 2 – 11, polymorphic information content 0.23 – 0.86, gene diversity 0.26 – 0.88, major allele frequency 0.09 – 0.50. All lines were divided into 10 groups G1–G10, in which all lines with unknown pedigree were included. Five typical heterotic groups were revealed. Three groups (G1, G2, and G7, totally included 695 lines, accounting for 91.69%) and two core heterotic types (G1×G2 and G7×G2) were consistently revealed according to typical lines, the genetic differences among different groups and between pairs of parental lines for 48 important commercial hybrids. Eighteen high-oil and 18 popcorn lines were distinctly classified in G7-9 and G8 groups. The namesake lines from different sources were classified into the same groups for 82 lines, while other 9 such lines were classified into different groups or subgroups. In addition, line improvement among typical groups was obviously reflected and proved to be the effective way in further maize breeding. These results were mainly consistent with inbred pedigree/heterotic groups, previous researches and breeding practice, and could provide useful germplasm information for all lines in further maize breeding and genetic research.

Keywords: maize, inbred line pedigree, germplasm characterization, heterotic type, genetic diversity, SSR markers

Introduction
Maize (Zea mays L) is the largest crop in the world due to its adaptability and productivity (Gerpacio and Pingali 2007; Shiri et al, 2014), which is widely used as food, feed and raw material in industry. Absolute heterosis utilization makes it very important to keep informed on the pedigree information of inbred lines and their genetic relationships for maize breeders and geneticists (Inghelandt et al, 2010). Although line pedigree could reflect the genetic relationships, the results were greatly influenced by artificial selection, domestication and environmental conditions (Yu and Buckler, 2006). In addition, the line pedigree or origin was unclear in some instance. Also, great genetic difference might exist among inbred lines selected from the same population with complicated backgrounds.

Molecular markers had been provided to be a useful method to reflect the genetic diversity among maize lines on DNA level (Li et al, 2002; Lu et al, 2009; Wang et al, 2008; Xie et al, 2008; Yang et al, 2010; Liu et al, 2012). Due to its high polymorphism, repeatability, reliability, easy utilization and low cost, simple sequence repeats (SSR) markers had been widely used in evaluating the genetic diversity and the relationship among different maize germplasms all over the world (Smith et al, 1990, 1997; Phelps et al, 1996; Taramino and Tinge, 1996; Lanza et al, 1997; Benchimol et al, 2000; Warburton et al, 2002; Barbosa et al, 2003; Liu et al, 2003; Reif et al, 2003; Li et al, 2004; Oliveira et al, 2004; Laborda et al, 2005; Reif et al, 2005; Stich et al, 2005; Bruel et al, 2006; Adetimirin et al, 2008; Aguiar et al, 2008; Wang et al, 2008; Warburton et al, 2008; Xie et al, 2008; Kuroda et al, 2009; Yang et al, 2010; Adeyemo et al, 2011; Cui et al, 2011; Terra et al, 2011; Liu et al, 2012; Gupta et al, 2014; Shiri et al, 2014; Fernandes et al, 2015; Pandey et al, 2015; Park et al, 2015). The divergence of lines evaluated by molecular markers could significantly reduce the labor and cost involved in line improvement and cross making (Guimarães et al, 2007; Van Inghelandt et al, 2010; Gupta et al, 2014; Fernandes et al, 2015; Park et al, 2015). Although recently developed SNP markers have some advantages in comparison with SSR markers (Tian et al, 2015; Wu et al, 2016), especially in the rapidly automatic analysis, some previous research had showed that SSR markers were better than SNP markers in clustering germplasm and measuring genetic distance (Jones et al, 2007; Hamblin et al, 2007; Inghelandt et al, 2010). Different numbers of groups were divided in previ-
ous research in China. In general, normal corn germplasms were considered having 5 heterotic groups, including Lancaster (Lancaster Surecrop; Wang et al., 1998, 2008; Xie et al., 2008; Lu et al., 2009; Liu et al., 2012), Reid (Wang et al., 1998, 2008; Xie et al., 2008; Lu et al., 2009; Liu et al., 2012), Tangsipingtou (TSP, improved lines from Sipingtou, a Chinese landrace; Wang et al., 1998, 2008; Xie et al., 2008; Lu et al., 2009; Liu et al., 2012), Lüdahonggu (LRC, Improved lines from «Lüda Reb Cob», a Chinese landrace; Wang et al., 1998; Xie et al., 2008; Lu et al., 2009; Liu et al., 2012) and other germplasms (Wang et al., 1998). Zhao et al. (1999) referred the inbreds selected from American hybrid P78599 as P heterotic group, which was also called PB (Li et al., 2003; Xie et al., 2008; Li and Wang, 2010; Liu et al., 2012) or Tem-tropic I (Yang et al., 2010) in later researches. Correspondingly, lines derived from other modern American hybrids were called PA (Li et al., 2003; Xie et al., 2008; Lu et al., 2009; Li and Wang, 2010). Some researchers referred the lines derived from American BSSS (Iowa Stiff Stalk Synthetic) as BSSS group (Xie et al., 2008; Lu et al., 2009). In addition, the analysis of germplasm diversity and their relationship with normal corn lines was also reported for high-oil corn and popcorn lines (Li et al., 2004; Yang et al., 2010; Cui et al., 2011).

In order to study the germplasm characteristics of inbred lines in Huanghuaihai maize growing region in China, a lot of lines were collected from several seed companies in our laboratory. But most of their pedigrees or origins were unclear, and some lines with the same name had several different sources. To reveal their pedigrees, 100 SSR markers were used to analyze their genotypes together with 307 lines with known pedigrees or origins in this study, including 269 normal corn, 20 high-oil corn and 18 popcorn inbred lines, and both parents for 48 commercial hybrids. Our first objective was to analyse the genetic diversity among all lines and to reveal the core germplasms and heterotic groups. The second objective was to obtain the origin message for unknown lines through their genetic relationship with lines having known pedigree and clear heterotic groups. This result could provide useful information for reasonable utilization of those lines and for further improvement in maize breeding and genetic research.

**Materials and Methods**

**Plant materials**

A panel of 758 maize inbred lines used in the present study included 307 lines with known pedigree or origin (269 normal corn lines, 20 high-oil corn lines and 18 popcorn lines) and 452 normal corn lines with unknown pedigree. Also, 91 line with the same but from 2~8 different sources were included. To reveal the consistency of classification result with breeding practice, the lines representing typical heterotic groups and the parental lines for 48 important commercial hybrids were included. All the inbred lines and their known pedigrees or origins were listed in Supplementary Table 1. They were collected or improved in our laboratory in Zhengzhou Sub-center of Chinese National Maize Improvement Center. The information about pedigrees or origins for most lines could be found in previous studies (Yuan et al., 2001; Li et al., 2002, 2004; Teng et al., 2004; Wang et al., 2008; Xie et al., 2008; Lu et al., 2009; Yang et al., 2010; Cui et al., 2011). All the high-oil lines were selected and provided by China Agricultural University (Jiang et al., 2005; Yang et al., 2010; Cui et al., 2011). And all the popcorn lines were selected in our laboratory.

All collected lines were planted at the Scientific Research and Education Center of Henan Agricultural University near Zhengzhou, Henan, China, in 2014. The rows were 4 m long with 0.67 m spacing between rows. Plots were planted by hand at a density of 75,000 plants ha⁻¹. Standard cultivation management practices were used, and the mixed lines were abandoned.

**SSR genotyping**

A total of 100 SSR primer pairs were chosen from MaizeGDB database (http://www.maizegdb.org) based on repeat units and bin locations. They were uniformly distributed throughout all 10 maize chromosomes (Supplementary Table 2). Total genomic DNA was extracted from fresh leaves from 10 plants of each line at the 6~7 leaf stage using the cetyltrimethylammonium bromide (CTAB) method with modifications according to Saghai-Maroof et al. (1984). The quantity and quality of DNA were estimated using a UV spectrophotometer (NanoVue, GE, UK) and 1% agarose gel.

DNA from the lines was amplified in 20 μl reaction volume with PTC-200 Peltier Thermal cycler (MJ Research). PCR amplification procedure was performed according to the following program: 95° for 5 min, 1 cycle; 95° for 60 s, 57° for 45 s, 72° for 60 s, 35 cycles; 72° for 5 min, 1 cycle. 5 μl of each amplified sample was separated with electrophoresis on 6% agarose gel. The electrophoresis time was adjusted according to each SSR marker’s size. The amplified fragments were detected by the silver staining method as described by Bassam et al. (1991). A ladder (Tiangen) of 50 bp DNA marker was used to determine allele size of each inbred.

For subsequent statistical analysis, polymorphic bands were scored as «bp» typed by SSR markers, and a binary matrix was obtained. In addition, the «bp» typed original data were transformed to several input file format for PowerMarker, Numerical Taxonomy Multivariate Analysis System (Ntsys-pc) and TASSSEL using DataTrans1.0 (Ge and Ren, 2011).

**Genetic diversity analysis**

The genetic similarity (GS) coefficients were calculated based on coefficient for similarity matching using SEL using DataTrans1.0 (Ge and Ren, 2011). The genetic similarity matrix was used to calculate the genetic distance (GD) matrix using 1-GS². The genetic distance matrix was used to find the relationship among the lines using UPGMA and Neighbor joining methods implemented in NTSYSpc v2.2.4 (Rohlf, 2000). The genetic similarity matrix was used to determine the relationship between lines.
ing, the GS matrix was used to construct a dendrogram based on an unweighted pair group methods
using arithmetic averages algorithm (UPGMA).

**Group classification and analysis**

The software package Ntsys-pc was employed to subdivide inbred lines into genetic groups according
to the GS coefficients. The number of groups was set to 10 according to the large quantity of inbred lines
and their known pedigree. The number of alleles, allele frequency, major allele frequency (MAF), gene
diversity and polymorphic information content (PIC) for each group were estimated using the software
package PowerMarker version 3.25 (Liu and Muse, 2005). PIC was calculated using the standard formula
described by Botstein et al (1980). For the groups including large number of lines, a re-sampling strategy
was also used to obtain genetic diversity of each subgroup. To easily understand, the GS among different
groups or subgroups, and between the parental lines for 48 commercial hybrids were converted to genetic
distance (GD) with GD = 1 - GS, accordingly.

**Results**

**Genetic variation and diversity among all inbred lines**

Among all the 758 inbred lines, a total of 384 alleles were detected across all the 100 SSR loci (Sup-
plementary Table 2). The number of alleles per locus ranged from 2 to 11, with an average of 3.84 (Figure
1a). Analysis of all alleles led to identifying 11.46% (44 alleles) as rare (frequency < 0.05), 77.34% (297 alleles)
as intermediate (frequency 0.05 ~ 0.5) and 11.20% (43 alleles) as abundant (frequency > 0.5; Figure 2).

The average polymorphic information content (PIC) value was 0.55, with a range of 0.23 ~ 0.86.
The average gene diversity among them was 0.61, ranging from 0.26 to 0.88. The average major allele
frequency (MAF) was 0.25, with a range of 0.09~0.50 (Figure 1b; Supplementary Table 2).

**Group classification of all inbred lines and cluster analysis**

All the 758 inbred lines were divided into ten groups according to GS = 0.697. A dendrogram using
UPGMA analysis showed that the clustering pattern was complex. This suggested that these inbred lines
represented a broad range of maize germplasms (Supplementary Figure 1). For groups G1 ~ G10, 331,
136, 5, 5, 1, 2, 228, 25, 8, and 5 lines were included, respectively. Obviously, G1, G2, and G7 were three major groups, in which 695 lines were totally included, accounting for 91.69%. Since so many lines were accounted in these three groups, subgroups were also classified for groups G1, G2, and G7.

For the largest group G1, the 331 lines were clustered in 9 subgroups, for which 190, 22, 13, 31, 3
33, 4, 23, and 12 lines were assigned to G1-1 ~ G1-9 subgroups when GS = 0.710, respectively. Accord-
ingly, the 190 lines in group G1-1 were divided into 10 sub-subgroups, 49, 46, 8, 5, 19, 4, 45, 4, 4, and
6 lines were assigned to G1-1-1 ~ G1-1-10 when GS = 0.724, respectively. Xun9058 and improved lines
from Zheng58 were clustered in G1-1-1 including WK858. Four famous lines B73, Ye478, Shen5003,
Tie7922, and other improved lines were included in G1-1-2. Most inbred lines in G1-1-7 were directly
elected from the original America hybrid P78599.

The 136 lines in group G2 were clustered in 8 sub-
groups when GS = 0.735, for which 34, 7, 3, 3, 65,
20, 3, and 1 lines were assigned to G2-1 ~ G2-8 sub-

![Figure 1](image-url)
Table 1 - Summary statistics for all 10 classified major groups.

<table>
<thead>
<tr>
<th>Groups</th>
<th>Line No.</th>
<th>MAF</th>
<th>Allele No.</th>
<th>Gene diversity</th>
<th>PIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>331</td>
<td>0.576</td>
<td>3.680</td>
<td>0.572</td>
<td>0.506</td>
</tr>
<tr>
<td>G2</td>
<td>136</td>
<td>0.522</td>
<td>3.660</td>
<td>0.521</td>
<td>0.464</td>
</tr>
<tr>
<td>G3</td>
<td>5</td>
<td>0.538</td>
<td>2.150</td>
<td>0.521</td>
<td>0.497</td>
</tr>
<tr>
<td>G4</td>
<td>17</td>
<td>0.657</td>
<td>2.460</td>
<td>0.402</td>
<td>0.413</td>
</tr>
<tr>
<td>G5</td>
<td>1</td>
<td>0.502</td>
<td>2.000</td>
<td>0.266</td>
<td>0.500</td>
</tr>
<tr>
<td>G6</td>
<td>2</td>
<td>0.238</td>
<td>1.343</td>
<td>0.171</td>
<td>0.785</td>
</tr>
<tr>
<td>G7</td>
<td>228</td>
<td>0.584</td>
<td>3.800</td>
<td>0.579</td>
<td>0.566</td>
</tr>
<tr>
<td>G8</td>
<td>25</td>
<td>0.453</td>
<td>3.680</td>
<td>0.446</td>
<td>0.506</td>
</tr>
<tr>
<td>G9</td>
<td>8</td>
<td>0.769</td>
<td>3.680</td>
<td>0.446</td>
<td>0.507</td>
</tr>
<tr>
<td>G10</td>
<td>5</td>
<td>0.503</td>
<td>3.680</td>
<td>0.329</td>
<td>0.511</td>
</tr>
<tr>
<td>Mean</td>
<td>75.8</td>
<td>0.534</td>
<td>3.013</td>
<td>0.425</td>
<td>0.526</td>
</tr>
</tbody>
</table>

MAF: major allele frequency; PIC: polymorphic information content.

For the genetic differences among the 10 different groups, the values of GDs were from 0.4072 to 0.9364, which were very high between G1 and G2 (0.9364), between G1 and G7 (0.9268) and between G2 and G7 (0.9177) (Supplementary Table 3). The lowest GDs were observed between G5 and both G6 and G9 (0.4072 and 0.4701), in which only lines with unknown pedigree were included. For the three major subgroups, great genetic differences were also existed within each subgroup. The values of GDs were from 0.7338 to 0.9503 among subgroups G1-1 - G1-9, with the highest between G1-1 and G1-4 and the lowest between G1-5 and G1-7 (Supplementary Table 4). For subgroups G2-1 – G2-8, GDs were from 0.5036 (between G2-4 and G2-8) to 0.8735 (between G2-1 and G2-5) (Supplementary Table 5). In subgroups G7-1~G7-9, the GD values were from 0.5934 (between G7-2 and G7-7) to 0.8971 (between G7-1 and G7-6) (Supplementary Table 6). Even among the 10 sub-subgroups in group G1-1 and G7-1, there also existed great genetic differences. GDs were from 0.7051 (between G1-1-6 and G1-1-9) to 0.9045 (between G1-1-1 and G1-1-2) among sub-subgroups G1-1-1-G1-1-10, and from 0.6754 (G7-1-3 / G7-1-6) to 0.9156 (G7-1-1 / G7-1-2) among sub-subgroups G7-1-1-G7-1-6 (data not showed).

Group classification and genetic distance for pairs of parental lines for commercial hybrids

The parental lines and their types of classified groups for 48 important commercial hybrids were listed in Table 2, including famous hybrids B73×Mo17, Yedan13, Zhendan958, Nongda108, Xundan20, Ludan981, Ludan50, Weike702, Yuyu22, and Xianyu335. Totally, 8 group types were revealed, which were 6 types between different major groups, G1×G2, G7×G2, G1×G3, G1×G7, G4×G2, and G7×G8, and 2 types between subgroups within the same major groups, G1×G1 and G7×G7. Clearly, G1×G2 and G7×G2 were the most important heterotic types, in which 21 and 10 hybrids were included. Hybrids Xundan20 and Weike702 belonged to G1×G2, Zhendan958, Ludan981, and Ludan50 belonged to G7×G2, and Yedan13 belonged to G1/G7×G2. B73×Mo17 belonged to G1×G3. Nongda108,
Xianyu335, and Yuyu22 belonged to G7×G7.

The genetic distances (GD) between each pairs of parental lines for all hybrids were high, with the average of 0.706. but there existed a great variance among different hybrids, from 0.628 for hybrid Ludan981 to 0.779 for hybrid Jiudan48. For the 9 major hybrids Xianyu335, Yuyu22, Ludan50, Weike702, Xundan20, B73×Mo17, Zhengdan958, Yedan13, and Nongda108, their parental GDs were 0.670, 0.674, 0.686, 0.707, 0.710, 0.710, 0.729, 0.745, and 0.753, respectively.

Grouping of inbred lines with unknown pedigree or origin and the namesake lines from different sources

For the total 452 lines with unknown pedigrees, 227, 64, 3, 12, 1, 2, 123, 7, 8, and 5 lines were included in groups G1~G10, respectively (Supplementary Table 1). Also, G1, G2, and G7 included the most lines as the tendency for the total lines, totally 414 lines, accounting for 91.57%. For the three main subgroups, 108, 17, 13, 25, 3, 33, 1, 20, and 7 lines in G1-1 – G1-9, 16, 26, 7, 4, 14, 4, 29, 4, 2, and 2 lines in sub-subgroups G1-1-1 – G1-1-10, 16, 7, 2, 3, 23, 10. and 3 in G2-1–G2-7, and 49, 2, 6, 14, 30, 17, 1, and 3 lines in G7-1 – G7-9, were included, respectively. Four major groups (G5, G6, G9, and G10) and six subgroups (G1-1-6, G1-1-8, G1-3, G2-4, G2-7, and G7-2) only included lines with unknown pedigrees.

For the 91 namesake lines from different sources, 68 lines were from two sources, 12 lines from three sources, 6 lines from four sources, 1 line from five sources, 3 lines from six sources, and 1 line from eight sources (Supplementary Table 1). The namesake lines from different sources for 82 lines were classified into the same groups, while other 9 such lines were classified into different groups.

Discussion

Core germplasm and heterotic groups revealed using molecular markers in maize

The revealing of core germplasms and their heterotic types played an important and basic role in maize successful breeding program. In this study, a total of 758 lines were classified into 10 groups using 100 pairs of SSR markers. This large quantity of lines represented a broad range of germplasm types, which included 720 normal corn, 20 high-oil and 18 popcorn lines. Compared with the results in previous research (Wang et al, 1998, 2008; Zhao et al, 1999; Li et al, 2003; Xie et al, 2008; Lu et al, 2009; Li and Wang, 2010; Yang et al, 2010; Liu et al, 2012), and considered the lines having known pedigree or origin, G1-1-1/G1-1-2, G2-1/G2-5, G7-5, G2-6, and G3 could be considered as the typical Reid, TSPT, PB, LRC, and Lancaster heterotic groups, respectively. In fact, Reid×TSPT, PB×TSPT were the most two important heterotic types at Huanghuaihai maize growing region in China (Li et al, 2003; Xie et al, 2008; Li and Wang, 2010; Yang et al, 2010; Liu et al, 2012).

Herein, large quantity of lines were included in G1, G2, and G7 (totally accounting for 91.69%), and the parental lines for most commercial hybrids belonged to the G1×G2, G7×G2 group types (accounting for 66.67%). Even for the 452 lines with unknown pedigrees, 414 lines were classified into G1, G2, and G7, accounting for 91.59%. In fact, very high GDs were observed between G1 and G2 (0.9364) and between G2 and G7 (0.9177). Clearly, this result showed that lines in groups G1, G2, and G7 were the core germplasm, and G1×G2 and G7×G2 were core heterotic group types, which was highly consistent with previous researches, breeding practice and inbred line pedigrees or heterotic groups.

In fact, Reid×Lancaster, Reid×PB, and Reid×LRC and Lan×LRC were also heterotic types in maize. Herein, G1-1-2×G3 for cross B73×Mo17 corresponded to the typical Reid×Lancaster heterotic group. G1-4×G7-3 for hybrid Lianyu15 was the typical Reid×PB heterotic group, G4×G2-6 for hybrid Nonghua101 was the typical Lan×LRC. G7-1-2×G8 for hybrid DH3119 was Reid×other cross type. The GDs between parental lines for these four hybrids were 0.710, 0.695, 0.690, and 0.727, respectively.

But for the other 12 commercial hybrids, their two pairs of parental lines were all included in the same groups but the different subgroups, for 7 hybrids within G1, and for 5 hybrids within G7. Since so many lines were included within both G1 (331 lines) and G7 (228 lines), great genetic diversity ex-
Table 2 - The types of groups, parental lines and their genetic distance for 48 important commercial hybrids.

<table>
<thead>
<tr>
<th>Group type</th>
<th>Subgroup type</th>
<th>Hybrid/Cross</th>
<th>Female inbred line ID</th>
<th>Male inbred line ID</th>
<th>GD</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1×G2</td>
<td>G1-1-1×G2-1</td>
<td>Xundan20/Xun908×Xun92-8</td>
<td>M15/5/88/174</td>
<td>M2/6/305/608/607</td>
<td>0.710</td>
</tr>
<tr>
<td>G1-1-1×G2-1</td>
<td></td>
<td>Xundan22/Xun908×Xun926</td>
<td>M15/5/18/174</td>
<td>M6/06</td>
<td>0.698</td>
</tr>
<tr>
<td>G1-1-1×G2-1</td>
<td>Lutian002/Zheng 58 Ga×Xun901</td>
<td>M103</td>
<td>M14/22</td>
<td></td>
<td>0.735</td>
</tr>
<tr>
<td>G1-1-1×G2-1</td>
<td>Yudan19×Xun926</td>
<td>M65/652</td>
<td>M606</td>
<td></td>
<td>0.722</td>
</tr>
<tr>
<td>G1-1-1×G2-1</td>
<td>Yudan9×Xun77</td>
<td>M65/650</td>
<td>M602</td>
<td></td>
<td>0.721</td>
</tr>
<tr>
<td>G1-1-1×G2-1</td>
<td>Xindan20/Xin 358×Xin77</td>
<td>M63/330</td>
<td>M602</td>
<td></td>
<td>0.859</td>
</tr>
<tr>
<td>G1-1-1×G2-1</td>
<td>Fenglan08×L505×Chnag7-2</td>
<td>M108</td>
<td>M20/486</td>
<td></td>
<td>0.688</td>
</tr>
<tr>
<td>G1-1-1×G2-1</td>
<td>Anyu574×Chnag7-2</td>
<td>M728</td>
<td>M20/486</td>
<td></td>
<td>0.688</td>
</tr>
<tr>
<td>G1-1×G2</td>
<td>G1-1×G2-1</td>
<td>Xindan22/Xun9058×Xun92-8</td>
<td>M10/5/88/174</td>
<td>M2/6/305/608/607</td>
<td>0.710</td>
</tr>
<tr>
<td>G1-1×G2</td>
<td>G1-2×G2-1</td>
<td>Xindan22/Xun9058×Xun926</td>
<td>M10/5/18/174</td>
<td>M6/06</td>
<td>0.698</td>
</tr>
<tr>
<td>G1-1×G2</td>
<td>G1-2×G2-1</td>
<td>Yudan9×Xun77</td>
<td>M65/650</td>
<td>M602</td>
<td>0.721</td>
</tr>
<tr>
<td>G1-1×G2</td>
<td>G1-2×G2-1</td>
<td>Xindan20/Xin 358×Xin77</td>
<td>M63/330</td>
<td>M602</td>
<td>0.859</td>
</tr>
<tr>
<td>G1-1×G2</td>
<td>G1-2×G2-1</td>
<td>Fenglan08×L505×Chnag7-2</td>
<td>M108</td>
<td>M20/486</td>
<td>0.688</td>
</tr>
<tr>
<td>G1-1×G2</td>
<td>G1-2×G2-1</td>
<td>Anyu574×Chnag7-2</td>
<td>M728</td>
<td>M20/486</td>
<td>0.688</td>
</tr>
<tr>
<td>G1-1×G2</td>
<td>G1-2×G2-1</td>
<td>Xindan22/Xun9058×Xun926</td>
<td>M10/5/18/174</td>
<td>M6/06</td>
<td>0.698</td>
</tr>
</tbody>
</table>

GD: genetic distance.

Both groups. This could be shown from the GDs among different subgroups G1-1 – G1-9 (0.7338 – 0.9503) in group G1 and G7-1 – G7-9 (0.5934 – 0.8971) in G7, and among different sub-subgroups G1-1-1 – G1-1-10 (0.7051 – 0.9045) in G1-1 and G7-1-1 – G7-1-6 (0.6754 – 0.9156) in G7-1. In fact, the GDs between their parental lines were all high, from 0.669 (G1-1-2×G1-2 for Anyu13) to 0.779 (G1-4×G1-1-10 for Yu82) in G1, and from 0.670 (G7-1-3×G7-4 for Xianyu335) to 0.753 (G7-5×G7-1-6 for Nongda108) in G7. In addition, from the pedigrees of those parental lines, G1-1-2×G1-1-10 for hybrid Yudan998, G1-1-3×G1-9 for hybrid Pudan6, G1-4×G1-1-10 for hybrid Jiudan48, and G7-5×G7-1-6 for hybrid Xianyu335 were all corresponded to the typical Reid×Lancaster heterotic group, while G1-1-2×G1-2 for hybrid Anyu13, G1-1-7×G1-1-9 for hybrid Jinhai5, G1-4×G1-1-7 for hybrid Liyu16, G7-7×G7-3 for hybrid Zheng93-1, G7-7×G7-5 for hybrid Yuyu22, and G7-5×G7-1-6 for hybrid Nongda108 were all the typical Reid×PB heterotic group. G1-1-2×G1-7 for hybrid Anyu12 was the typical Reid×PB/LRC heterotic group.

In previous research, different results were obtained about the correlation between GD and the performance of hybrids (Smith et al, 1990; Stuber et al, 1992; Lanza et al, 1997; Bruel et al, 2006; Guimarães et al, 2007; Fernandes et al, 2015). But the correspondence of GDs among different groups with heterotic types was commonly observed in maize breeding in China (Zhao et al, 1999; Li et al, 2000; Yuan et al, 2001; Teng et al, 2004; Liu et al, 2012). Therefore, it was very important to classify groups for all inbred lines and to reveal their heterotic types.

Further germplasm improvement at Huanghuaihai maize growing region in China

62 ~ M18
Maydica electronic publication - 2017
Successive breeding depended on the continuous germplasm improvement, which needed large genetic diversity among different germplasms. Due to domestication and modern breeding, the genetic diversity in maize has been increasingly narrowed compared with its wild relatives (Choukan et al., 2004; Shiri et al., 2014). Maize breeding faces unique challenges resulting from the narrow genetic background of commercial cultivars (Choukan et al., 2004; Shiri et al., 2014; Park et al., 2015). Narrow genetic diversity is problematic when breeding for adaptation to biotic and abiotic stress (Shiri et al., 2014). In order to broaden genetic variation for use in future maize breeding, the genetic relationships among all maize germplasms should be investigated.

In our present study, the complex dendrogram with 10 major groups and many subgroup was observed, which brought valuable information on genetic diversity among all the lines. Although the allele frequencies, PIC values, and gene diversity values for all groups showed that those inbred lines contain extensive genetic variation, several core lines or hybrids were clearly used extensively in modern hybrid breeding, especially for the three major groups (G1, G2, and G7), such as Chang7-2, P78599. Therefore, enhancing the genetic base of elite germplasm through introgression of exotic germplasm was important to exploit heterotic potential (Ron-Parra and Hallauer, 1997). In addition to the direct utilization for newly introduced germplasm, such as America hybrid P78599, the successfully indirect way should be adapted according to their heterosis with the typical lines in major heterotic groups.

In addition, the cross improvement for lines between different typical groups could clearly be seen in this study, such as TSPT×LRC (in G2), TSPT×Lancaster (in G2 and G3), Reid×PB (in G1 and G7), etc. Especially, both groups G1 and G7 included many improved lines from the cross of Reid×PB lines. Five subgroups in G1 (G1-1-1, G1-1-7, G1-2, G1-4, G1-8) and two subgroups in G7 (G7-1-1, G7-1-2) consistently included Reid×P78599 related germplasm. The result of group classification for namesake lines from different sources.

Group classification for lines with unknown pedigree and their utilization

Simultaneous cluster analysis for lines in the absence of pedigree with lines having clear pedigree or origin could bring valuable information for their efficient utilization. In this study, 452 lines with unknown pedigree or origin were included, which were classified into all the 10 major groups. Their importance in further germplasm improvement could be seen from their classification result and their genetic relationship with lines having known pedigrees. The first way was to use them to improve the lines with known pedigree or origin within the same group, and to make crosses between lines with known pedigree or origin in their corresponding heterotic groups. For the 10 subgroups, G1-1-6, G1-1-8, G1-3, G1-6, G2-2, G2-3, G2-4, G2-7, and G7-2, in which only lines with unknown pedigree or origin were included, those lines could be used to improve the lines with known pedigree or origin within the same major groups.

For the four groups (G5, G6, G9, and G10) only included lines with unknown pedigree, lines in those groups could be utilized to improve lines in any other groups according to their GDs with other groups. Considering the great improvement potential from G1×G7, it would be favorable to use them to improve lines in G2.

Group classification for namesake lines from different sources

The inconsistency of group classification with pedigree was commonly found in several previous researches (Li et al., 2004; Wang et al., 2008; Xie et al., 2008; Inghelbandt et al., 2010; Yang et al., 2010; Cui et al., 2011). Several possible reasons were considered as follows: i) lines developed from germplasms with complicated backgrounds (such as hybrids, landraces and populations); ii) differences in criterion and environment for selection; iii) sampling effects and genetic drift; iv) misnamed and miscellaneous lines. In all such cases, it was necessary and very useful to use molecular makers to obtain clear knowledge about genetic relationship among lines both in breeding (inbred development and cross making) and in theoretical research.

In this study, 91 lines with the same name but from 2–8 different sources were included. Their re-
ality should be made clear in ahead of utilization. The result of classification showed that most namesake lines from different sources were clustered in the same group, totally accounting for 90.11%. But about 9.89% such lines were clustered in different groups. Unfavorable use of such lines might result in the germplasm confusion between heterotic groups, which would lead to fruitless labor in further line improvement and cross making.

Two major reasons might lead to namesake lines from various sources being classified into different groups. The first one might be line nickname, for which the line name was not consistent with its real germplasm. In this case, such lines might be classified into different major groups, such as W26 (M181/591, G1-1-7/G7-5), ZWM1 (M347/435, G4/G1-2), BX111 (M535/662, G7-6/G1-2), IAS09 (M538/664, G7-6/G1-2), KY7M (M539/666, G7-6/G1-2), IAS01 (M184/537/663, G1-1-2/G7-1-1/G1-1-2), and ZXW2 (M106/350/391, G4/G4/G7-4) in this study. The second reason might be the late selection during seed keeping in different units, in which such lines might be classified into different subgroups within the same major groups, such as D340 (M69/521, G1-2/G1-8) and Ji223 (M144/145/522/523, G7-1-1/G7-3/G7-1-1/G7-3). So, line reality keeping was also important in practical breeding or genetic research.

Conclusions

To conclude, this study revealed three core groups and two core heterotic patterns through the classification of a large quantity of lines with broad germplasm types. Accordingly, the 452 lines with absence of pedigree and the 91 lines with 2–8 different sources were clearly clustered. This result could provide useful information for the efficient utilization of all lines in further line improvement and cross making in maize breeding.

Acknowledgements

We greatly thank China Agricultural University for providing us the high-oil maize lines.

References


Gerpacio VR, Pingali PL, 2007. Tropical and subtropical maize in Asia: Production systems, constraints and research priorities. Mexico, DF


Jiang HY, Chen SJ, Gao LF, Xing JM, Song TM, Dai...
core heterotic germplasm and characterizing of maize inbred lines using SSRs


Li XH, Yuan LX, Li XH, Zhang SH, Li MS, Li WH, 2003. Heterotic grouping of 70 maize inbred lines by SSR markers. Scientia Agricultura Sinica 36: 622-627


Phelps TL, Hall AE, Buckner B, 1996. Microsatellites repeat variation within the γT gene of maize and teosinte. The J Heredity 87: 396-399


in European elite maize germplasm investigated with SSRs. Theor Appl Genet 111: 723-730


De Freitas Terra TF, Wiethölter P, de Souza Almeida CC, dos Anjos e Silva SD, Bered F, de Melo Sereno MJC, Neto JFB, 2011. Genetic variability in maize and teosinte populations estimated by microsatellites markers. Cienc Rural 4: 205-211


