GMR

Molecular identification of variety purity in a cotton hybrid with unknown parentage using DNA-SSR markers

X.Q. Fu[1], F.X. Yang[1], X.K. Lu[1], X.G. Wang[1], B.X. Yang[2], F.J. Liu[1], Y. Liu[1] and J. Peng[1]

[1]Institute of Cotton Research, Chinese Academy of Agricultural Sciences, State Key Laboratory of Cotton Biology, Key Laboratory for Cotton Genetic Improvement, Ministry of Agriculture, Anyang, Henan, China
[2]Institute of Handan Agricultural Research, Handan, Hebei, China

Corresponding author: Jun Peng
E-mail: jun_peng@126.com

Genet. Mol. Res. 16 (3): gmr16039799
Received August 14, 2017
Accepted August 28, 2017
Published September 27, 2017
DOI http://dx.doi.org/10.4238/gmr16039799

ABSTRACT. Molecular identification of hybrid purity is difficult in regional trials of cotton varieties and hybrid trials. In particular, the molecular detection of hybrid purity has not yet been reported in the case of unknown parentage. In this study, we screened 5000 pairs of primers and chose 17 pairs of core simple sequence repeat (SSR) primers to determine the F1 purity of Han6402. The results showed that the purity based on SSR markers reached 100%. Twelve of the 17 pairs of primers exhibited co-dominant banding patterns, and 5 showed non-co-dominant banding patterns. Moreover, we constructed an F1 SSR fingerprinting profile that enabled the identification of the authenticity of Han 6402. Using these primers, we subsequently detected 44 individual F2 seedlings, and the results exhibited different extents of separation, in which the majority of genotypes were heterozygous with co-dominance at most of the loci that differed from each other. The results validated

the underlying heterozygous status of the F2 population at the molecular level. Therefore, we conclude that the set of core SSR primers can be used for the laboratory identification of the authenticity and purity of cotton hybrids, not only for distinguishing Fl hybrids or segregating F2 populations, but also for detecting volunteer seeds as fake F1 hybrids in the cotton hybrid industry, based on the hybrid fingerprinting.

Key words: Cotton; Core SSR primers; SSR-based purity; DNA fingerprinting; Molecular identification

INTRODUCTION

Compared to the regular cotton variety, hybrid cotton has the advantages of greater incremental range, better fiber quality, higher stress resistance, and, in particular, stronger growth in early stage and good resistance to seedling disease. Therefore, hybrid cotton has been rapidly popularized and utilized in cotton production, and hybrid cotton is appreciated by cotton growers. However, in recent years, F2 hybrids faked as F1 hybrids or F1 hybrids mixed with large amounts of parent seeds constantly occur in the cotton seed market, which have caused growers to suffer great losses (Wu et al., 2001). Therefore, identifying the authenticity and purity of cotton seeds remains an urgent problem.

Usually, authenticity and purity identification of cotton seeds involves regular distinctness, uniformity and stability (DUS) field tests, which take a long time to conduct the investigation and sensitive to environmental conditions. In the 1990s, with the rapid development of molecular biology, the utilization of molecular markers made it possible to identify the authenticity and purity in the laboratory.

Simple sequence repeat (SSR) markers have the advantages of good repeatability, easy operation and co-dominance; these markers are rich in quality and involve relatively established technology (Ali et al., 2011). SSRs have been widely applied to authenticate and identify the purity of crops such as corn and rice (Li et al., 1999; Wang et al., 2003; Islam et al., 2012). When a hybrid is detected using SSR markers, the parents should be provided, and then primers are selected on the condition of the known parent. This method is very effective in confirming the purity and authenticity of a plant (Singh et al., 2016).

However, in regional trials of national and provincial new varieties, the participant is not required to provide the parent materials of tested varieties. Hence, a purity test is performed on cases whereby the parentage is unknown. Our research team has gathered experience from years of purity tests involving tested varieties in regional trials of national cotton and has screened a series of core primers that have been applied to more than 700 cotton materials; these primers obtained good results. In the identification of the hybrids, SSR primers amplified in the test material display co-dominance or non-co-dominance in the absence of parentage information or materials. Non-co-dominance suggests that the SSR locus does not differ between the two parents of the corresponding material. Obviously, more primers showing co-dominance indicates a high SSR diversity of the parent, a distant genetic relationship and better heterosis.

By constructing fingerprinting profiles of both F1 individuals of Han6402 and the F2 population (44 plants) and comparing the profiles among plants, we then constructed fingerprinting profiles of varieties using core SSR primers, thereby enabling the identification

of hybrid cotton varieties from the conventional varieties and distinguishing F1 plants from the segregating F2 population.

MATERIAL AND METHODS

Cotton material

The Han6402 F1 material was provided by the Institute of National Cotton Regional Trial and was planted in the east test field of the Institute of Cotton Research of the Chinese Academy of Agricultural Sciences on April 27, 2016. Twenty-four plants were sampled from the field, and Han6402 F2 seeds were harvested in October 2016. The seeds were sterilized with 0.1% $HgCl_2$ and sowed in a growth chamber in March 2016. Ultimately, 44 seedlings were harvested.

DNA extraction

By combining the method of Paterson (Paterson et al., 1993) with the modified CTAB method (Porebski et al., 1997), DNA was extracted from true leaves in 2016 and from seedling cotyledons in 2012. Seventeen pairs of core SSR primers (Table 1) were applied to detect purity and to construct the fingerprinting profile of the F1 plants (Cotton Research Institute of Chinese Academy of Agricultural Sciences, 2011) as well as the fingerprinting profile of 44 isolates of the F2 population.

Table 1. Primer sequences for 17 pairs of core SSR primers used in the test.

No.	Primer name	Forward primers	Reverse primers
1	NAU1269	TACCTGAAACCCAAAATGGT	ACGCTGTTATAGGGCTCATC
2	NAU1186	AATGGTCCTGCTCCAGATT	AATCGTCGTCGTCGAATTAT
3	NAU1187	AACAAGAGCCAAGGTTCATC	GGATGCTGTATAGGGCTCAT
4	NAU 2026	GAATCTCGAAAACCCCATCT	ATTTGGAAGCGAAGTACCAG
5	NAU 1233	TTCGGGAAAGTTAGAGGAGA	TCCTCAGAGCTCGGAATAGT
6	DPL0431	CTATCACCCTTCTCTAGTTGCGTT	ATCGGGCTCACAAACATCA
7	NAU1102	ATCTCTCTGTCTCCCCCTTC	GCATATCTGGCGGGTATAAT
8	NAU1255	CATGCAAATCCATGCTAGAG	GGTTTCTTTGGTGGTGAAAC
9	NAU868	GGCAAAACCATAAGGGTAAC	TAGCGTGAGATTGTGGCTTA
10	NAU2343	GCTTTGCTTTGGAATGAGAT	TACTGCAACCCCTCACACT
11	NAU1085	AGTCGCCCCTTCTCTAATTT	TGTAAACCGAACTCGTTGTG
12	NAU2274	TCCTCGGATTATCAAAACCT	TGAAGAGGACATTGATGACG
13	CS62	GATGGCTACCTCCCTTTGTA	CGTAAGGAAGCCTAGCAAAA
14	NAU1071	ACCAACAATGGTGACCTCTT	CCCTCCATAACCAAAAGTTG
15	NAU1103	GGAGCCAGAAGTTGAGAAAA	TTCGGCTTCTGCTTTTACTT
16	NAU1369	TGGCAGAGATGAATGTAAGC	GGTAACGGATGGAAAATCAC
17	NAU2277	GAACTAGCCACATGATGCAC	TTGTTGAGGCATTAGTTTGC

PCR amplification

The mixture for the process of amplification included 1 µL DNA template, 1 µL buffer (containing $MgCl_2$), 0.3 µL dNTPs (10 mM), 0.5 µL forward primer (10 µM), 0.5 µL reverse primer (10 µM), 0.2 µL Taq polymerase (2.5 U/µL), and 6.5 µL ddH_2O. The program used was 94°C for 3 min (pre-degeneration); 30 cycles of 94°C for 30 s, 56°C for 45 s, 72°C for 45 s. After the 30 cycles, the reaction was held at 4°C.

Polyacrylamide gel electrophoresis

Amplification products were detected using 8.0% polyacrylamide gel electrophoresis. Two microliters of sample was added per well. The main steps included electrophoresis for 1 h at 200 V, fixation for 10 min, penetration for 12 min, and development for 5 min. Stop buffer was used to stop the reaction.

Identification of heterozygosis and homozygosis of genes

We selected 1825 pairs of polymorphic primers from more than 5000 pairs of primers and selected 17 pairs of core primers from the 1825 pairs of SSR primers. These 17 pairs can be used for the SSR-based purity identification of hybrid cotton in China. The bands are clear and easy to distinguish, with good polymorphism.

The bands of primers showing co-dominance have three types, which are defined as types I, II, and III. Types I and II are simple band types and are the same as the parent type, and type III is a co-dominant band type. The presence of type III bands for different primers indicates the genes at these loci are heterozygous. In this case, the cotton plants are determined to be heterozygous. The absence of type III amplification bands of all the primers suggests that the genes of these cotton lines are relatively homozygous.

RESULTS

Purity detection of Han6402 F1 individuals

Using 17 pairs of core SSR primers, the purity of F1 SSR markers of Han6402 was determined, and the results showed that the purity reached 100%. The first 12 pairs in Table 2 exhibited co-dominance, and the last 5 pairs displayed non-co-dominance. Figures 1, 2, 3, and 4 show 24 F1 SSR amplification bands in the presence of the primers of NAU1085, NAU1102, NAU2026, and CS62, respectively, which were consistent. This result indicated that there was no fake F1 hybrid, suggesting that the purity of the F1s was coincident at the molecular level. Among the primers, NAU1085, NAU1102, and NAU2026 exhibited co-dominance in the F1 individuals, whereas CS62 showed non-co-dominance.

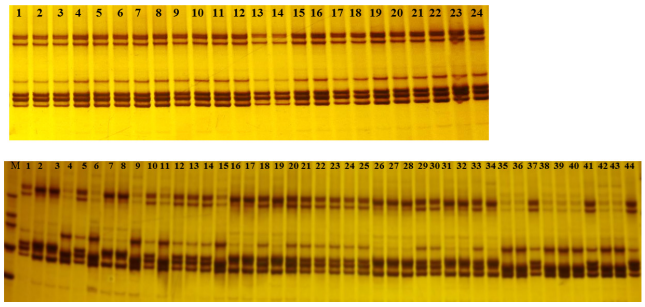

Figure 1. Banding patterns of F_1 hybrids and F_2 segregants generated with the SSR primer NAU1085. *Lane M*, DNA marker (500 bp); banding patterns of hybrid F_1 individuals 1-24 (above), assigned as type III; banding Patterns of F_2 segregating individuals 1-44 (below), assigned as types I, II and III. Type I: 4, 6, 9, 11, 15, 35, 36, 38, 39, 40, 42, 43. Type II: 2, 3, 7, 8, 16, 17, 19, 26, 27, 28, 31, 32, 34. Type III: 1, 5, 10, 12, 13, 14, 18, 20, 21, 22, 23, 24, 25, 29, 30, 33, 37, 41, 44.

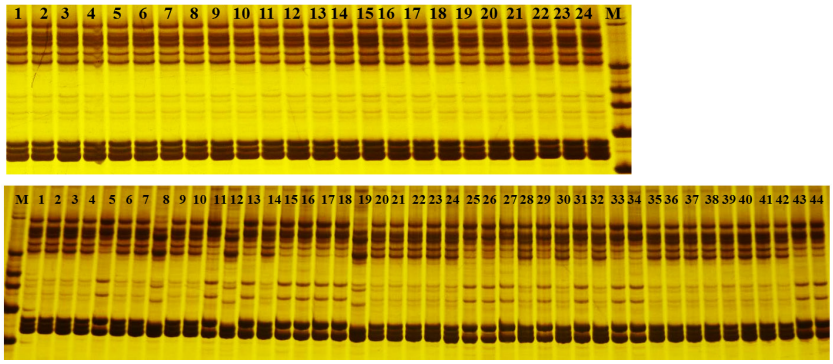

Figure 2. Banding patterns of F_1 hybrids and F_2 segregants generated with the SSR primer NAU1102. *Lane M*, DNA marker; banding patterns of hybrid F_1 individuals 1-24 (above), assigned as type III; banding patterns of F_2 segregating individuals 1-44 (below), assigned as types I, II, and III. Type I: 8, 12, 19. Type II: 5, 11, 13, 15, 16, 17, 18, 25, 26, 27, 29, 31, 33, 34, 43, 44. Type III: 1, 2, 3, 4, 6, 7, 9, 10, 14, 20, 21, 22, 23, 24, 28, 30, 32, 35, 36, 37, 38, 39, 40, 41, 42.

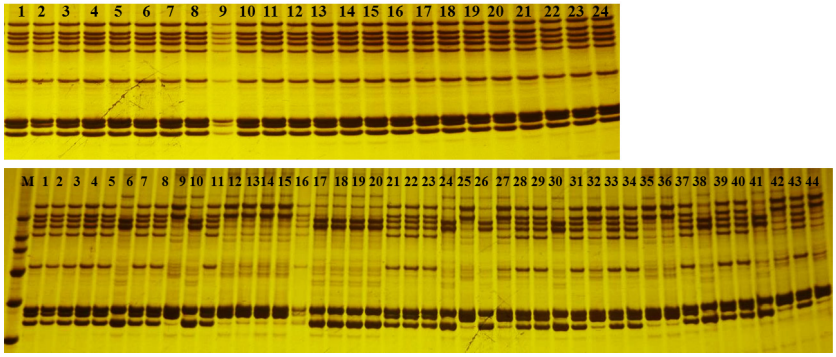

Figure 3. Banding patterns of F_1 hybrids and F_2 segregants generated with the SSR primer NAU2026. *Lane M*, DNA marker; banding patterns of hybrid F_1 individuals 1-24 (above), assigned as type III; banding patterns of F_2 segregating individuals 1-44 (below), assigned as types I, II, and III. Type I: 6, 10, 17, 18, 19, 20, 24, 26, 30, 38, 41. Type II: 9, 12, 13, 14, 15, 25, 27, 32, 35, 36, 42, 44. Type III: 1, 2, 3, 4, 5, 7, 8, 11, 16, 21, 22, 23, 28, 29, 31, 33, 34, 37, 39, 40, 43.

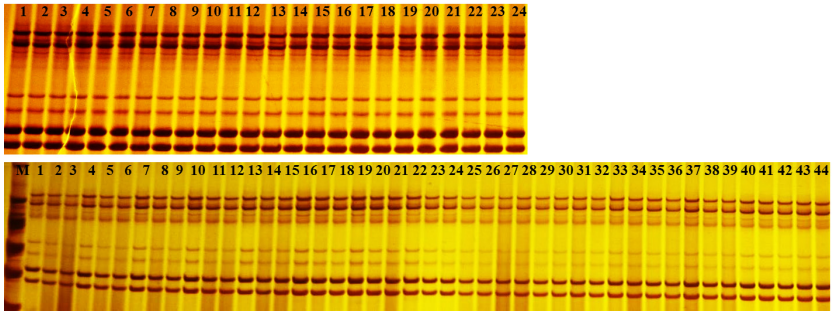

Figure 4. Banding patterns of F_1 hybrids and F_2 segregants generated with the SSR primer CS62. *Lane M*, DNA marker; banding patterns of hybrid F_1 individuals 1-24 (above), assigned as type III; banding patterns of F_2 segregating individuals 1-44 (below), assigned as type II.

Construction of the SSR fingerprinting profile of the Han6402 F1 plants

Once the SSR marker-based purity was determined, the fingerprinting profiles of 17 pairs of SSR primers in the F1 individuals were constructed. The electrophoretic band information was transformed into corresponding digital codes (Table 2) to form a unique identity mark of the corresponding variety. Such codes can serve as the foundation for the identification of authenticity and can be used to distinguish the plant from other varieties.

Molecular identification of Han6402 F2 individuals and determination of the profile and codes

Through amplification using 12 pairs of primers showing co-dominance, 44 F2 cotton seedlings all showed separation of band spectra. The three pairs in Figures 1, 2, and 3 had a high separation in F2 cotton seedlings. Five pairs showing non-co-dominance in the F1 seedlings exhibited consistent amplification patterns across 43 cotton seedlings, and cotton seedling No. 14 showed co-dominance through amplification using NAU1103 and NAU2277. This co-dominance may be caused by natural pollination with other varieties or mixtures of false F1 hybrids. As shown in Figure 4, 44 F2 cotton seedlings had no separation when CS62 was used, which was consistent with the non-co-dominance in the F1 plants. Cotton seedling No. 17 did not exhibit co-dominant bands when using 17 pairs of primers. This cotton line may be a homozygous line or a foreign homozygous variety, or co-dominance was not detected due to insufficient SSR primers. The other 42 had co-dominant bands when different primers were used for amplification, which were all heterozygous plants. The number of primers showing co-dominance ranged from 3 to 10, and each cotton seedling was different from the F1 seedlings; there were also differences between the 42 cotton seedlings. The amplification patterns I, II and III of individual F2 cotton seedlings were assigned as digital codes 1, 2 and 3 (Table 2), respectively; thus, the fingerprinting profiles of the F2 individuals were constructed.

In the table, in order to present the results easily, we defined band patterns types I, II, and III as types 1, 2, and 3. First line represents the number of individulals, and the last line represents the total of co-dominance markers.

Verification of heterozygous status of the F2 population at the molecular level

Five pairs of primers showing non-co-dominance were used for amplification in the F2 population. The results showed stable band spectra but with variation among individuals. Twelve pairs of primers with co-dominance were used for amplification in the F2 population, and the separation rate of the "parent 1:parent 2:hybrid" is shown in Table 3. Because of the small number of individuals analyzed, the ratio was not strictly 1:1:2 but was close. One SSR locus was separated into three genotypes, but for multiple loci, the heterozygosis among separated plants was very high.

Differences between F2 and conventional cotton populations

When primers showing co-dominance were adopted for the amplification of F1 hybrids, each plant exhibited co-dominance with uniform type III bands. When primers showing non-co-dominance were adopted, each plant showed consistent simple band patterns. When different primers were used, the band of each F2 population was inconsistent, and the

vast majority of the F2 seedlings were heterozygotes. However, each of the conventional cotton populations showed simple type I or type II bands without co-dominant band 3 when different primers were employed. F1, F2 and conventional cotton populations can be distinguished according to the structures of the bands of different plants. The doped components in cotton populations can be preliminarily identified by comparing the bands of individual plants.

Table 2. Digital coding transformed from the banding patterns of F_1 hybrids and F_2 separating individuals of Han6402.

Generation	No. of individuals	NAU 1269	NAU 1186	NAU 1187	NAU 2026	NAU 1233	DPL 0431	NAU 1102	NAU 1255	NAU 868	NAU 2343	NAU 1085	NAU 2274	CS62	NAU 1071	NAU 1103	NAU 1369	NAU 2277	Number of Co-dominance markers
F_1		3	3	3	3	3	3	3	3	3	3	3	3	2	2	2	2	2	12
F_2	1	1	3	1	3	3	2	3	1	2	3	3	3	2	2	2	2	2	7
	2	1	2	3	3	1	2	3	1	2	1	2	2	2	2	2	2	2	3
	3	3	2	3	3	3	3	3	2	2	2	2	2	2	2	2	2	2	6
	4	3	1	1	3	3	3	3	3	1	2	1	3	2	2	2	2	2	7
	5	2	2	2	3	2	1	2	2	3	1	3	2	2	2	2	2	2	4
	6	2	1	3	1	2	1	3	3	2	3	1	1	2	2	2	2	2	4
	7	1	3	3	3	1	2	3	2	3	1	2	1	2	2	2	2	2	5
	8	1	3	3	3	2	1	1	2	3	3	2	1	2	2	2	2	2	5
	9	2	1	1	2	2	1	3	1	3	2	1	3	2	2	2	2	2	3
	10	2	2	1	1	1	2	3	3	1	1	3	3	2	2	2	2	2	4
	11	3	3	3	3	1	2	2	2	3	2	1	3	2	2	2	2	2	7
	12	3	2	1	2	3	3	1	3	3	3	3	1	2	2	2	2	2	7
	13	3	3	1	2	1	2	2	1	3	2	3	3	2	2	2	2	2	5
	14	2	3	2	2	2	2	3	1	3	2	3	2	2	2	3	2	3	8
	15	3	1	3	2	3	3	2	1	2	3	1	3	2	2	2	2	2	6
	16	1	3	2	3	3	3	2	3	2	1	2	2	2	2	2	2	2	5
	17	1	2	2	1	2	1	2	1	2	1	2	2	2	2	2	2	2	0
	18	2	2	3	1	2	1	2	3	3	3	3	2	2	2	2	2	2	5
	19	1	3	2	1	2	1	1	3	3	3	2	3	2	2	2	2	2	5
	20	1	3	1	1	3	3	3	1	1	1	3	2	2	2	2	2	2	5
	21	3	2	2	3	1	2	3	1	3	2	3	3	2	2	2	2	2	6
	22	3	1	3	3	1	2	3	3	2	3	3	3	2	2	2	2	2	8
	23	3	2	2	3	1	1	3	1	3	3	3	1	2	2	2	2	2	6
	24	1	3	3	1	1	1	3	3	3	3	3	3	2	2	2	2	2	9
	25	1	3	2	2	3	3	2	3	3	3	3	2	2	2	2	2	2	8
	26	3	1	3	1	2	1	2	3	3	2	2	2	2	2	2	2	2	4
	27	3	1	2	2	3	3	2	3	2	3	2	3	2	2	2	2	2	6
	28	3	3	2	3	3	3	3	2	3	3	2	3	2	2	2	2	2	10
	29	3	1	1	3	1	2	2	3	3	3	3	3	2	2	2	2	2	8
	30	2	3	3	1	2	1	3	2	1	2	3	2	2	2	2	2	2	5
	31	2	3	3	3	1	2	2	2	3	3	2	3	2	2	2	2	2	6
	32	2	3	2	2	3	3	3	3	2	3	2	3	2	2	2	2	2	8
	33	2	3	1	3	3	3	2	2	3	3	3	3	2	2	2	2	2	8
	34	3	3	2	3	3	3	2	3	3	1	2	3	2	2	2	2	2	8
	35	3	1	2	2	2	2	3	3	3	3	1	2	2	2	2	2	2	6
	36	3	1	1	2	3	3	3	2	1	1	1	3	2	2	2	2	2	6
	37	3	3	3	3	2	2	3	3	3	2	3	3	2	2	2	2	2	9
	38	1	3	1	1	2	2	3	3	3	3	1	3	2	2	2	2	2	6
	39	3	1	3	3	1	2	3	1	1	3	1	1	2	2	2	2	2	5
	40	1	1	3	3	1	2	3	1	1	3	1	2	2	2	2	2	2	5
	41	2	2	3	1	3	2	3	3	2	3	3	3	2	2	2	2	2	7
	42	3	2	3	2	3	3	3	2	1	3	1	2	2	2	2	2	2	7
	43	1	2	1	3	3	3	2	3	3	1	1	2	2	2	2	2	2	6
	44	1	3	1	2	3	3	2	3	3	2	3	3	2	2	2	2	2	8

Table 3. Separate ratio of SSR markers of the 17 pairs of core primers.

Types of the banding pattern	NAU 1269	NAU 1186	NAU 1187	NAU 2026	NAU 1233	DPL 0431	NAU 1102	NAU 1255	NAU 868	NAU 2343	NAU 1085	NAU 2274	CS62	NAU 1071	NAU 1103	NAU 1369	NAU 2277
1	14	12	13	11	13	11	3	12	8	10	12	6	0	0	0	0	0
2	11	12	13	12	12	17	16	11	11	11	13	15	44	44	43	44	43
3	19	20	18	21	18	16	25	21	25	23	19	23	0	0	1	0	1
Other					1												
Total	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44

DISCUSSION

On the premise of good purity of F1 seedlings, the heterozygous status of the F2 seedlings was investigated at the molecular level. In the case of the use of multiple primers for amplification, F2 isolates were highly heterozygous at the genetic level. The differences in multiple gene loci can lead to inconsistencies in phenotypic traits, poor purity in the field, and the separation of economic characters, causing inconsistencies in yield and quality. The mass application of F2 plants in production causes hybrid cotton to fail in providing its advantages, leading to negative influences on the promotion of hybrid cotton.

The rapid laboratory identification of the authenticity of the cotton and the purity detection of SSR markers are helpful for seed companies to control seed quality and for seed regulators to conduct effective management. The construction of fingerprinting profiles of cotton varieties is helpful for breeders to protect the property rights of cotton varieties. At various levels in the process of regional cotton trials, there are a few breeders rushing to use low-generation material for tests. Such varieties were once identified as new varieties, but they easily degrade and cause losses to the farmers if promoted and applied. Therefore, it is necessary to eliminate low-purity varieties though SSR purity detection. Certain breeders especially use F1 plants as a substitute for conventional cotton in trials, resulting in unfair competition or the use of repeated varieties for trials. This situation can only be effectively controlled by comparing the fingerprints of tested material. We expect to establish a set of core SSR primers to identify each tested material, thereby building a platform of fair competition for testing varieties.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by grants from the Special Funds for Basic Research of the Basic Scientific Research Institutes of Central Public Welfare (#1610162016018).

REFERENCES

Ali S, Ijaz A, Zafar A, Ashraf A, et al. (2011). Investigation of simple sequence repeats (SSR) marker-assisted genetic diversity among upland Bt- and non Bt-cotton varieties. *Afr. J. Biotechnol.* 10: 15222-15228. https://doi.org/10.5897/AJB11.894

Cotton Research Institute of Chinese Academy of Agricultural Sciences (2011). A novel kit to distinguish the cotton normal and hybrid varieties. China Patent ZL 2009 1 0235514.4.

Islam MN, Molla MR, Rohman MM, Hasanuzzaman M, et al. (2012). DNA fingerprinting and genotyping of cotton varieties using SSR markers. *Not Bot Horti Agrobo* 40: 261-265.

Li YH, Xiao H, Zhang CQ, Hu GC, et al. (1999). Genetic variation of main parents of hybrid rice in China was revealed with simple sequence repeat markers. *Zhiwuxue Tongbao* 41: 1061-1066.

Paterson AH, Brubaker CL and Wendel JF (1993). A rapid method for extraction of cotton (*Gossypium spp*.) genomic DNA suitable for RFLP or PCR analysis. *Plant Mol. Biol. Report.* 11: 122-127. https://doi.org/10.1007/BF02670470

Porebski S, Bailey LG and Baum BR (1997). Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol. Biol. Report.* 15: 8-15. https://doi.org/10.1007/BF02772108

Singh N, Choudhury DR, Tiwari G, Singh AK, et al. (2016). Genetic diversity trend in Indian rice varieties: an analysis using SSR markers. *BMC Genet.* 17: 127. https://doi.org/10.1186/s12863-016-0437-7

Wang FG, Zhao JR, Guo JL and Liu LZ (2003). Series of research on establishing DNA fingerprinting pool of Chinese new maize cultivars I. The establishment of a standard SSR system fitting for maize cultivars' identification. *Yumi Kexue* 11: 3-6.

Wu YT, Zhang TZ, Guo WZ and Yin JM (2001). Detecting polymorphism among upland cotton (*Gossypium hirsutum* L.) cultivar sand their roles in seed purity of hybrids with SSR markers. *Mianhua Xuebao* 13: 131-133.