GENETIC VARIATION WITHIN AND AMONG POPULATIONS OF MAIZE INBRED B37 REVEALED BY SSR MARKERS

A. Kostova1, E. Todorovska1, N. Christov1, K. Hristov2, J.-M. Jacquemin3, Y. Muhovski3, A.I. Atanassov1
AgroBioInstitute, Sofia, Bulgaria1
Institute of Forage Crops, Pleven, Bulgaria2
CRAW, Department Biotechnology, Gembloux, Belgium3

ABSTRACT
Historically important public US inbred line B37 continues to play an important role in maize (Zea mays L.) improvement in many different breeding programs. Since its original release it has undergone numerous seed increases in diverse programs that may lead to genetic changes due to artificial selection, drift mutations etc. Our objective was to estimate the level of genetic diversity among and within two B37 inbred populations from different sources in Bulgaria (Maize Research Institute (MRI) - Kneja and Institute of Forage Crops-Pleven) using 10 SSR markers. The set of SSR markers detected in average 1.9 different allelic variants per locus for two B37 populations, with mean expected heterozygosity (He)=0.170 for population from MRI-Kneja and 0.046 for populations from IFC - Pleven. Cluster analyses indicated that both populations could be distinguished from each other. The observed new allelic variants at microsatellite loci bnlg1917, dupssr21 and bnlg1017 could be successfully used as markers for distinguishing of the maintained by different breeding programs populations of the inbred B37.

Introduction
The most widely recognized and utilized inbreds in the USA fall into one of three heterotic groups (Reid Yellow Dent, Lancaster Surecrop and miscellaneous heterotic groups). The most commonly used heterotic group is Reid Yellow Dent, of which the most utilized source population is Iowa Stiff Stalk Synthetic (BSSS). The Lancaster Surecrop group consists of material that is more flinty (harder starch in endosperm) in nature than Reid Yellow Dent (1), (2), (3). The other miscellaneous heterotic category comprises of inbred lines developed from crosses between the two major heterotic groups, between adapted and exotic germplasm, or derived from distantly related material such as Minnesota 13 (4).

Historically important public US inbred line B37 continues to play an important role in maize improvement in many different breeding programs all over the world. For example, among inbreds available from US foundation seed companies in 1999, only 82 out 381 inbreds (22%) had genetic backgrounds other than eight widely used inbreds: B14, B37, B73, B84, Mo17, C103, Oh 43 and H99 (MBS Inc. 1999). In Bulgaria B37 was introduced in 1969 and ever since have been maintained by different breeding programs.

Once released, inbred lines have been maintained for decades through periodic seed increases in breeding programs and at germplasm repositories. Effects of artificial selection regimes, natural selection in maintenance environments, drift, migration
(contamination), and mutations could lead to genetic changes (5), (6), (3). Such genetic changes would be influenced by the frequency of regeneration, methods used for regeneration, unintentional outcrossing, and addition of newer versions of the same inbred from other source.

Variation among and within inbreds can be investigated by means of phenotypic and genetic measurements. Quantitative character studies of a long-time inbred lines detected genetic changes larger than those expected by breeders (7), (8), and (9). These changes could affect yields of hybrid combinations after several cycles of regeneration. Genetic changes were not constant across inbred lines and sources, and were thought to have resulted from residual heterozygosity, mutation, or a combination of both (8). Busch and Russell (10) studied 31 sublines from two inbred lines and reported biological changes in more than one character in a subline. They attributed such changes either to pleiotropism or mutation at more than one locus. Russell and Vega (9) reported that genetic changes either occurring in inbred lines were independent and occurred continuously. The studies by Busch and Russell (10) and Russel and Vega (9) revealed that mutations had little practical importance, particularly for expressed characteristics like yield, but were biologically significant and may have negative implications for usage after years of regeneration. Higgs and Russell (11) showed that inbred lines from six different sources had significant differences in traits such as plant height, silking date, ear height and grain yield. This was attributed to maintenance, residual heterozygosity, or mutation.

Smith and Smith (12), (13) studied associations among 18 Lancaster and BSSS-derived inbred lines using electrophoretic, allozymic, and pedigree data. They were able to identify uniquely 79% of the Lancaster lines.

In a study involving 148 US maize inbred lines, Mumm and Dudley (3) used 46 restriction fragment polymorphisms (RFLP) markers to cluster all the inbred lines in two major heterotic groups. They were also able to identify subgroups within the major heterotic groups. Although some discrepancies were observed, their data tended to correlate with pedigree data.

More recently, random amplified polymorphic DNA (RAPD), simple sequence repeats (SSR) and amplified fragment length polymorphism (AFLP) analyses have been used in maize diversity studies (14), (6). In a study of 33 inbred lines, SSRs produced twice as much information as AFLPs and RAPDs, and 40% more than RFLPs in terms of numbers of alleles per locus (14). The main advantages of SSR markers are their high information content (PIC), codominant inheritance, locus specificity, extensive genome coverage and simple detection.

Many studies have reported genetic diversity and relatedness of maize inbred lines at molecular level (15), (16), (17), (5), (18). However, only few investigations were devoted to genetic variation within and among identically named inbred lines maintained by different programs Gethi (4). This approach is complementary to phenotypic measures in quantifying genetic changes because it may show variation in DNA that is not be phenotypically expressed. The objective of this study was to estimate the level of genetic diversity both among and within inbred line B37 from two different sources by means of SSR markers.

Materials and Methods

Plant material and DNA isolation

Seeds of B37 maize inbreds representing two populations were obtained from different sources – gene banks of the Institute of Forage Crops (IFC) - Pleven and Maize Research Institute (MRI) - Kneja, Bulgaria. Both populations were imported to Bulgaria from two US sources and maintained
in Bulgaria by different breeders. The population named Population 2 was imported 1969 from Illinois while Population 1 had arrived 15-20 years later from a different source. The two B37 populations were selected and analyzed using SSR markers.

Leaf tissue samples were obtained from 10 plants of the first population (P1) and 14 plants from the second population (P2). DNA was isolated from 2 weeks old seedlings employing a modified CTAB procedure (19).

**Microsatellite markers**

Ten SSR markers were chosen from MaizeGDB (www.maizegdb.org) (**Table 2**). The detailed protocol for maize SSR amplification is available at The Asian Maize Biotechnology Network (AMBIONET) (http://www.cimmyt.org/ambionet). Fragment analysis was carried out on ALF ExpressII sequencer and fragment sizes were determined using the computer program Allele Locator v. 1.03 (Amersham Biosciences).

**Polymerase chain reaction and fragment analyses**

PCR reaction was conducted in a volume of 10 μl each containing 1x PCR buffer (500mM KCl, 15 mM MgCl₂ and 100mM Tris-HCl (pH 9.0), 0.25μM of each fluorescently labelled-Cy5 forward and unlabelled reverse primers and 0.2 mM dNTPs, 0.4U of Taq polymerase (Amersham Biosciences), and 100ng DNA template.

The amplification reaction consisted of a denaturing step of 2 min at 94 ºC, followed by 30 cycles began with 94 ºC for 2 min, annealing reaction of 1 min at 56 ºC and terminated at 72 ºC for 1 min.

The detailed protocol for maize SSR amplification is available at The Asian Maize Biotechnology Network (AMBIONET) (http://www.cimmyt.org/ambionet).

Fragment analysis was carried out on automated laser fluorescence (ALF ExpressII) sequencer (Amersham Biosciences). 1/20 or 1/10 of each amplification reaction was mixed with 2 μl of formamide loading buffer containing labeled internal size standards. After denaturation at 95 ºC for 5 min, and cooling down on ice, samples were loaded onto a standard sequencing gel (Reprogel™, Amersham Biosciences).

Fragment sizes were determined using the computer program Allele Locator v. 1.03 (Amersham Biosciences) by comparison with internal size standards with length 50, 100, 150, and 200 bp amplified from pUC 19 plasmid. To reduce the variance in the estimate of fragment sizes between runs, several control samples corresponding to the same SSR locus were repetitively run in all gels.

**Data Analysis**

Allele frequencies, expected and observed heterozigosity were calculated using the software PowerMarker v. 3.25 (20).

The MICROSAT software (21) was used for calculation of genetic distances in using log transformed proportion of shared alleles (Dps). The distance matrix obtained from MICROSAT was processed with FITCH from the PHYLIP package (22) and the resulting phenogram was visualized using TREEVIEW (23). In FITCH the J option was used to randomize the input order of samples.

**Results and Discussion**

In this study, we compared two B37 populations (P1 and P2) obtained from different sources and maintained by different programs. Twenty four individual plants (10 from P1 and 14 from P2) were analyzed at 10 SSR loci. A total of 19 alleles were detected across all loci in both inbred populations (**Table 1**). The appearance of null alleles was observed in one out of 10 studied microsatellite loci in this study which was confirmed by 2nd amplification/or using other DNA sample from the same plant material.

The number of alleles per locus varied between 1 for bnlg1237, umc1014 and phi059 to 3 for dupssr21 and bnlg1917 with average number of 1.9 alleles per lo-
TABLE 1

Chromosomal location, repeat type, allele number, expected (He) and observed heterozygosity (Ho) estimates of the 10 microsatellite loci studied

<table>
<thead>
<tr>
<th>No</th>
<th>SSR locus</th>
<th>Bin location</th>
<th>Repeat</th>
<th>No of alleles</th>
<th>He</th>
<th>Ho</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total P1 P2</td>
<td>Total P1 P2</td>
<td>Total P1 P2</td>
</tr>
<tr>
<td>1</td>
<td>bnlg1484</td>
<td>1.03</td>
<td>(AG)19</td>
<td>2 2 1</td>
<td>0.451</td>
<td>0.172</td>
</tr>
<tr>
<td>2</td>
<td>umc1035</td>
<td>1.06</td>
<td>(CT)19</td>
<td>2 2 1</td>
<td>0.148</td>
<td>0.297</td>
</tr>
<tr>
<td>3</td>
<td>bnlg1017</td>
<td>2.02</td>
<td>(AG)18</td>
<td>2 1 1</td>
<td>0.466</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>dupssr21</td>
<td>2.05</td>
<td>(AG)10</td>
<td>3 1 2</td>
<td>0.629</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>bnlg1917</td>
<td>4.10</td>
<td>(AG)26</td>
<td>3 2 1</td>
<td>0.532</td>
<td>0.362</td>
</tr>
<tr>
<td>6</td>
<td>bnlg1237</td>
<td>5.05-5.06</td>
<td>(AG)29</td>
<td>1 1 1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>bnlg2305</td>
<td>5.7</td>
<td>(AG)21</td>
<td>2 2 1</td>
<td>0.240</td>
<td>0.414</td>
</tr>
<tr>
<td>8</td>
<td>umc1014</td>
<td>6.04</td>
<td>(GA)12</td>
<td>1 1 1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>bnlg1732</td>
<td>6.05</td>
<td>(AG)15</td>
<td>2 2 1</td>
<td>0.270</td>
<td>0.451</td>
</tr>
<tr>
<td>10</td>
<td>phi059</td>
<td>10.02</td>
<td>ACC</td>
<td>1 1 1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td></td>
<td></td>
<td>1.9 1.5 1.1</td>
<td>0.274</td>
<td>0.170</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td></td>
<td>19 15 11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

cus (Table 1). The average number of alleles obtained in the present study is higher than those reported in previous maize diversity studies. Gethi (4) observed SSR variations in six important U.S maize inbred lines (B37, CM105, Mo17, Oh43, W153R and Wf19) with 44 SSR markers and reported an average 1.38 alleles per locus in inbred populations from 5 different sources maintained under the same name.

*Ho* within B37 inbreds in our study was low (0.063) as would be expected from the inbred nature of the materials studied. The *Ho* value for our sample of maize inbreds including both B37 populations was higher as compared to the mean *Ho* value of 0.0081 reported by Gethi (4) within six inbreds investigated with 44 SSRS. Our study showed that the mean *Ho* of Population 2 (0.007) is 20 times less than that of Population 1 (0.14). This information is crucial for breeders when uniform and genetically stable materials have to be developed, maintained and further used in breeding programs.

The *He* value for the 24 individual plants including both B37 populations studied at 10 microsatellite loci varied from 0 in bnlg1237, umc1014 and phi059 loci to 0.466 for locus umc1017 with an average 0.274. Among the six inbreds studied by Gethi (4) the lowest *He* (0.0317) was reported for B73 inbreds and highest for Wf9 (0.1139) while the mean *He* for all inbreds was 0.0672.

In order to study the genetic variation among B37 populations originating from MRI-Kneja and IFC-Pleven, the *He* values were separately calculated for each population.

*He* value for Population 1 (P1) varied from 0 in loci bnlg1017, dupssr21, bnlg1237, umc1014 and phi059 to 0.451 in the locus bnlg1732 with a mean value of 0.170. The gene diversity in Population 2 was lower than that of Population 1. Among the 10 microsatellite loci studied in Population 2 allele variation was observed only at the locus dupssr21 with *He*=0.064.

Three SSR loci - umc1014, phi059 and bnlg1237 were monomorphyc between the two populations. In addition to above mentioned loci, the Population 1 was also monomorphyc at SSR loci dupssr21 and bnlg1017, while the Population 2 was monomorphyc for all studied SSRs with the
exception of dupssr21. (Table 2)

The observed new allelic variants at microsatellite loci bnlg1917, dupssr21 (Fig. 1) and bnlg1017 could be successfully used as markers for distinguishing of the maintained by different breeding programs populations of the inbred B37.

In order to further characterize the structure and grouping of 24 analyzed B37 inbreds a dendrogram was build. The resulting dendrogram (Fig. 2) shows that plants from Population 1 and 2 were grouped into two different clusters. The dendrogram clearly shows that the Population 1 is more diverse than the second one. Most of Population 1 (B37-1- ) plants, with the exception of B37-1-8 and B37-1-9 which show identical genome composition, are grouped in several different subclusters. The individuals of Population 2 are separated into two main subclusters. Each of subclusters is composed of plants having identical genome composition at the 10 microsatellite loci studied. The grouping of Population 2 plants into two subclusters is due to the presence of two different allelic variants at dupssr21 locus (Table 2). Only one plant (B37-2-10), which was heterozygous at dupssr21, was separately placed forming the root of both subclusters thus
Fig. 1. Electrophoregram of PCR products amplified with primer pairs corresponding to the microsatellite locus dupssr21 in two B37 inbred populations indicating the higher level of homogeneity of Population 2.

The study showed that the origin and selection practices of inbred lines are potentially important source of genetic variation that may influence the performance of produced hybrids after years of inbred line maintenance with periodic seed increases. Our results also demonstrate that the selection by phenotype is not sufficient to preserve the genome composition of an initially released inbred during long term maintenance including numerous seed regenerations. In addition to phenotype selection the use of SSR markers will allow for better preservation of the authentic inbreds and will reduce the heterogeneity of the breeding materials.

**REFERENCES**

Fig. 2. Dendrogram of two B37 populations genotyped at 10 microsatellite loci.