

Sequence-related amplified polymorphisms (SRAPs) reveal genetic diversity and variation regions in upland cotton (*Gossypium hirsutum* L.) in China

Crop Breeding and Applied Biotechnology
18: 155-160, 2018
Brazilian Society of Plant Breeding.
Printed in Brazil
<http://dx.doi.org/10.1590/1984-70332018v18n2a22>

Jinna Hou¹, Sufang An¹, Yongjuan Liu¹, Li Lei² and Baoquan Li^{1*}

Abstract: *The lack of genetic diversity is an obstacle for genetic improvement of upland cotton in China; thus, new technologies must be developed to produce more polymorphic molecular markers associated with agricultural traits of the existing resources for breeding. A highly efficient and economical technology of sequence-related amplified polymorphism (SRAP) molecular markers with an automated fragment analyzer ABI3500xl was developed, to detect genetic diversity in upland cotton. Using this new strategy, we easily screened for polymorphisms with 7,872 pairs of SRAP primers, and detected 504 polymorphic markers. Of all these, 165 were used for genetic diversity analysis in 128 upland cotton varieties collected nationwide in China. Our method combined a traditional molecular marker development technology with an economical and easy operation strategy for breeders.*

Key words: *SRAP, genetic diversity, high efficient molecular marker, upland cotton.*

INTRODUCTION

Upland cotton (*Gossypium hirsutum* L., $2n=AADD=52$) is the most important textile fiber crop in the world. It is one of the four wild *Gossypium* species that were independently domesticated by aboriginal domesticators more than 5000 years ago and transformed into fiber plants (Smith and Stephens 1971). Due to the high yield, long fiber and wide environmental adaptability, upland cotton became the most common *Gossypium* species on the market, accounting for more than 90% of the world production (Dhivya et al. 2016). The planting area of upland cotton reached up to four million hectares in 2014, with an average yield of 1463.25 kg ha⁻¹ in China (National Bureau of Statistics of People's Republic of China, 2014, <http://www.stats.gov.cn/>). Genetic improvement (*e.g.* hybridization and transgenic introgression for insect resistance) plays a key role in the great success of cotton products.

In spite of the significant achievements of upland cotton breeding, the genetic diversity is extremely low in China, according to recent studies (Wendel 1992, Ahmad et al. 2012). The low genetic diversity can be explained by three reasons: 1) upland cotton originated from South America (Basu 1996) and the absence of wild varieties and naturally occurring germplasm in China, resulting in genetic barriers against the introduction of wild resources into the available

***Corresponding author:**
E-mail: lbq308@126.com

Received: 20 December 2016
Accepted: 01 July 2017

¹ Henan Academy of Agricultural Sciences, Crop Designing Centre, Zhengzhou, Henan, 450002, China

² University of Minnesota, Department of Agronomy & Plant Genetics, St. Paul, MN, 551085

germplasm; 2) most of the cotton varieties planted in China were imported from the United States, resulting in simple genetic origins; and 3) the creation of insect-resistant plants by genetic engineering restricted the core germplasm collection to a few elite genotypes.

A low level of genetic diversity is a major obstacle for genetic research and improvement of upland cotton in China. Firstly, under low genetic diversity only few useful marker can be identified to construct the genetic linkage map. An integrated map covering 7,424 molecular markers was constructed from 28 mapping populations of tetraploid cotton (Yu et al. 2010). Among those 28 mapping populations, 21 were derived from *G. hirsutum* x *G. barbadense*, one from *G. hirsutum* x *G. tomentosum*, and six from two *G. hirsutum* parents. The latter six populations with parent *G. hirsutum* contributed with only 574 markers to the integrated map (Yu et al. 2010). The restricted number of markers is a limitation for dissecting the genetic basis of complex agricultural traits and map-based gene cloning. Secondly, low genetic diversity hampers the screening of molecular markers associated with agricultural traits. The linkage disequilibrium (LD) of upland cotton was found to be about 3-4 cM and 1 cM is equal to a physical distance of 450 kb in upland cotton. In molecular marker-assisted selection, the LD block may harbor several antagonistic traits. More markers will be needed to break LD. Thirdly, low allele variation limits the gene source for genetic improvement, e.g., for choosing parents and predicting the heterosis level.

To accurately evaluate genetic diversity in plant species, especially in crop cultivars with low genetic diversity, researchers have been trying to explore more molecular markers to evaluate genetic diversity. There are many types of molecular marker, such as AFLP, TRAP, SSR, SNP (Vos et al. 1995, Powell et al. 1996, Zhang and Stewart 2000, Chen and Du 2006, Fang et al. 2010, Wang et al. 2012a, Klabunde et al. 2016) and Sequence-Related Amplified Polymorphism (SRAP) markers, a PCR system based on primer pair amplification, prior to the amplification of open reading frames and even distribution in the genome (Li and Quiros 2001). Compared with other molecular markers, SRAP has advantages for being more economical and efficient than SSR and SNP, simpler than AFLP, and more reproductive than RAPD. A set of upstream and downstream primers can be combined randomly for PCR amplification, producing plenty of primer pairs for SRAP. These characteristics of SRAP facilitate polymorphism detection among genotypes with a low level of genetic diversity, as for example in upland cotton.

In this study, SRAP technology was used to screen polymorphisms or genetic markers in a panel of 128 upland cotton varieties from the two main cotton-producing regions of China. We are trying to detect more molecular markers among upland cotton varieties with highly efficient and lower cost technology. With these, the genetic diversity of upland cotton varieties was evaluated. Moreover our results will provide more molecular markers for assisting breeding and accelerate the process of upland cotton breeding.

MATERIAL AND METHODS

Plants material

A panel of 128 upland cotton varieties was collected from two main ecological cotton-producing areas in China: the Yellow and Yangtze river regions. These varieties include typical cultivars bred by major cotton breeding institutes in China, which inherited the accumulated genetic diversity of cotton germplasm since the 1950s. The varieties are described in Table 1. All experiments were carried out in Zhengzhou (lat 34° 44' N, long 113° 42' E, alt 110 m asl), China.

DNA extraction

The seed coats of upland cotton plants were removed after the seeds were immersed in H₂O overnight at room temperature. Then the cotyledons were isolated and ground with CTAB preheated to 65 °C in a mortar. From the ground

Table 1. Summary of information of varieties in this study

Ecological area	Province of origin	Number of varieties
Yellow river region	Beijing	2
	Hebei	18
	Henan	39
	Shandong	14
	Shanxi	2
	Anhui	6
Yangtze river region	Hubei	19
	Hunan	4
	Jiangsu	2
	Jiangxi	2
Northwest of China	Xinjiang	7
Unknown		13
Total number of varieties		128

cotyledons, genomic DNA was extracted according to Zhang and Stewart (2000), with some modifications. The DNA quality was tested in agarose gel and manually adjusted to a final concentration of 50 ng μL^{-1} .

Polymorphism detection among varieties

One hundred and twenty-eight DNA samples were pipetted in 384-well PCR plates with three replications, and eight samples randomly selected in rows were used as template. Nine upstream primers of SRAP were labeled with different dyes to detect the PCR products with ABI3500xl. The PCR products with distinct bands and unbiased distribution among eight samples were considered polymorphic.

Genotyping varieties with polymorphism revealed by SRAP primers

Primer pairs with polymorphic amplification were used to amplify DNA samples of a total of 128 varieties. At least two PCR products derived from different fluorescence primers were mixed when genotyping with ABI3500xl to make the procedure more effective and economical. To avoid residues from different dyes used to label different primers, SRAP products with varied length or at least divergent in polymorphic bands were mixed for electrophoresis on ABI3500xl. For dominant markers, 1 and 0 were assigned to presence and absence of amplified fragments, respectively; for co-dominant markers, 1 and 0 represented polymorphic bands with different length.

Analyzing genetic diversity

Genetic similarities were calculated using the Numerical Taxonomy Multivariate Analysis System software package NTSYS-pc Version 2.10e. Cluster analysis was performed using the UPGMA (Unweighted Pair-Group Method based on Arithmetic Means) obtained from genetic distances. The UPGMA tree was constructed using the SHAN tool of NTSys-pc 2.10e software

RESULTS

Polymorphisms among varieties indicate lack of genetic diversity

Although SRAP markers are more efficient to detect polymorphisms, and 504 polymorphic markers were detected using 7,872 pairs of SRAP primers, the ratio of polymorphism was only 6.4%. This result is consistent with previous research and indicates low genetic diversity among upland cotton varieties in China. We chose 72 primer pairs with higher amplification efficiency and reproducibility (Figure 1A), which produced 165 polymorphic fragments. More than 60% of the primer pairs generated 1-2 polymorphisms, about 20% produced three, and 10% amplified more than four (Table 2). Using these primers, 128 varieties were genotyped. Most SRAPs in upland cotton are dominant (Figure 1B) and a few co-dominant among the different varieties (Figure 1C).

Table 2. Information of polymorphisms produced by different primer pairs

No. of polymorphisms produced by each primer pair	No. of primer pairs	Percentage in all tested primer pairs (%)
1	29	40.28
2	16	22.22
3	14	19.44
4	7	9.72
5	4	5.56
6	1	1.39
8	1	1.39
Total	165	100

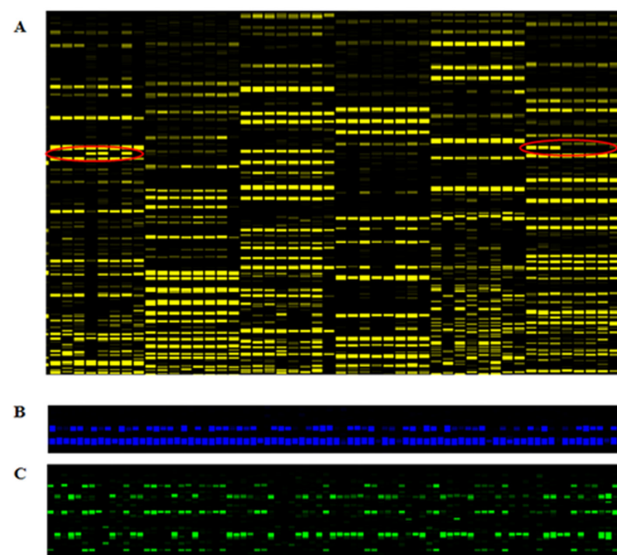


Figure 1. Polymorphism screening and cultivar genotyping. A. Polymorphism screening with eight DNA upland cotton cultivars. Different primer pairs have distinct band profiles. Cycles indicate the polymorphisms with higher amplification efficiency, selected to genotype the 128 varieties. B and C. Show amplification profiles of dominant and co-dominant polymorphism of SRAPs, respectively.

Genetic diversity of upland cotton collection

The genetic diversity of 128 cultivars was evaluated with 165 polymorphic markers. The genetic similarity coefficients among different cultivars were between 0.48 and 0.80 (Figure 2). The similarity between Lumianyan37 and Xiangzhamian8 was highest (0.80), and second highest between Zhongmianyan30 and Zhongmiansuo41 (0.79) (Figure 2). All these cultivars were bred for insect resistance with transgenic Bt gene, and the high genetic similarity may be due to lines with a similar genetic background. In contrast, the similarity coefficients between Xinluzao19 and Huazhamian4, from the Autonomous Region of Uyghur in the Xinjiang province in northwestern China and the Yangtze valley in the Jiangxi Province, respectively, were the lowest (0.48) (Figure 2).

To understand the genetic relationships of the 128 varieties, we performed cluster analysis. The varieties were divided into four clusters with a threshold of 0.52 (Figure 2). Cluster I, II, and IV included 36, 57 and 34 varieties, respectively, while cluster III contained only one (Xinluzao19). Each cluster contained varieties from different cultivation regions (Yellow River Valley, Yangtze River Valley and northwestern China) and different cotton types (hybrids and conventional).

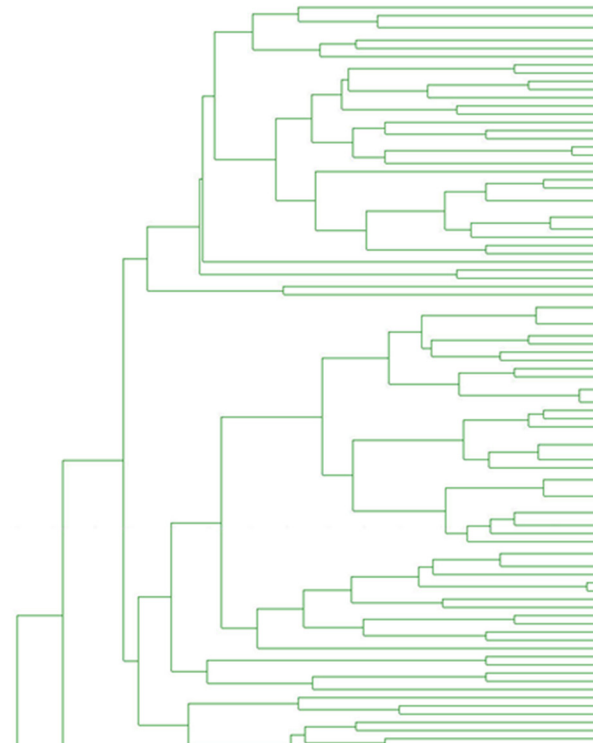


Figure 2. Genetic diversity in 128 upland cotton varieties based on SRAPs.

Divergent SRAP fragments among three clusters

To demonstrate the significant divergence of SRAP-related genomic regions among the three larger clusters (I, II and IV), 10 varieties from each were randomly selected to construct a genotype graph using 165 polymorphic fragments (Figure 3). Three sets of polymorphic markers (Figure 3) were selected to analyze the characteristic of the genomic regions that accumulated variation during breeding selection. Set A displays a similar genotype profile between cluster I and II, but differs from cluster IV, while set B shows a similar genotype profile between cluster I and IV, but differs from cluster II. Rich recombination among markers in set C was observed in the clusters. Markers from each cluster prefer to linkage with others in the same set except cluster I in set

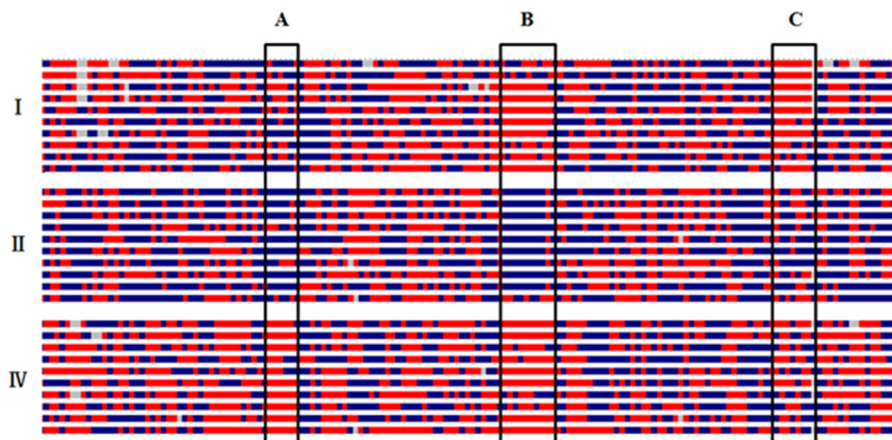


Figure 3. Genotype graph constructed from SRAPs of three clusters. Bars represent the varieties of each cluster. Red and blue background colors show different alleles at the same polymorphic site. Framed regions indicate significantly divergent main regions among clusters.

Sequence-related amplified polymorphisms (SRAPs) reveal genetic diversity and variation regions in upland cotton (*Gossypium...*

A and cluster IV in set C. Primer sequences of those markers were aligned to a genome sequence of *G. hirsutum* TM-1 in the database (<https://www.cottongen.org>), and most of them were located around the microsatellite sequence. Those results indicate that the microsatellite regions prior to accumulate most genetic diversity during long period of selection with similar goals.

DISCUSSION

A number of research projects have characterized genome sequences of cotton species, for example, the genome of allotetraploid upland cotton “TM-1” (Zhang et al. 2015). Owing to genomic complexity and limited germplasm available in modern breeding, detection of diversities for genetic improvement is still a challenging task for breeders (Borba et al. 2005, Guo et al. 2010, Kawuki et al. 2011). In our research, we propose the SRAP technology combined with four types of fluorescence evaluated on ABI 3500xl as a more efficient and economical way to develop polymorphic markers in upland cotton. One pair of SRAP primers amplifies and detects 1-8 polymorphic loci, i.e., with higher efficiency than SSR. Since SRAP primers from upstream and downstream groups should randomly combine with each other in amplification, the number of primer pairs is much higher than for other types of primers. In addition, since the nucleotide sequences of gene coding regions were highly conserved among different plant species, the same set of primers can be used in different crops. The amplicons with primers labeled by four types of fluorescences can be mixed and detected at the same time using ABI 3500xl, which can highly improve efficiency and reduce costs. SRAP first amplifies the open reading frames and the markers are distributed evenly in the genome. Those characteristics can improve the usage of molecular markers to explore the genetic basis of agronomic traits for genetic improvement in crops.

Considering the low genetic diversity of upland cotton cultivars and germplasm, we propose several strategies of genetic improvement in China: developing new technologies for producing more polymorphic molecular markers to explore the genetic basis of key agronomic traits in the available germplasm; dissecting and transferring important genes controlling valuable agronomic traits from wild cotton species into elite lines. Molecular markers such as SRAPs and others, e.g., SSR and SNP, need to be combined with other technologies to develop a high-throughput and low-cost marker detection method (Ahmad 2004). Decades of intensive artificial selection for similar goals in upland cotton in China have severely depleted the genetic diversity available for breeding. Refer to elite gene dissecting and transferring, more polymorphisms should be detected for gene cloning and selection for break the linkage drag of harmful traits.

The genomes of both diploid and tetraploid *Gossypium* species were sequenced, and the draft genome sequence of upland cotton and its two ancestors were published. However, there are a series of gaps in those genomes (Yu et al. 2010, Wang et al. 2012b, Zhang et al. 2015). The SRAP fragments are evenly distributed in the whole genome, but those fragment sequences cannot be mapped into the known physical map, due to gaps in the genomes. Thus, SRAP product fragments will provide a sufficient amount of sequence information for filling the gaps.

ACKNOWLEDGMENTS

The study was financially supported by Program of Financial budget of Henan province 2015 (No. ycy20157816).

REFERENCES

- Ahmad MQ, Khan SH and Azhar FM (2012) Decreasing level of genetic diversity in germplasm and cultivars of upland cotton (*Gossypium hirsutum*) in Pakistan. *Journal of Agriculture & Social Sciences* 8: 92-96.
- Ahmad R (2004) Genotyping of peach and nectarine cultivars with SSR and SRAP molecular markers. *Journal of the American Society for Horticultural Science American Society for Horticultural Science* 129: 204-210.
- Basu AK (1996) Current genetic research in cotton in India. *Genetica* 97: 279-290.
- Borba TCDO, Brondani RV, Rangel PHN and Brondani C (2005) Evaluation of the number and information content of fluorescent-labeled SSR markers for rice germplasm characterization. *Crop Breeding and Applied Biotechnology* 5: 157-165.
- Chen G and Du XM (2006) Genetic diversity of source germplasm of Upland cotton in China as determined by SSR marker analysis. *Acta Genetica Sinica* 33: 733-745.
- Dhivya K, Sathish S, Balakrishnan N, Udayasuriyan V and Sudhakar D (2016) Genetic engineering of cotton with a novel cry2AX1 gene to impart insect resistance against *Helicoverpa armigera*. *Crop Breeding and Applied Biotechnology* 16: 205-212.
- Fang DD, Xiao JH, Canci PC and Cantrell RG (2010) A new SNP haplotype

- associated with blue disease resistance gene in cotton (*Gossypium hirsutum* L.). **Theoretical and Applied Genetics** **120**: 943-53.
- Guo XY, Chen ZH, Zhu YF, Wang AG and Wu C (2010) Characteristic of Tuxpeno germplasm and its utilization in maize breeding in China. **Guizhou Agricultural Sciences** **1**: 4-6.
- Kawuki RS, Ferguson M, Labuschagne MT, Herselman L, Orone J, Ralimanana I, Bidiaka, Lukombo S, Kanyange MC, Gashaka, Mkamillo, Gethi J and Obiero H (2011) Variation in qualitative and quantitative traits of cassava germplasm from selected national breeding programmes in sub-Saharan Africa. **Field Crops Research** **122**: 151-156.
- Klabunde GHF, Junkes CFO, Tenfen SZA, Dantas ACM, Furlan CRC, Mantovani A, Denardi F, Boneti JI and Nodari O (2016) Genetic diversity and apple leaf spot disease resistance characterization assessed by SSR markers. **Crop Breeding and Applied Biotechnology** **16**: 189-196.
- Li G and Quiros CF (2001) Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in Brassica. **Theoretical and Applied Genetics** **103**: 455-461.
- Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S and Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. **Molecular Breeding** **2**: 225-238.
- Smith CE and Stephens SG (1971) Critical identification of Mexican archaeological cotton remains. **Economic Botany** **25**: 160-168.
- Vos P, Hogers R, Bleeker M, Reijans M, Lee TVD, Hornes M, Frijters A, Pot J, Peleman J and Kuiper M (1995) AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Research** **23**: 4407-14.
- Wang B, Nie Y, Lin Z, Zhang X, Liu J and Bai J (2012a) Molecular diversity, genomic constitution, and QTL mapping of fiber quality by mapped SSRs in introgression lines derived from *Gossypium hirsutum* X *G. darwinii* Watt. **Theoretical and Applied Genetics** **125**: 1263-74.
- Wang K, Wang Z, Li F, Ye W, Wang J, Song G, Yue Z, Cong L, Shang H, Zhu S, Zou C, Li Q, Yuan Y, Lu C, Wei H, Gou C, Zheng Z, Yin Y, Zhang X, Liu K, Wang B, Song C, Shi N, Kohel RJ, Percy RG, Yu JZ, Zhu Y, Wang J and Yu Shuxun (2012b) The draft genome of a diploid cotton *Gossypium raimondii*. **Nature Genetics** **44**: 1098-103.
- Wendel JF (1992) Genetic Diversity in *Gossypium hirsutum* and the origin of upland cotton. **American Journal of Botany** **79**: 1291-1310.
- Yu J, Kohel RJ and Smith CW (2010) The construction of a tetraploid cotton genome wide comprehensive reference map. **Genomics** **95**: 230-40.
- Zhang J and Stewart JM (2000) Economical and rapid method for extracting cotton genomic DNA. **Journal of Cotton Science** **4**: 193-201.
- Zhang T, Hu Y, Jiang W, Fang L, Guan X, Chen J, Zhang J, Saski CA, Scheffler BE, Stelly DM, Hulse-Kemp AM, Wan Q, Liu B, Liu C, Wang S, Pan M, Wang Y, Wang D, Ye W, Chang L, Zhang W, Song Q, Kirkbride RC, Chen X, Dennis E, Llewellyn DJ, Peterson DG, Thaxton P, Jones DC, Wang Q, Xu X, Zhang H, Wu H, Zhou L, Mei G, Chen S, Tian Y, Xiang D, Li X, Ding J, Zuo Q, Tao L, Liu Y, Li J, Lin Y, Hui Y, Cao Z, Cai C, Zhu X, Jiang Z, Zhou B, Guo W, Li R and Chen ZJ (2015) Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. **Nature Biotechnology** **33**: 531-537.

