Genetics and Molecular Research

Construction and characterization of a bacterial artificial chromosome library for the allotetraploid *Gossypium tomentosum*

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

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

Corresponding authors: F. Liu / K.B. Wang
E-mail: liufcri@163.com / wkbcri@163.com

Genet. Mol. Res. 14 (4): 16975-16980 (2015)
Received February 18, 2015
Accepted June 25, 2015
Published December 15, 2015
DOI http://dx.doi.org/10.4238/2015.December.15.3

ABSTRACT. *Gossypium tomentosum* is a wild allotetraploid species with the $(AD)_5$ genome. It is characterized by many useful traits including finer fiber fineness, drought tolerance, and *Fusarium* and *Verticillium* resistance. We constructed the first bacterial artificial chromosome library for *Gossypium tomentosum*. With high quality and broad coverage, this library includes 200,832 clones, with an average insert size of about 122 kb and fewer than 3% empty clones. Our library is approximately 10-fold the size of the $(AD)_5$-genome (2400 Mb) and provides a 99.7% probability of isolating genes of interest or their sequences. Seven of eight simple sequence repeats markers that are located on five different chromosomes and linked with resistance to *Verticillium* wilt could amplify the 50 superpools and obtained one to five hits. This high capacity library will be an important genomic resource for classifying and analyzing the evolution of allotetraploid cotton species as well as for isolating disease-resistance and drought-tolerance

genes.

Key words: Wild cotton; *Gossypium tomentosum*; Drought tolerance; Bacterial artificial chromosome library; *Verticillium* wilt resistance; SSR marker

INTRODUCTION

Cotton (*Gossypium* spp) is the leading fiber crop worldwide. There are 50 recognized *Gossypium* species (Stewart, 1995; Ma et al., 2008) that are distributed throughout the tropical and subtropical regions of the world (Fryxell, 1979). Among the 50 species of *Gossypium*, 45 are diploid (2n = 2x = 26) and five are allotetraploid (2n = 4x = 52) (Fryxell, 1979; Stewart, 1995; Brubaker et al., 1999; Zhang et al., 2005). The diploid *Gossypium* species are further grouped into eight diploid genomes (A, B, C, D, E, F, G, and K) and the five allotetraploid species are grouped into one tetraploid genome (AD) (Stewart, 1995). There are four cultivated species of cotton around the world: two diploids, *G. herbaceum* and *G. arboreum*, and two allotetraploids, *G. barbadense* and *G. hirsutum*. The majority of the world's fiber production comes from allotetraploids (Stewart, 1995). *Gossypium hirsutum*, also called upland cotton, is the most cultivated cotton because of its wider adaptability in new environments, which results in higher productivity. Alternatively, *G. barbadense*, the second most cultivated cotton, is grown for its high fiber quality (Zhang et al., 2005).

The genetic diversity in cultivated plants for breeding programs is sometimes limited, especially when gene pools are not easily accessible. Wild cotton has long been used as a genetic resource to introduce new traits that increase the potential of cotton cultivars (Stewart, 1995). *Gossypium tomentosum*, a wild allotetraploid species with the (AD)3 genome, is closely related to *G. hirsutum*, but is quite different from the cultivated *G. hirsutum* and *G. barbadense*. It has many excellent traits, including finer fibers, drought tolerance, and *Fusarium* and *Verticillium* resistance. Although its tolerance to insects and diseases is comparable to that of the other cultivars, and other characteristics match those of *G. tomentosum*, few reports have been conducted on this species, and little research has focused on constructing a bacterial artificial chromosome (BAC) library.

A BAC library is an important resource for genomics studies, such as genetic and physical mapping, and genome sequencing (Shizuya et al., 1992; Özdemir et al., 2004; Hajime et al., 2005; Xia et al., 2005; Zhang et al., 2010; Shi et al., 2011). Nevertheless, a BAC library for *G. tomentosum* has not yet been reported. We optimized our construction system for allotetraploid cotton and established a BAC library for this wild cotton species.

MATERIAL AND METHODS

G. tomentosum materials were obtained in the field at National Wild Cotton long-term *in vivo* nursery, Sanya, China. HMW DNAs extraction and digestion, DNAs ligation and transformation were prepared according to the methods described by Gao et al. (2013) with some modifications. All the SSR primers for screening BAC super pools were reference form Cotton Marker Database (CMD) (http://www.cottonmarker.org/) and Synthesized by Shanghai Sangon Biological Engineering co., LTD.

RESULTS AND DISCUSSION

Nuclei DNA fragment concentrations were >600 kb. The main band was clear, and DNA

was seldom degraded or mechanically broken (Figure 1). The fragments of digested DNA were approximately 100, or 300 kb long when 1.0 or 1.5 U of *Hind*III was used (Figure 2). Therefore, we selected 1.5 U of *Hin*dIII for large-scale DNA digestion.

After digestion, the DNA was treated with two rounds of electrophoresis to remove the smaller fragment (Figure 3); then, the DNA concentration (Figure 4) during ligation and transformation was estimated. Examination of the insert size of BAC clones revealed that the average insert size was 110 kb (Figure 5), and fewer than 3% of the clones were empty.

To validate the genome coverage and value of our BAC library, we selected seven SSR markers for screening 50 BAC super pools that were located on chromosomes A1-a05, A1-a07, A1-a08, A1-a09, and A1-a11. These markers are closely linked or even cover the resistance region (Luo et al., 2001; Luo and Rod, 2003; Yang et al., 2008). We applied the screening method described by Wu et al. (2010) and utilized the genomic DNA of CCRI 12, Hai 7124, and *G. tomentosum* as controls. All SSR markers successfully amplified the positive clones, and one to five hits were obtained (Table 1).

We report here the successful construction of the first BAC library for *G. tomentosum*, which is an important wild allotetraploid cotton. In total, 200832 clones were found, with an average insert size of 110 kb and fewer than 3% of clones were empty. Seven SSR markers located on five different chromosomes were selected for screening 15% of this library, and positive clones were obtained for each chromosome. The BAC library was about 10 times larger than the size of the *G. tomentosum* genome, which is 2440 Mb. Using the formula N = ln(1-p)/ln(1 - f/G) (Clarke and Carbon, 1976), we determined that there was a 99.7% probability of isolating a gene of interest (Gao et al., 2013). This BAC library will serve as an important genomics resource for investigating the classification, evolution, and relationships among cotton species. It can also be used as a foundation for further integration of physical chromosomes and molecular genetics mapping.

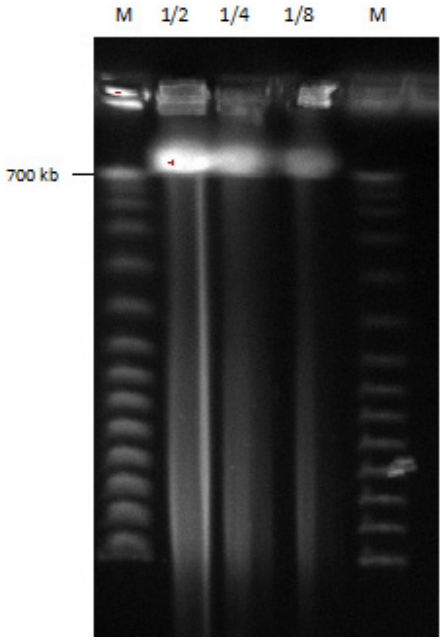

Figure 1. The quality of high molecular weight (HMW) DNA tested by PFGE.

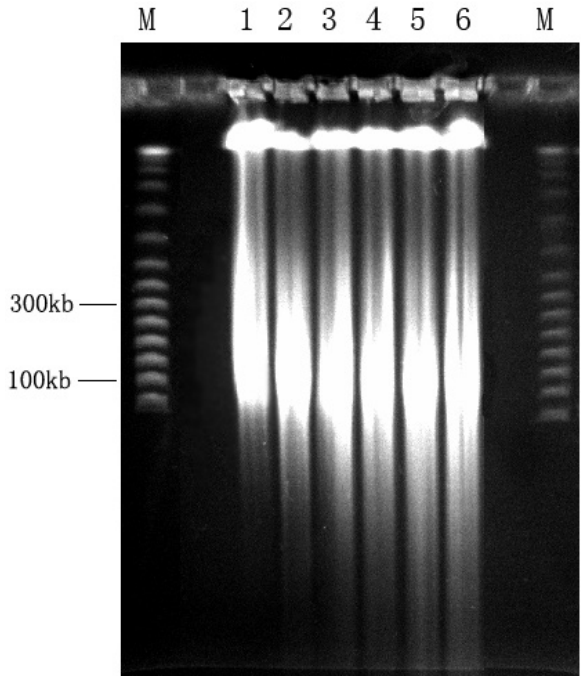

Figure 2. Determination of optimized enzyme dosage for HMW DNA. M: PFGE marker; *lanes 1* to *6* represent 0.5 U, 1.0 U, 1.5 U, 3 U, 4 U, and 5 U of Hind III, respectively.

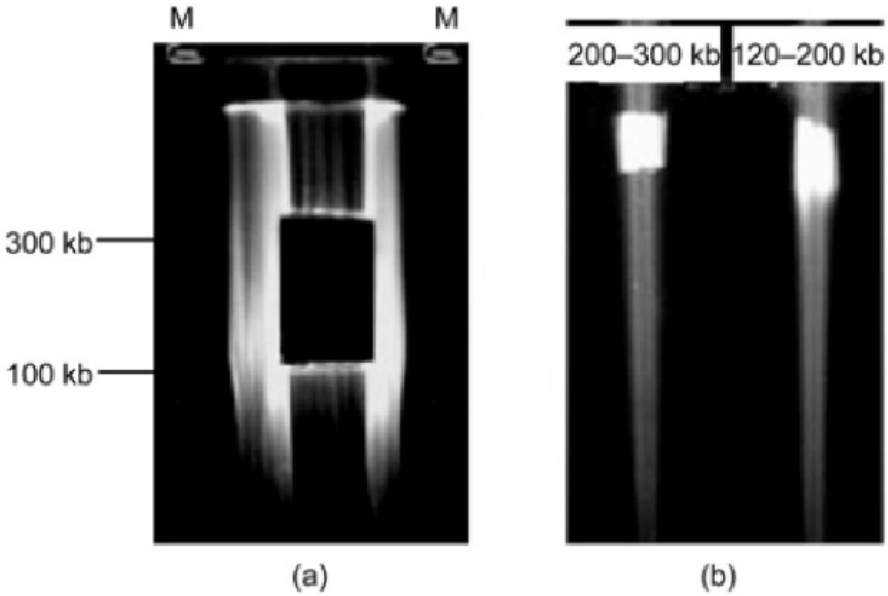

Figure 3. The first (**a**) and second (**b**) size selections of the partial digestion products of genomic DNA.

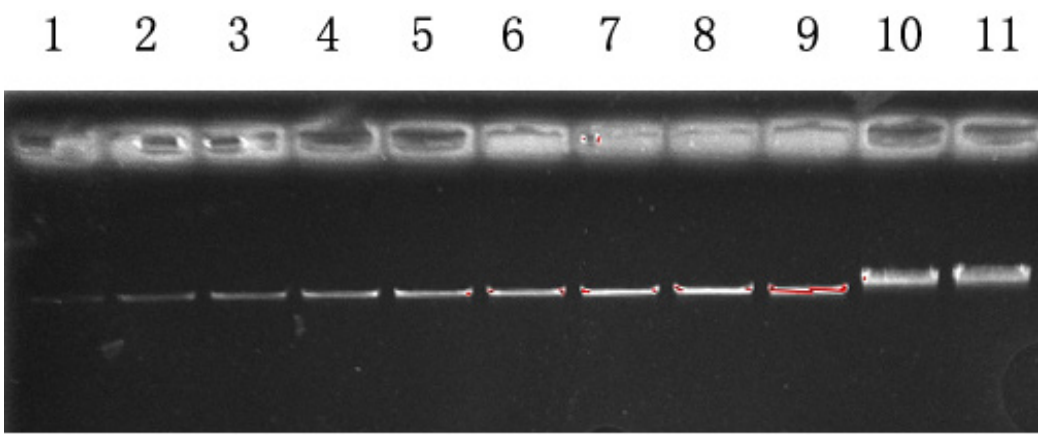

Figure 4. Detection of HMW DNA concentrations. *Lanes 1-9* = λ DNA with concentration shown as ng/ml; *lanes 10* and *11* = 120 and 200 kb HMW DNA, respectively.

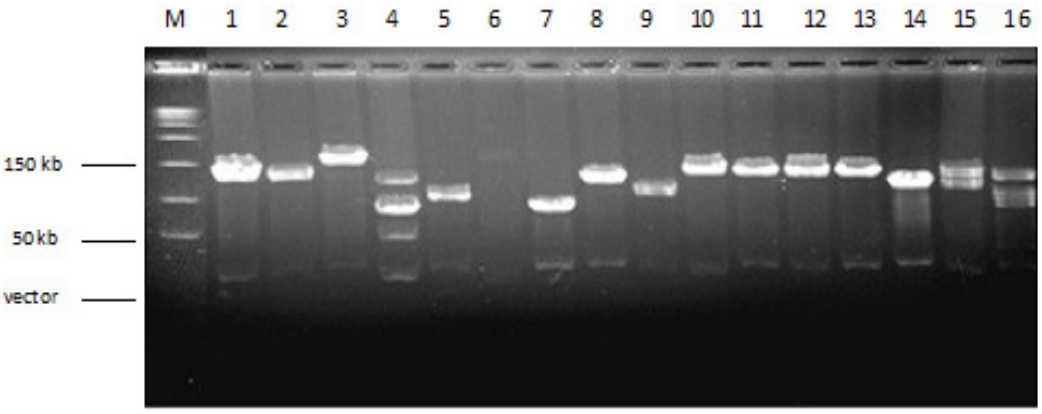

Figure 5. Insert size of BAC clones. *Lanes 1-14*, single BAC clones digested by Not I. *Lanes 15* and *16*, mixed BAC clones digested by NotI.

Table 1. Results of screening the BAC library with SSR markers linked to Verticillium wilt resistance.

Number	SSR marker	Chro. No.	Size of amplification (bp)	No. of hits from super pool
1	NAU3607	A_{t-a}05	150	5
2	NAU2121	A_{t-a}05	200	1
3	NAU1048	A_{t-a}07	250	3
4	NAU3964	A_{t-a}08	120	2
5	NAU3201	A_{t-a}08	170	4
6	BNL3582	A_{t-a}09	140	5
7	BNL3147	A_{t-a}11	110	3

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

Research supported by the National Basic Research Program of China (#2010CB126000), the National Key Technology Support Program of China (#2013BAD01B03), and the Foundation of National Plan of China (#2013AA102601).

REFERENCES

Brubaker CL, Bourland FM and Wendel JF (1999). The origin and domestication of cotton. In: Cotton: origin, history, technology, and production (Smith W and Cothren T, eds.). John Wiley & Sons, Inc, 3-31.

Clarke L and Carbon J (1976). A colony bank containing synthetic *Col El* hybrid plasmids representative of the entire *E. coli* genome. *Cell* 9: 91-99.

Fryxell P (1979). The natural history of the cotton tribe (Malvaceae, tribe Gossypieae). Texas A&M University Press, College Station.

Gao HY, Wang XF, Liu F, Peng RH, et al. (2013). Construction of a bacterial artificial chromosome library for *Gossypium herbaceum var. africanum*. *Chin. Sci. Bull*. 58: 3199-3201.

Hajime M, Takashi K and Sonoe OY (2005). Construction of a bacterial artificial chromosome (BAC) library of *Coprinus cinereus*. *Mycoscience* 46: 49-53.

Luo MZ, Wang YH, David F, Tarek J, et al. (2001). Melon bacterial artificial chromosome (BAC) library construction using improved methods and identification of clones linked to the locus conferring resistance to melon Fusarium wilt (*Fom-2*). *Genome* 44: 154-162.

Luo MZ and Rod AW (2003). An improved method for plant BAC library construction. In: Methods in molecular biology (Grotewold E, ed.). Humana Press, New York, 3-19.

Ma X, Zhou B, Lu Y, Guo W, et al. (2008). Simple sequence repeat genetic linkage maps of A-genome diploid cotton (*Gossypium arboreum*). *J. Integr. Plant Biol.* 50: 491-502.

Özdemir N, Horn R, and Friedt W (2004). Construction and characterization of a BAC library for sunflower (*Helianthus annuus L.*). *Euphytica* 138: 177-183.

Shi X, Zeng HY, Xue YD, Luo MZ (2011). A pair of new BAC and BIBAC vectors that facilitate BAC/BIBAC library construction and intact large genomic DNA insert exchange. *Plant Methods* 7: 33.

Shizuya H, Birren HB, Kim UJ, Mancino V, et al. (1992). Cloning and stable maintenance of 300 kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci. U.S.A.* 89: 8794-8797.

Stewart JM (1995). Potential for crop improvement with exotic germplasm and genetic engineering. In: Challenging the future (Constable G and Forrester NW, eds.). Proceedings of the World Cotton Research Conference-1, Melbourne, 313-327.

Wu Q, Cheng H, Liu F, Wang SF, et al. (2010). Screen of FISH marker of chromosomes at *Gossypium* D genome species. *Chin. Sci. Bull*. 55: 2099-2105.

Xia ZJ, Hiroko S, Satoshi W, Shiji K, et al. (2005). Construction and characterization of a BAC library of soybean. *Euphytica* 141: 129-137.

Yang C, Guo WZ and Li G (2008). QTLs mapping for wilt resistance at seedling and maturity stages in *Gossypium barbadense* L. *Plant Sci.* 174: 290-298.

Zhang JF, Lu YZ and Yu SX (2005). Cleaved AFLP (cAFLP), a modified amplified fragment length polymorphism analysis for cotton. *Theor. Appl. Genet.* 111: 1385-1395.

Zhang XJ, Zhang Y, Chantel S, Zhang HB, et al. (2010). Construction and characterization of a bacterial artificial chromosome (BAC) library of Pacific white shrimp, *Litopenaeus vannamei*. *Mar. Biotechnol*. 12: 141-149.