Chinese Science Bulletin

Development of *Gossypium barbadense* chromosome segment substitution lines in the genetic standard line TM-1 of *Gossypium hirsutum*

WANG Peng, DING YeZhang, LU QiongXian, GUO WangZhen & ZHANG TianZhen†

National Key Laboratory of Crop Genetics & Germplasm Enhancement, Cotton Research Institute, Nanjing Agricultural University, Nanjing 210095, China

Chromosome segment substitution lines (CSSL) consist of a battery of near-isogenic lines that have been developed and cover the entire genome of some crops. With the exception of one homozygous chromosome segment transferred from a donor parent, the remaining genome of each CSSL line is the same as the recipient parent. It is an ideal material for genome research and particularly QTL mapping. In the present study, we first developed one set of CSSL lines using *G. hirsutum* acc. TM-1 (the genetic standard), as the recipient parent and *G. barbadense* cv. Hai7124 as the donor parent using molecular assisted-selection in BC_5S_{1-3} generations. The CSSL consisted of 330 different lines, in which 1—4 different lines had the same or overlapping substituted segments. The genetic length of the substituted segments covered 5271.9 cM with an average segment distance of 10.9 cM, 1.5 times the total genetic length of Upland cotton (3514.6 cM). The substituted segments of each line varied in length, ranging from 3.5 cM for the shortest segment to 23.2 cM in the longest segment. Our CSSL have not yet to cover the entire tetraploid cotton genome, due to the absence of some donor parent interval segments.

cotton, chromosome segment substitution lines, fiber qualities, leaf morphology, QTL

Cotton (*Gossypium* spp.) is one of the most important cash crops worldwide. China produces 24% of the world's cotton and consumes 38%, most of which is used to manufacture cotton yarn and cotton fabric[1]. During 2006, the export of textiles in China grossed 144 billion dollars[2]. Improvements in the standard of living in the country and the elimination of textile quotas are expected to rapidly increase the consumption and export of textile products.

Upland and Sea Island cotton are two cultivated tetraploid cotton varieties. Upland cotton (*Gossypium hirsutum* L.) is characterized by high yield and moderate fiber quality performance. Sea Island cotton (*G. barbadense* L.) is characterized by low yield, increased fiber fineness and strength. Sea Island cotton is used as the raw material for fine count yarn. Cotton varieties developed in China are not adequate for fine count and coarse yarns, and in particular 60-count yarn[3]. Therefore, combining the merits of high yield in *G. hirsutum* and good fiber qualities in *G. barbadense* will improve the cotton fiber qualities of Upland cotton. However, years of breeding programs have proven that it is difficult to transfer fiber quality genes from Sea Island cotton into Upland cotton.

Fiber qualities including fiber length, strength, fineness and uniformity are inherited quantitatively, under polygene control and influenced by environment. Rong et al.[4] reported approximately 224 quantitative trait loci (QTLs) for fiber qualities. QTL were mapped using different populations (F_2, BC and RIL) of crosses between

Received December 18, 2007; accepted March 17, 2008
doi: 10.1007/s11434-008-0220-x
†Corresponding author (email: cotton@njau.edu.cn)
Supported by the National Natural Science Foundation of China (Grant No. 30730067), Program for Changjiang Scholars and Innovative Research Team in University (Grant No. IRT0432) and the 111 Project (Grant No. B08025).

G. hirsutum and *G. barbadense*. However, the study demonstrated low repeatability, indicating that the stability of QTLs could not be consistently verified. In addition, this form of interspecific mapping of populations has not been successful in plants cultivated for many years at multiple locations.

Chromosome segment substitution lines (CSSL), including a set of near-isogenic lines spanning the whole genome of one crop can be developed by crosses between the donor and the recipient parents, backcrossing to the donor parent and assisted by molecular markers. One homozygous chromosome segment exists from the donor parent, but the reminder of the genome remains identical to the recipient parent. Therefore, it is an ideal material for genome research and particularly for QTL mapping. Fine mapping of QTLs using CSSL has been reported in tomato, rape and rice, among others. Paterson et al.[5] constructed contiguous introgression lines by RFLPs to perform fine mapping of QTLs. Eshed and Zamir[6] mapped six tomato quality QTLs using CSSL. Yamamoto et al.[7] developed near-isogenic lines and finely mapped three heading date QTLs in rice. In rice, Liu et al.[8] localized 77 QTLs for 18 agronomic traits using 12 single segment substitution lines.

In this study, we developed one set of Sea Island cotton CSSL in the genetic standard line, TM-1 in Upland cotton, by crossing, backcrossing and MAS. The development of this set of CSSL serves as the basis for fine mapping of targeted QTL. CSSL enabled us to investigate interaction between QTLs and QTL cloning and function validation. In addition, this study offered rich germplasm resources for breeding to develop new cotton cultivars with high yield and good fiber quality.

1 Materials and methods

1.1 Marker selection

Our laboratory previously constructed a microsatellite-based, gene-rich linkage map. The map consisted of 1790 loci in 26 linkage groups, covering 3425.8 cM with an average distance of 1.91 cM between markers[9,10]. Based on the map, we selected 330 SSR markers with an average interval of 10 cM between two markers for this study. Not more than 20 markers were present on chromosome 5 and at least eight markers were detected on chromosomes 1, 2 and 8. The shortest genetic distance between two markers was 3 cM and the farthest 25 cM.

1.2 Development of CSSL

The Southern Plains Agricultural Research Center, USDA-ARS, College Station/Texas, USA kindly provided the TM-1 genetic standard from Upland cotton[11]. The variety exhibits moderate fiber quality performance with an average fiber length of 28.56 mm, strength of 30.53 cN/text, and fineness of 5.4. The Cotton Research Institute, Nanjing Agricultural University (NAU), China, developed and provided Hai-7124, long staple Sea Island cotton. The variety is known for superior fiber qualities with an average fiber length of 31.24 mm, strength of 36.25 cN/text, and fineness of 4.2.

Development of CSSL: The F_1 was generated by a cross using TM-1 as the recipient parent and Hai7124 as the donor parent. The F_1 was produced in summer, 2001 at Jiangpu Cotton Breeding Station, NAU. The TM-1× Hai7124 F_1 progeny were backcrossed with TM-1 in winter 2001 in Sanya, Hainan Province, generating BC_1 (Figure 1). In 2002, BC_1 were subsequently backcrossed with TM-1 to produce BC_2 in Nanjing. Backcrosses were continued systematically until reaching a BC_5 generation. In 2004, BC_5 was self-pollinated to produce BC_5S_1 in Nanjing. BC_5S_1 was then planted and the genomic DNA of individual plants was extracted in Nanjing. MAS was conducted to identify individual BC_5S_1 genotypes having 1—2 donor chromosome segments in the TM-1 background. The BC_5S_1 heterozygous individuals were self-pollinated to generate homozygous BC_5S_{2-3}. In this way, additional CSSL were produced, which might be absent in BC_5S_1.

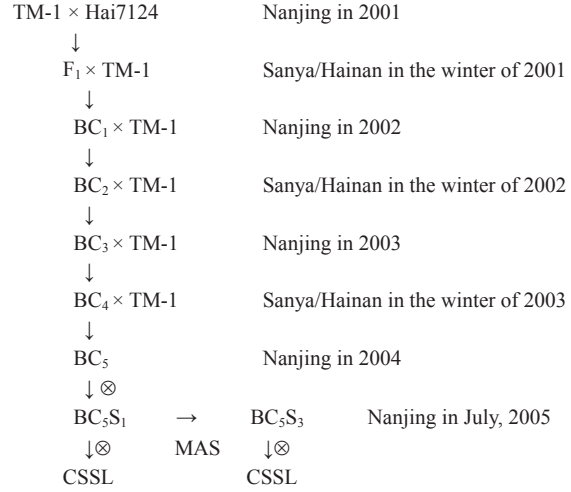

Figure 1 Flow chart of construction of CSSL in cotton.

1.3 Construction of DNA pools and CSSL genotypic analysis

Three DNA pools were constructed for increased efficiency in CSSL genotypic analysis. Genomic DNAs from 2100 BC_5S_1 individuals were extracted and used to construct the third DNA class pool. The second class pool was constructed from the same row of mixed 803 BC_5S_1 individuals. Genomic DNAs from individuals of 93 BC_4 families of the same parents in the second class pool were mixed to construct the first class pool. The following protocol was followed to ensure that we obtained the same genomic DNA content from each individual plant: two leaves from the same position on the plant were collected; three holes were separately cut to remove leaf tissue; the tissue was subsequently ground with liquid nitrogen.

The genotypic compositions of CSSL were analyzed by GGT software (http://www.dpw.wau.nl/pv/pub/ggt/).

1.4 Phenotypic trait evaluation

BC_5S_1 seeds were planted at intervals in twenty rows, including one row of TM-1 as a control. Seedlings were transplanted to the field for phenotype scoring in 2005.

(i) Investigation of leaf morphology. The leaf shape of BC_5S_1 individuals was qualitatively evaluated during flowering.

(ii) Fiber quality trait measurements. Three fiber quality traits were measured, including fiber length, strength and fineness from two parents and BC_5S_1 individuals. Traits were measured and recorded at Henan Cotton Quality Supervision, Inspection and Test Center, Zhengzhou, Henan, China.

2 Results and discussion

2.1 CSSL development in cotton

Two methods are employed to develop CSSL. The first can be described as follows: Two parents are crossed with subsequent backcrossing with the recurrent parent to generate a BC_3. An extensive MAS is initiated until SSL overlapping the entire genome are developed. The methodology is analogous to that of Hao et al., where CSSL were developed that carried the overlapping chromosome segments of the entire wild rice genome[12]. The second approach in the development of CSSL is the use of recombinant inbred lines (RIL). A specific RIL is chosen and backcrossed with the donor parent, assisted by molecular markers. Applying this method, Kubo et al. generated a series of *indica* CSSL from a *japonica* background in rice[13]. RILs are a timesaving approach to develop CSSL because the genetic constitution of each line in the RIL is known.

However, problems in the development of CSSL by the two methods described above are evident. For example, the substituted segments are longer (>20 cM), so several linked QTLs may be present in the interval. In addition, the fine mapping and map-based cloning necessary to construct the second mapping population is timely. Furthermore, more than three different substituted segments in addition to the targeted substituted segment may be present in the CSSL. Due to these limitations, we improved the CSSL development method. In the present study, a co-dominant SSR marker was used to develop the CSSL. This approach is more advantageous than AFLP markers[14,15] because SSR makers are quick and easy to identify, and demonstrate increased levels of polymorphism. In our CSSL development, BC_1—BC_5 were not selected until BC_5S_1 were produced. The single homozygous and short substituted segment lines were more likely generated and subsequently observed because we increased the number of backcrosses and developed a larger BC_5S_1 population (i.e. 2100 individuals).

We constructed three types of DNA pools to decrease the amount of screening required. The first class DNA pool samples were screened with anchored SSR markers based on our saturated genetic map. The second class DNA pool samples were screened according to the screening results of the first class pool; and the third class DNA pool, namely individual DNAs of BC_5S_1 were screened according to the results of the second class pool screening. Consequently, genotypes of each individual were identified to construct the CSSL. This method, compared with the conventional screening procedure, decreased the work demand 10-fold, not only reducing costs but also accelerating the speed of construction.

2.2 CSSL phenotypes

Fiber quality in Hai7124 was characterized as possessing a longer fiber with increased fineness and strength. TM-1 exhibited a moderate fiber quality. Sea Island cotton leaf morphology, relative to upland cotton revealed a longer and narrower leaf lobe and larger angles between the two lobes. The CSSL fiber quality was normally distributed (Figure 2(a)—(c)). CSSL-50, CSSL-85

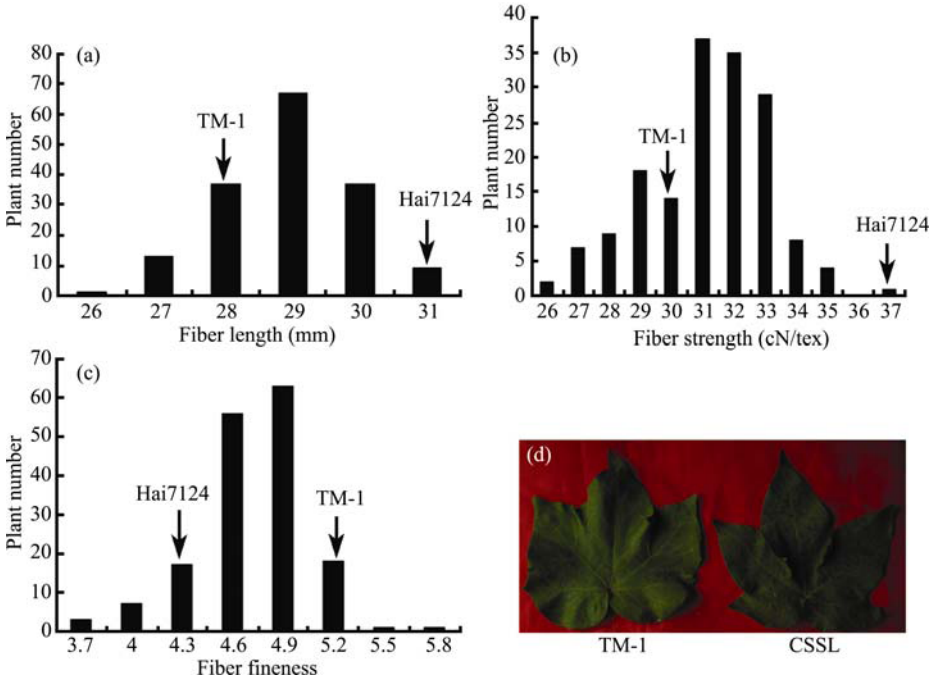

Figure 2 Phenotype of fiber qualities and leaf morphology in CSSL (2005). (a) Fiber length; (b) fiber strength; (c) fiber fineness; and (d) leaf morphology.

and CSSL-90 leaf morphology was typical of Hai7124 (Figure 2(d)), and revealed introgression in the Hai7124 chromosome.

2.3 Molecular evaluation of CSSL

Two thousand and one hundred individuals comprised BC_5S_1 and SSR anchored markers were used to evaluate each individual genotype. Eight hundred plants were identified lacking a substituted donor segment from Hai7124, 220 plants showed 1－2 homozygous segments substituted from Hai7124, 600 plants with 1－2 heterozygous segments substituted from Hai7124, 320 plants exhibited 3－4 heterozygous segments substituted from Hai7124, and 160 plants with more than four heterozygous segments substituted from Hai7124. Furthermore, in 2007, 110 individuals exhibited 1－2 homozygous segments in BC_5S_3. These individuals were identified based on MAS from 200 heterozygous BC_5S_2.

In the present research, we developed 330 *G. barbadense* CSSL from *G. hirsutum* TM-1 genetic standard line using 220 individuals selected from BC_5S_1 and 110 from BC_5S_2. The BC_5S_1 and BC_5S_3 had only one and/or two homozygous segments, which subsequently constituted our CSSL. Among these CSSL, 1－4 different lines might have overlapping segments. One hundred and forty eight lines with different segments were selected and illustrated to depict chromosome segment substitution (Figure 3). The total length of the substituted chromosome segments, which spanned 5271.9 cM with an average segment distance of 10.9 cM, was 1.5 times the total length of tetraploid cotton (3514.6 cM)[10]. The chromosome substituted segment lengths varied among lines (the shortest, 3.5 cM; the longest, 23.2 cM).

2.4 CSSL summary

To date, an entire suite of CSSL that cover the cotton genome have not been developed, which certainly impacts QTL mapping because some intervals lack segments of donor parents (Figure 3). We provide the following two explanations. First, following five backcrosses, some segments were completely replaced by TM-1 and second, many heterozygous individuals were generated because self-pollination was only performed once. We are presently perfecting the methodology so the entire tetraploid cotton genome can be covered by CSSL. This will undoubtedly improve QTL mapping accuracy. Furthermore, we plan to self-pollinate heterozygous BC_5S_1 individuals to generate BC_5S_3. We expect to construct one set of CSSL spanning the entire *G. barbadense* cv. Hai7124 genome.

CSSL are valuable forms of data because they de-

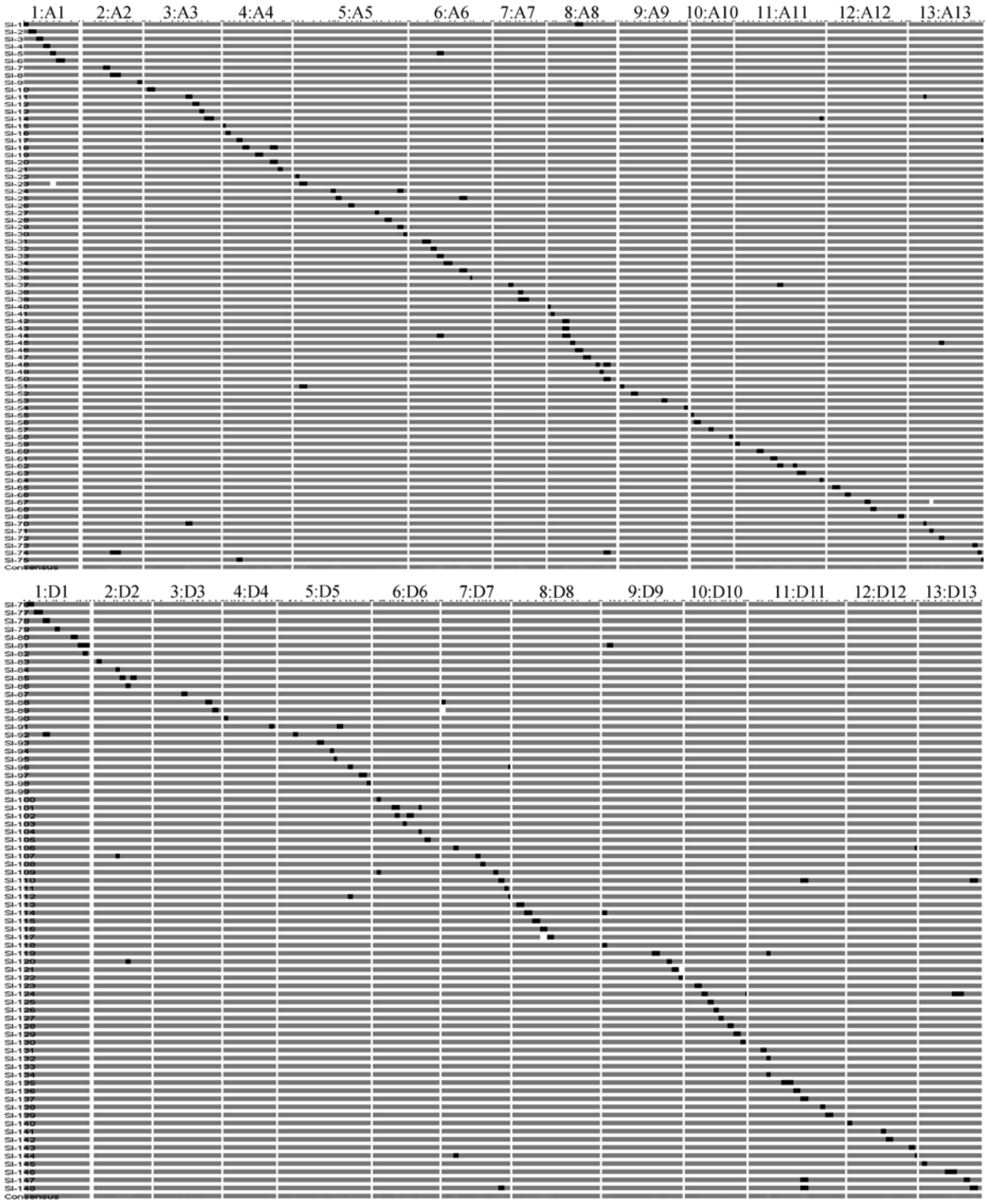

Figure 3 Graphical representation of genotypes of CSSL. Chromosome segment substitution lines having the same substitution segment were represented by one row. The number on the top of the chart is chromosome number: the number on the left end of each row is the number of CSSL; □, receptor genotype; ■, homozygous donor genotype.

crease the interaction effects between QTLs anchored in different segments and subsequently divide the QTL into single Mendelian factors, resulting in higher precision QTL mapping. Therefore, CSSL are the ideal materials for QTL mapping. The large cotton tetraploid genome and complex inheritance results in nominal QTL tagging accuracy and stability. CSSL is one means to resolve the problems inherent in QTL mapping. However, the de-

velopment of CSSL is time-consuming, and CSSL had not previously been developed in cotton. The CSSL and methodology reported here serve to enhance the accuracy and stability of QTL mapping and provide a foundation to promote research in QTL fine mapping, conduct extensive map-based cloning of QTLs, enhance research in the interaction effects between QTLs and provide abundant germplasm for molecular breeding to ultimately develop new cultivars with good fiber quality and disease resistance.

1 Ma S P. Definitude cotton position and development domestic production. Market Cotton Jute China (in Chinese), 2007, 151(4): 25－29

2 Hao X Y, Bian H B. The Effect of the 242 Inform of United States of America. China WTO Tribune (in Chinese), 2007, 51(9): 78－79

3 Xiang S K, Yu N, Hu Y C, et al. On the fiber quality of our country. Cotton Sci Sin (in Chinese), 1999, 11(1): 1－10

4 Rong J K, Feltus A F, Waghmare V N, et al. Meta-analysis of polyploid cotton QTL shows unequal contributions of subgenomes to a complex network of genes and gene clusters implicated in lint fiber development. Genetics, 2007, 176: 2577－2588

5 Paterson A H, Deverna J W, Lanini B, et al. Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes in an interspecies cross of tomato. Genetics, 1990, 124: 735－742

6 Eshed Y, Zamir D. A genomic library of *Lycopersicon pennellii* in *L. esculentum*: A tool for fine mapping of genes. Euphytica, 1994, 79: 175－179

7 Yamamoto T, Kuboki Y, Lin S Y, et al. Fine mapping of quantitative trait loci *Hd-1*, *Hd-2* and *Hd-3*, controlling heading date of rice, as single Mendelian factors. Theor Appl Genet, 1998, 97: 37－44

8 Liu G M, Li W T, Zeng R Z, et al. Development of single segment substitution lines (SSSLs) of subspecies in rice. Chin J Rice Sci (in Chinese), 2003, 17(3): 201－204

9 Han Z G, Wang C B, Song X L, et al. Characteristics, development and mapping of *Gossypium hirsutum* derived EST-SSRs in allotetraploid cotton. Theor Appl Genet, 2006, 112: 430－439

10 Guo W Z, Cai C P, Wang C B, et al. A microsatellite-based, gene-rich linkage map reveals genome structure, function and evolution in *Gossypium*. Genetics, 2007, 176: 527－541

11 Kohel R J, Lewis C F, Richmond T R. Texas marker-1: Description of a genetic standard for *Gossypium hirsutum* L. Crop Sci, 1970, 10: 670－671

12 Hao W, Jin J, Sun S Y, et al. Construction of chromosome segment substitution lines carrying overlapping chromosome segments of the whole wild rice genome and identification of quantitative trait loci for rice quality. J Plant Physiol Mol Biol (in Chinese), 2006, 32(3): 354－362

13 Kubo T, Nakamura K, Yoshimura A. Development of a series of *indica* chromosome segment substitution lines in *japonica* background of rice. Rice Genet Newsletter, 1999, 16: 104－106

14 Howell P M, Marshall D F, Lydiate D J. Towards developing intervarietal substitution lines in *Bassica napus* using maker-assisted selection. Genome, 1996, 39: 248－358

15 Temnykh S, Par W D, Ayres N, et al. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L). Theor App Genet, 2000, 100: 697－712