Breeding Science 61: 160–167 (2011)
doi:10.1270/jsbbs.61.160

Development and linkage mapping of unigene-derived microsatellite markers in *Brassica rapa* L.

Yu Ge[†1)], Nirala Ramchiary[†2)], Tao Wang[1)], Cui Liang[1)], Na Wang[1)], Zhe Wang[1)], Su-Ryun Choi[2)], Yong Pyo Lim[2)] and Zhong Yun Piao*[1)]

[1)] *Department of Horticulture, Shenyang Agricultural University*, Shenyang 110-866, China
[2)] *Department of Horticulture, Chungnam National University*, Daejeon 305-765, Republic of Korea

Brassica rapa plants are highly important as vegetables, sources of oilseeds and fodder crop. Here, we developed 450 unigene derived microsatellite (UGMS) markers in *B. rapa* using unigenes downloaded from the National Center for Biotechnology Information database. Of the 450 UGMS primer pairs, 428 (95.1%) produced repeatable and reliable amplifications of expected size in at least one parental line of *B. rapa*, and 70 UGMS markers gave 72 polymorphic loci between the two contrasting parental lines. Cross-species transferability analysis of these 70 polymorphic UGMS markers in five other cultivated *Brassica* species showed varying transferability rates ranging from 82.9% in *B. nigra* to 97.1% in *B. juncea* and *B. napus*, and overall 53 UGMS markers amplified targets in all five species. The *B. rapa* linkage map was constructed using the 72 UGMS polymorphic loci and 154 previously developed SSRs. The newly developed UGMS markers and linkage map in this study would help in future studies to better understand the organization and evolution of *Brassica* genomes with respect to unigenes, in addition to mapping, tagging and cloning of economically important trait QTL/gene(s) and marker-assisted breeding in *Brassica* crops.

Key Words: *Brassica rapa*, UGMS markers, cross-species transferability, linkage map.

Introduction

During the past two decades tremendous progress has been made in plant molecular genetics with the development of DNA markers, especially in the areas of genetic mapping, comparative genome mapping and evolutionary studies, as well as tagging and marker-assisted selection of economically important quantitative trait loci (QTL)/gene(s). Of the many DNA markers that have been developed, abundant simple sequence repeats (SSRs), the repetitions of nucleotide motifs of 1 to 6 bases, are widely preferred in plant genetics and breeding because of their wide genomic distribution, co-dominant nature, high reproducibility, multi-allelic nature, and suitability for high-throughput analysis. SSR markers have been developed in many plant species from random genomic DNA sequences and used for genetic analysis and comparative genomic studies. However, the traditional SSR marker development method is time-consuming, laborious and expensive. Additionally, the developed SSR markers are anonymous and less cross-species transferable. Therefore, the scientific community is actively searching for new, quick and easy alternative ways to develop SSRs in many plant species using coding sequences of expressed genes (Varshney *et al.* 2004).

Recent advances in several plant genome sequencing and related projects have allowed continuous acquisition of large amounts of DNA sequence information in the form of expressed sequence tags (ESTs) and genome survey sequences (GSS) of many plant species including *Brassica* crops which are being deposited in the National Center for Biotechnology Information (NCBI) database. ESTs are considered ideal candidates for rapid and economical SSR marker development using computational approaches since a large number of SSRs are found in coding regions (Morgante *et al.* 2002). As a result, several efforts were made by many researchers to develop SSRs from ESTs in *Brassica napus* (Kaur *et al.* 2009), peanut (Hong *et al.* 2010), *Hordeum vulgare* (Castillo *et al.* 2008), cassava (Raji *et al.* 2009), pearl millet (Senthivel *et al.* 2008) and many other plant species (Varshney *et al.* 2005). However, due to the problem of sequence redundancy yielding multiple sets of markers at the same locus, random EST sequences are being assembled into unique gene sequences called unigenes (http://www.ncbi.nlm.nih.gov/), and SSRs are designed from these transcribed genes in order to have unique identities and positions in the genome (Parida *et al.* 2006, 2010, Sharma *et al.* 2009). The availability of increasingly large unigene databases now makes systematical and easy identification of microsatellites from a large number of plants possible, greatly facilitating the study of functional diversity in the natural populations and many germplasm collections as well as comparative

Communicated by T. Sato
Received December 16, 2010. Accepted March 13, 2011.
*Corresponding author (e-mail: zypiao@syau.edu.cn)
[†] These authors contributed equally to this work

mapping and evolutionary studies (Ishikawa *et al.* 2008, Parida *et al.* 2006, 2010).

Brassica rapa is an important diploid species with diverse morphological subspecies including Chinese cabbage, pakchoi, flowering Chinese cabbage, turnip and broccoletto, as well as oilseeds that include yellow and brown Sarson. It is also the diploid progenitor parent which contributed the A genome to the widely grown oilseed crops *B. juncea* and *B. napus*. Furthermore, due to the comparatively smaller genome size among the cultivated *Brassica* crops, an inbred line Chiifu-401 of Chinese cabbage (*B. rapa* L. ssp. *pekinensis*) was chosen by the Multinational *Brassica rapa* Genome Sequencing Project as a model for *Brassica* A genome sequencing in 2003 (Yang *et al.* 2005). The high economic importance of *B. rapa* as a vegetable and source of oilseeds and fodder crop led to the development of several molecular markers and construction of several genetic linkage maps in this species during the last two decades (Choi *et al.* 2007, Kim *et al.* 2006, Kim *et al.* 2009, Suwabe *et al.* 2002, 2004, 2006). With the exception of Kim *et al.* (2006) who used sequence tagged sites (STS) and Li *et al.* (2009) who used single nucleotide polymorphisms (SNP) markers from expressed genes in their studies, the majority of the above-mentioned studies used anonymous markers and the genetic linkage exhibited with QTL often may be only due to a physical linkage and statistical relationship.

Therefore, the development of EST-SSR markers from coding regions of the *B. rapa* genome would facilitate of the discovery of new genetic markers from the known biochemical and physiological properties of the candidate gene products in relation to phenotypes. Recently, Parida *et al.* (2010) developed 347 unigene-derived microsatellite (UGMS) markers in the *B. rapa* genome, although there still are undoubtedly many more that remain to be identified. So, considering the advantages and utility of UGMS markers, in the present study, we attempted to develop more UGMS markers from unique and transcribed gene regions and map in the *B. rapa* genome along with previously developed markers.

Materials and Methods

Plant materials

The 154 F_2 mapping population was derived by crossing two diverse Chinese cabbage lines '501 and 601'. The maternal parent (501) is a microspore culture-derived double-haploid (DH) line with a small leaf head. The paternal parent (601) is an inbred line with a heavy leaf head. For evaluating cross-species transferability, five accessions belonging to the *Brassica* species were used which included a single accession of *B. nigra* (24018), *B. oleracea* (24003), *B. juncea* (25001), *B. napus* (264017) and *B. carinata* (25076) along with *B. rapa* (501). All accessions were supplied by Korea National Plant Research Resource Bank in Chungnan National University, Korea. All six accessions used in the cross-species transferability analysis were grown in a field of Shenyang Agriculture University in China. Young leaves from the mapping population, F_1, two parental lines, and five accessions were collected from 3-week-old seedlings. Plant materials were stored at −80°C after flash freezing in liquid nitrogen until DNA extractions were performed. DNA extractions were performed following the procedure by Guillemaut and Laurence (1992) with minor modifications.

Sequence sources and trimming of ESTs

A total of 182,703 *B. rapa* ESTs from various tissues and developmental stages were collected in November, 2008 from NCBI (http://www.ncbi.nlm.nih.gov/). The EST-trimmer (http://pgrc.ipk-gatersleben.de/misa/download/est_trimmer.pl/) was used to remove the poly A or poly T stretches, and EST sequences shorter than 100 bp were excluded, while those longer than 700 bp were retained. The remaining sequences were compared with non-redundant vector databases (ftp://ftp.ncbi.nih.gov/pub/UniVec/) of *Arabidopsis* (http://www.arabidopsisi.org/) by using the Seqclean program (ftp://occams.dfci.harvad.edu/pub/bio/tgi/software/seqclean.tar.gz), and non nuclear/organellar sequences were removed.

Unigene prediction, mining of UGMS and functional annotation of UGMS containing genes

To eliminate duplicates, the trimmed sequences were assembled into clusters by using CAP3 software with parameters of minimum match percentage of 95 and a minimum overlap of 40 bases and further compared by BLAST with E-value $< 1e^{-10}$ to obtain non-redundant unigenes. SSR-containing unigenes were identified using the microsatellite identification tool program (http://pgrc.ipk-gatersleben.de/misa/). The UGMS have either mononucleotides (18 times) or dinucleotide (6 times), tri-, tetra-, penta-, and hexanucleotide (5 times) repeats. For the compound repeats, the maximum interruption of 100 nucleotides was set as default. Frequencies of repeat motifs were calculated for each of the different genic regions accounting for sequence complementarities (eg. (T) n = (A) n, (CT) n = (AG) n, and (CTT) n = (AAG) n). For gene ontology (GO) annotation of SSR-containing unigenes in *B. rapa*, the sequences from predicted coding sequence (CDS) regions were compared with the downloaded complete GO data of *Arabidopsis* (http://www.arabidopsis.org/). In addition, BAC clones harboring UGMS markers were identified from the NCBI (http://www.ncbi.nlm.nih.gov/) through homology search (Supplemental Table 1).

UGMS and SSR genotyping

A total of 450 UGMS markers were developed and designated as "sau_um", to indicate Shenyang Agricultural University and unigene-based microsatellite. All primer pairs were synthesized by Genscript Biotech (Nanjing, China). Furthermore, 651 previously developed SSR primer pairs were used as potential anchor markers to facilitate cross-referencing with other mapping studies. Primer pairs with the prefix "MB" were developed by Lagercrantz *et al.* (1993), "MR" by Uzunova and Ecke (1999), "BRAS" and

"CB" by Piquemal *et al.* (2005), "Ra", "Ol", Na" and "Ni" by Lagercrantz *et al.* (1993) and Lowe *et al.* (2002, 2004), "BRMS" by Suwabe *et al.* (2002, 2006), "BrFLC" by Kim *et al.* (2006), "ENA", "nia" and "cnu" by Choi *et al.* (2007) and by Kim *et al.* (2009), "BnGMS" by Cheng *et al.* (2009) and FITO markers were obtained by Iniguez-Luy *et al.* (2008).

PCR amplifications were carried out in 10 μL reaction volumes, each containing 0.5 unit of Taq polymerase, 1× Taq buffer, 250 μM of dNTPs, 5 pmol of each primer, 2.0 mM of $MgCl_2$, and 15 ng of template DNA. The reaction mixtures were amplified using the following PCR conditions: 5 min at 94°C, followed by 30–35 cycles with 30 s of DNA denaturation at 94°C, 30–45 s of annealing at the appropriate temperature, and 30 s of extension at 72°C, and a final extension at 72°C for 5 min in a BIO-RAD iCycler. Amplified fragments were separated on 6% denaturing polyacrylamide gels. After electrophoresis, the gels were stained as previously described (Sanguinetti *et al.* 1994). Markers showing more than one polymorphic loci were denoted by letters after the primer name (for example, sau_um030 detected two loci which were designated as sau_um030A or sau_um030B).

Map construction

Genetic maps were constructed using JoinMap version 4 (Stam 1993, Van Ooijen and Voorrips 2001). LOD scores 4.0 to 6.0 was used to assign the markers into linkage groups (LGs), and Kosambi's (1944) mapping function was used to convert the recombination value into the map distance. The threshold for goodness-of-fit was set to ≤5.0, a recombination frequency of <0.4 and minimum logarithm of odds scores of 2.0. The map was drawn using Mapchart 2.1 (Voorrips 2002).

BLAST analysis and map alignment with the Arabidopsis genome

The unigene sequences containing SSRs were aligned with the *Arabidopsis thaliana* genome sequences using the BLAST program of NCBI (http://www.ncbi.nlm.nih.gov/). We regarded the sequences as homologous loci of *A. thaliana* genome with a threshold value of $E \leq 1e^{-10}$. The regions having conserved colinearity with *A. thaliana* were regarded as homologous syntenic regions.

Results

Development and characterization of UGMS markers in B. rapa

To develop UGMS markers in *B. rapa*, we analyzed a total of 38,753 unigenes assembled from 182,703 EST sequences in another experiment (data not shown). Analyses using these unigenes identified 5,537 microsatellite motifs in 4,881 genes with a density of 1 per 4.76 kb of EST sequences. Comparison of the relative distribution of microsatellite motifs in 5′ untranslated region (UTR), CDS and 3′ UTR region showed that 5′ UTR (44.8%) had the highest number of microsatellite motifs followed by coding sequences (28.5%) and 3′ UTR (26.7%). Of the 4,881 unigenes containing one or more SSRs, 450 were randomly selected to design primer pairs. Among the primer pairs deigned, trinucleotide repeats were the highest (229, 50.9%) followed by di- (151, 33.6%), compound (56, 12.4%), mono- (11, 2.4%), tetra- (2, 0.4%) and penta-nucleotide repeats (1, 0.2%) (Table 1). The compound microsatellites either contain the same repeat motif that is interrupted by a short non-repetitive sequence or consist of more than two repeat motifs. All the UGMS primer pairs were evaluated for successful PCR amplification by taking the genomic DNA of two contrasting parents used to produce F_2 mapping population. Of the 450 UGMS primer pairs, 428 (95.1%) produced repeatable and reliable amplifications of expected size in at least one parental line of *B. rapa*, while 22 (4.9%) primer pairs completely failed and thus were excluded from further analysis. In order to search BAC clones that contains unigenes, we aligned the sequences of the BACs to the unigenes sequence. Respective corresponding BAC clone was identified with the unigene, with their sequences showing the highest homology. The detailed information of BAC clones ID are listed in Supplemental Table 1. Gene ontology search by comparing UGMS containing genes with *Arabidopsis* proteome database revealed that many of the unigenes belonged to functional category of genes which were involved in transcription factor activity (6.4%), ATP binding (6.4%), protein binding (4.4%), DNA binding 11 (4.4%), or other molecular functions (51.8%). However, 31.1% of the unigenes belonged to genes whose functional categories were unknown (Supplemental Table 1). Of the 450 unigenes used to design primers, 315 showed homology with *A. thaliana* genes, while 135 were unique for *B. rapa*. The detailed information of homologuous loci in *B. rapa* and *A. thaliana* are listed in Supplemental Table 1.

Polymorphism survey and construction of linkage map

A total of 428 newly developed UGMS markers showing at least one fragment amplification in the *B. rapa* genome and 651 previously developed SSR markers were used for surveying polymorphisms between the two parental lines of the *B. rapa* F_2 mapping population. Of these, 70 UGMS (16.4%) gave 72 polymorphic loci, and the 158 previously developed public SSR markers (24.3%) produced polymorphic banding patterns between the parental lines. The majority of the polymorphic UGMS and SSR markers were co-dominant, and only a few (6 UGMS and 10 SSR) markers showed dominant inheritance. The polymorphic markers (226 of 1,079 loci for the mapping population) selected based on clarity of the bands and ease of scoring were used to genotype the entire mapping population for construction of a linkage map. Leaf hairiness (hairless 501 and hairy 601) was a phenotypic trait segregated in the F_2 mapping population and included in linkage analysis.

Linkage analysis with 70 UGMS (72 marker loci), 154

Table 1. Characteristics of *Brassica rapa* unigene derived microsatellite markers

Motifs[a]	Total	5′UTR		CDS		3′UTR	
		number	%motif[b]	number	%motif	number	%motif
Mononucleotide	11	8	72.7	2	18.2	1	9.0
Dinucleotide	151	117	77.5	23	15.2	11	7.3
AG	27	20	74.1	7	25.9	–	–
TC	45	36	80.0	7	15.6	2	4.4
GA	17	14	82.4	3	17.6	–	–
CT	42	33	78.6	4	9.5	5	11.9
Others	20	14	70.0	2	10.0	4	20.0
Trinucleotide	229	60	26.2	159	69.4	10	4.4
AAG	10	4	40.0	6	60.0	–	–
AGA	14	6	42.9	8	57.1	–	–
TTC	12	6	50.0	6	50.0	–	–
TCT	22	9	40.9	12	54.5	1	4.5
GAA	14	6	42.9	7	50.0	1	7.1
GAT	11	1	9.1	10	90.9	–	–
CCT	16	1	6.3	15	93.7	–	–
CTC	12	4	33.3	8	67.7	–	–
CTT	14	5	35.7	9	64.3	–	–
Others	104	18	17.3	78	75.0	8	7.7
Tetranucleotide	2	2	100	–	–	–	–
Pentanucleotide	1	–	–	1	100	–	–
Composite	56	–	–	–	–	–	–

[a] Mono-, di-, tri-, tetra-, and penta-, composite-nucleotide repeats.

[b] Frequency of repeat motif was calculated by the percent of specific motif in that repeat type within the same genic region.

SSRs and one phenotypic marker gave 10 LGs of *B. rapa* with a total genome coverage of 933.9 cM, and the average distance between adjacent loci was 4.11 cM (Table 2 and Fig. 1). Most of the markers segregated following the expected 1 : 2 : 1 Mendelian segregation ratio in the F_2 mapping population, and only 31 (13.7%) markers deviated significantly ($P < 0.01$) from this ratio. A few markers showing skewed segregation ratios were clustered on A7 and A9 (Fig. 1).

The 10 LGs were designated as A1–A10 according to international unified *Brassica* LG nomenclature based on multiple anchor markers mapped on each LG (referred to as R by Choi *et al.* 2007, Suwabe *et al.* 2006 or A by Kim *et al.* 2009). The marker order in the current map was in general agreement with previously published linkage maps (Kim *et al.* 2009, Lowe *et al.* 2004, Piquemal *et al.* 2005, Suwabe *et al.* 2006). LG A9 (126.2 cM) was the longest, LG A4 (69.2 cM) was the shortest, and the markers density varied across LGs for this map (Table 2). Based on the estimated DNA content of 529 Mbp for *B. rapa* (Johnston *et al.* 2005), the estimated physical average spacing between markers was 1.77 Mbp in the current map. The UGMS marker loci were evenly distributed on most of the *B. rapa* LGs, especially on A3, A6 and A9 with a total of 10, 12 and 10 UGMS marker loci, respectively. LGs A1, A2 and A4 had only 4 UGMS marker loci, while the remaining LGs had 6–9 UGMS marker loci (Fig. 1).

Alignment of B. rapa linkage map with A. thaliana chromosomes

Alignment of the current genetic map with the *Arabidopsis* genome was conducted using a total of 147 markers (48 UGMS and 100 SSR). The five chromosomes of *A. thaliana* were divided into many segments and distributed to various regions of the *B. rapa* map (Fig. 1). However, there were regions where collinearity was interrupted by the presence of markers showing hits to multiple regions of the *Arabidopsis* genome. For instance, the upper part of A6 had markers with sequence similarity to *Arabidopsis* chromosomes 1, 2, 3 and the lower half of A3 had markers showing homology with *Arabidopsis* chromosomes 1, 2 and 3. There were 40 large putative syntenic regions containing at least two markers and conserved collinearity with *A. thaliana* chromosomes. The total length of 84 collinear segments with *A. thaliana* chromosomes in *B. rapa* was 627.7 cM (67.2%). This coverage reflected an average length of 6.7 cM per collinear region for the A genome. The longest large putative collinear regions were located in A5 (38.7 cM in length and spanning seven loci). The shortest large putative collinear region (1.6 cM in length and spanning two loci) was identified in A6.

Cross-species transferability across Brassica species

In order to assess the transferability rates of the *B. rapa* UGMS markers, one accession from each of five other cultivated *Brassica* species, *B. nigra*, *B. oleracea*, *B. napus*, *B. juncea* and *B. carinata* were used. The 70 UGMS markers which were mapped in the *B. rapa* genome in the present

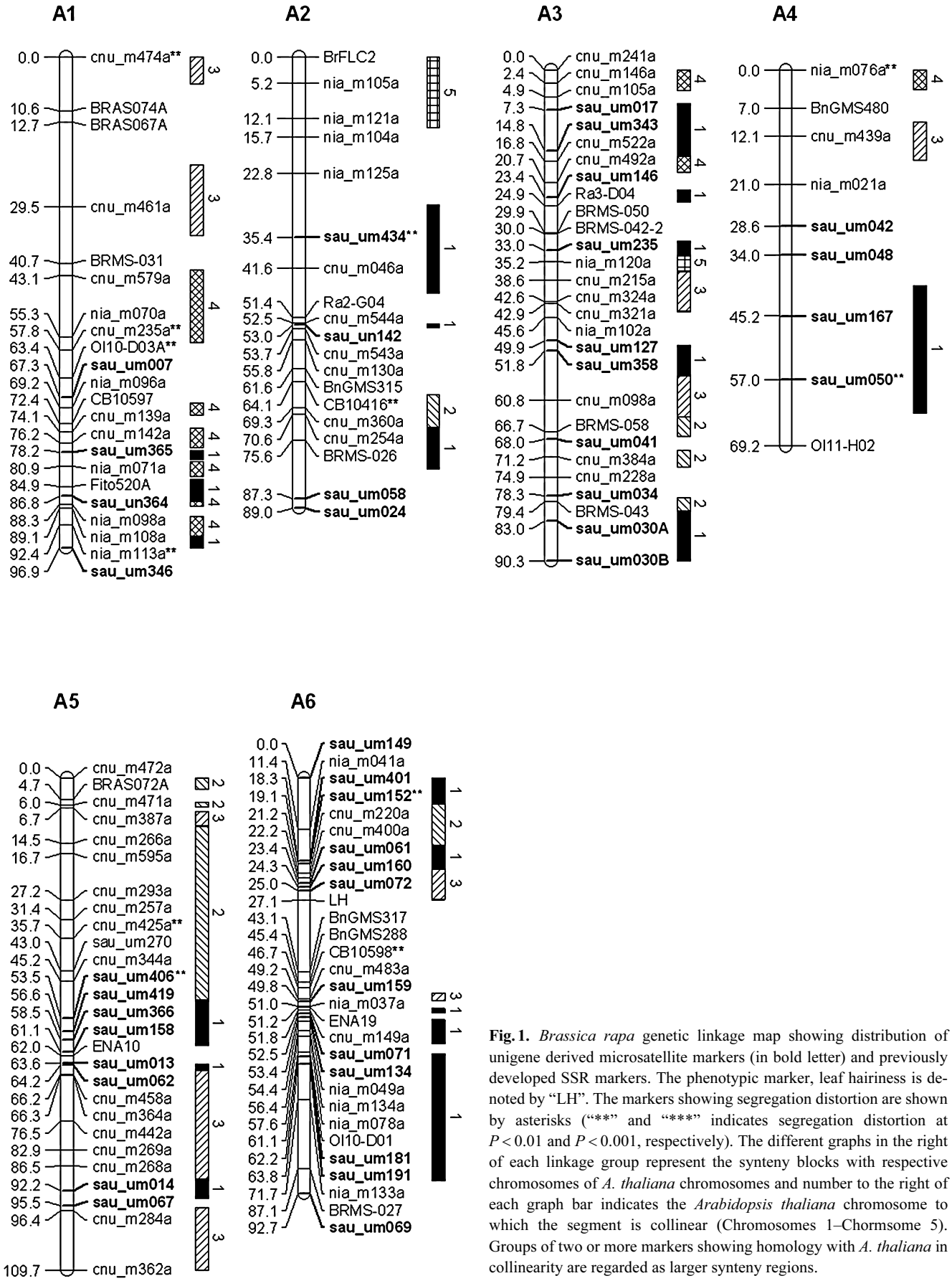

Fig. 1. *Brassica rapa* genetic linkage map showing distribution of unigene derived microsatellite markers (in bold letter) and previously developed SSR markers. The phenotypic marker, leaf hairiness is denoted by "LH". The markers showing segregation distortion are shown by asterisks ("**" and "***" indicates segregation distortion at $P < 0.01$ and $P < 0.001$, respectively). The different graphs in the right of each linkage group represent the synteny blocks with respective chromosomes of *A. thaliana* chromosomes and number to the right of each graph bar indicates the *Arabidopsis thaliana* chromosome to which the segment is collinear (Chromosomes 1–Chormsome 5). Groups of two or more markers showing homology with *A. thaliana* in collinearity are regarded as larger synteny regions.

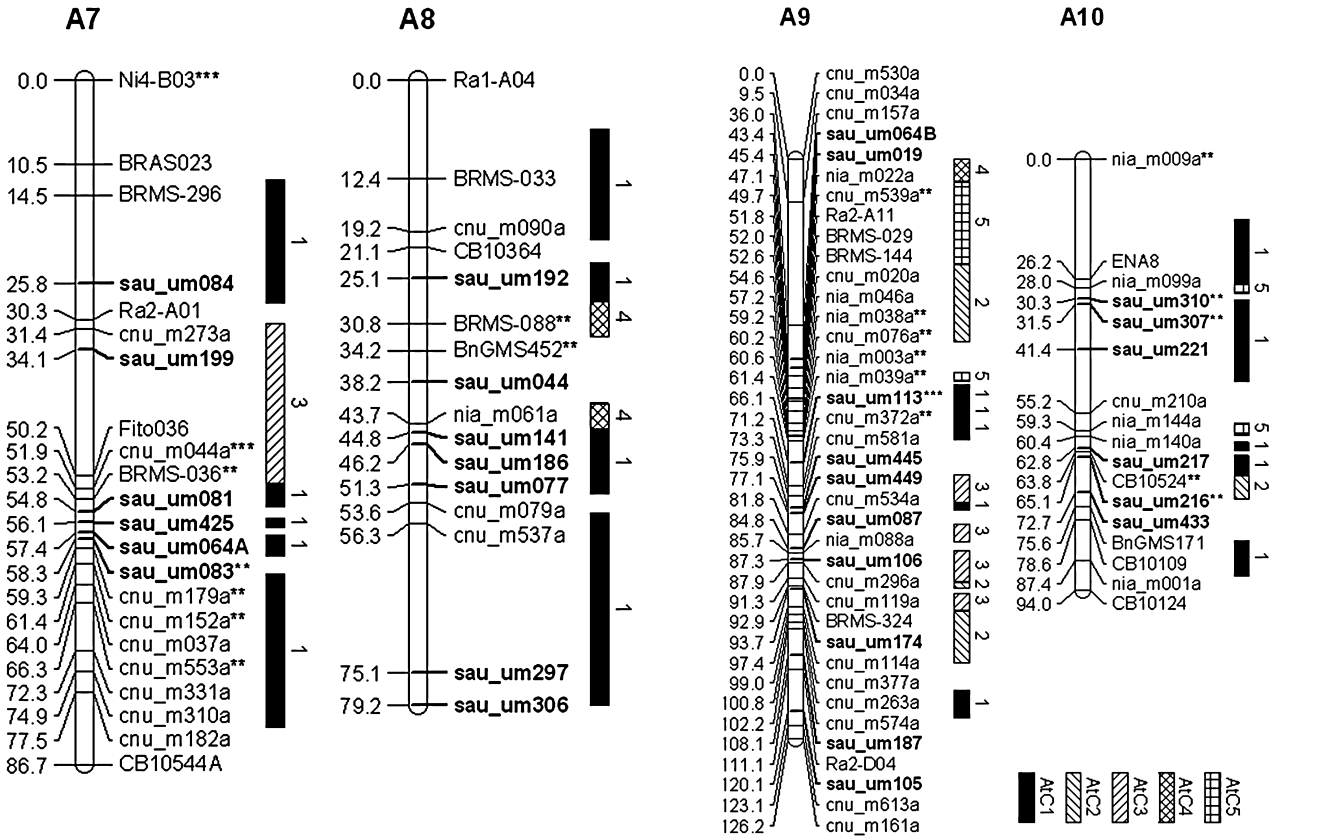

Fig. 1. (continued)

study were used for the transferability analysis so that these markers could be used for comparative mapping and evolutionary studies in other *Brassica* species. PCR amplification of 70 these UGMS markers showed varying degrees of transferability. The transferability rate was 82.9% for *B. nigra*, 91.4% for *B. oleracea*, while it was 97.1% for *B. napus*, *B. carinata* and *B. juncea*. Of the 70 UGMS markers used for PCR amplification in the 5 related *Brassica* species, 53 UGMS markers (75.7%) amplified fragments in all the species (Supplemental Table 2).

Discussion

UGMS markers have been recently developed in order to use unique functional genes in the plant genomes as molecular markers (Parida *et al.* 2006, 2010). This has reduced the redundancy, i.e., multiple copies of genes or repeats of the same DNA sequences, which is a general problem when SSRs are designed from random genomic DNA or EST sequences. To make proper use of this advantage, in another experiment, we downloaded a total of 182,703 *B. rapa* EST sequences from NCBI database and assembled into 38,753 unigenes (data not shown). Characterization of these unigenes gave a total of 5,557 UGMS from 4,887 unigenes. In the present study, however, we initially develop only 450 UGMS markers. Of the 450 UGMS primer pairs designed, 428 (95.1%) primer pairs successfully amplified DNA fragments of expected sizes in the *B. rapa* genome. Several UGMS markers were designed from *B. rapa* and *B. napus* earlier (Parida *et al.* 2010). However, due to the large genome size containing many more functional unigenes in *Brassica* species, representation of the developed UGMS were few. Therefore, we developed additional 450 new UGMS markers using the unigene sequence information which was not used in the earlier study (Parida *et al.* 2010).

The newly developed *B. rapa* UGMS markers displayed a low level of polymorphism (16.4%) between the two parental lines compared to 24.3% polymorphism detected by

Table 2. Details of *Brassica rapa* linkage map developed using UGMS and SSRs

Linkage group	Length (cM)	Number of loci	Average marker interval (cM)
A1	96.9	22	4.40
A2	89.0	19	4.68
A3	90.3	28	3.23
A4	69.2	9	7.69
A5	109.7	27	4.06
A6	92.7	29	3.20
A7	86.7	22	3.94
A8	79.2	16	4.95
A9	126.2	38	3.32
A10	94.0	17	5.53
Total	933.9	227	4.11

genomic SSRs. This observation is noteworthy as SSRs located in the coding regions are under strong selection pressure and therefore generally accumulate only a few mutations. However, despite the lower level of polymorphism, the UGMSs are preferable over genomic SSRs as these are associated with the coding regions of the genome, which may represents "true genetic diversity" that would directly assist in "perfect" marker-trait associations. The low level of polymorphism was also reported in pearl millet with EST-SSRs compared to genomic SSR markers (Senthilvel *et al.* 2008). In our study, UGMS markers mapped to the terminal of *B. rapa* LGs increased the size of respective LGs to a great extent. For example, the size of A8 largely increased from 56 cM to 79 cM at one end of LGs, while A4 extended from 21 cM to 69 cM, nearly twice the length of this LG. The lengths of LGs A3 and A6 also increased by 11 cM and 17 cM, respectively, towards the bottom part of the LGs.

As we used previously developed markers (BAC derived SSRs from Kim *et al.* 2009, and other genomic SSRs mentioned in materials and methods section), we could combine SSR markers from different studies into a single genetic map besides mapping of our newly developed UGMS markers into the *B. rapa* genome. When the current map was compared with previously published *B. rapa* genetic maps, the arrangement of markers in the same order was observed. The length of the newly developed map containing 227 markers (including 72 UGMS loci, 154 genomic SSRs and 1 phenotypic marker) was 933.9 cM which was slightly shorter than the maps of Iniguez-Luy *et al.* (2009), Kim *et al.* (2009) and Li *et al.* (2009), where map lengths were 1,125.3, 1,123.3, and 1,396 cM, repectively. However, Soengas *et al.* (2007) produced a shorter map (663 cM), and longer map lengths have been generated in other studies, including 1,850 cM (Song *et al.* 1991), 1,876 cM (Chyi *et al.* 1992), and 1,785 cM (Teutonico and Osborn 1994). Additionally, we compared the relative lengths of each linkage group with the corresponding linkage groups of Iniguez-Luy *et al.* (2009), Kim *et al.* (2009), and Li *et al.* (2009). Our linkage map indicated that A4 and A8 were markedly shorter than the corresponding linkage groups in other maps. The observation of different map lengths in different studies are attributed to different mapping population size, the number and type of markers, recombination frequencies, scoring errors, and the software employed for analyses. However, of these many reasons, we believe that further addition of markers is required to improve the *B. rapa* map developed in the current study.

Brassica and *A. thaliana*, which diverged 14.5–20.4 MYA from a common ancestor, belongs to the *Brassicaceae* family (Blanc *et al.* 2003). Many comparative genetic mapping studies between *Brassica* species and *A. thaliana* have highlighted the complexity of the *Brassica* genome with its extensive replicated nature, as well as frequent appearance of chromosomal rearrangements (Kim *et al.* 2009, Parkin *et al.* 2005, Suwabe *et al.* 2006). In the present study, comparison of *B. rapa* map with *A. thaliana* using 147 UGMS and SSR markers loci revealed majority large syntenic regions showing consistency with previously reported syntenic regions (Choi *et al.* 2007, Iniguez-Luy *et al.* 2009, Kim *et al.* 2009, Li *et al.* 2009, Suwabe *et al.* 2006). However, we could not detect all the previously reported conserved chromosomal segments in *B. rapa* genome due to the lack of molecular markers. We envision that further addition of markers to create maps with greater density would help to identify all the genomic blocks in the *Brassica* A genome.

The newly developed 450 UGMS markers in the present study could be used by the *Brassica* research community along with the genetic map information in the same or other *Brassica* species since they are highly transferable. The use of previously developed SSRs in our genetic map as anchor markers would also help to identify specific genomic regions in earlier maps for fine mapping, tagging and marker-assisted selection of important traits loci/QTL for economically important traits in their studies (Cheng *et al.* 2009, Lowe *et al.* 2004, Piquemal *et al.* 2005, Suwabe *et al.* 2006). We anticipate that by releasing the information provided in this study, our newly developed UGMS markers and linkage map would help to better understand the organization and evolution of *Brassica* genomes with respect to unigenes, in addition to mapping, tagging and cloning of economically important trait QTL/gene(s) and marker-assisted breeding in *Brassica* crops.

Acknowledgements

This study was supported by grants from the National Natural Scientific Foundation of China, project no. 30771468 and Research Fund for the Doctoral Program of Higher Education of China, project no. 20092103110006.

Literature Cited

Blanc, G., K. Hokamp and K.H. Wolfe (2003) A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. Genome Res. 13: 137–144.

Castillo, A., H. Budak, R.K. Varshney, G. Dorado, A. Graner and P. Hernandez (2008) Transferability and polymorphism of barley EST-SSR markers used for phylogenetic analysis in *Hordeum chilense*. BMC Plant Biol. 8: 97.

Cheng, X.M., J. Xu, S. Xia, J. Gu, Y. Yang, J. Fu, X. Qian, S. Zhang, J. Wu and K. Liu (2009) Development and genetic mapping of microsatellite markers from genome survey sequences in *Brassica napus*. Theor. Appl. Genet. 118: 1121–1131.

Choi, S., G. Teakle, P. Plaha, J. Kim, C. Allender, E. Beynon, Z. Piao, P. Soengas, T. Han, G. King *et al.* (2007) The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project. Theor. Appl. Genet. 115: 777–792.

Chyi, Y.-S., M.E. Hoenecke and J.L. Sernyk (1992) A genetic linkage map of restriction fragment length polymorphism loci for *Brassica rapa* (syn. *campestris*). Genome 35: 746–757.

Guillemaut, P. and M.D. Laurence (1992) Isolation of plant DNA: a fast, inexpensive and reliable method. Plant Mol. Biol. Rep. 10: 60–65.

Hong, Y., X. Chen, X. Liang, H. Liu, G. Zhou, S. Li, S. Wen, C.C. Holbrook and B. Guo (2010) A SSR-based composite genetic linkage map for the cultivated peanut (*Arachis hypogaea* L.) genome.

BMC Plant Biol. 10: 17.

Iniguez-Luy, F.L., A. Vande Voort and T.C. Osborn (2008) Development of a set of public SSR markers derived from genomic sequence of a rapid cycling *Brassica oleracea* L. genotype. Theor. Appl. Genet. 117: 977–985.

Iniguez-Luy, F.L., L. Lukens, M.W. Farnham, R.M. Amasino and T.C. Osborn (2009) Development of public immortal mapping populations, molecular markers and linkage maps for rapid cycling *Brassica rapa* and *B. oleracea*. Theor. Appl. Genet. 120: 31–43.

Ishikawa, G., T. Nakamura, T. Ashida, M. Saito, S. Nasuda, T.R. Endo, J.Z. Wu and T. Matsumoto (2008) Localization of anchor loci representing five hundred annotated rice genes to wheat chromosomes using PLUG markers. Theor. Appl. Genet. 118: 499–514.

Johnston, J.S., A.E. Pepper, A.E. Hall, Z.J. Chen, G. Hodnett, J. Drabek, R. Lopez and H.J. Price (2005) Evolution of genome size in Brassicaceae. Ann. Bot. London 95: 229–235.

Kaur, S., N.O.I. Cogan, G. Ye, R.C. Baillie, M.L. Hand, A.E. Ling, A.K. Mcgearey, J. Kaur, C.J. Hopkins, M. Todorovic *et al.* (2009) Genetic map construction and QTL mapping of resistance to blackleg (*Leptosphaeria maculans*) disease in Australian canola (*Brassica napus* L.) cultivars. Theor. Appl. Genet. 120: 71–83.

Kim, H.R., S.R. Choi, J.N. Bae, C.P. Hong, S.Y. Lee, M.J. Hossain, D.V. Nguyen, M.N. Jin, B.-S. Park, J.-W. Bang *et al.* (2009) Sequenced BAC anchored reference genetic map that reconciles the ten individual chromosomes of *Brassica rapa*. BMC Genomics 10: 432.

Kim, J.S., T.Y. Chung, G.J. King, M.N. Jin, T.-J. Yang, Y.-Mn. Jin, H.-I. Kimand and B.-S. Park (2006) A Sequence-Tagged Linkage Map of *Brassica rapa*. Genetics 174: 29–39.

Kosambi, D.D. (1944) The estimation of map distance from recombination values. Ann. Eugen. 12: 172–175.

Lagercrantz, U., H. Ellegren and L. Andersson (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Res. 21: 1111–1115.

Li, F., H. Kitashiba, K. Inaba and T. Nishio (2009) A *Brassica rapa* linkage map of EST-based SNP markers for identification of candidate genes controlling flowering time and leaf morphological traits. DNA Res. 16: 311–323.

Lowe, A.J., A.E. Jones, A.F. Raybould, M. Trick, C.L. Moule and K.J. Edwards (2002) Transferability and genome specificity of a new set of microsatellite primers among *Brassica* species of the U triangle. Mol. Ecol. Notes 2: 7–11.

Lowe, A.J., C. Moule, M. Trick and K. Edwards (2004) Efficient large-scale development of microsatellites for marker and mapping applications in *Brassica* crop species. Theor. Appl. Genet. 108: 1103–1112.

Morgante, M., M. Hanafey and W. Powell (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat. Genet. 30: 194–200.

Parida, S.K., K.A.R. Kumar, V. Dalal, N.K. Singh and T. Mohapatra (2006) Unigene derived microsatellite markers for the cereal genomes. Theor. Appl. Genet. 112: 808–817.

Parida, S.K., D.K. Yadava and T. Mohapatra (2010) Microsatellites in *Brassica* unigenes: relative abundance, marker design, and use in comparative physical mapping and genome analysis. Genome 53: 55–67.

Parkin, I.A.P., S.M. Gulden, A.G. Sharpe, L. Lukens, M. Trick, T.C. Osborn and D.J. Lydiate (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics 171: 765–781.

Piquemal, J., E. Cinquin, F. Couton, C. Rondeau, E. Seignoret, I. Doucet, D. Perret, M.J. Villeger, P. Vincourt and P. Blanchard (2005) Construction of an oilseed rape (*Brassica napus* L.) genetic map with SSR markers. Theor. Appl. Genet. 111: 1514–1523.

Raji, A.A.J., J.V. Anderson, O.A. Kolade1, C.D. Ugwu1, A.G.O. Dixon and I.L. Ingelbrecht (2009) Gene-based microsatellites for cassava (*Manihot esculenta* Crantz): prevalence, polymorphisms, and cross-taxa utility. BMC Plant Biol. 9: 118.

Sanguinetti, C.J., E. Dias Neto and A.J. Simpson (1994) Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. Biotechniques 17: 914–921.

Senthilvel, S., B. Jayashree, V. Mahalakshmi, P.S. Kumar, S. Nakka, T. Nepolean and C.T. Hash (2008) Development and mapping of simple sequence repeat markers for pearl millet from data mining of expressed sequence tags. BMC Plant Biol. 8: 119.

Sharma, R.K., P. Bhardwaj, R. Negi, T. Mohapatra and P.S. Ahuja (2009) Identification, characterization and utilization of unigene derived microsatellite markers in tea (*Camellia sinensis* L.). BMC Plant Biol. 9: 53.

Soengas, P., P. Hand, J.G. Vicente, J.M. Pole and D.A.C. Pink (2007) Identification of quantitative trait loci for resistance to *Xanthomonas campestris* pv. *campestris* in *Brassica rapa*. Theor. Appl. Genet. 114: 637–645.

Song, K.M., J.Y. Suzuki, M.K. Slocum, P.H. Williams and T.C. Osborn (1991) A linkage map of *Brassica rapa* (syn. *campestris*) based on restriction fragment length polymorphism loci. Theor. Appl. Genet. 82: 296–304.

Stam, P. (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J. 3: 739–744.

Suwabe, K., H. Iketani, T. Nunome, T. Kage and M. Hirai (2002) Isolation and characterization of microsatellites in *Brassica rapa* L. Theor. Appl. Genet. 104: 1092–1098.

Suwabe, K., H. Iketani, T. Nunome, A. Ohyama, M. Hirai and H. Fukuoka (2004) Characteristics of microsatellites in *Brassica rapa* genome and their potential utilization for comparative genomics in *Cruciferae*. Breed. Sci. 54: 85–90.

Suwabe, K., H. Tsukazaki, H. Iketani, K. Hatakeyama, M. Kondo, M. Fujimura, T. Nunome, H. Fukuoka, M. Hirai and S. Matsumoto (2006) Simple sequence repeat-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. Genetics 173: 309–319.

Teutonico, R.A. and T.C. Osborn (1994) Mapping of RFLP and qualitative trait loci in *Brassica rapa* and comparison to the linkage maps of *B. napus*, *B. oleracea*, and *Arabidopsis thaliana*. Theor. Appl. Genet. 89: 885–894.

Uzunova, M.I. and W. Ecke (1999) Abundance, polymorphism and genetic mapping of microsatellites in oilseed rape (*Brassica napus* L.). Plant Breed. 118: 323–326.

Van Ooijen, J.W. and R.E. Voorrips (2001) JoinMap Version 3.0: software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen.

Varshney, R.K., S. Nayak, B. Jayashree, K. Eshwar, H.D. Upadhyaya and D.A. Hoisington (2004) Development of cost-effective SNP assays for chickpea genome analysis and breeding. J. SAT Agric. Res. 3: 1.

Varshney, R.K., A. Graner and M.E. Sorrells (2005) Genetic microsatellite markers in plants: features and applications. Trends in Biotech. 23: 48–55.

Voorrips, R.E. (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J. Hered. 93: 77–78.

Yang, T.J., J.S. Kim, K.B. Lim, S.J. Kown, J.A. Kim, M. Jin, J.Y. Park, M.N. Lim, H.I. Kim, S.H. Kim *et al.* (2005) The Korea *Brassica* Genome Project: A glimpse of the *Brassica* gemone based on comparative genome analysis with *Arabidopsis*. Comp. Funct. Genom. 6: 138–146.