GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

# Genetic diversity and relationship analysis of *Gossypium arboreum* accessions

**F. Liu[1,2], Z.L. Zhou[2], C.Y. Wang[2], Y.H Wang[2], X.Y. Cai[2], X.X. Wang[2], Z.S. Zhang[1] and K.B. Wang[2]**

[1]College of Agronomy and Biotechnology, Southwest University, Chongqing, China
[2]State Key Laboratory of Cotton Biology, China/Cotton Research Institute, Chinese Academy of Agricultural Sciences, Anyang, China

Corresponding author: K.B. Wang
E-mail: wkbcri@cricaas.com.cn

Genet. Mol. Res. 14 (4): 14522-14529 (2015)
Received June 28, 2015
Accepted September 5, 2015
Published November 18, 2015
DOI http://dx.doi.org/10.4238/2015.November.18.15

**ABSTRACT.** Simple sequence repeat techniques were used to identify the genetic diversity of 101 *Gossypium arboreum* accessions collected from India, Vietnam, and the southwest of China (Guizhou, Guangxi, and Yunnan provinces). Twenty-six pairs of SSR primers produced a total of 103 polymorphic loci with an average of 3.96 polymorphic loci per primer. The average of the effective number of alleles, Nei's gene diversity, and Shannon's information index were 0.59, 0.2835, and 0.4361, respectively. The diversity varied among different geographic regions. The result of principal component analysis was consistent with that of unweighted pair group method with arithmetic mean clustering analysis. The 101 *G. arboreum* accessions were clustered into 2 groups.

**Key words:** Cluster analysis; Genetic diversity; *Gossypium arboreum* L.; Simple sequence repeat

 

## INTRODUCTION

Asiatic cotton (*Gossypium arboreum* L.) is a type of cultivated diploid cotton originating in the Indian subcontinent and has been cultivated for 5000 years. More than 2000 years ago, Asiatic cotton was introduced into China through Myanmar, Vietnam; after long-term natural and manual selection, this cotton has been developed into a variety of unique Asiatic cotton varieties in the southwest area of China (Shen, 1993). Because of its unique properties, including early maturity, wide adaptability, stress resistance, and disease and insect resistance, Asiatic cotton contains an important gene pool for improving cotton quality and stress tolerance (Stanton et al., 1992; Bie et al., 2001; Xu et al., 2001; Mehetre et al., 2003; Kebede et al., 2007). However, because of low yield, Asiatic cotton was gradually abandoned and replaced with upland cotton (*G. hirsutum* L.) and Sea Island cotton (*G. barbadense* L.), although it is still used as an important resource for cotton improvement. Zhou (2011a) and Zhou et al. (2013) analyzed the genetic diversity of *G. arboreum* cultivars collected from different geographical locations in China. Their results revealed the abundant genetic diversity of Asiatic cotton at the molecular level. This study employed simple sequence repeat (SSR) markers to analyze the genetic diversity of and relationships between 68 Asiatic cotton accessions collected from the southwest of China (Yunnan, Guangxi, and Guizhou provinces) and 33 Asiatic cotton accessions imported from India and Vietnam.

## MATERIAL AND METHODS

### Materials

Among the 101 Asiatic cotton materials used in this study (Table 1), 68 accessions were collected in 2002 (Song et al., 1999; Liu et al., 2003; Wang et al., 2003a,b) from the southwest of China (Yunnan, Guangxi, and Guizhou provinces) and 33 accessions were imported from India (21) and Vietnam (12) by the Institute of Cotton Research of Chinese Academy of Agricultural Sciences.

### Genomic DNA extraction

Genomic DNA was extracted from the young leaves of each cultivar using the cetyltriethyl ammnonium bromide DNA extraction method as described by Song et al. (1999), with minor modifications.

### Genotype testing of SSR markers

A total of 26 SSR markers were selected and used in this study based on previous studies (Zhou, 2011b; Fu, 2012; Wang et al., 2012; Chen, 2013). The polymerase chain reaction (PCR) volume was 10 μL and included 0.5 μL 10 mM dNTPs, 0.8 μL 5 μM forward and reverse primers, 0.1 μL 5 U/μL *Taq* polymerase, 1.0 μL 10X PCR buffer (including 15 mM $Mg^{2+}$), 25 ng genomic DNA, and 7.1 μL $ddH_2O$. The PCR was performed on a TP 600 Thermal Cycler (Takara, Shiga, Japan). The PCR program was as follows: 95°C for 3 min; 30 cycles of denaturing at 94°C for 45 s, annealing at 57°C for 45 s, extension at 72°C for 1 min; and hold at 4°C. Electrophoresis-based detection of PCR products was conducted in an 8% poly-

acrylamide gel using a DYC-Z30 electrophoresis apparatus at 200 V for 45 min, followed by silver staining. SSR-amplified polymorphic bands were encoded as "0" for absence, "1" for presence, and "-" for missing data, which included blurred or vague bands.

**Table 1.** Asiatic cotton materials and their origins.

| No. | Name | Origin | No. | Name | Origin | No. | Name | Origin |
|---|---|---|---|---|---|---|---|---|
| 1 | AK235 | India | 35 | Dianya23 | China | 69 | Qianya29-1 | China |
| 2 | B2.III.4 | India | 36 | Dianya3 | China | 70 | Qianya29-2 | China |
| 3 | B212 | India | 37 | Dianya4 | China | 71 | Qianya2 | China |
| 4 | Bac Ai | India | 38 | Dianya7 | China | 72 | Qianya30-1 | China |
| 5 | Hasonbinh | India | 39 | Guiya1 | China | 73 | Qianya30-2 | China |
| 6 | JKCDL-1 | India | 40 | Guiya2 | China | 74 | Qianya31 | China |
| 7 | JKCDL-1-1 | India | 41 | Guiya3 | China | 75 | Qianya132 | China |
| 8 | JKCDL-12 | India | 42 | Guiya4 | China | 76 | Qianya33 | China |
| 9 | JKCDL-14 | India | 43 | Guiya5 | China | 77 | Qianya34 | China |
| 10 | JKCDL-15 | India | 44 | Guiya6 | China | 78 | Qianya35 | China |
| 11 | JKCDL-16 | India | 45 | Guiya7 | China | 79 | Qianya36 | China |
| 12 | JKCDL-19 | India | 46 | Guiya8 | China | 80 | Qianya37 | China |
| 13 | JKCDL-19 | India | 47 | Guiya9 | China | 81 | Qianya38 | China |
| 14 | JKCDL-4 | India | 48 | Guiya10 | China | 82 | Qianya3 | China |
| 15 | JKCDL-5 | India | 49 | Guiya11 | China | 83 | Qianya40 | China |
| 16 | JKCDL-5 | India | 50 | Qianya12-1 | China | 84 | Qianya4 | China |
| 17 | JKDESI-L1 | India | 51 | Qianya12-2 | China | 85 | Qianya5 | China |
| 18 | JKCDL-7 | India | 52 | Qianya13 | China | 86 | Qianya6 | China |
| 19 | K-7 | India | 53 | Qianya14 | China | 87 | Qianya7 | China |
| 20 | Lucngan | India | 54 | Qianya15 | China | 88 | Qianya8 | China |
| 21 | Nghe An | India | 55 | Qianya16 | China | 89 | Qianya9 | China |
| 22 | Dianya1 | China | 56 | Qianya17 | China | 90 | VHJKL-100 | Vietnam |
| 23 | Dianya10 | China | 57 | Qianya18 | China | 91 | VHJKL-101 | Vietnam |
| 24 | Dianya12 | China | 58 | Qianya1 | China | 92 | VHJKL-101-1 | Vietnam |
| 25 | Dianya13 | China | 59 | Qianya20 | China | 93 | VHJKL-1-1 | Vietnam |
| 26 | Dianya14 | China | 60 | Qianya21 | China | 94 | VHJKL-1-2 | Vietnam |
| 27 | Dianya15 | China | 61 | Qianya22 | China | 95 | VHJKL-2-1 | Vietnam |
| 28 | Dianya17 | China | 62 | Qianya23-1 | China | 96 | VHJKL-2-2 | Vietnam |
| 29 | Dianya18 | China | 63 | Qianya23-2 | China | 97 | VHJKL-78 | Vietnam |
| 30 | Dianya19 | China | 64 | Qianya24 | China | 98 | VHJKL-80 | Vietnam |
| 31 | Dianya20-1 | China | 65 | Qianya25 | China | 99 | VHJKL-91 | Vietnam |
| 32 | Dianya20-2 | China | 66 | Qianya26 | China | 100 | VHJKL-95 | Vietnam |
| 33 | Dianya21 | China | 67 | Qianya28 | China | 101 | VHJKL-99 | Vietnam |
| 34 | Dianya21 | China | 68 | Qianya29 | China | | | |

## Data analysis

The polymorphic information content, genetic diversity index, and Shannon's information index were analyzed using the Popgene_32 software. Cluster analysis and principal component analysis were conducted using unweighted pair group method with arithmetic mean with the NTSYS-pc_2.1 software.

## RESULTS

### Analyzing genetic diversity of SSR markers

A total of 26 SSR markers were selected to determine the genetic diversity of polymorphisms after preliminary selection of 514 SSR markers based on PCR amplification of 10 materials. A total of 103 polymorphic bands were amplified. The amplification efficiency of each pair of SSR primers differed slightly from 2 to 6, with an average of 3.96 polymorphic

fragments per primer pair. A total of 101 accessions of Asiatic cotton were amplified by 26 pairs of SSR primers and identified 0.45-0.69 polymorphisms with an average of 0.59, genetic diversity index of 0.1586-0.4504, with an average of 0.2835, and Shannon's information index of 0.2775-0.6423, with an average of 0.4361. This suggests that the genetic diversity is rich in Asiatic cotton (Table 2).

**Table 2.** Polymorphysim information of SSR primers.

| No. | Primer name | NPL | PIC | *H* | *I* |
|---|---|---|---|---|---|
| 1 | BNL1034 | 4 | 0.57 | 0.3000 | 0.4547 |
| 2 | BNL2569 | 5 | 0.50 | 0.1636 | 0.2775 |
| 3 | BNL2634 | 5 | 0.50 | 0.2764 | 0.4207 |
| 4 | BNL3254 | 6 | 0.45 | 0.2637 | 0.4175 |
| 5 | BNL4108 | 6 | 0.52 | 0.2788 | 0.4332 |
| 6 | NAU0895 | 4 | 0.57 | 0.2541 | 0.3916 |
| 7 | NAU0935 | 2 | 0.69 | 0.4413 | 0.6332 |
| 8 | NAU0943 | 5 | 0.52 | 0.3038 | 0.4685 |
| 9 | NAU1346 | 3 | 0.64 | 0.3649 | 0.5486 |
| 10 | NAU2156 | 3 | 0.65 | 0.3426 | 0.5091 |
| 11 | NAU2679 | 3 | 0.65 | 0.2330 | 0.3673 |
| 12 | NAU3468 | 3 | 0.66 | 0.3745 | 0.5563 |
| 13 | MON_CGR5142 | 3 | 0.64 | 0.3113 | 0.4648 |
| 14 | MON_CGR5152 | 3 | 0.65 | 0.4445 | 0.6350 |
| 15 | MON_CGR5161 | 4 | 0.56 | 0.2939 | 0.4422 |
| 16 | MON_CGR5166 | 6 | 0.49 | 0.2100 | 0.3313 |
| 17 | MON_CGR5334 | 2 | 0.69 | 0.1586 | 0.2928 |
| 18 | MON_CGR5350 | 3 | 0.65 | 0.4504 | 0.6423 |
| 19 | MON_CGR5587 | 4 | 0.57 | 0.2758 | 0.4212 |
| 20 | MON_CGR5793 | 2 | 0.69 | 0.2438 | 0.4087 |
| 21 | MON_CGR5807 | 2 | 0.69 | 0.3714 | 0.5575 |
| 22 | MON_CGR5873 | 4 | 0.57 | 0.2823 | 0.4356 |
| 23 | MON_CGR6682 | 5 | 0.63 | 0.2039 | 0.3348 |
| 24 | JES156 | 5 | 0.51 | 0.2810 | 0.4401 |
| 25 | MUSS049 | 5 | 0.53 | 0.2750 | 0.4270 |
| 26 | Gh70 | 6 | 0.67 | 0.2587 | 0.4079 |
| Mean | | 3.96 | 0.59 | 0.2835 | 0.4361 |

The polymorphism information content differed among the cotton collected from different regions. Shannon's information index was low in Chinese samples compared to in Indian and Vietnamese samples; the gene diversity index showed the same pattern as Shannon's information index (Table 3). This suggests that Asiatic cotton had less genetic diversity in China than that in India and Vietnam.

**Table 3.** Polymorphysim information of Asiatic cotton among different origin areas.

| Origin | Southwest China | India | Vietnam | Total |
|---|---|---|---|---|
| Shannon *I* | 0.4025 | 0.4514 | 0.4525 | 0.4361 |
| Nei *H* | 0.2651 | 0.2954 | 0.3001 | 0.2835 |

## Cluster analysis

Based on the analysis using 26 SSR markers, the unweighted pair group method with arithmetic mean cluster showed that the coefficients of the 101 Asiatic cotton accessions were similar, ranging from 0.64 to 0.96, with an average of 0.76. The similarity coefficient of qianya 5 and qianya 6 was the largest (0.96), indicating that these 2 accessions are genetically related

to each other. The 101 accessions were divided into 2 groups based on their similarity coefficients. The 1st group consisted of 21 accessions, including 14 Asiatic cottons from India and 7 Asiatic cottons from Vietnam. The second group consisted of 80 accessions, including 7 Asiatic cottons from India, 5 Asiatic cottons from Vietnam, and 68 Asiatic cottons from the southwest of China.

The 2nd group of 80 accessions was further divided into 5 subgroups with a similarity coefficient of 0.69, in which the 2nd, 3rd, and 4th sub-groups included only 1 Asiatic cotton, including Qianya 10 (from the southwest of China), Guiya 1 (from the southwest of China), and VHJKL-2-1 (from Vietnam), respectively. The 5th sub-group consisted of 3 accessions, including Nghe An (from India), Qianya 34 (from the southwest of China), and Qianya 14 (from the southwest of China). The remaining 74 accessions belonged to the 1st subgroup, including 6 accessions from India, 4 accessions from Vietnam, and 64 accessions from the southwest of China. The results showed that the relationship were relatively consistent between molecular clusters and the geographical origins for the 101 Asiatic cotton samples, most Indian cotton and Vietnam cotton were clustered into 1 group, and the Chinese cotton (from different provinces in southwest China) and some of the Vietnam cotton were clustered into another large group. The different accessions collected from different provinces in southwest China (Guangxi, Guizhou, and Yunnan) overlapped with each other; they were clustered into some Indian groups and subgroups, and they could not be definitively classified (Figure 1).

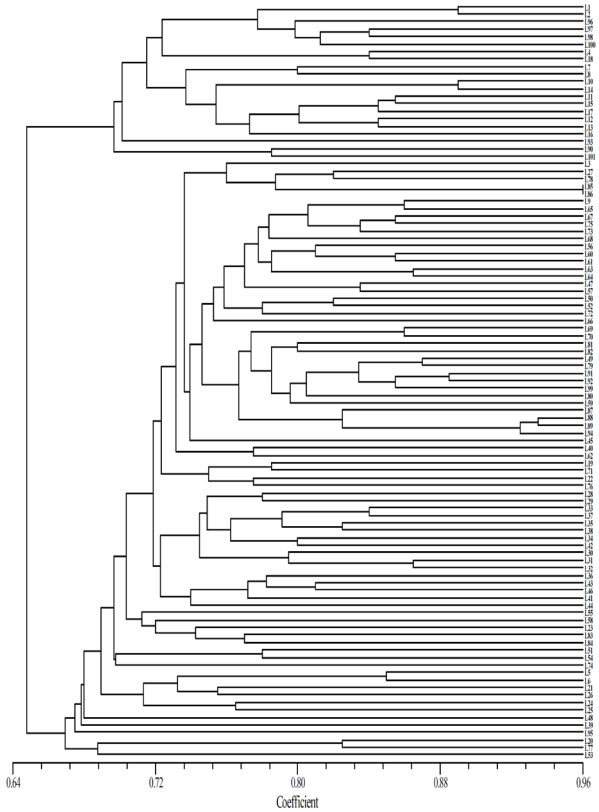


**Figure 1.** Clusters of 101 accessions based on SSR.

## Principal component analysis

Principal component analysis of the SSR molecular markers of the 101 accessions showed that the 101 Asiatic cotton accessions could be divided into 2 groups. The 1st group, a total of 21 Asiatic cottons, including 14 accessions from India and 7 accessions from Vietnam, revealed relatively close relationships. The 2nd group of 80 Asiatic cottons, including 7 accessions from India, 5 accessions from Vietnam, and all accessions from the southwest of China, showed relatively high divergence from each other (Figure 2).

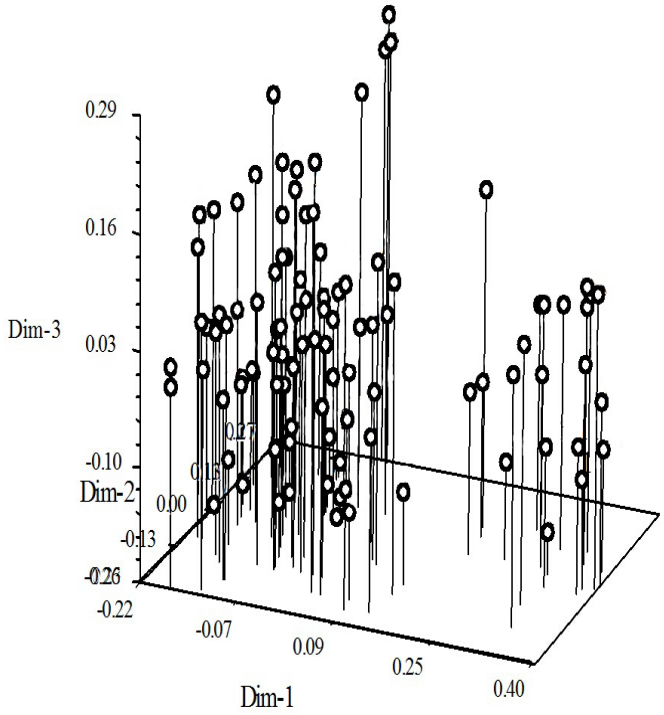


**Figure 2.** Principal component analysis of 101 accessions.

# DISCUSSION

## Genetic diversity of Asiatic cotton

Currently, there are two contradictory opinions regarding the genetic diversity of Asiatic cotton. It has been suggested that Asiatic cotton shows low genetic diversity. The results of previous studies revealed close genetic relationship and low genetic diversity among different Asiatic cotton germplasms (Dong, 2007; Deosarkar et al., 2010; Rahman et al., 2008). In contrast, it has been suggested that Asiatic cotton shows high levels genetic diversity. Asiatic cotton was found to have richer genetic diversity within species and higher genetic variation than upland cotton (Liu et al., 2006; Sinha et al., 2008; Kantartzi et al., 2009; Zhou, 2011a,b). In this study, polymorphism information content, gene diversity, and Shannon's information

index results revealed rich intra-specific genetic diversity in Asiatic cotton. The reason for these inconsistent results may be that different materials were studied or that different molecular markers were used for different locations on the chromosome (Yeh and Boyle, 1997; Rohlf, 2000).

## Geographic origin and classification of Asiatic cotton

Clustering analysis of 101 Asiatic cotton accessions showed that most Asiatic cottons, which have similar geographical origins, had close genetic relationships and were classified together; this suggests a causal relationship between the cluster grouping and geographical origin. A hundred and one Asiatic cotton accessions distributed in different geographic areas in the same sub-groups derived from the UPGMA cluster method, which may be associated with its wide introduction and spread of Asiatic cotton. This experiment shows that most India and Vietnam Asiatic cotton accessions were clustered in the same group; additionally, all Chinese accessions as well as some Indian and Vietnamese accessions, were clustered in another group, which was further divided into sub-groups. This study demonstrates that Asiatic cotton, which originated in India and then spread to Vietnam, is gradually spread throughout the southern and southwest regions of China.

In this study, Asiatic cotton showed varying diversity depending on its geographical origins. Genetic relationships between Asiatic cotton from different regions appear to be closer than those in the same region, indicating the presence of transitional types between different regions; therefore, Asiatic cotton cannot be effectively classified based on geography alone.

Our results revealed abundant genetic diversity within Asiatic cotton species. Genetic relationships were close between most of the Indian accessions and some Vietnam accessions. A small number of Indian accessions, some Vietnamese accessions, and all Chinese accessions were loosely clustered into one category.

## Conflicts of interest

The authors declare no conflict of interest.

## ACKNOWLEDGMENTS

Research co-supported by the National Science and Technology Support Plan of China (#2013AA102601) and the National Key Technology Support Program of China (#2013BAD01B03).

## REFERENCES

Bie S, Kong FL, Zhou YY, Zhang GM, et al. (2001). Genetic diversity analysis of elite cotton varieties in three main cotton regions in China by RAPD and its relation with agronomic characteristics. *Sci. Agri. Sin.* 34: 597-603.

Chen HD (2013). Transcriptome sequencing under drought stress, EST-SSR development and high density genetic linkage map construction in *Gossypium Darwinii*. Doctoral thesis, Chinese Agricultural Academy Sciences, Beijing, China, 47-48.

Deosarkar DB, Gupta VK and Gaikwad AR (2010). Genetic diversity studies in intra-specific desi cotton (*G. arboreum*) through DNA marker. *J. Cotton Res. Dev*. 24: 133-137.

Dong W (2007). The analysis of genetic diversity and SSR abundance in cotton germplasm. Doctoral thesis, Chinese Agricultural Academy Sciences, Beijing, China, 18-19.

Fu XQ (2012). Studying of SSR Markers-Assisted Breeding on Cotton. Proceedings of 2012 Annual Meeting and the 8th

Congress of China Cotton Association (in Chinese), 68-71.
Kantartzi SK, Ulloa M, Sacks E and Stewart JM (2009). Assessing genetic diversity in *Gossypium arboreum* L. cultivars using genomic and EST-derived microsatellites. *Genetica* 136: 141-147.
Kebede H, Burrow G, Dani RG and Allen RD (2007). A-genome cotton as a source of genetic variability for Upland cotton (*Gossypium hirsutum* L). *Genet. Resour. Crop Evol.* 54: 885-895.
Liu DQ, Guo XP, Lin ZX, Nie Y, et al. (2006). Genetic diversity of Asiatic cotton (*Gossypium arboreum* L.) in China evaluated by microsatellite analysis. *Genet. Resour. Crop Evol.* 53: 1145-1152.
Liu JD, Wang KB, Song GL, Su L, et al. (2003). Investigation report of cotton germplasm resources in Guangxi Province. *China Cotton* (in Chinese) 30: 16-18.
Mehetre SS, Aher AR and Gawande VL (2003). Induced polyploidy in *Gossypium*: a tool to overcome interspecific incompatibility of cultivated tetraploid and diploid cottons. *Curr. Sci*. 84: 1510-1512.
Rahman M, Yasmin T, Tabbasam N, Ullah I, et al. (2008). Studying the extent of genetic diversity among *Gossypium arboreum* L. genotypes/cultivars using DNA finger printing. *Genet. Resour. Crop Evol*. 55: 331-339.
Rohlf FJ (2000). Statistical power comparisons among alternative morphometric methods. *Am. J. Phys. Anthrop*. 111: 463-478.
Shen DZ (1993). Growth period and maturity of Asiatic cotton geographic family. *China Cotton* (in Chinese) 01: 10-11.
Sinha DP, Amudha J and Balasubramani G (2008). Studies of genetic diversity among drought tolerant tetraploid and diploid cotton by RAPD and ISSR markers. *Indian J. Crop Sci*. 3: 71-75.
Song GL, Cui RX and Wang KB (1999). Genetic Diversity Analysis of Australia Cotton Species with RAPD markers. *Cotton Sci.* (in Chinese) 11: 65-69.
Stanton MA, McD Stewart J and Tugwell NP (1992). Evaluation of *Gossypium arboreum* L. germplasm for resistance to thrips. *Genet. Resour. Crop Evol*. 39: 89-95.
Wang F, Kuang M and Xu HX (2012). Research Progress of SSR Molecular Markers Technology Applied in Cotton Variety Identification. Proceedings of 2012 Annual Meeting and the 8th Congress of China Cotton Association (in Chinese), 27-30.
Wang RH, Hu SA, Du J, Song GL, et al. (2003a). Investigation and collection report of cotton germplasm resources in Yunnan Province. *China Cotton* (in Chinese) 30: 10-13.
Wang ZH, Liu F, Wang KB, Song GL, et al. (2003b). Investigation report of cotton germplasm resources in Guizhou Province. *China Cotton* (in Chinese) 30: 23-25.
Xu QH, Zhang XL and Nie YC (2001). Genetic diversity evaluation of cultivars (*G. hirsutum* L.) from the Changjiang river valley and the Tellow river valley by RAPD markers. *Yi Chuan Xue Bao* 28: 683-690.
Yeh FC and Boyle TJB (1997). Population genetic analysis of co-dominant and dominant markers and quantitive traits. *Belgian J. Bot*. 129: 157.
Zhou ZL (2011a). Phenotypic diversity analysis of *G. arboreum* L. germplasm resources conserved in China. *J. Plant Genet. Resour.* 12: 881-889.
Zhou ZL (2011b). Genetic diversity analysis of *G. arboreum* L. conserved in China. Master's thesis, Chinese Agricultural Academy Sciences, Beijing, China, 31-35.
Zhou ZL, Du XM, Pan ZE and He SP (2013). Study on genetic diversity of *G.arboreum* L. germplasm resources with SSR markers. *Cotton Sci.* (in Chinese) 25: 217-226.