RESEARCH NOTE

Genetic diversity based on SSR markers in maize (*Zea mays* L.) landraces from Wuling mountain region in China

YAO QI-LUN[1,2]*, FANG PING[1], KANG KE-CHENG[2] and PAN GUANG-TANG[2]

[1]*Department of Life Science, Yangtze Normal College, Fuling Chongqing 408003, People's Republic of China*
[2]*Maize Research Institute, Sichuan Agricultural University, Yaan, Sichuan Province 625014, People's Republic of China*

Introduction

Genetic diversity of maize (*Zea mays* L.) plays a key role in maize breeding (William and Michael 2002). Knowledge of the amount and the distribution of genetic variation within and among maize landraces will provide a guide for predicting the degree of inheritance, variation, and level of heterosis, that are essential for maize breeding (Duan *et al*. 2006). For several decades, maize breeders have focused on short-term breeding. This has resulted in the development of a narrow genetic base for commercial maize hybrids (Darrah and Zuber 1986). A survey that was conducted in late 1970s and mid 1980s on inbred lines showed that some of the inbreds continued to contribute substantially to hybrids marketed in the United States. And the pedigrees of most hybrids are derivatives of 6–8 inbred lines (Darrah and Zuber 1986; James *et al*. 2002). In China, the parenthood of 91.6% hybrids consists of about 20 elite inbred lines (Li *et al*. 2002). With such a restrictive base, maize may not contain all the desirable and favourable alleles to maintain selection progress.

Wuling mountain region covers Hunan, Hubei, Chongqing, and Guizhou provinces in China, where farmer-saved maize landraces are still grown. The development of molecular markers provides a means of assessing genetic diversity at the DNA level (Reif *et al*. 2003). In particular, SSR markers are potentially useful for large-scale DNA fingerprinting of maize genotypes due to a high level of polymorphism detected (Smith *et al*. 1997), automated analysis systems (Sharon *et al*. 1997), and high accuracy and repeatability (Heckenberger *et al*. 2002). No attempts have been made so far to characterize maize landraces in Wuling mountain region using SSR markers. These original landraces represent a vast array of germplasm, which can be used in maize breeding for improving disease and pest resistance, nutritional quality, and other traits of interest (Rong *et al*. 2002). Thus, the objectives of this study were: (i) to reveal genetic diversity of maize landraces from this region with SSR markers, (ii) estimate the genetic structure and associations among individuals of landraces, and (iii) evaluate genetic variation of SSR loci within landraces.

A total of 286 alleles were detected among the landraces using bulk DNA samples and 45 SSR loci distributed on 10 chromosomes of maize. Clustering analysis based on the genetic similarity coefficients separated the landraces into five groups. The landraces collected from the same region were mostly grouped together. A total of 357 alleles were detected in 180 individuals of 12 landraces. Estimation of the mean number of allele A, the effective allelic number A_e, the observed heterozygosity H_o, and expected heterozygosity H_e were 7.93, 4.03, 0.70, and 0.39, respectively. An obvious genetic deviation from Hardy–Weinberg expectation was observed both among and within landraces, and considerable genetic variation was revealed within, as compared to among landraces. The obtained results suggested that 180 individuals genotypes could construct the maize land-race core collection of Wuling mountain region.

Materials and methods

Plant materials

The plant materials used in this study consisted of 124 maize landraces from Wuling mountain region, including Hunan 30, Hubei 32, Chongqing 35 and Guizhou 27, respectively. They were planted in Plant Garden in Life Science Department of Yangtze Normal University at Fuling of Chongqing City, China. The landrace numbers, origins and germplasm characteristics are presented in table 1 of electronic supplementary material at http://www.ias.ac.in/jgenet/. For 124 landraces, genomic DNA was isolated from a bulk sample of

*For correspondence. E-mail: yql641@yahoo.com.cn.

Keywords. maize landraces; genetic diversity; SSR; Wuling mountain region.

15 individual plants to analyse genetic diversity of landraces. Based on their geographic distribution in the four provinces, 12 out of 124 landraces were chosen to isolate genomic DNA of a single plant to obtain information on the within-landrace genetic variability. For each of 12 landraces, 15 individual plants were separately chosen to isolate genomic DNA of an individual sample. In total, 180 individual DNA samples were used.

DNA isolation and SSR analysis

Genomic DNA was isolated from the third-fresh leaf following CTAB procedure described by Scott (1998) with minor modifications. PCR amplification was performed in a PTC-220 thermalcycler programmed for 35 cycles of 1 min at 95°C, 2 min at 55°C, and 2 min at 72°C, followed by a 10 min extension at 72°C. The PCR amplification products were separated on 6% (w/v) denatured polyacrylamide gel and visualized by silver staining.

SSRs data scoring and analysis

The SSR bands were scored as present (1) or absent (0), each of which was treated as an independent character. Genetic diversity analyses were conducted on the basis of the scores. The statistical methods and formulae used are described below.

(i) *The GS index of genetic similarity:* $\mathrm{GS} = 2N_{ij}/(N_i + N_j)$, where N_{ij} is the number of SSR alleles common to landraces i and j, while N_i and N_j are the total numbers of SSR alleles observed for landraces i and j, respectively. The dendrogram was constructed by the unweighted pair-group method with arithmetic mean (UPGMA) clustering with the computer software NTSYS-pc version 2.10 (Rohlf 1998) (Exter Software, Setauket, New York).

(ii) *The mean number of alleles A:* $A = \sum_{i=1}^{n} A_i/n$, where A_i is the number of alleles at ith locus.

(iii) *The effective allelic number* A_e: $A_e = \sum_{i=1}^{n} A_{ei}/n = \sum_{i=1}^{n} \left(1/\sum_{j=1}^{m} q_j^2\right)/n$, where A_{ei} is the effective allelic number at ith allele, and q_j the frequency of the jth locus.

(iv) H_o *is the observed heterozygosity:* $H_o = \sum_{i=1}^{n} H_{oi}/n = \sum_{i=1}^{n} \left(1 - \sum_{j=1}^{m} q_{ij}^2\right)/n$, where H_{oi} represents the observed heterozygosity of the ith allele, and q_{ij} is the frequency of the jth homozygous allele at ith locus.

(v) *The expected heterozygosity, also known as the index of gene diversity:* $H_e = \sum_{i=1}^{n} H_i/n = \sum_{i=1}^{n} \left(1 - \sum_{j=1}^{m} q_{ij}^2\right)/n$, where H_i is the expected heterozygosity of the ith allele, and q_{ij} refers to the frequency of the jth homozygous allele at ith allele.

(vi) *Wright fixation index, defined as inbreeding coefficient* F: $F = 1 - H_o/H_e$. It ranges theoretically from -1 to 1. It is one only if landraces are genetically heterozygous. F_{IT}, F_{IS} and F_{ST} are Wright's F-statistics (Wright 1943; Tao and Ren 2004). F_{IT} and F_{IS} were defined as genetic deviation from Hardy–Weinberg expectation within and among landraces, respectively. When F_{IT} and F_{IS} are 0, landraces are at Hardy–Weinberg equilibrium. F_{ST}, ranging from 0 to 1, is an estimate of gene differentiation between landraces, which represents genetic variation among landraces (Nei 1973). F_{ST} is 0 if there is no genetic variation among landraces.

(vii) *Genetic distances D between each pair of landraces were estimated by the modified Rogers' distance (MRD) as follows:* $D = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{1}{2}\left(p_{ij}^X - p_{ij}^y\right)^2$, where p_{ij}^X and $q^Y ij$ are the frequencies of ith allele at jth allele in landraces X and Y, respectively. Associations between individuals were graphically depicted by principal components analysis (PCA) based on MRD estimates.

The A, A_e, H_o, F, F_{IT}, F_{IS}, F_{ST} and D were computed with the POPGENE software version 1.2 (Yeh *et al*. 1997) (Alberta, Edmonton, Canada). The PCA was performed with the statistical-software R.

Results

Genetic diversity among 124 maize landraces for bulk DNA samples

With 45 SSR loci distributed uniformly on 10 chromosomes of maize, a total of 286 alleles with 96.8% polymorphism were detected among 124 landraces for bulk DNA samples. For each SSR locus, the number of alleles ranged from three to nine corresponding to an average of 6.4 per locus, revealing a high level of genetic diversity of maize landraces in Wuling mountain region. The dendrogram based on SSR markers could distinguish 124 landraces, which were distinctly clustered into five groups. The biggest group (group I) consisted of 28, 15 and 2 landraces from Hunan, Hubei and Guizhou, respectively. Group II included 34 landraces, of which 15 were from Hubei and 19 from Guizhou. Twenty-four landraces from Guizhou and Chongqing clustered into group III, and 13 landraces from Chongqing into group IV. In addition, eight landraces from all the four regions were less related with their geographic origins, and formed group V. Comparison of the landraces within the same region revealed that most of landraces could be grouped together and had similarity coefficients of over 0.40. Thus, genetic relationships among landraces in Wuling mountain region tended to associate with their geographic origins.

Genetic variation of SSR loci within maize landraces

Table 1 shows individual locus statistics for SSR analysis: a total of 357 alleles were observed at the 45 SSR loci. Allelic diversity at SSR loci varied greatly from one locus to another, and was high on average. The number of alleles A ranged from 3 to 12 with an average of 7.93. In comparison with those among 124 landraces for bulk DNA samples, the alleles detected in 12 landraces for individual DNA samples were higher at the same 45 loci. A_e was 4.03 on average and ranged from 2.62 to 7.45. H_e exhibited a range of

Table 1. The genetic variation of SSR loci and landrace genetic structure.

Locus	Chromosome	A	A_E	H_E	H_O	F_{IS}	F_{IT}	F_{ST}	F
bnlg1429	1.02	3	2.91	0.65	0.33	0.40	0.50	0.06	0.51
umc1719	1.04	12	6.03	0.79	0.46	0.39	0.48	0.05	0.47
bnlg1023	1.06	5	3.92	0.68	0.56	0.16	0.27	0.08	0.24
phi308707	1.10	10	7.45	0.82	0.66	0.22	0.24	0.11	0.21
phi083	2.04	8	4.10	0.71	0.37	0.45	0.57	0.07	0.53
nc031	2.05	9	4.22	0.72	0.41	0.47	0.66	0.14	0.59
nc133	2.05	7	3.15	0.64	0.24	0.29	0.36	0.13	0.43
phi090	2.08	7	4.61	0.76	0.49	0.45	0.51	0.14	0.42
phi127	2.08	8	3.42	0.64	0.37	0.38	0.44	0.12	0.47
phi029	3.04	7	3.90	0.71	0.42	0.41	0.53	0.16	0.52
umc1501	3.05	8	4.84	0.76	0.39	0.46	0.57	0.15	0.64
phi102228	3.06	8	3.46	0.68	0.32	0.52	0.61	0.18	0.60
phi046	3.08	9	3.31	0.65	0.26	0.43	0.45	0.20	0.48
phi072	4.00	8	4.26	0.72	0.39	0.47	0.57	0.22	0.56
phi096	4.04	7	3.39	0.67	0.33	0.30	0.43	0.16	0.41
Phi095	4.05	10	3.95	0.72	0.40	0.48	0.59	0.26	0.62
phi092	4.08	8	3.75	0.69	0.44	0.28	0.44	0.13	0.48
phi076	4.11	8	4.48	0.72	0.29	0.21	0.28	0.17	0.34
nc130	5.00	8	4.59	0.71	0.47	0.49	0.62	0.18	0.57
phi109188	5.03	7	3.84	0.68	0.49	0.36	0.38	0.07	0.44
phi008	5.03	7	4.26	0.70	0.34	0.26	0.38	0.06	0.40
umc1447	5.03	8	3.75	0.68	0.44	0.44	0.46	0.16	0.49
umc1332	5.04	7	3.65	0.69	0.45	0.37	0.56	0.18	0.55
phi075	6.00	7	3.81	0.69	0.38	0.17	0.27	0.11	0.27
phi031	6.04	8	4.40	0.70	0.37	0.35	0.42	0.09	0.45
bnlg1617	6.05	7	3.40	0.67	0.50	0.70	0.79	0.25	0.78
phi299852	6.07	7	3.86	0.69	0.51	0.34	0.45	0.07	0.41
umc1545	7.00	9	3.95	0.70	0.56	0.40	0.50	0.08	0.51
phi057	7.01	9	3.60	0.68	0.41	0.39	0.63	0.15	0.60
phi034	7.02	9	4.99	0.79	0.38	0.14	0.22	0.07	0.29
phi069	7.05	9	2.62	0.62	0.29	0.45	0.55	0.08	0.52
phi051	7.05	9	3.81	0.65	0.55	0.50	0.63	0.26	0.62
umc1304	8.02	10	3.70	0.64	0.35	0.41	0.59	0.11	0.59
phi115	8.03	7	3.69	0.63	0.28	0.42	0.52	0.05	0.52
phi014	8.04	8	4.81	0.74	0.32	0.47	0.63	0.19	0.63
Umc1333	8.05	8	5.26	0.71	0.39	0.48	0.58	0.14	0.59
umc1161	8.06	7	3.48	0.65	0.36	0.51	0.64	0.10	0.67
umc1297	9.00	8	2.90	0.58	0.25	0.33	0.40	0.08	0.41
phi065	9.03	8	3.71	0.71	0.31	0.53	0.64	0.13	0.60
phi108411	9.05	6	3.89	0.72	0.28	0.42	0.53	0.11	0.54
umc2359	9.07	8	4.24	0.73	0.47	0.46	0.43	0.14	0.50
Phi041	10.00	7	3.64	0.69	0.28	0.54	0.67	0.19	0.57
umc1432	10.02	9	4.35	0.73	0.40	0.49	0.53	0.07	0.51
phi062	10.04	9	4.07	0.70	0.41	0.48	0.48	0.06	0.50
umc1877	10.07	9	3.92	0.66	0.31	0.60	0.63	0.08	0.61
Mean		7.93	4.03	0.70	0.39	0.41	0.50	0.13	0.50

variation from 0.58 at *umc1297* to 0.82 at *phi308707* with a mean of 0.70. H_o averaged 0.39 and ranged from 0.23 (*nc133*) to 0.66 (*phi308707*). The most polymorphic loci, *umc1719* and *phi308707*, revealed a high level of genetic variation.

The genetic structure of maize landraces

The F, F_{IT}, F_{IS} and F_{ST} were estimated to analyse the genetic structure. F varied from 0.24 (*bnlg1023*) to 0.78 (*bnlg1617*), implying that landraces had a typical mixed-mating system and were deficient in heterozygotes. F_{IS} average was 0.46, varying from 0.14 (*phi034*) to 0.70 (*bnlg1617*), and F_{IT} was 0.50 on average, ranging from 0.22 to 0.79 at corresponding loci. This suggested that an obvious genetic deviation from Hardy–Weinberg expectation occurred among and within landraces. F_{ST} was 0.13 on average, implying that the among landrace genetic variation accounted

for only 13% of the total genetic variation, whereas the within landrace genetic variation accounted for 87%.

Associations among individuals of maize landraces

Associations among 180 individuals of 12 landraces were investigated by the PCA method. The location of individuals was defined by the first principal component (PC1) and second principal component (PC2). The PC1 and PC2 expained 5.3% and 4.5%, respectively. The PCA method yielded an obvious separation among landraces from different regions (figure 1). Most individuals within a region were grouped more closely. Forty-five individuals from Chongqing were located quite distant from the remaining individuals. The individuals from the three landraces within Guizhou were more distant, indicating that higher genetic diversity resided in the landraces from Guizhou.

Discussion

Among 124 maize landraces for the bulked DNA samples, 6.4 alleles per locus were detected using 45 SSR loci (primers). Liu *et al*. (2005) detected an average of 4.1 alleles with 50 SSR primers in 38 waxy-maize landraces. Wu *et al*. (2004) reported an average of 5.4 alleles in popcorn landraces and with 61 SSR primers. Hence, the total number of alleles per locus was higher in this study than previously reported, which suggested a broad genetic base of maize landraces in Wuling mountain region. The genetic variability of maize landraces has been affected by various factors throughout their evolutionary history. Outcrossing and fitness-relevant mutations generate intrapopulation diversity, whereas direct natural or human selection and bottleneck effects lead to an increase in interpopulation diversity (Dreisigacker *et al*. 2005). In this study, the within landrace genetic variation (87%) was higher than that of the among landraces (13%).

The observed and expected heterozygosities within and among landraces showed obvious deviations from Hardy–Weinberg expectations, reflecting from heterozygote deficiency. This finding is expected, since farmers are used to renewing maize landraces from year-to-year, with seeds taken from a small number of ears. Consequently, mass selection, which was practiced before development of hybrids, could have led to the deficit of heterozygous individuals. Since inbreeding noticeably affects the length and diameter of the ear, it is likely that farmers unconsciously selected ears from the most heterozygous plants, which prevents genetic drift among landraces and maintains a high level of genetic diversity within landraces.

The development of a core collection, which identifies a small number of unique germplasms to represent the genetic diversity present in a large collection of germplasm, is a very important research area. This research of genetic diversity by SSR marker analyses will provide important methods and technologies for the construction of a maize landrace core collection of Wuling mountain region. Assuming a population with 20,000 polymorphic loci and two alleles per locus, Lawrence *et al*. (1995) concluded that about 172 plants are sufficient to conserve nearly all alleles at a 5% frequency. Generally, 300 to 400 alleles are required to reflect stable relationships between accessions and effectively establish core collections (Zhang *et al*. 2002). In our study, a total of 357 alleles were detected in 180 individuals of 12 landraces by use of 45 SSRs, which can differentiate individual genotypes of maize landraces. Thus, the 180 individual genotypes sampled on the basis of their geographic origins were sufficient to construct the maize landrace core collection of Wuling mountain region.

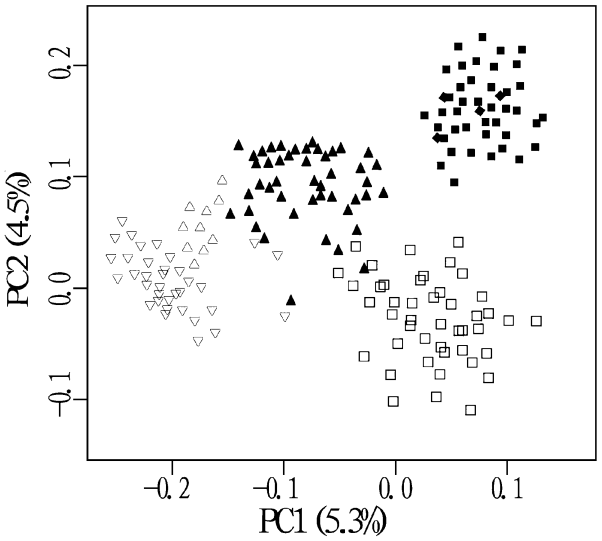

Figure 1. Principal components analysis of individuals from 12 maize landraces. Open and filled triangles represent individuals from Hunan and Hubei, respectively; open squares refer to individuals from Guizhou, and filled squares refer to individuals from Chongqing.

Acknowledgements

This work was supported by the National High Technology Research and Development Programme of China (863 program 2004BA525B04) and Programme for Changjiang scholars and innovative research team in University (IRTO453) of China.

References

Darrah L. L. and Zuber M. S. 1986 The United States farm corn germplasm base and commercial breeding strategies. *Crop Sci*. **26**, 1109–1113.

Dreisigacker S., Zhang P., Warburton M. L., Skovmand B., Hoisington K. and Milchinger A. E. 2005 Genetic diversity among and within CIMMYT wheat landrace accessions investigated with SSRs and implications for plant genetic resources management. *Crop Sci*. **45**, 653–661.

Duan Y. P., Chen W. G., Li M. S., Li X. H., Liu X., Tian Q. Z., Bai L. and Zhang S. H. 2006 The genetic diversity among 27 maize populations based on SSR data. *Sci. Agric. Sinica* **39**, 1102–1113.

Heckenberger M., Melchinger A. E., Ziegle J. S., Joe L. K., Hauser J. D., Hutton M. and Bohn M. 2002 Variation of DNA fingerprints among accessions within maize inbred lines with regard to the identification of essentially derived varieties. I. Genetic and technical sources of variation in SSR data. *Mol. Breed.* **10**, 181–191.

James G. G., Joanne A. L., Kendall R. L., Margaret E. S. and Stephen K. 2002 SSR variation in important U.S. maize inbred lines. *Crop Sci*. **42**, 951–957.

Lawrence M. J., Marshall D. F. and Davies P. 1995 Genetics of genetic conservation. I. Sample size when collecting germplasm. *Euphytica* **84**, 89–99.

Li Y., Du J., Wang T., Shi Y., Song Y. and Jia J. J. 2002 Genetic diversity and relationships among Chinese maize inbred lines revealed by SSR markers. *Maydica* **47**, 93–101.

Liu Y. J., Huang Y. B., Rong T. Z., Tian M. L. and Yang J. P. 2005 Comparative analysis of genetic diversity in land-races of waxy maize from Yunnan and Guizhou using SSR markers. *Sci. Agric. Sinica* **4**, 648–653.

Nei M. 1973 Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* **70**, 3321–3323.

Reif F. C., Melchinger A. E., Xia X. C., Warburton M. L., Hoisington D. A., Vasal S. K. *et al*. 2003 Genetic distance based on simple sequence repeats and heterosis in tropical maize populations. *Crop Sci*. **43**, 1275–1282.

Rong T. Z., Li W. S., Yang K. C., Zhang S. K., Tang H. J. and Pan G. T. 2002 *Maize breeding in southwest China*. Agricultural Press, Beijing.

Scott O. R. 1998 Extraction of DNA from plant tissue. *Plant Mol. Biol. Manual* **6**, 1–6.

Sharon E. M., Kresovich S., Jester C. A., Hernandez C. J. and Szewc-McFadden A. K. 1997 Application of multiplex PCR and fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. *Crop Sci*. **37**, 617–624.

Smith J.-C., Chen E. C. L., Shu H., Smith O. N., Wall S. J., Senior M. L. *et al*. 1997 An evaluation of 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.

Tao L. and Ren J. 2004 *A quantitative approach to the study of evolutionary ecology*. Forestry Press, Beijing.

William S. K. and Michael R. C. 2002 *Essentials of genetics*. Higher Education Press, Beijing.

Wright S. 1943 Isolation by distance. *Genetics* **28**, 114–138.

Wu Y. S., Zheng Y. L., Sun R., Wu S. Y., Gu H. B. and Bi Y. H. 2004 Genetic diversity of waxy corn and popcorn land-races in Yunnan by SSR markers. *Acta Agrono. Sinica* **30**, 36–42.

Yeh F. C., Yang R. C., Boyle T. B. J., Ye Z. H. and Mao J. X. 1997 POPGENE, version 1.2, Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton.

Zhang X. Y., Li C. W., Wang L. F., Wang H. M., You G. X. and Dong Y. S. 2002 An estimation of the minimum number of SSR alleles needed to reveal genetic relationships in wheat varieties. I. Information from large scale planted varieties and cornerstone breeding parents in Chinese wheat improvement and production. *Theor. Appl. Genet*. **106**, 112–117.

Received 26 October 2007, in revised form 3 November 2007; accepted 23 June 2008

Published on the Web: 3 October 2008