

Genetic Diversity in CIMMYT Nontemperate Maize Germplasm: Landraces, Open Pollinated Varieties, and Inbred Lines

M. L. Warburton,* J. C. Reif, M. Frisch, M. Bohn, C. Bedoya, X. C. Xia, J. Crossa,
J. Franco, D. Hoisington, K. Pixley, S. Taba, and A. E. Melchinger

ABSTRACT

CIMMYT is the source of improved maize (*Zea mays* L.) breeding material for a significant portion of the nontemperate maize growing world. Landraces which did not serve as sources for improved maize germplasm may contain untapped allelic variation useful for future breeding progress. Information regarding levels of diversity in different germplasm would help to identify sources for broadening improved breeding pools and in seeking genes and alleles that have not been tapped in modern maize breeding. The objectives of this study were to examine the diversity in maize landraces, modern open pollinated varieties (OPVs), and inbred lines adapted to nontemperate growing areas to find unique sources of allelic diversity that may be used in maize improvement. Twenty-five simple sequence repeat markers were used to characterize 497 individuals from 24 landraces of maize from Mexico, 672 individuals from 23 CIMMYT improved breeding populations, and 261 CIMMYT inbred lines. Number of alleles, gene diversity per locus, unique alleles per locus, and population structure all differ between germplasm groups. The unique alleles found in each germplasm group represent a great reservoir of untapped genetic resources for maize improvement, and implications for hybrid breeding are discussed.

M.L. Warburton, C. Bedoya, J. Crossa, K. Pixley, and S. Taba, The International Maize and Wheat Improvement Center (CIMMYT, Int). Apdo. Postal 6-641, 06600 Mexico D.F., Mexico; J.C. Reif, A.E. Melchinger, and M. Frisch, Institute of Plant Breeding, Seed Science and Population Genetics, Univ. of Hohenheim, 70593 Stuttgart, Germany; M. Bohn, Crop Science Dep., 101 Turner Hall, Univ. of Illinois, 1102 S. Goodwin, Urbana, IL 61801; J. Franco, Facultad de Agronomía, Universidad de la República, Av. Garzón 780 CP 12900, Montevideo, Uruguay; D.A. Hoisington, the International Crops Research Institute for the Semi-arid Tropics, Patancheru, Hyderabad 502-324 Andhra Pradesh, India; X.C. Xia, Institute of Crop Breeding and Cultivation, Chinese Academy of Agricultural Sciences, Zhongguancun South Street 12, 100081, Beijing, China. Received 25 Feb. 2007. *Corresponding author (mwarburton@cgiar.org).

Abbreviations: CMLs, CIMMYT maize inbred lines; G_{ST} , coefficient of gene differentiation; H_s , gene diversity between individuals within each population; H_t , total gene diversity; OPVs, open pollinated varieties; PCoA, principal coordinate analysis; RFLP, random fragment length polymorphism; SSRs, simple sequence repeats or microsatellites.

MAIZE (*ZEA MAYS* L.) was domesticated about 9000 yr ago in Mexico from tropical teosinte (*Zea mays* ssp. *parviglumis*) (Beadle, 1939; Doebley, 2004). Molecular analyses suggest a single domestication event (Matsuoka et al., 2002) that reduced the diversity present in maize compared to teosinte (Eyre-Walker et al., 1998; Vigouroux et al., 2002). Following domestication, mutation generated new alleles, while recombination created novel allele combinations. Furthermore, postdomestication gene flow from teosinte presumably increased the existing genetic base of maize (Doebley, 2004). The genetic variation of domesticated maize populations can be reduced or restructured by genetic drift

Published in Crop Sci. 48:617-624 (2008).

doi: 10.2135/cropsci2007.02.0103

© Crop Science Society of America

677 S. Segoe Rd., Madison, WI 53711 USA

All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher.

and selection, both natural and artificial, by early farmers. This has eventually resulted in a large number of landraces adapted to the specific environmental conditions of their habitats and desired uses by humans.

During the past century, the existing landraces were the bases for developing modern open pollinated varieties (OPVs). Open pollinated varieties have begun to replace landraces in the developing world; although worldwide about half of the nontemperate maize-growing area is still sown with landraces, this is a decreasing trend (Taba et al., 2005). In the last 20 yr, hybrids are now replacing the OPVs, such that 65% of the global acreage was sown to hybrids in 1999 (Aquino et al., 2000). The International Maize and Wheat Improvement Center (CIMMYT) is the source of maize breeding material for a significant portion of the nontemperate maize growing world. During the past 40 yr, CIMMYT has had a tremendous impact on maize breeding and production in subtropical and tropical environments (Vasal et al., 1999; Morris, 2001). In developing countries, 59% of public and 58% of private maize varieties (hybrids and OPVs) sold in 1998 contained CIMMYT or CIMMYT-related maize germplasm. CIMMYT inbred lines (CMLs) and OPVs are bred to contain considerable diversity and are then taken by National Agriculture Research Programs and selected for further adaptation in their own particular environment(s). CIMMYT inbred lines are chosen from OPVs and other breeding populations, which were in turn created by mixing many different landrace varieties. Landraces which did not serve as sources for improved maize germplasm may contain untapped allelic variation useful for future breeding progress.

Changes in genetic diversity following the replacement of landraces by improved germplasm and during ongoing hybrid breeding have been investigated based on molecular markers for U.S. and European germplasm. All surveys revealed a significant reduction in diversity (Dubreuil et al., 1999; Tenaillon et al., 2001; Duvick et al., 2004; Reif et al., 2005). Diversity present in subtropical and tropical improved germplasm and landraces of maize

Table 1. Summary of average simple sequence repeat (SSR) statistics of the three germplasm groups. CIMMYT maize inbred lines (CMLs) include all lines grouped into populations consisting of tropical (white or yellow), subtropical, midaltitude, and highland germplasm.[†]

Group	No. of alleles	Hs	Ht	G _{ST}	Unique alleles
Landraces	7.84a [‡]	0.48a	0.61a	0.21a	1.4
Populations	8.44a	0.54b	0.61a	0.11b	1.9
CMLs	8.52a	0.61c	0.65b	0.06c	1.3

[†]Hs is the gene diversity between individuals within each population (landrace, OPV, or CML grouped into populations); Ht is the total gene diversity across all populations within each group; G_{ST} is the relative differentiation of the populations within each group, and unique alleles appear in only one landrace, population, or CML.

[‡]Significance tests: values followed by the same letters are not different at the 0.01 probability level (Wilcoxon signed rank test).

has also been measured (Reif et al., 2004; Xia et al., 2004, 2005; Reif et al., 2006). These studies suggest that traditional farmer's landrace varieties may be a good source of new allelic diversity for improving the diversity of the CIMMYT (and other) improved inbred lines. A study of the levels of latent genetic diversity in inbred lines, OPVs, and landraces from the center of origin of maize (and thus one of the most important center of diversity as well) will show the potential to use landraces to identify unique allelic diversity for inbred line improvement.

The objectives of this study were to examine the levels of diversity and population structure in maize landraces, modern OPVs, and inbred lines adapted to the tropics, subtropics, midaltitude, and highlands of nontemperate growing areas, and see if significant sources of allelic diversity exist in the germplasm groups for future maize improvement.

MATERIALS AND METHODS

Plant Materials

A total of 497 individuals from 23 landraces of maize from Mexico were chosen to represent the diversity of germplasm and agro-ecosystems from the center of maize domestication. Detailed information about the landraces is published elsewhere (Reif et al., 2006), with the exception of Jala and Conico Norte, which do not appear in this study because they did not group into a single population in the study by Reif et al. (2006). Studied as well were 672 individuals from 23 OPVs and improved breeding populations (collectively referred to as OPVs) of the CIMMYT maize breeding program, including OPVs adapted to tropical, subtropical, and temperate areas. Detailed information is published in Reif et al. (2004). Finally, 261 CMLs adapted to tropical (155 inbreds), subtropical (73 inbreds), midaltitude (22 inbreds), and highland (11 inbreds) growing conditions were chosen for this study and a detailed description is published in Xia et al. (2004, 2005). All germplasm can be found in Supplementary Table 1.

Molecular Marker Analysis

The 1430 individual plants were genotyped in the Applied Biotechnology Center at CIMMYT-Mexico. Details of the protocols were described in Warburton et al. (2002). Briefly, DNA was extracted by the CTAB method and 25 simple sequence repeat (SSR) loci were amplified by polymerase chain reaction with fluorescent-labeled primers (Supplementary Table 2). Twenty-five loci were used as these were the only ones in common between all three sets of germplasm analyzed in this study. Amplified fragments were size-separated on an ABI377 automatic DNA sequencer (Applied Biosystems, Foster City, CA) and classified with GENESCAN 3.1 (PerkinElmer/Applied Biosystems, Foster City, CA) and GENOTYPER 2.1 (PerkinElmer/Applied Biosystems) software programs.

Statistical Analysis

The number of alleles per locus for each of the three germplasm groups (landraces, OPVs, CMLs) was determined, and group-specific alleles were identified. For further analysis of the

population structure within the CMLs, some were grouped into populations based on the OPV they were derived from (these groups are denoted as CML-Pop). Only CMLs selected from the OPVs in this study were included in the CML-Pops study, and only CML-Pops with more than four individuals were analyzed together (63 total) (see Supplementary Table 1 for more information). Total gene diversity (H_t) across all populations, gene diversity between individuals within each population (H_s) of the three germplasm groups, and coefficient of gene differentiation (G_{ST}) were all calculated according to Nei (1987). G_{ST} is the relative differentiation of the populations. Significant differences between H_s , H_t , and G_{ST} values between germplasm groups were tested by a Wilcoxon signed rank test (Hollander and Wolfe, 1973). Relationships among the landraces, OPVs and CML-Pops were analyzed by applying: (i) classification, using average linkage (UPGMA) clustering based on the modified Rogers distances (Wright 1978), and (ii) ordination by applying principal coordinate analysis (PCoA) (Gower, 1966). All analyses were performed with the software Plabsoft (Maurer et al., 2004), which is implemented as an extension to the statistical software R (Ihaka and Gentleman, 1996).

A total of 209 CMLs, which represent the full range of CMLs produced by CIMMYT but not the very closely related sister lines, were analyzed together with 497 individuals from the 23 landraces to determine the genetic contribution of each landrace to the CMLs, individually and as a group. The analyses were conducted using the Structure program (Pritchard et al., 2000) with the admixture model and assigning each individual from the populations to their known population, but allowing the CMLs to vary. The number of clusters k varied from 24 to 32. This was done to see if all CMLs fall into the predetermined groups defined by the 23 landraces; if not, those CMLs who are not genetically close enough to the landraces to cluster with them will fall into the “extra” groups represented by between 1 and 9 alternate clusters. A total of 250,000 replications were run after a burn-in period of 25,000. The results were visualized using a graph generated with the Distruct program (Rosenberg, 2002).

RESULTS AND DISCUSSION

Relationships among Landraces, OPVs, and CMLs

The PCoA revealed a clear separation of the improved germplasm (CMLs and OPVs) from the landraces (Fig. 1). This can be explained by (i) nonsimilar selection pressure for landraces and improved germplasm, since landraces were selected over a long time by farmers who generally employed a low selection pressure on only cob and kernel characteristics following harvest, and by natural selection, whereas CMLs and OPVs were selected following intense selection pressure for a wide range of agronomic characters; (ii) drift during the establishment or improvement of the improved germplasm, which has not been widely studied but is expected to play a strong role especially since bottlenecks would occur during inbreeding; (iii) mutation, although mutation rates in most species, even one as

Table 2. Comparison of the number of CIMMYT maize inbred lines (CMLs) that clustered with each of the 23 landraces (observed relationship according to Structure when $k = 23$) compared to the number of CMLs with the same landrace in their pedigree (expected relationship).[†]

Name	No.	%	Importance as parent in CIMMYT breeding pools
Arrocillo Amarillo	1	0.48	Little to none
Bolita	5	2.38	Minimal
Cacahuacintle	1	0.48	Little to none
Celaya 2	20	9.52	Very
Chalqueno (1 and 2)	0	0.00	Minimal
Chapalote	3	1.43	Little to none
Comiteco	2	0.95	Minimal
Conico	0	0.00	Little to none
Harinoso de Ocho-10 Hilleras	0	0.00	Little to none
Maiz Dulce	2	0.95	Little to none
Nal-Tel	0	0.00	Minimal
Olotillo Blanco	5	2.38	Moderate
Oloton	2	0.95	Moderate
Palomero Toluqueno	0	0.00	Little to none
Pepitilla	4	1.90	Minimal
Reventador	1	0.48	Little to none
Tabloncillo	3	1.43	Moderate
Tehua	3	1.43	Little to none
Tepecintle	64	30.48	Very
Tuxpeno	12	5.71	Very
Zapalote Chico	0	0.00	Moderate
Zapalote Grande	10	4.76	Minimal
Other	71	34.29	
Total	209		

[†]The number of CMLs that grouped with each landrace according to Structure falls in the No. column, and the percent overall variation at the marker level of the CMLs that was similar to each landrace is shown in the % column. The importance in the pedigree of the CIMMYT breeding pool (last column) has been estimated a priori to the results of the current study by CIMMYT breeders.

mutable as maize, have been demonstrated to be exceedingly low per generation. In addition, SSRs do change rapidly, but not so quickly that they are unable to distinguish individual maize plants (Smith et al., 1997), and thus mutation rates probably contribute little to the separation of groups in this graph; and (iv) the fact that many of the parental landraces of the improved germplasm were not characterized in this study. The CIMMYT OPVs and breeding populations routinely list dozens of landraces in their pedigrees. Pop25, for example, is composed of white flint selections from crosses among germplasm from Mexico, Colombia, the Caribbean, Central America, India, Thailand, and the Philippines. The diversity included from other sources show up in the OPVs and not in the landraces in this study, which were only from Mexico.

Inbreeding of OPVs could lead to a severe shift in the allele frequencies due to a high amount of sublethal alleles, which when present in homozygous state would

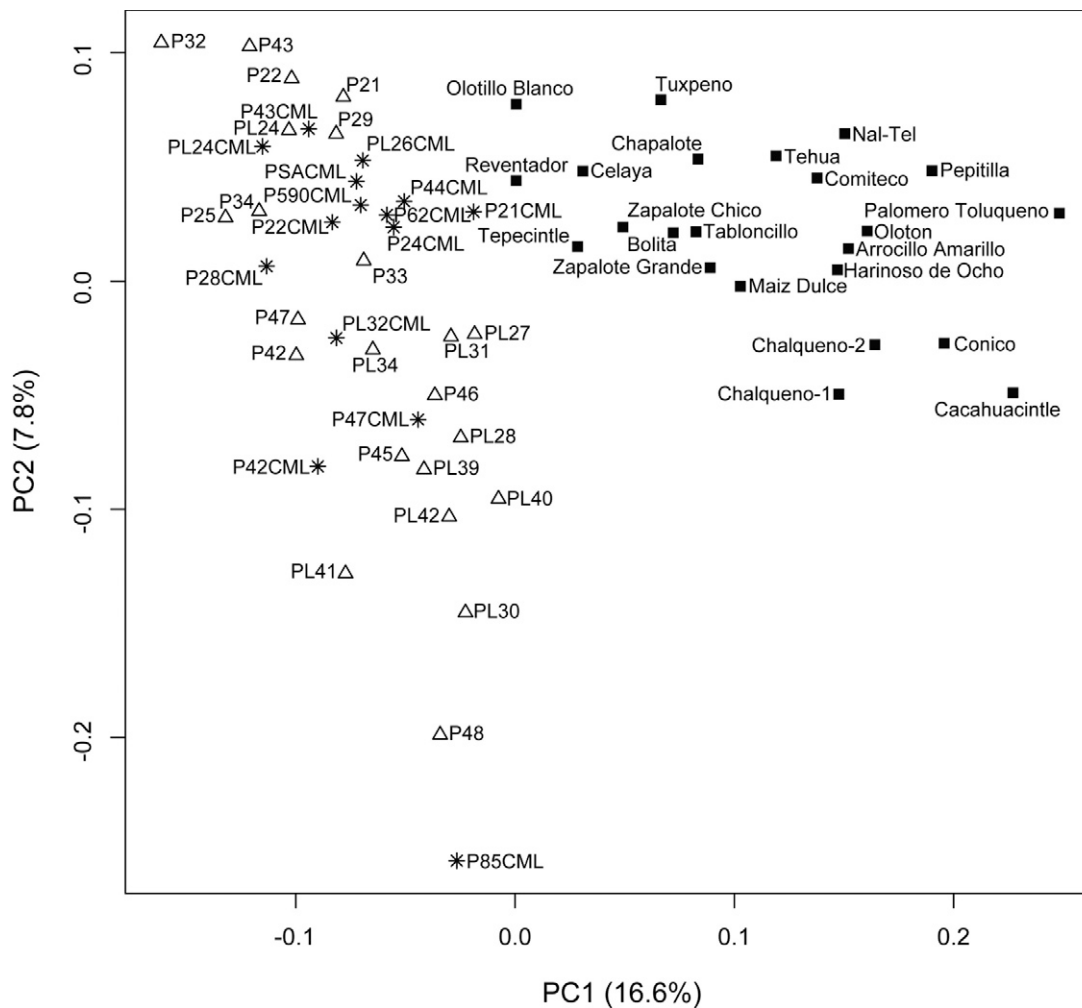


Figure 1. Principal coordinate analysis based on 25 simple sequence repeat (SSR) markers scored on 23 maize landraces (filled squares), 23 improved CIMMYT open pollinated varieties (OPVs) (open triangles), and 63 improved CIMMYT inbred lines derived from 15 of the OPVs (asterisks). The first two principal coordinates are shown in this biplot.

greatly reduce the fitness of the plant carrying them, and reduce the frequency of these alleles and any linked to them. However, CMLs extracted from the OPVs clustered closely to the OPVs and not to the landraces or in a separate cluster, showing no tendency for change due to drift (Fig. 1). Genetic distance measurements, such as modified Rogers distance employed in this study, are more influenced by the alleles of major frequency, which the CMLs were more likely to inherit, than those of minor frequency, which may have been lost following inbreeding and selection.

Comparisons between Landraces and OPVs

The landraces contain a high number of unique alleles that are not present in the OPVs (1.4 alleles per locus on average, Table 1). The presence of so many unique alleles in the landraces is most likely explained by the large numbers of landraces that were not fully exploited as parents, and is an indication that variation for agronomic traits is present in the landraces for future maize improvement (Table 2). Unfortunately, this genetic variation is

often masked in poor agronomic backgrounds. Furthermore, combining many landraces into a single population increases the risk of losing rare alleles, which are exactly the alleles lost as the germplasm suffers potential bottlenecks due to selection and introduction of maize into new areas (via migration or commercial activities). Differences in allele frequencies can be seen between the landraces and the OPVs. A PCoA of both groups (landraces and OPVs) clearly distinguishes the landraces from the OPVs on the first axis (which accounts approximately 17% of the variation in both Fig. 1 and the Supplemental Fig. 1). The cause(s) of the differences are probably multiple and possibly simultaneous, including selection, drift, mutation, and introgression of novel exotic germplasm not characterized here into the OPVs.

The G_{ST} showed a tremendous difference in landraces as compared to OPVs (Table 1). This can be explained by the breeding methodology used at CIMMYT, particularly after 1974 (Vasal et al., 1999). Germplasm from different racial complexes was mixed and more than 100 breeding populations were established to capitalize on the combining ability

(additive gene effects) of different germplasm sources for intrapopulation improvement. While this procedure created huge amounts of within population variation for further selection in specific growing conditions and subsequent release as an OPV (CIMMYT, 1998; Warburton et al., 2002), it was suboptimal with regard to conserving the differentiation between the populations, which can be detrimental to hybrid breeding programs. This genetic diversity between populations becomes important when switching from intrapopulation to interpopulation improvement as has happened at CIMMYT with the initiation of a hybrid breeding program. With hybrid breeding, the maximum divergence among populations is desired, because of an expected increase of heterosis with increasing genetic divergence of the parental populations (Falconer, 1989).

Comparison between OPVs and CMLs

A slightly higher H_t and number of alleles per SSR are seen in the CMLs when compared to the OPVs (Table 1). This may reflect a sampling bias in this study, because only 23 of the more than 140 CIMMYT OPVs and breeding populations were characterized. Because of the need to characterize multiple individuals per population to adequately sample all the variation within each population, it was not feasible to study a larger number of OPVs. The slightly higher H_t values in the CMLs may also be due to additional source germplasm not included in this study, either local or exotic from various diverse geographic regions used to develop some of the CMLs (e.g., Pop590, from which some of the CMLs were extracted, contains temperate germplasm from DeKalb). The high number of unique alleles present in the OPVs (1.9) but not in the other two germplasm groups indicate that it may be worthwhile to return to the OPVs to try to extract more of the diversity they contain, either by the creation of new inbreds or via allele mining using association mapping.

When the CMLs and the OPVs were analyzed together, no clear separation was seen between the two groups of germplasm (Fig. 1). Comparisons can be made between the OPVs and the CMLs derived from OPVs in this study (the CML-pops). CML-pops drawn from a particular OPV do not always cluster closest to that OPV, an indication of the high diversity but low differentiation between the OPVs. In addition, the separations that are seen between a CML-pop and its parental OPV can be attributed to genetic selection by the breeders during inbreeding, and to loss of alleles, especially those at low frequency. This will result in a large potential for genetic drift to diverge the CMLs from OPVs and breeding populations from which they derived. Drift is also probably the major explanation for the large decline in G_{ST} seen when moving from OPVs to CMLs (Table 1), although

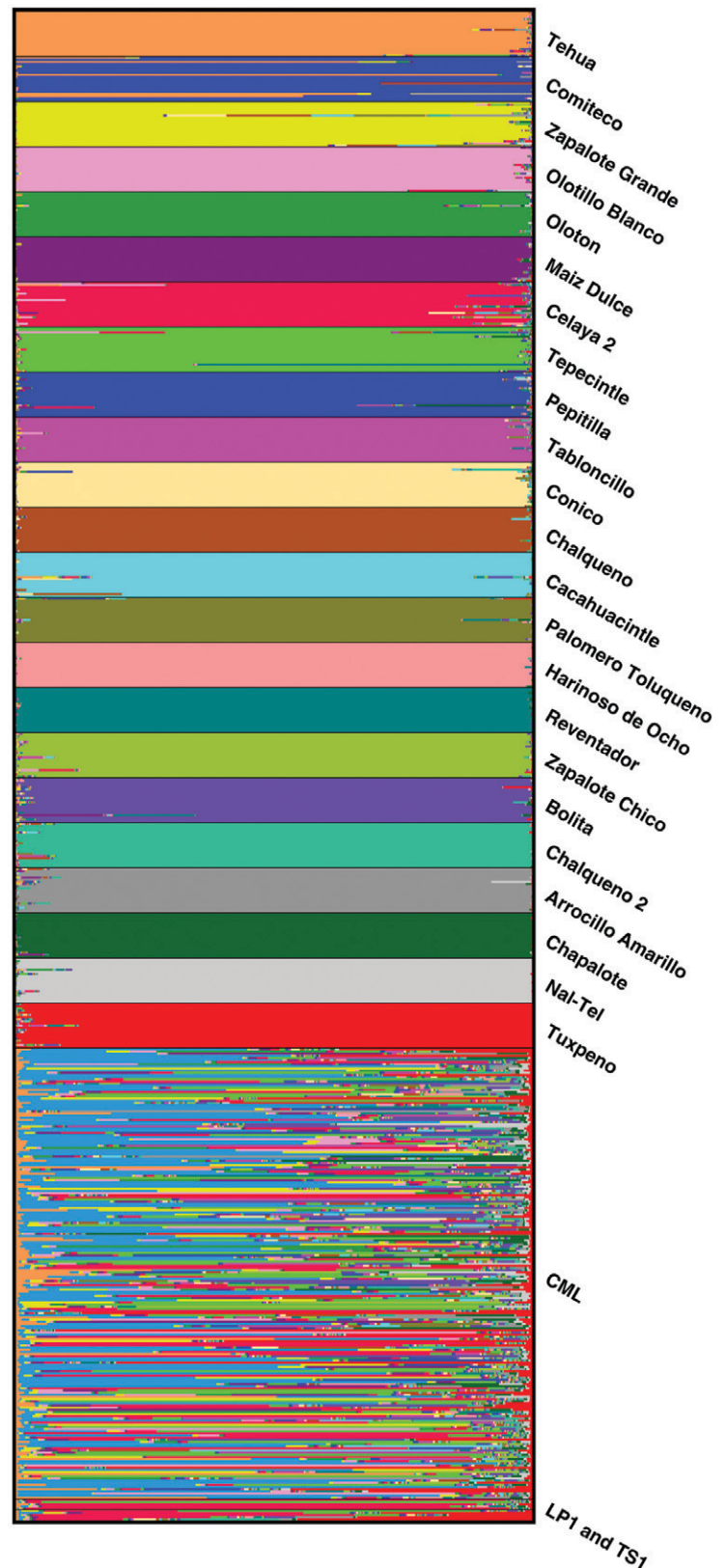


Figure 2. Population structure in the CIMMYT maize lines (CMLs) and 21 individuals each from 23 maize landraces analyzed in this study by the program Structure and visualized with the program Distruct. Each vertical bar represents one individual or inbred line, which is partitioned into up to k colored segments, which represents the individual's estimated membership in each of the k clusters ($k = 24$ in this example). The CMLs were not constrained by cluster, nor were the control inbreds (called LP1 and TS1).

sampling and selection probably explain part of the differences. Unlike the case with the OPVs, most CMLs were not formed by admixture, (e.g., they were drawn from one OPV, and not inter-OPV crosses).

When the CMLs were classified based on the 25 SSRs used in this study, no clear patterns of relationships could be seen (data not shown). This is corroborated by the low G_{ST} value for the CMLs (Table 1). This was also the case in past studies of the CMLs using many more SSR markers (Warburton et al., 2002; Xia et al., 2004; 2005) and random fragment length polymorphism (RFLP) markers (Warburton et al., 2005). The lack of clear structure found among the CMLs reflects CIMMYT's breeding methodology of selecting the CMLs from OPVs and breeding populations, which had themselves been formed by mixing many different germplasm sources. Although the OPVs and breeding populations formed by this method have a very wide genetic base and can take advantage of intrapopulation diversity for maximum heterosis within each OPV, improvement of populations for extraction of CMLs for hybrid development has been impeded by the lack of clear heterotic groups in the CIMMYT OPVs and breeding populations. Despite this, and the loss of some rare alleles, the CMLs encompass a vast array of diversity and have been used to create many highly productive hybrids. Many of the newest CIMMYT breeding populations (created after 2002) are now formed using known heterotic patterns and reciprocal recurrent selection, which ensures that these patterns are not mixed and lost.

Comparison of CMLs to Landraces

When the Structure results of the analysis of the CMLs with the landraces are studied, it can be seen that many of the CMLs contain variation from multiple landraces, many of which are not represented in this study (Fig. 2). These results were expected, considering the mixed origins of the OPVs and breeding populations from which the CMLs were extracted and the many generations that have passed since these populations were formed. However, it was unexpected that so many of the CMLs were apparently not mixed, as their pedigree would suggest, but looked very like only one of the landraces. Six of the 209 CMLs had a 90% or more probability of belonging to only one landrace, and 40 had a 75% or more probability of belonging to only one landrace (Fig. 2). One hundred thirty-eight CMLs were clustered by Structure into one of the populations defined by the landraces (Table 2). This indicates considerably less mixing in the CIMMYT OPVs and breeding populations has occurred since their formation than might have been expected. However, some of this clustering is an artificial effect caused by setting the total numbers of clusters within Structure to 24 (one more than the number of landraces). When we increased the number of clusters to 28, the optimal number according to the program, only 34 CMLs still

clustered within landraces (data not shown). This is still a much larger number than expected given the complicated pedigree of the CMLs.

The number of CMLs that grouped with each landrace according to Structure when k was set to 24 and the percent overall variation at the marker level of the CMLs that was similar to each landrace are found in Table 2. There are several reasons why the variation of any given landrace would show up in many CMLs. The most obvious would be the number of times each landrace was used in the formation of the OPVs. It is unfortunately very difficult to determine what percentage of any given landrace went into the formation of each OPV. Pedigrees of each OPV routinely list more than 50 landraces, synthetics, crosses, lines, and populations that went into its formation. General trends as to the importance of each landrace in the formation of each of the OPVs can be obtained from CIMMYT breeders, as indicated in Table 2 (S. Taba, unpublished data, 2006). These data were an independent estimation of the breeders, compiled without knowledge of the marker results. The Structure results are very similar to what would be expected based on the breeders' estimations. The few cases where this is not true provide some interesting points. For example, landraces that were used fairly often in the formation of the OPVs, but whose variation are not reflected in any of the CMLs (such as Zapalote Chico), may have been poor parents and had their variation selected out during inbred development. Landraces that were not important in the formation of the CMLs (either in the pedigree or the marker analysis) may contain alleles of use to the breeders that may be masked in a particularly unsuitable background. These are unlikely to be found using classical breeding techniques, and new ideas for gene identification and allele mining may be more helpful in tapping these alleles.

Consequences for Use of the Diversity Present in the CIMMYT Germplasm for Breeding

The molecular marker studies of CIMMYT maize germplasm suggest that the CMLs cover a considerable amount of the variation present in the entire nontemperate maize gene pool. In contrast, temperate inbreds usually contain less diversity than temperate OPVs, and certainly less than temperate and tropical landraces (Liu et al., 2003; Duvick et al., 2004). In addition to containing an impressive amount of allelic diversity, the CMLs have the added advantage of being fixed genotypes, which makes them a valuable source for association mapping studies. They will be quite useful as an association mapping panel, because they do not show a distinct population structure and it is likely that linkage disequilibrium will decay rapidly (Remington et al., 2001). Nevertheless, the many unique alleles found only in the landraces indicates that

there is considerable variation left to exploit from the landraces for the improvement of future OPVs and inbreds. This variation must be further mined by generating core subsets of these landraces, using methodologies that ensure no loss of allelic diversity, and screening them extensively for phenotypes of interest and for new alleles of previously characterized genes. These core subsets are being formed by various groups, including the Generation Challenge Program; more information on the core and obtaining seeds and data can be found at <http://www.generationcp.org/subprogramme1.php>.

Acknowledgments

This research was supported by funds from the German “Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung” Projekt No. 98.7860.4-001-01. The authors wish to thank Salvador Ambriz, Emilio Villordo, Leticia Diaz, and Ana Gomez, for their excellent technical assistance, and all CIMMYT maize breeders without whose technical and intellectual input this paper would never have been written.

References

- Aquino, P., F. Carrion, R. Calvo, and D. Flores. 2000. Selected maize statistics. p. 45–57. *In* P. Pingali (ed.) 1999–2000 World maize facts and trends: Meeting world maize needs—Technological opportunities and priorities for the public sector. CIMMYT, Mexico, DF.
- Beadle, G.W. 1939. Teosinte and the origin of maize. *J. Hered.* 30:245–247.
- CIMMYT. 1998. A complete listing of maize germplasm from CIMMYT. Maize Program Special Report. CIMMYT, Mexico, DF.
- Doebley, J.F. 2004. The genetics of maize evolution. *Annu. Rev. Genet.* 38:37–59.
- Dubreuil, P., C. Rebourg, M. Merlino, and A. Charcosset. 1999. The DNA-pooled sampling strategy for estimating the RFLP diversity of maize populations. *Plant Mol. Biol. Rep.* 17:123–138.
- Duvick, D., J. Smith, and M. Cooper. 2004. Long-term selection in a commercial hybrid maize breeding program. *In* J. Janick (ed.) *Plant breeding reviews*. Vol. 24, Part 2: Long term selection: Crops, animals, and bacteria. John Wiley & Sons, New York.
- Eyre-Walker, A., R.L. Gaut, H. Hilton, D.L. Feldman, and B.S. Gaut. 1998. Investigation of the bottleneck leading to the domestication of maize. *Proc. Natl. Acad. Sci. USA* 95:4441–4446.
- Falconer, D.S. 1989. *Introduction to quantitative genetics*. 3rd ed. John Wiley & Sons, New York.
- Gower, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 53:325–338.
- Hollander, M., and D.A. Wolfe. 1973. *Nonparametric statistical inference*. John Wiley & Sons, New York.
- Ihaka, R., and R. Gentleman. 1996. A language for data analysis and graphics. *J. Comput. Graph. Stat.* 5:299–314.
- Liu, K., M. Goodman, S. Muse, J.S. Smith, E. Buckler, and J. Doebley. 2003. Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. *Genetics* 165:2117–2128.
- Matsuoka, Y., Y. Vigouroux, M.M. Goodman, J. Sanchez Garcia, E. Buckler, and J. Doebley. 2002. A single domestication for maize shown by multilocus microsatellite genotyping. *Proc. Natl. Acad. Sci. USA* 99:6080–6084.
- Maurer, H.P., A.E. Melchinger, and M. Frisch. 2004. Plabsoft: Software for simulation and data analysis in plant breeding. p. 359–362. *In* XVIIth EUCARPIA General Congr., Tulln, Austria. 8–11 Sept. 2004. Kluwer Academic Publ., Dordrecht, the Netherlands.
- Morris, M.L. 2001. Assessing the benefits of international maize breeding research: An overview of the global maize impacts study. p. 25–34. *In* P.L. Pingali (ed.) 1999–2000 World maize facts and trends: Meeting world maize needs—Technological opportunities and priorities for the public sector. CIMMYT, Mexico, DF.
- Nei, M. 1987. *Molecular evolutionary genetics*. Columbia Univ. Press, New York.
- Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Reif, J.C., M.L. Warburton, S. Taba, D. Hoisington, J. Crossa, J. Franco, X. Xia, J. Muminović, M. Bohn, M. Frisch, and A.E. Melchinger. 2006. Grouping of accessions of Mexican landraces of maize revisited with SSR markers. *Theor. Appl. Genet.* 113:177–185.
- Reif, J.C., X.C. Xia, A.E. Melchinger, M.L. Warburton, D.A. Hoisington, D. Beck, M. Bohn, and M. Frisch. 2004. Genetic diversity determined within and among CIMMYT maize populations of tropical, subtropical, and temperate germplasm by SSR markers. *Crop Sci.* 44:326–334.
- Reif, J.C., P. Zhang, S. Dreisigacker, M.L. Warburton, M. van Ginkel, D. Hoisington, M. Bohn, and A.E. Melchinger. 2005. Trends in genetic diversity during the history of wheat domestication and breeding. *Theor. Appl. Genet.* 110:859–864.
- Remington, D.L., J.M. Thornsberry, Y. Matsuoka, L.M. Wilson, S.R. Whitt, J. Doebley, S. Kresovich, M.M. Goodman, and E.S. Buckler. 2001. Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc. Natl. Acad. Sci. USA* 98:11479–11484.
- Rosenberg, N.A. 2002. *Distruct*: A program for the graphical display of STRUCTURE results. Available at rosenberglab.bioinformatics.med.umich.edu/distruct.html (verified 9 Jan. 2008).
- Smith, J.S.C., E.C.L. Chin, H. Shu, O.S. Smith, S.J. Wall, M.L. Senior, S.E. Mitchell, S. Kresovich, and J. Ziegler. 1997. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): Comparisons with data from RFLPs and pedigree. *Theor. Appl. Genet.* 95:163–173.
- Taba, S., H.L. Shands, and S.A. Eberhart. 2005. The growth of CIMMYT’s maize collection with the introduction of Latin American maize landrace accessions through the cooperative regeneration project. p. 1–8. *In* S. Taba (ed.) *Latin American Maize Germplasm Conservation: Regeneration, In Situ Conservation, Core Subsets and Prebreeding*, Proc. of a Workshop Held at CIMMYT. 7–10 Apr. 2003. CIMMYT, Mexico, DF.
- Tenaillon, M.I., M.C. Sawkins, A.D. Long, R.L. Gaut, J.F. Doebley, and B.S. Gaut. 2001. Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Proc. Natl. Acad. Sci. USA* 98:9161–9166.
- Vasal, S.K., H. Cordova, S. Pandey, and G. Srinivasan. 1999. Tropical maize and heterosis. p. 363–373. *In* J.C. Coors and S. Pandey (ed.) *The genetics and exploitation of heterosis in crops*. ASA, CSSA, and SSSA, Madison WI.
- Vigouroux, Y., M. McMullen, C.T. Hittinger, K. Houchis, L. Shulz, S. Kresovich, Y. Matsuoka, and J. Doebley. 2002.

- Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. *Proc. Natl. Acad. Sci. USA* 99:9650–9655.
- Warburton, M.L., J.M. Ribaut, J. Franco, J. Crossa, P. Dubreuil, and F.J. Betrán. 2005. Genetic characterization of 218 elite CIMMYT inbred maize lines using RFLP markers. *Euphytica* 142:97–106.
- Warburton, M.L., X.C. Xia, J. Crossa, J. Franco, A.E. Melchinger, M. Frisch, M. Bohn, and D.A. Hoisington. 2002. Genetic characterization of CIMMYT maize inbred lines and open pollinated populations using large scale fingerprinting methods. *Crop Sci.* 42:1832–1840.
- Wright, S. 1978. *Evolution and genetics of populations*. Vol. 4. Univ. of Chicago Press, Chicago, IL.
- Xia, X.C., J.C. Reif, D.A. Hoisington, A.E. Melchinger, M. Frisch, and M.L. Warburton. 2004. Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers: I. Lowland tropical maize. *Crop Sci.* 44:2230–2237.
- Xia, X.C., J.C. Reif, A.E. Melchinger, M. Frisch, D.A. Hoisington, D. Beck, and M.L. Warburton. 2005. Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers: II. Subtropical, tropical mid-altitude, and highland maize inbred lines and their relationships with elite U.S. and European maize. *Crop Sci.* 45:2573–2582.