Genetics and Molecular Research

Collinearity analysis of allotetraploid *Gossypium tomentosum* and *Gossypium darwinii*

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

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

Corresponding authors: K.B. Wang / Z.S. Zhang
E-mail: wkbcri@163.com / zhangzs@swu.edu.cn

Genet. Mol. Res. 15 (3): gmr.15038391
Received January 7, 2016
Accepted March 11, 2016
Published August 5, 2016
DOI http://dx.doi.org/10.4238/gmr.15038391

ABSTRACT. *Gossypium tomentosum* and *G. darwinii* are wild allotetraploid cotton species, characterized by many excellent traits, including fiber fineness, drought tolerance, and *Fusarium* and *Verticillium* wilt resistance. Based on the construction of F_2 linkage groups of *G. hirsutum* x *G. tomentosum* and *G. hirsutum* x *G. darwinii*, two genetic linkage maps were compared. As a result, we found a total of seven inverted fragments on chr02, chr05, chr08, chr12, chr14, chr16, and chr25, and three translocated fragments on chr05, chr14, and chr26. In addition, comparison of the inverted and translocated fragments revealed that the orientation of four of seven markers in *G. tomentosum* were consistent with *G. hirsutum* or *G. raimondii*. The orientation of one of seven inverted markers of *G. darwinii* was consistent with *G. hirsutum*, and the orientation of one of three translocated markers of *G. tomentosum* was consistent with *G. raimondii.* These results

indicate that, in comparison to *G. darwinii*, *G. tomentosum* has a closer genetic relationship to *G. hirsutum*. These findings will be important for our understanding on the genome structure of *G. tomentosum* and *G. darwinii*, and set the scene for further in-depth genome research such as fine mapping, tagging genes of interest from wild relatives, and evolutionary study.

Key words: Cotton germplasm resource; Allotetraploid cotton; *Gossypium tomentosum*; *Gossypium darwinii*; Linearity relationship

INTRODUCTION

Cotton is a natural white fibrous agricultural product, which is of great economic importance as a raw material for the textile industry, and has a wide variety of uses in the paper industry, in home fixtures, medical supplies, chemicals, and oil (Ensminger et al., 1990). Cotton (*Gossypium* spp) is one of the most expansively grown species around the globe, and is grown in tropical and subtropical regions between the latitudes 36°S and 46°N (Reller and Gerstenberg, 1997) in more than 80 countries (Fryxell, 1979; Smith, 1995). The *Gossypium* genus consists of 50 species (Fryxell, 1992; Stewart, 1995; Ma et al., 2008), which includes five allotetraploid species [2n = 4x = 52; $(AD)_1$ to $(AD)_5$] and 45 diploid species (2n = 2x = 26, A through G and K) (Fryxell, 1979; Stewart, 1995; Brubaker et al., 1999; Zhang et al., 2005). The evolution of domesticated cotton has involved various steps, which provide scientists with insight into the role of polyploidy in the diversification process and permit the investigation of unique traits. In addition, it can further help geneticists to utilize germplasm from wild relatives. Allopolyploids are a great tool that can be used to understand the evolutionary process of two diploid but different sized genomes that evolved simultaneously. Studies on the evolution and diversity of the *Gossypium* genus provide basic knowledge on the morphological diversity of the genus and plant biology, which can help in the better utilization of genetic resources (Wendel et al., 2009).

More than 30 genetic maps have already been published in cotton, and most of them are based on interspecific crosses of domesticated tetraploid species, namely *G. hirsutum* and *G. barbadense* (Jiang et al., 1998; Zhang et al., 2002; Nguyen et al., 2004; Rong et al., 2004; Guo et al., 2007; He et al., 2007; Lacape et al., 2003, 2009). Interspecific tetraploid genetic maps are useful for understanding genome structure and for exploring the genetic basis of important agronomic characters. In addition, they provide a basis for identifying new DNA markers for further high density maps (Guo et al., 2007; Zhang et al., 2008; Yu et al., 2011; Kalivas et al., 2011; Tu et al., 2014).

Wild cotton has long been used as a genetic resource to introduce new traits that increase the potential of cultivated cotton species (Stewart, 1995). *Gossypium darwinii*, a wild allotetraploid species with an $(AD)_5$ genome, is closely related to *Gossypium barbadense*, but is quite different from the cultivated *Gossypium hirsutum*. It has many excellent traits, including finer fiber fineness, drought tolerance, and *Fusarium* and *Verticillium* wilt resistance. *Gossypium tomentosum* is endemic to Hawaiian Islands, and it has many unique agronomic traits such as insect-pest resistance, drought tolerance, salt tolerance, heat tolerance, nectarilessness, and lint color (Liu et al., 2015). Many genetic studies, such as those investigating genomic and genetic structure/organization, have been conducted on the two cultivated tetraploids, *G. hirsutum* and *G. barbadense*. However, very little is known about the genomic architecture,

gene transfer, or introgression of unique traits from the other three tetraploids.

In the present study, simple sequence repeat (SSR) genetic maps were developed from two F_2 populations, *G. hirsutum* x *G. tomentosum* (Chen et al., 2015) and *G. hirsutum* x *G. darwinii* (Kashif et al., 2015). The linearity relationship was compared between the two genetic maps, which will serve as an indispensable genomic resource for genome structure study, comparative genomic analysis, fine mapping, and map-based gene cloning of important traits.

MATERIAL AND METHODS

A linkage map was reconstructed from two previously constructed maps (Chen et al., 2015; Kashif et al., 2015). SSR markers uniformly distributed on different chromosomes were selected and checked after downsizing. The JoinMap 4.0 software was used for linkage analysis and map construction. Kosambi mapping function was used to convert recombination frequencies into map distances (cM). The Mapchart 2.2 software was used to draw the genetic map.

Linkage map assignment was established by common markers, which were already anchored as previously described (Wang et al., 2012; Li et al., 2014, 2015; Zhang et al., 2015). Chromosomal nomenclature was used as described by Guo et al. (2008), i.e., SSR loci anchored on chromosome (chr) 1-13 were designated to the A sub-genome (At), whereas loci confined to chr 14-26 were designated to the D sub-genome (Dt).

RESULTS

Comparing the linear relationship of the two genetic linkage maps, we found that most SSR markers present were collinear between the two maps. Meanwhile, part of the non-linear relationship appeared on the individual chromosomes between the genetic linkage maps of *G. hirsutum* x *G. tomentosum* and *G. hirsutum* x *G. darwinii*, which include seven inverted and three translocated fragments (Figure 1A-J).

Seven inversions were found on seven different chromosomes between the genetic maps of *G. tomentosum* and *G. darwinii*. The first inversion was on chr02 (MON_COT043-HAU2643-HAU0248-HAU2923), the second on chr05 (MON-DPL0384-TMB1586-BNL2448), the third on chr08 (CGR6748-HAU0709-BNL3257), the fourth on chr12 (MON_CGR5158-NAU943-BNL1673), the fifth on chr14 (NAU3691-MON_CGR5818-TMB1513), the sixth on chr16 (CM56-JESPR102-NAU3676-GH56), and the seventh on chr25 (MON_CGR5665-NAU2388-BNL3103). Three translocations were found on three different chromosomes between the genetic maps of *G. tomentosum* and *G. darwinii*, respectively: on chr05 the NAU2001-NAU1200-NAU3569 fragment underwent a parallel translocation with the MUSB0592-NAU1127 fragment; on chr26, the MON-CGR5678-MON-DPL0391 fragment underwent a parallel translocation with the MON-CGR6318-MON-CGR5802 fragment; and on chr14, the TMB0803-NAU2987 fragment underwent a parallel translocation with the MON_C2_0079-HAU0551 fragment.

Referring to the published *G. raimondii* genome sequencing data (D_5, Wang et al., 2012) and the latest *G. hirsutum* genome sequencing data (AD_1, Zhang et al., 2015), we compared the inversions and translocations on the genetic linkage map with the physical position and arrangement on the corresponding genome chromosome. The result showed that, among the inversions, the fragment orientation of *G. darwinii* on chr05 is consistent with the physical direction of the *G. raimondii* (D_5) and *G. hirsutum* (AD_1) genome sequences.

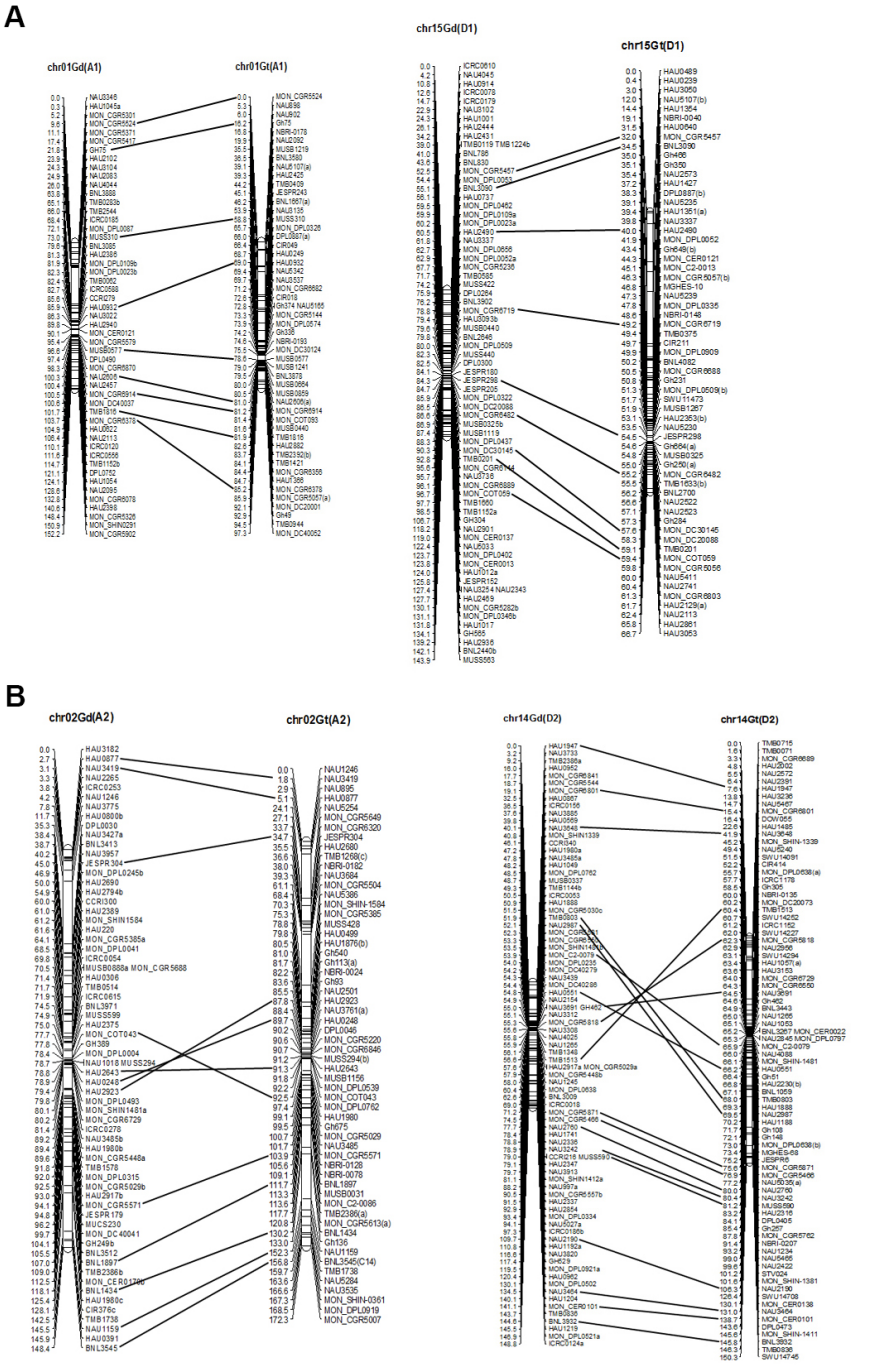
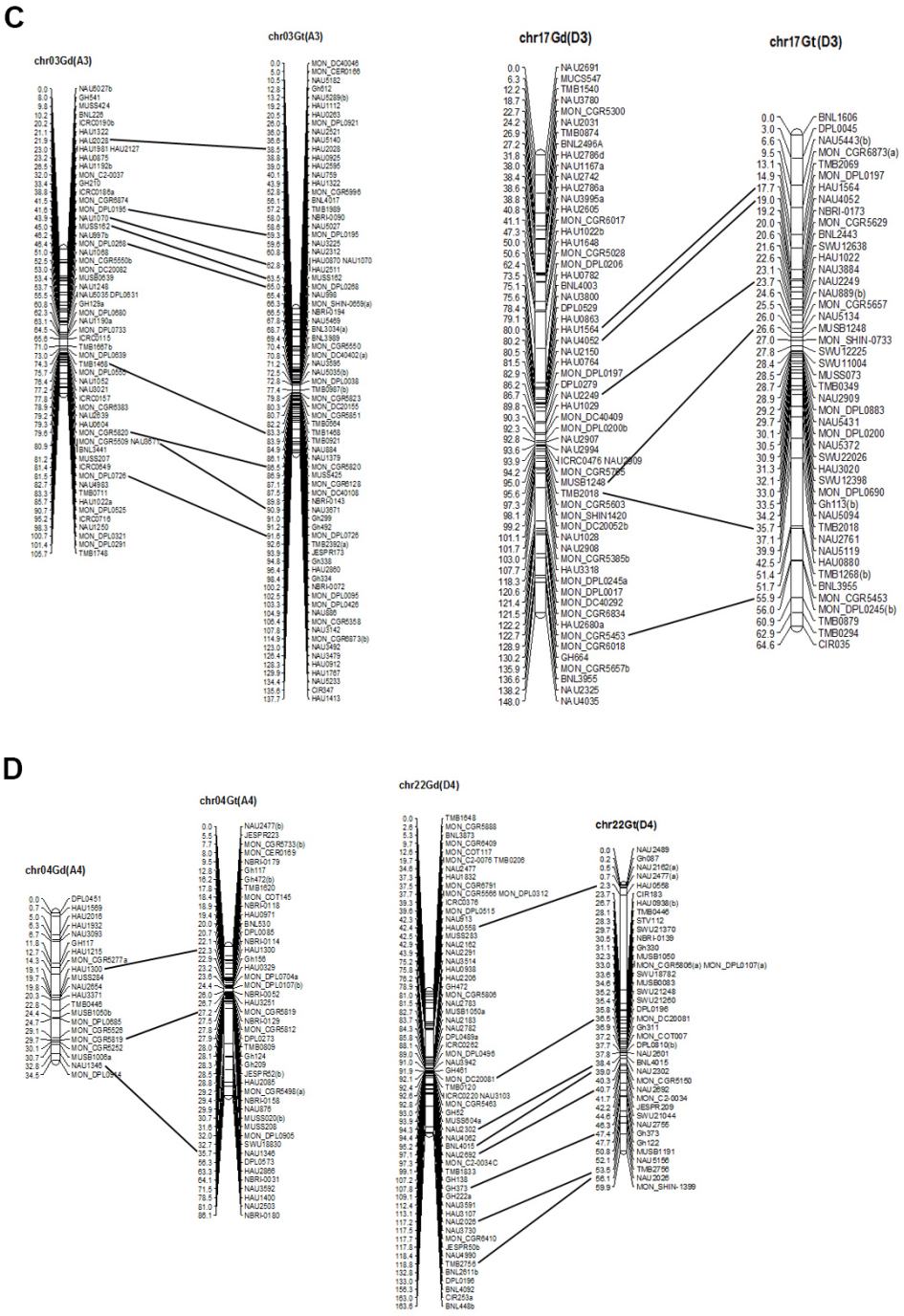


Figure 1. A.-J. Genetic map and collinearity comparison of *Gossypium tomentosum* (GT) and *Gossypium darwinii* (Gd). Chr01 to chr26 refer to the Chromsome from 01 to 26, respectively.

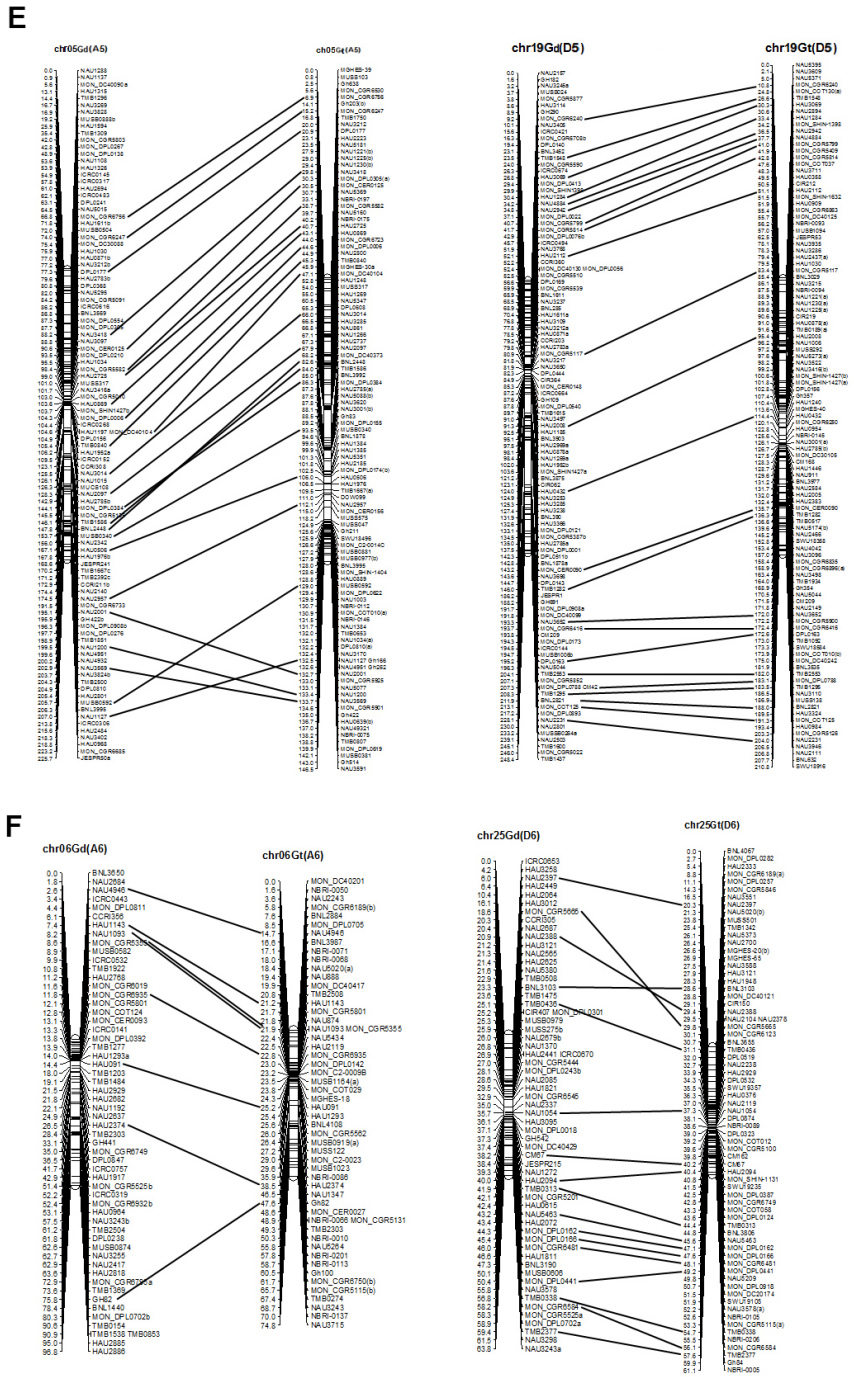
Continued on next page

Figure 1. Continued.



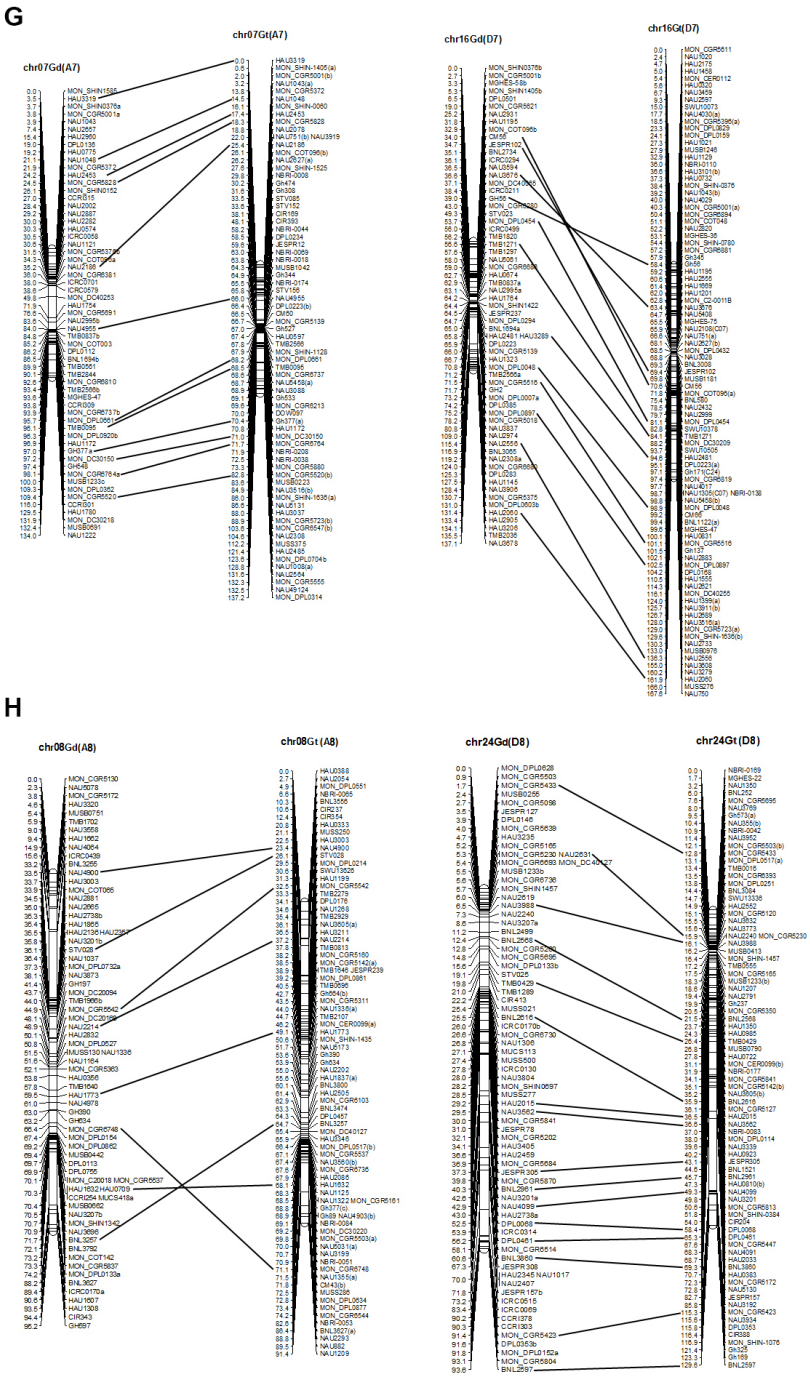
Continued on next page

Figure 1. Continued.



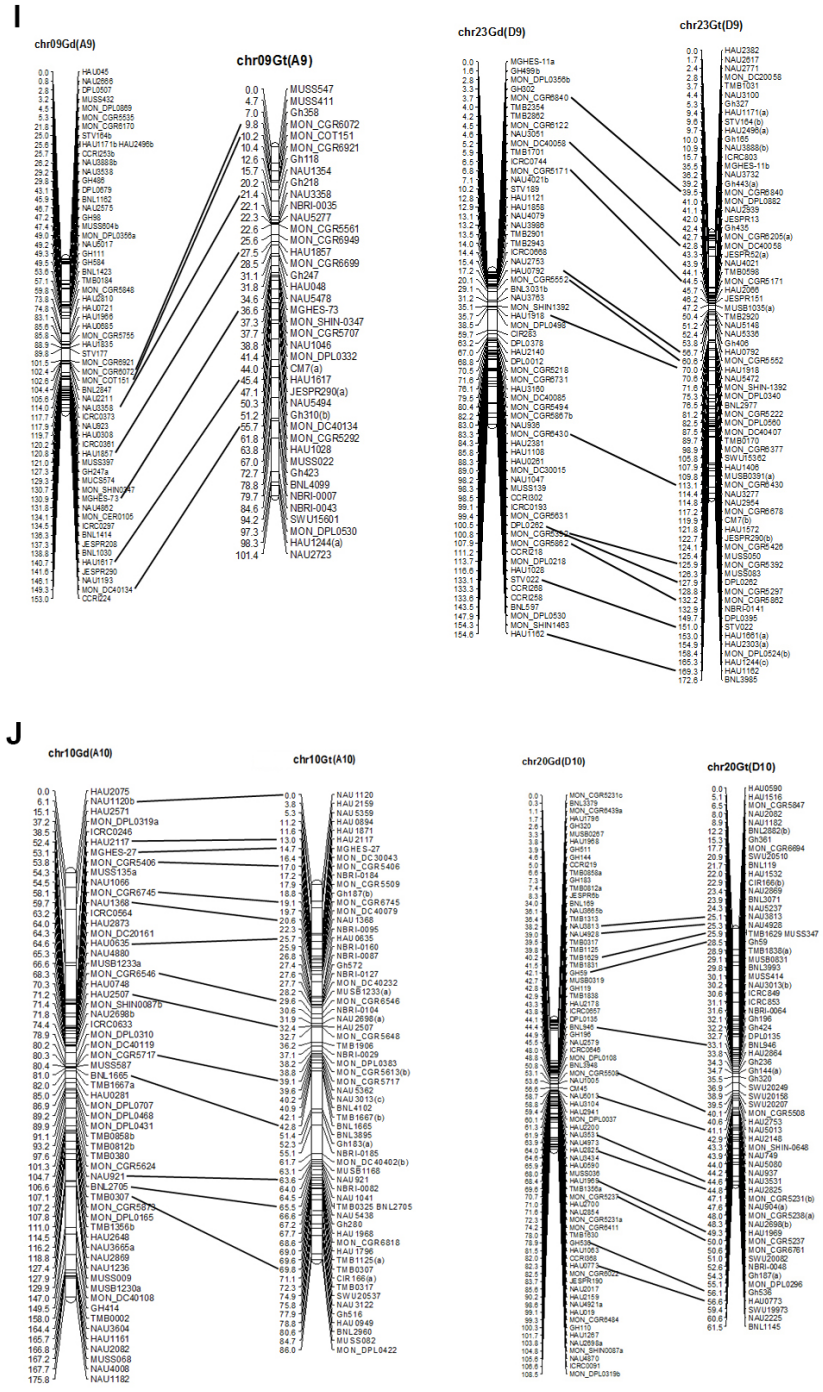
Continued on next page

Figure 1. Continued.



Continued on next page

Figure 1. Continued.



The fragment orientations of *G. tomentosum* on chr08, chr12, and chr25 are consistent with the physical directions of *G. raimondii* (D_5). The fragment orientations of *G. tomentosum* on chr08, chr12, and chr16 are consistent with the physical directions of *G. hirsutum* (AD_1). Among the translocation, just the fragment on chr26 corresponding physical position was found on *G. raimondii* (D_5) (Table 1).

Table 1. Comparison between the location of the genetic linkage fraction and the genome physical position.

Linkage group	Markers	Location on *G. darwinii* (genetic position, cM)	Location on *G. tomentosum* (genetic position, cM)	Corresponding location on *G. raimondii* (physical position, Mb)	Corresponding location on *G. hirsutium* (physical position, Mb)	Events
chr02	MON_COT043-HAU2643-HAU0248-HAU2923	77.7-78.8-78.9-79.4	92.5-91.3-89.7-87.8			Inversion
chr05	MON_DPL0384-TMB1586-BNL2448	144.1-146.1-147.8	86.3-84-82.6	21.4-21.6	37.2-58.2	Inversion
chr08	MON_CGR6748-HAU1632-BNL3257	66.4-70.3-71.7	71.1-68.1-64.7	6.5-31.3-46.1	85.6-58.9	Inversion
chr12	MON_CGR5158-NAU943-BNL1673	107.3-108-111	111.2-109.2-107.1	44.2-42.6	93.5-41.7	Inversion
chr16	CM56-JESPR102-NAU3676-GH56	34-34.7-36.6-39	70.6-69.4-63.4-58.4		47.1-39.5	Inversion
chr25	MON_CGR5665-NAU2388-BNL3103	18.6-20.9-23.3	29.8-29.4-28.6	42.9-40.6		Inversion
chr05	NAU2001-NAU1200-NAU3569-MUSB0592-NAU1127	195.1-199.5-202.9-205.7-207	132.7-133.4-133.7-129-132.5			Translocation
chr26	MON_CGR5678-MON_DPL0391-MON_CGR6318-MON_CGR5802	34.9-35.3-37.4-38.2	31.5-33-29.6-30	17.1-31.1-10.2-10.8		Translocation
chr14	TMB0803-NAU2987-MON_C2_0079-HAU0551-NAU3691-MON_CGR5818-TMB1513	51.9-52.1-53.9-54.8-55 -55.3-56.6	68-69.5-65.9-66.2-64.5-62.3-60.4			Inversion and translocation

DISCUSSION

Reciprocal translocations and inversion

In the construction of *Gossypium* genetic linkage maps, researchers have repeatedly found reciprocal translocations between homologous chromosomes (Rong et al., 2004; Wang et al., 2006; He et al., 2007; Yu et al., 2012). Very few studies have focused on different linear relationships between homologous chromosomes in cotton. In this study, using *G. hirsutum*, we constructed allotetraploid genetic linkage maps of *G. hirsutum* x *G. tomentosum* and *G. hirsutum* x *G. darwinii*. Analysis of linear relationships at homologous positions revealed seven inversions (on chr02, chr05, chr08, chr12, chr14, chr16, and chr25) and three translocations (on chr05, chr14, and chr26), which were found separately on the two genetic linkage maps. When comparing the locations of the inversions and translocations with the physical positions of *G. hirsutum*, we found the same pattern of inversions at the corresponding positions on chr02, chr05, chr12, and chr14 in *G. hirsutum*, consistent with the translocation on the corresponding position of chr05 in *G. hirsutum*. For the first time, we report an inversion on the homologous chromosome fragments of chr02 and chr14, and describe the translocations on chr05 and chr14. In this study, we also found an inversion and translocation on the same fragment, which indicates that this is an active genetic region. At present, we are unable to explain the roles of the reciprocal translocation and inversion in the structural genes and geographical

distribution of different cotton species. Further research is needed to confirm whether the inversion occurred before or after the translocation.

Genetic mapping coupled with physical alignment of the genome

A lot of previous research has focused on the origin and evolution of allotetraploid cotton. Findings reported by Fryxell (1992) and Wendel et al. (2009) support the idea that *G. tomentosum* and *G. hirsutum* have a close relationship, similarly *G. darwinii* and *G. barbadense* also have a close relationship. Limited research has been performed to determine the linear relationship of the chromosome genetic linkage between *G. tomentosum* and *G. darwinii.* Following the sequencing of the cotton genome, studying the relationship between the genetic linkage map and the physical map became feasible and convenient. Using the F_2 population derived from interspecific crosses, we constructed two linkage maps of allotetraploid cotton *G. hirsutum* x *G. tomentosum* and *G. hirsutum* x *G. darwinii* separately. By referring to the latest published *G. hirsutum* genome sequencing data (Zhang et al., 2015), the linear relationship between the two wild allotetraploid cotton genetic linkage maps was compared based on genome-wide SSR markers. The results show that there is a good linear relationship between the genetic linkage maps of *G. hirsutum* x *G. tomentosum* and those of *G. hirsutum* x *G. darwinii*. Furthermore, part of the non-linear relationship appeared on the individual chromosomes between the two genetic linkage maps, including seven inverted and three translocated fragments. Compared with the sequence data for *G. hirsutum* (Zhang et al., 2015), we noted that among the seven inverted fragments, the orientation of just one marker (chr05), which comes from *G. darwinii*, is consistent with the physical arrangement on *G. hirsutum*, while the orientations of four markers (chr08, chr12, chr16, and chr25), which derive from *G. tomentosum*, are consistent with the physical position of *G. hirsutum* or *G. raimondii.* Among the three translocated fragments, the orientation of one marker (chr26), which comes from *G. tomentosum,* is consistent with the physical arrangement of *G. raimondii.* These results indicate that, in comparison to *G. darwinii*, *G. tomentosum* has a closer genetic relationship with *G. hirsutum*.

Conflicts of interests

The authors declare no conflict of interests.

ACKNOWLEDGMENTS

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

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.

Chen H, Khan MK, Zhou Z, Wang X, et al. (2015). A high-density SSR genetic map constructed from a F2 population of *Gossypium hirsutum* and *Gossypium darwinii. Gene* 574: 273-286. http://dx.doi.org/10.1016/j.gene.2015.08.022

Ensminger ME, Oldfield JE and Heinemann WW (1990). Excerpts with reference to cottonseed components. Ensminger Publishing Company, USA.

Fryxell PA (1979). The natural history of the cotton tribe. Texas A&M University.
Fryxell PA (1992). A revised taxonomic interpretation of *Gossypium* L. Malvaceae. *Rheedea* 2: 108-165.
Guo W, Cai C, Wang C, Han Z, et al. (2007). A microsatellite-based, gene-rich linkage map reveals genome structure, function and evolution in *Gossypium*. *Genetics* 176: 527-541. http://dx.doi.org/10.1534/genetics.107.070375
Guo W, Cai C, Wang C, Zhao L, et al. (2008). A preliminary analysis of genome structure and composition in *Gossypium hirsutum*. *BMC Genomics* 9: 314. http://dx.doi.org/10.1186/1471-2164-9-314
He DH, Lin ZX, Zhang XL, Nie YC, et al. (2007). QTL mapping for economic traits based on a dense genetic map of cotton with PCR-based markers using the interspecific cross of *Gossypium hirsutum* x *Gossypium barbadense*. *Euphytica* 153: 181-197. http://dx.doi.org/10.1007/s10681-006-9254-9
Jiang C, Wright RJ, El-Zik KM and Paterson AH (1998). Polyploid formation created unique avenues for response to selection in *Gossypium* (cotton). *Proc. Natl. Acad. Sci. USA* 95: 4419-4424. http://dx.doi.org/10.1073/pnas.95.8.4419
Kalivas A, Xanthopoulos F, Kehagia O and Tsaftaris AS (2011). Agronomic characterization, genetic diversity and association analysis of cotton cultivars using simple sequence repeat molecular markers. *Genet. Mol. Res.* 10: 208-217. http://dx.doi.org/10.4238/vol10-1gmr998
Kashif RK, Zhou ZL, Chen HD, Wang XX, et al. (2015). Genome wide SSR high density genetic map construction from an interspecific cross of *G. hirsutum* x *G. tomentosum*. *Front. Plant Sci.*, in press.
Lacape JM, Nguyen TB, Thibivilliers S, Bojinov B, et al. (2003). A combined RFLP-SSR-AFLP map of tetraploid cotton based on a Gossypium hirsutum x Gossypium barbadense backcross population. *Genome* 46: 612-626. http://dx.doi.org/10.1139/g03-050
Lacape JM, Jacobs J, Arioli T, Derijcker R, et al. (2009). A new interspecific, *Gossypium hirsutum* x *G. barbadense*, RIL population: towards a unified consensus linkage map of tetraploid cotton. *Theor. Appl. Genet.* 119: 281-292. http://dx.doi.org/10.1007/s00122-009-1037-y
Li F, Fan G, Lu C, Xiao G, et al. (2015). Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nat. Biotechnol.* 33: 524-530. http://dx.doi.org/10.1038/nbt.3208
Li FG, Fan GY, Wang KB, Sun FM, et al. (2014). Genome sequence of the cultivated cotton Gossypium arboretum. *Nat. Genet.* 46: 467-572. http://dx.doi.org/10.1038/ng.2987
Liu F, Wang YH, Gao HY, Wang CY, et al. (2015). Construction and characterization of a bacterial artificial chromosome library for the allotetraploid Gossypium tomentosum. *Genet. Mol. Res.* 14: 16975-16980. http://dx.doi.org/10.4238/2015.December.15.3
Ma XX, Zhou BL, Lü YH, Guo WZ, et al. (2008). Simple sequence repeat genetic linkage maps of A-genome diploid cotton (*Gossypium arboreum*). *J. Integr. Plant Biol.* 50: 491-502. http://dx.doi.org/10.1111/j.1744-7909.2008.00636.x
Nguyen TB, Giband M, Brottier P, Risterucci AM, et al. (2004). Wide coverage of the tetraploid cotton genome using newly developed microsatellite markers. *Theor. Appl. Genet.* 109: 167-175. http://dx.doi.org/10.1007/s00122-004-1612-1
Reller A and Gerstenberg J (1997). White gold, where do you go? Status quo and prospects of cotton usage. *Gaia* 1: 35-51.
Rong J, Abbey C, Bowers JE, Brubaker CL, et al. (2004). A 3347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (*Gossypium*). *Genetics* 166: 389-417. http://dx.doi.org/10.1534/genetics.166.1.389
Smith CW (1995). Cotton (*Gossypium hirsutum* L.). Chapter 6. In: Crop Production: Evolution, History, and Technology. John Wiley and Sons, Inc., New York, 287-349.
Stewart JM (1995). Potential for crop improvement with exotic germplasm and genetic engineering. In: G. Constable, Forrester N W. Challenging the future. Proceedings of the World Cotton Research Conference-1 Melbourne, Australia, CSIRO. 313-327.
Tu JL, Zhang MJ, Wang XQ, Zhang XL, et al. (2014). Genetic dissection of upland cotton (Gossypium hirsutum) cultivars developed in Hubei Province by mapped SSRs. *Genet. Mol. Res.* 13: 782-790. http://dx.doi.org/10.4238/2014.January.31.4
Wang K, Song X, Han Z, Guo W, et al. (2006). Complete assignment of the chromosomes of *Gossypium hirsutum* L. by translocation and fluorescence in situ hybridization mapping. *Theor. Appl. Genet.* 113: 73-80. http://dx.doi.org/10.1007/s00122-006-0273-7
Wang K, Wang Z, Li F, Ye W, et al. (2012). The draft genome of a diploid cotton Gossypium raimondii. *Nat. Genet.* 44: 1098-1103. http://dx.doi.org/10.1038/ng.2371
Wendel JF, Brubaker C, Alvarez I, Cronn R, et al. (2009). Plant Genetics and Genomics: Crops and Models, Genetics and Genomics of Cotton: Evolution and Natural History of the Cotton Genus. Springer Science Business Media. *LLC* 3: 3-22.
Yu JZ, Kohel RJ, Fang DD, Cho J, et al. (2012). A high-density simple sequence repeat and single nucleotide polymorphism genetic map of the tetraploid cotton genome. *G3 (Bethesda)* 2: 43-58. http://dx.doi.org/10.1534/g3.111.001552

Yu Y, Yuan D, Liang S, Li X, et al. (2011). Genome structure of cotton revealed by a genome-wide SSR genetic map constructed from a BC1 population between *gossypium hirsutum* and *G. barbadense. BMC Genomics* 12: 15. http://dx.doi.org/10.1186/1471-2164-12-15

Zhang J, Guo W and Zhang T (2002). Molecular linkage map of allotetraploid cotton (*Gossypium hirsutum* L. x *Gossypium barbadense* L.) with a haploid population. *Theor. Appl. Genet.* 105: 1166-1174. http://dx.doi.org/10.1007/s00122-002-1100-4

Zhang J, Lu Y and Yu S (2005). Cleaved AFLP (cAFLP), a modified amplified fragment length polymorphism analysis for cotton. *Theor. Appl. Genet.* 111: 1385-1395. http://dx.doi.org/10.1007/s00122-005-0070-8

Zhang T, Hu Y, Jiang W, Fang L, et al. (2015). Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat. Biotechnol.* 33: 531-537. http://dx.doi.org/10.1038/nbt.3207

Zhang Y, Lin Z, Xia Q, Zhang M, et al. (2008). Characteristics and analysis of simple sequence repeats in the cotton genome based on a linkage map constructed from a BC_1 population between *Gossypium hirsutum* and *G. barbadense. Genome* 51: 534-546. http://dx.doi.org/10.1139/G08-033