

Development of InDel markers for the restorer gene *Rf1* and assessment of their utility for marker-assisted selection in cotton

Jianyong Wu  · Meng Zhang · Xuexian Zhang · Liping Guo · Tingxiang Qi · Hailin Wang · Huini Tang · Jinfa Zhang · Chaozhu Xing

Received: 14 August 2017 / Accepted: 19 October 2017 / Published online: 25 October 2017
© Springer Science+Business Media B.V. 2017

Abstract The cytoplasmic male sterility (CMS) system is ideal for exploiting heterosis in crops such as cotton. However, CMS-D2, which is based on *Gossypium harknessii* cytoplasm, is still not widely used for cotton production. In this study, we developed an efficient marker-assisted selection method based on insertion/deletion (InDel) markers that can identify restorer lines carrying *Rf1*. Whole-genome resequencing was first completed for restorer [N(*Rf1Rf1*)] and maintainer [N(*rf1rf1*)] lines with normal fertile cytoplasm (N). Comparisons with the TM-1 reference genome sequence resulted in the identification of 292,065 and 183,657 InDels for the restorer and maintainer lines, respectively. Most of the InDels in

the restorer line were distributed on Chromosome_D05, which carries *Rf1*. Of the 12 InDel markers near the *Rf1* target region that were further validated, four co-dominant markers (i.e., InDel-1891, InDel-3434, InDel-7525, and InDel-9356) co-segregated with *Rf1*, as verified by a segregation analysis in an F₂ population. We subsequently used InDel-1891 to determine the allele status at the *Rf1* locus in a backcross scheme for transferring *Rf1*. In this study, we developed new markers to increase the marker density in the *Rf1* target region, which will be useful for the fine mapping of *Rf1*. The development of convenient and inexpensive co-segregating InDel markers will facilitate the marker-assisted selection of restorer lines carrying *Rf1*.

Electronic supplementary material The online version of this article (<http://doi.org/10.1007/s10681-017-2043-9>) contains supplementary material, which is available to authorized users.

Keywords CMS-D2 · *Rf1* · Whole-genome resequencing · InDel · Marker-assisted breeding

J. Wu (✉) · M. Zhang · X. Zhang · L. Guo · T. Qi · H. Wang · H. Tang · C. Xing (✉)
State Key Laboratory of Cotton Biology, Institute of Cotton Research Chinese Academy of Agricultural Sciences, Key Laboratory for Cotton Genetic Improvement, Ministry of Agriculture, Anyang 455000, China
e-mail: dr.wujianyong@live.cn

C. Xing
e-mail: chaozhuxing@126.com

J. Zhang
Department of Plant and Environmental Sciences, New Mexico State University, Las Cruces, NM 88003, USA

Introduction

Cotton is an important fiber crop worldwide. As in other crops, heterosis has been used to improve cotton yield, quality, and stress resistance, and more than 90% of the cotton hybrids grown in China were produced by artificial emasculation and pollination (Yu et al. 2016). However, this procedure is time-consuming, labor-intensive, and costly. Additionally, there is no guarantee that the resulting hybrid seeds

will be pure, which is an important limiting factor associated with hybrid seed production. In other crops, cytoplasmic male sterility (CMS) systems as an ideal way for producing hybrid seeds have been widely used to exploit heterosis during breeding. In cotton, CMS-D2 and CMS-D8 are the two main CMS systems, with *Rf1* and *Rf2* as the restorer genes, respectively (Meyer 1975; Weaver and Weaver 1977; Zhang and Stewart 2001a, b). Considering the importance of the restoration systems, random amplification of polymorphic DNA (RAPD), simple sequence repeat (SSR), sequence tagged site (STS), target region amplified polymorphism (TRAP), and cleaved amplified polymorphic sequence (CAPS) markers linked to *Rf1* have been developed for molecular mapping studies (Guo et al. 1998; Lan et al. 1999; Liu et al. 2003; Feng et al. 2005; Yin et al. 2006; Yang 2009; Wang et al. 2009; Wu et al. 2014). These markers may be useful for the molecular marker-assisted selection (MAS) of restorer lines. However, the application of most of these marker types requires complex experimental processes and/or the markers exhibit low sensitivity. Thus, additional markers that are sensitive and simple to use must be developed to enable MAS for restorer lines.

The InDel markers are based on the insertions and/or deletions of DNA fragments at the same genomic locus. Insertions/deletions occur frequently and are widely distributed throughout genomes (Weber et al. 2002). These markers, especially those with large insertions/deletions, can be designed as fragment-length polymorphic markers, and are easily distinguished by agarose gel electrophoresis following a polymerase chain reaction (PCR) amplification. Because they are highly accurate and stable (Jander et al. 2002), InDel markers have been widely used in many areas, including analyses of germplasm resources (Lu et al. 2009; Lv et al. 2017), genetic mapping, map-based cloning (Srivastava et al. 2016; Ye et al. 2016; Singh et al. 2017; Zhang et al. 2017; Liu et al. 2017), and marker-assisted breeding (Hayashi et al. 2006; Zhao et al. 2017).

In cotton, sequences from InDels have been used for phylogenetic analyses (Grover et al. 2008), genetic mapping (Li et al. 2014b; Wang et al. 2015), and analyses of tetraploid cotton genetic resources (Shen et al. 2017). However, most of these InDels have not been validated, and no PCR-based InDel markers have been developed to track a specific trait of interest

during cotton breeding. The completely sequenced genomes of three cotton species (Paterson et al. 2012; Wang et al. 2012; Li et al. 2014a, 2015; Zhang et al. 2015) may be useful for developing genome-wide or gene-derived InDel markers based on whole-genome resequencing. Thus, in this study, a restorer (ZBR) line and its near-isogenic maintainer (ZB) line of the CMS-D2 system were re-sequenced, and InDel markers in the *Rf1* target region were identified according to the results of a comparative InDel analysis. The InDel markers tightly linked to *Rf1* were subsequently used for molecular breeding of restorer lines.

Materials and methods

Plant materials and DNA extraction

The ZBR restorer line (i.e., R line derived from the backcross offspring of the ZB maintainer line through eight generations of backcrossing) of CMS-D2 cotton and its near-isogenic ZB maintainer line (i.e., B line) with Upland cotton (AD1) cytoplasm were developed and provided by the Institute of Cotton Research (ICR), Chinese Academy of Agricultural Science, Anyang, Henan, China. Plants were grown at the Cotton Research Farm at the ICR. Young leaves were collected from individual plants and immediately frozen in liquid nitrogen. The leaf samples were stored at -70°C until analyzed. Genomic DNA was isolated from the leaves using a Plant Genomic DNA kit (Tiangen Biotech, Beijing, China). The concentration of the purified DNA was determined using the NanoDrop 2000C Spectrophotometer (Thermo Scientific, Wilmington, DE, USA), while the quality was evaluated by 1% agarose gel electrophoresis.

Whole-genome resequencing and data analysis

Using 5 μg genomic DNA as the template, paired-end sequencing libraries with 250 and 300 bp inserts were constructed for the R and B lines, respectively, according to the manufacturer's instructions (Illumina, San Diego, CA, USA). The whole-genome sequencing was completed at Biomarker Technologies (Beijing, China) using the HiSeq 4000 system according to the manufacturer's instructions (Illumina). The raw reads were filtered by removing low-quality reads (i.e., Q value < 20), adapter sequences, and reads with $> 5\%$

ambiguous bases. The clean reads were then mapped onto the TM-1 cotton reference genome (<http://mascotton.njau.edu.cn/info/1054/1118.htm>) using the Burrows–Wheeler Aligner software package (Li and Durbin 2009).

Identification of InDels

InDels were called using the default parameters of the Genome Analysis Toolkit program (McKenna et al. 2010). The distribution of the InDels in the R and B lines was visualized in maps constructed with the default settings of the Circos software package (<http://circos.ca/>). The InDels on Chromosome (Chr)_D05 carrying the restorer gene *Rfl* (Wu et al. 2014) were further filtered using the following criteria: minimum read depth: < 10 (R line) or < 5 (B line); average base quality: < 30. The identical InDels detected in both lines as compared to the reference sequences in the TM-1 genome were removed, and the InDels detected only in the R line near the *Rfl* target region were further validated.

InDel marker development and validation

To identify and validate the restorer-specific InDel markers for future applications, 12 InDels with insertions/deletions longer than 40 bp were used to design flanking primers with the Primer Express software (Applied Biosystems, Foster City, CA, USA). Primers specific for an approximately 300 bp region flanking an InDel marker were commercially synthesized (Tianyi Huiyuan Biotechnology, Beijing, China) (Online Resource 1). A 20 μ l PCR mixture was prepared consisting of 1 \times reaction buffer, 2.0 mM MgCl₂, 0.2 mM dNTPs, 0.5 mM each primer, 1 U Taq DNA polymerase (Takara, Kusatsu, Shiga, Japan), and 50 ng DNA template. The PCR amplification conditions were as follows: 25 cycles of 94 °C for 30 s, 56 °C for 30 s, and 72 °C for 1 min. The PCR products were analyzed by 2.0% agarose gel electrophoresis.

The R and B lines as well as the commercial three-line F₁ hybrid CRI83 were used to identify polymorphic InDel markers. We then assessed whether these markers could be used to identify restorer lines containing *Rfl* in different genetic backgrounds (i.e., H46, H80, ZR, and DR) and normal lines lacking *Rfl* (i.e., P1–4). The polymorphic markers were

subsequently used to characterize the genotype of each plant in an BC₃F₁ population (Wu et al. 2014). According to the genotype and phenotype of each plant, the genetic distances between *Rfl* and the associated markers were calculated.

Marker-assisted breeding of restorer lines

The utility of the InDel markers for marker-assisted selection was determined in a segregating population. First, the restorer line Zhonghui 46 [N(*RflRfl*)] was crossed with the recurrent parent P16 [N(*rflrfl*)], which produces fine fibers. Beginning in the BC₁F₁ generation, co-dominant InDel marker InDel-1891 was used to track the restorer gene in each generation, and the other markers were used for further verification. Only those individuals verified by the markers were chosen as the female parent for successive backcrosses. In the BC₅F₂ population 120 individuals were randomly selected, and then InDel-1891 was used to do segregation analysis, the individuals that were verified by the markers as homozygous at the restorer gene locus were test-crossed with the sterile line ZBA S(*rflrfl*) to determine the segregation of the fertility phenotype in the offspring under field conditions.

Results

Whole-genome sequencing and mapping

A total of 238,495,943 and 102,560,619 300 and 250 bp paired-end raw reads generated for the R and the B line, respectively. After filtering the raw reads, 237,038,051 and 102,560,619 clean reads were obtained, representing 59,728,373,610 bases (i.e., approximately 22-fold sequencing depth of the tetraploid Upland cotton genome) and 30,456,504,300 bases (i.e., approximately 12-fold sequencing depth) for the two lines, respectively. About 98.6% and 99.2% of these clean reads were mapped to the TM-1 reference genome, which covered more than 98.1 and 92.8% of the complete R line and B line genomes, respectively (Table 1).

Identification of InDels on chromosome_D05

Based on the TM-1 reference genome sequence, 292,065 and 183,657 InDels were identified in the R

Table 1 Summary of the sequence data and mapping statistics

Species	R line	B line
Total reads	238,495,943	102,560,619
Clean reads	237,038,051 (99.39%)	101,031,801 (98.51%)
Clean bases	59,728,373,610	30,456,504,300
Sequencing depth (fold)	22	12
Total reads mapped	228,528,385 (96.41%)	95,838,766 (94.86%)
Genome coverage (%)	98.12%	92.83%

and B lines, respectively. The genome distribution of these InDels is presented in Fig. 1 and listed in Online Resource 2. Most of the InDels in the R line were detected on Chr_D05, which carries the restorer gene *Rfl*. Following an additional filtering step, 29,131 InDels (approximately 10% of the total number of InDels) were detected on Chr_D05 of the R line (Online Resource 3). In contrast, the corresponding chromosome in the B line contained only 1956 InDels (approximately 1% of the total number of InDels) (Online Resource 4). There were 1253 InDels that were common to both the R and B lines, as compared to TM-1. These results suggest that a substantial portion of the DNA from the *Rfl* donor parent was retained at the *Rfl* locus even after eight generations of targeted backcrossing.

Validation of InDel markers

Twelve InDel markers located near the *Rfl* target region of Chr_D05 were further validated. There were no differences in the PCR product size for eight markers. However, four of these InDels (Online Resource 5), which were named InDel-1891, InDel-3434, InDel-7525, and InDel-9356 (i.e., named after the marker type followed by the final four numbers of the nucleotide position), were polymorphic. As compared to the maintainer B line, the *Rfl*-carrying R line had a 103 nt insertion in the InDel-1891 allele, a 40 nt deletion in the InDel-3434 allele, a 95 nt deletion in the InDel-7525 allele, and a 71 nt deletion in the InDel-9356 allele, they were verified as co-dominant markers in the R line, B line, and the three-line F₁ hybrid CRI83 according to PCR amplifications with the flanking primers (Fig. 2). We confirmed that the four polymorphic InDel markers were able to distinguish between lines containing *Rfl* in different genetic backgrounds and those without *Rfl* (Fig. 2). Additionally, the results of genetic analyses in the BC₈F₁

population revealed that all of the four InDel markers co-segregated with *Rfl* (i.e., based on 409 progenies in the segregating population), indicating a genetic distance below 0.25 cM between the *Rfl* gene and these four InDel markers.

Marker-assisted selection of restorer lines

As the size difference between the two alleles in the InDel-1891 marker locus was the highest to distinguish B and R lines among the four InDel markers, the co-dominant marker InDel-1891 was used for marker-assisted selection of restorer lines exhibiting improved agronomic traits. The BC₁F₁ individuals that produced two PCR products (Fig. 2) were considered heterozygous for the restorer gene [*N(RfIRfl)*], and were chosen for successive backcrossing in each generation from BC₁F₁ to BC₅F₁. One hundred and twenty individual BC₅F₂ plants were randomly analyzed using InDel-1891 (see Fig. 3 as an example). The agarose gel electrophoresis results revealed three different banding patterns. Plants that produced a single large PCR product were considered homozygous for the restorer gene allele [*N(RfIRfl)*], while plants with a single small PCR product were considered to lack the restorer gene allele (*rflrfl*). Plants that produced both fragments were considered heterozygous at the restorer gene locus [*N(RfIrfI)*]. The segregating ratio followed a 1 (*RfRf*):2 (*RfIrf*):1 (*rflrfl*) (26 *RfRf*: 67 *RfIrf*: 27 *rflrfl*, $\chi^2_{0.05} = 1.65 < 5.99$) ratio, as expected for a co-dominant marker. And those plants with restorer gene chosen for further analysis in each generation were all double confirmed by the other three InDel markers.

The fertility results were verified via a progeny test. Ten *N(RfIRfl)* homozygous BC₅F₂ plants as male were test-crossed with CMS-D2 line ZBA to produce 300 plants in the field for the fertility investigation. All of the 300 offspring were genetically heterozygous for

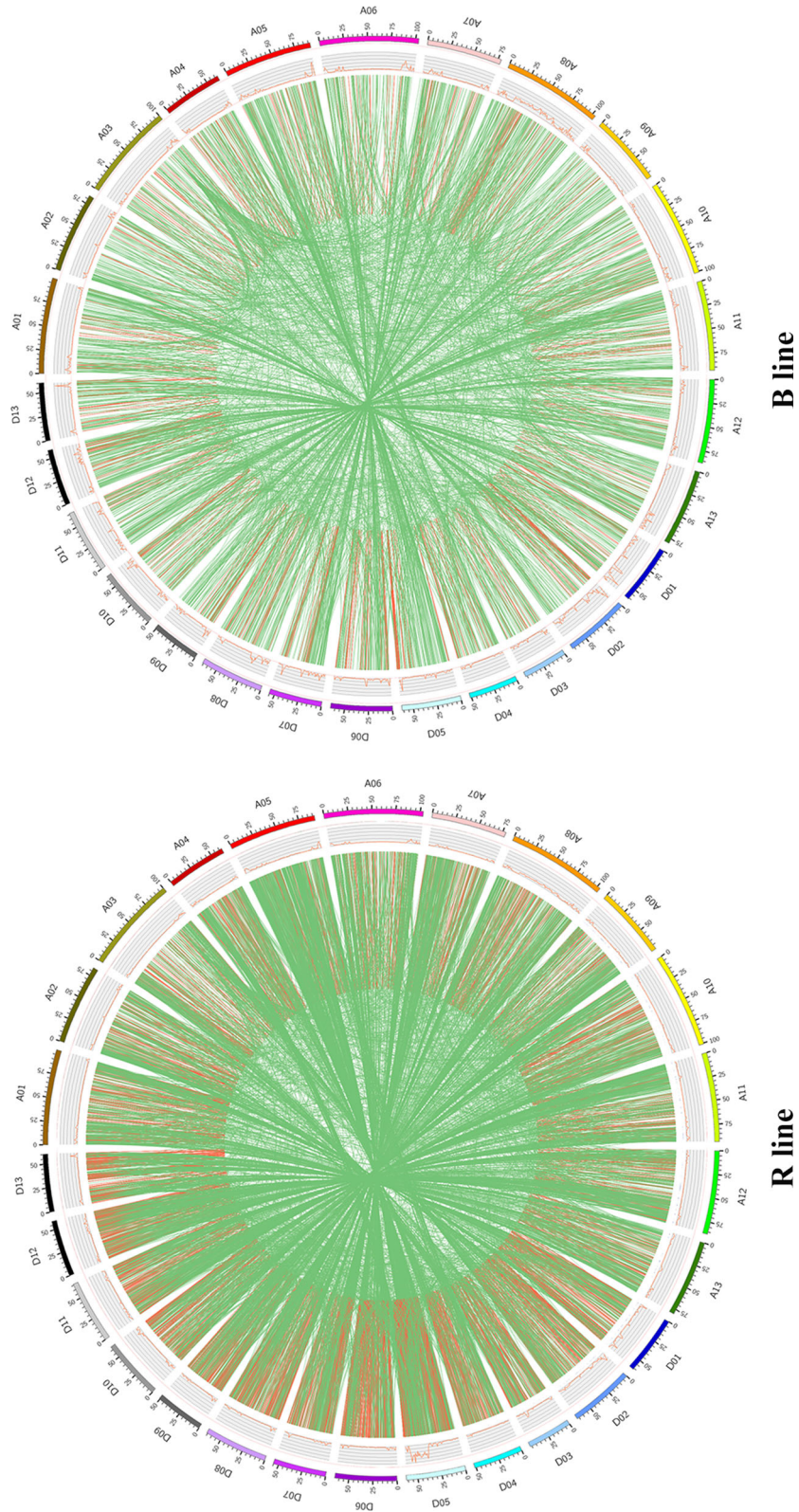


Fig. 1 Genome-wide distribution of InDels in the R and B lines

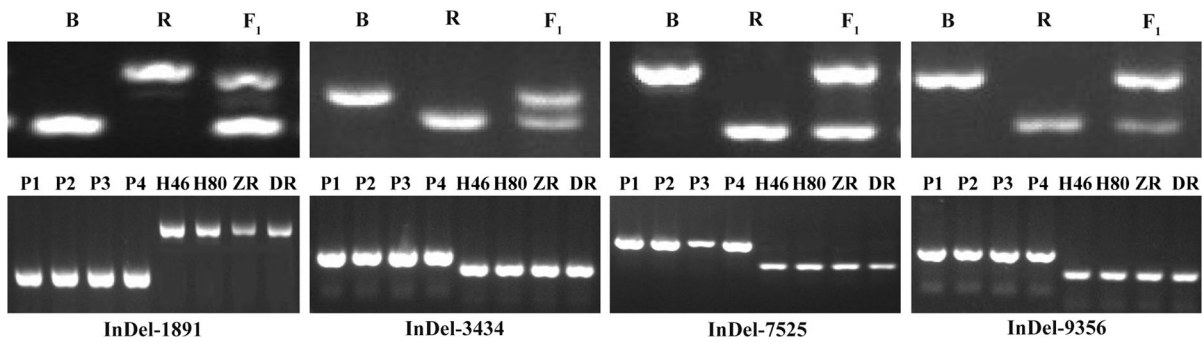


Fig. 2 An analysis of polymorphic InDel markers based on a PCR amplification and 2.0% agarose gel electrophoresis. *B* maintainer line ZB, *R* restorer line ZBR, *F*₁ three-line hybrid CRI83, *P*₁, *P*₂, *P*₃, and *P*₄ lines without *Rf1*, *H46*, *H80*, *ZR*, and *DR* restorer lines

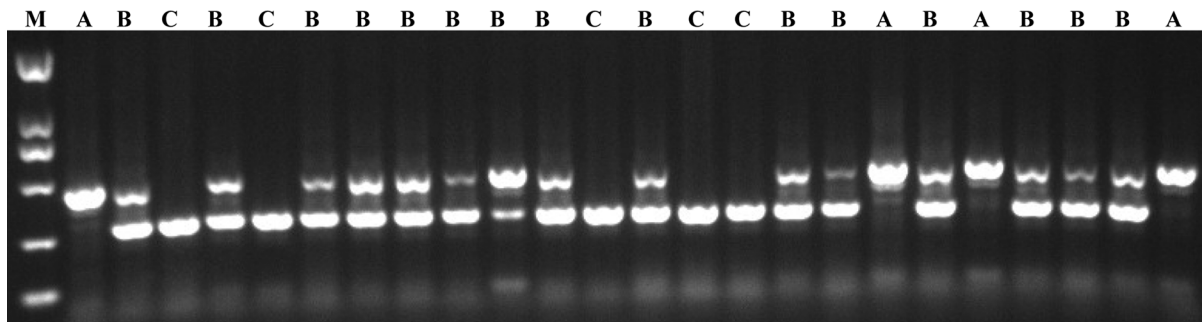


Fig. 3 BC₅F₂ plants were screened with InDel-1891. *M* marker, *A* *Rf1* homozygous plants, *B* *Rf1* heterozygous plants, *C* plants lacking the restorer gene *Rf1*

Rf1, and they were also fertile according to the field testing results, suggesting that the newly developed InDel markers can be used for marker-assisted breeding of restorer lines carrying *Rf1*.

Discussion

Similar to its use in other crops, the CMS system may be ideal for exploiting heterosis in cotton. However, because of the limited availability of suitable restorer line resources, it is still not widely used. Cotton breeders have been attempting to improve the characteristics of restorer lines using traditional breeding methods. Beginning with the BC₁F₁ generation, plants in each generation must be backcrossed with a sterile line, and then planted in the field to assess the fertility of the offspring to confirm the restorer gene remains in the genome. Thus, this procedure is labor-intensive, time-consuming, and expensive. Furthermore, if the restorer line is contaminated by material lacking *Rf1*, some sterile F₁ plants will be produced, which will

considerably decrease cotton yields. To improve the efficiency of restorer line breeding, different types of molecular markers tightly linked with *Rf1* have been developed (Guo et al. 1998; Lan et al. 1999; Liu et al. 2003; Feng et al. 2005; Yin et al. 2006; Yang 2009; Wang et al. 2009; Wu et al. 2014). However, there are disadvantages to these markers, which limit their utility. For example, some simple sequence repeat markers are nonspecific or very weak. Single nucleotide polymorphism (SNP) markers can be used for high-throughput analyses (e.g., microarray hybridization, allele-specific PCR detection, and primer extension; (Kim and Misra 2007; Garvin et al. 2010; Byers et al. 2012), but these procedures are costly. The results generated by SNP markers may be verified with CAPS markers (Thiel et al. 2004) and allele-specific PCR primers (Liu et al. 2012), but only in low-throughput analyses. Genotyping-by-sequencing is a high-throughput option, but it is very expensive for cotton, which has a large genome. In contrast, InDel markers are highly accurate, stable, distributed throughout the genome, and occur frequently (Weber

et al. 2002). With the exception of SNPs, InDels are the most common marker in the rice genome, with one InDel every 953 bp (Shen et al. 2004). The InDel markers, especially those with large insertions/deletions, can be easily amplified and analyzed on agarose gels. Thus, these markers are easy to use and inexpensive. However, developing genome-wide InDel markers is difficult for crops with an incompletely sequenced genome. Because the cotton genome has been fully sequenced (Paterson et al. 2012; Wang et al. 2012; Li et al. 2014a, 2015; Zhang et al. 2015), InDel markers may be a viable option for genome-wide investigations. In this study, whole-genome resequencing was used to identify the InDels in a restorer line. Combining the molecular mapping results for *Rfl*, we focused on the InDels located on Chr_D05. Finally, 29,131 InDels were identified in the R line. Four of 12 InDel markers located near the *Rfl* target region could be used to track *Rfl* and assess the allele status at the *Rfl* locus. The application of these markers may considerably improve the marker-assisted breeding of cotton restorer lines.

Draft genome sequences usually contain several assembly errors and gaps. Therefore, to identify reliable InDel markers, a maintainer line was also resequenced and used to minimize the number of errors. In this study, we identified only 1956 InDels on Chr_D05, including 1253 that were also in the R line. When using draft cotton genome sequences to identify InDel markers, it is important that an efficient method of verifying the utility of many InDel markers is developed.

Acknowledgements Funding was provided by Natural Science Foundation of Henan Province (Grant No. 162300410328).

References

- Byers RL, Harker DB, Yourstone SM et al (2012) Development and mapping of SNP assays in allotetraploid cotton. *Theor Appl Genet* 124:1201–1214
- Feng CD, Stewart JM, Zhang JF (2005) STS markers linked to the *Rfl* fertility restorer gene of cotton. *Theor Appl Genet* 110:237–243
- Garvin MR, Saitoh K, Gharrett AJ (2010) Application of single nucleotide polymorphisms to non-model species: a technical review. *Mol Ecol Resour* 10:915–934
- Grover CE, Yu Y, Wing RA et al (2008) A phylogenetic analysis of indel dynamics in the cotton genus. *Mol Biol Evol* 25:1415–1428
- Guo W, Zhang T, Pan J, Kohel R (1998) Identification of RAPD marker linked with fertility-restoring gene of cytoplasmic male sterile lines in upland cotton. *Chin Sci Bull* 43:52–54
- Hayashi K, Yoshida H, Ashikawa I (2006) Development of PCR-based allele-specific and InDel marker sets for nine rice blast resistance genes. *Theor Appl Genet* 113:251–260
- Jander G, Norris SR, Rounsley SD et al (2002) *Arabidopsis* map-based cloning in the post-genome era. *Plant Physiol* 129:440–450
- Kim S, Misra A (2007) SNP Genotyping: technologies and biomedical applications. *Annu Rev Biomed Eng* 9:289–320
- Lan TH, Cook CG, Paterson AH (1999) Identification of a RAPD marker linked to a male fertility restoration gene in cotton (*Gossypium hirsutum* L.). *J Agric Genom* 4:1–5
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinforma Oxf Engl* 25:1754–1760
- Li F, Fan G, Wang K et al (2014a) Genome sequence of the cultivated cotton *Gossypium arboreum*. *Nat Genet* 46:567–572
- Li X, Gao W, Guo H et al (2014b) Development of EST-based SNP and InDel markers and their utilization in tetraploid cotton genetic mapping. *BMC Genomics* 15:1046
- Li F, Fan G, Lu C et al (2015) Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nat Biotechnol* 33:524–530
- Liu L, Guo L et al (2003) Inheritance and fine mapping of fertility restoration for cytoplasmic male sterility in *Gossypium hirsutum* L. *Theor Appl Genet* 106:461–469
- Liu J, Huang S, Sun M et al (2012) An improved allele-specific PCR primer design method for SNP marker analysis and its application. *Plant Methods* 8:34
- Liu XP, Gao BZ, Han F-Q et al (2017) Genetics and fine mapping of a purple leaf gene, *BoPr*, in ornamental kale (*Brassica oleracea* L. var. *acephala*). *BMC Genomics* 18:230
- Lu BR, Cai X, Xin J (2009) Efficient indica and japonica rice identification based on the InDel molecular method: its implication in rice breeding and evolutionary research. *Prog Nat Sci* 19:1241–1252
- Lv H, Wang Q, Han F et al (2017) Genome-wide indel/SSR scanning reveals significant loci associated with excellent agronomic traits of a cabbage (*Brassica oleracea*) elite parental line “01–20”. *Sci Rep* 7:41696
- McKenna A, Hanna M, Banks E et al (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303
- Meyer VG (1975) Male sterility from *Gossypium harknessii*. *J Hered* 66:23–27
- Paterson AH, Wendel JF, Gundlach H et al (2012) Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. *Nature* 492:423–427
- Shen YJ, Jiang H, Jin JP et al (2004) Development of genome-wide DNA polymorphism database for map-based cloning of rice genes. *Plant Physiol* 135:1198–1205
- Shen C, Jin X, Zhu D, Lin Z (2017) Uncovering SNP and indel variations of tetraploid cottons by SLAF-seq. *BMC Genomics* 18:247

- Singh VK, Khan AW, Saxena RK et al (2017) Indel-seq: a fast-forward genetics approach for identification of trait-associated putative candidate genomic regions and its application in pigeonpea (*Cajanus cajan*). *Plant Biotechnol J* 15:906–914
- Srivastava R, Singh M, Bajaj D, Parida SK (2016) A High-Resolution InDel (Insertion-Deletion) markers-anchored consensus genetic map identifies major QTLs governing pod number and seed yield in chickpea. *Front Plant Sci* 7:1362
- Thiel T, Kota R, Grosse I et al (2004) SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. *Nucleic Acids Res* 32:e5
- Wang F, Yue B, Hu JG et al (2009) A Target Region Amplified Polymorphism marker for fertility restorer gene *Rf(1)* and chromosomal localization of *Rf(1)* and *Rf(2)* in cotton. *Crop Sci* 49:1602–1608
- Wang K, Wang Z, Li F et al (2012) The draft genome of a diploid cotton *Gossypium raimondii*. *Nat Genet* 44:1098–1103
- Wang H, Jin X, Zhang B et al (2015) Enrichment of an intraspecific genetic map of upland cotton by developing markers using parental RAD sequencing. *DNA Res* 22:147–160
- Weaver DB, Weaver JB (1977) Inheritance of pollen fertility restoration in cytoplasmic male-sterile Upland cotton. *Crop Sci* 17:497–499
- Weber JL, David D, Heil J et al (2002) Human diallelic insertion/deletion polymorphisms. *Am J Hum Genet* 71:854–862
- Wu J, Cao X, Guo L et al (2014) Development of a candidate gene marker for *Rf1* based on a PPR gene in cytoplasmic male sterile CMS-D2 Upland cotton. *Mol Breed* 34:1–10
- Yang L (2009) Map-based Cloning of Fertility restoring gene of CMS and analysis of PPR gene family in cotton. Nanjin Agricultural University, Nanjing
- Ye XL, Hu FY, Ren J et al (2016) Fine mapping and candidate gene analysis of *Brtril*, a gene controlling trichome development in Chinese cabbage (*Brassica rapa* L ssp *pekinensis*). *Genet Mol Res.* <https://doi.org/10.4238/gmr15048924>
- Yin J, Guo W, Yang L et al (2006) Physical mapping of the *Rf1* fertility-restoring gene to a 100 kb region in cotton. *Theor Appl Genet* 112:1318–1325
- Yu S, Fan S, Wang H et al (2016) Progresses in research on cotton high yield breeding in China. *Sci Agric Sin* 49:3465–3476
- Zhang JF, Stewart JM (2001a) Inheritance and genetic relationships of the D8 and D2-2 restorer genes for cotton cytoplasmic male sterility. *Crop Sci* 41:289–294
- Zhang JF, Stewart JM (2001b) CMS-D8 restoration in cotton is conditioned by one dominant gene. *Crop Sci* 41:283–288
- Zhang T, Hu Y, Jiang W et al (2015) Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat Biotechnol* 33:531–537
- Zhang H, Wu J, Dai Z et al (2017) Allelism analysis of Br Rfp locus in different restorer lines and map-based cloning of a fertility restorer gene, Br Rfp1, for pol CMS in Chinese cabbage (*Brassica rapa* L.). *Theor Appl Genet* 130:539–547
- Zhao J, Zhang S, Dong J et al (2017) A novel functional gene associated with cold tolerance at the seedling stage in rice. *Plant Biotechnol J* 15:1141–1148