# Fine mapping and physical mapping of a seed coat color gene in Chinese cabbage (*Brassica rapa* L.)

**Jiefu Zhang[1,2], Yuxiang Yuan[3], Xiaowei Zhang[3], Jianfeng Geng[3], Peter B. E. McVetty[1], Sylvie Cloutier[4], Genyi Li[1]**

[1] *Department of Plant Science, University of Manitoba, Winnipeg, R3T2N2, Canada*
[2] *Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China*
[3] *Henan Academy of Agricultural Sciences, Zhengzhou 450001, China Email: fsz@jaas.ac.cn*
[4] *Cereal Research Center, Agriculture Agri-Food Canada, Winnipeg, R3T 2N2, Canada*

**Abstract**

Two DH lines in Chinese cabbage (*B. rapa*) were used as parents to develop a double haploid (DH) population that segregated for seed coat color trait. The segregation data suggested that one Mendelian locus controled the seed coat color trait in this population. After testing 1100 primer combinations, 14 SRAP markers were found to be linked to the seed coat color gene. The sequences of two closely linked SRAP molecular markers matched with annotated genes in Arabidopsis. BAC clones anchoring these two markers were selected to confirm the gene order and colinearity between *B. rapa* and Arabidopsis. With Arabidopsis genome sequence as a reference, chromosome walking to the yellow-seeded gene was performed. SNP and SCAR molecular markers were developed and, together with the original SRAP molecular markers, a fine genetic map was constructed. This map was aligned with its corresponding Arabidopsis physical map, showing that the gene order in this syntenic region containing the seed coat color gene in *B. rapa* is the same as that in Arabidopsis. The closest molecular marker, SNP24650, was 0.2 cM away from the seed coat color gene. The next step is to find the candidate gene and to confirm its function through complementary transformation.

**Key words:** Fine mapping; physical mapping; yellow-seeded gene; Chinese cabbage

## Introduction

Yellow-seeded varieties in oilseed *Brassica* crops, such as 'Yellow Sarson' in *B. rapa*, yellow-seeded *B. napus*, *B. juncea* and *B. carinata*, have inherent advantages over their dark-seeded counterparts in both oil and meal quality (Stringam et al., 1974 ). Yellow seeds have a significantly thinner seed coat than black seeds, thereby leading to lower hull proportion and higher oil and protein content in *Brassica* crops. Additionally some other advantages of yellow seeds involvelighter meal color and lower fiber content in the meal. Consequently yellow seeds result in a better feed value for livestock (Tang et al., 1997; Meng et al. 1998).

The inheritance of the seed coat color trait in *Brassica* species has been analyzed. In *B. rapa*, Ahmed and Zuberi (1971) reported that a single gene is responsible for the dominant brown seed color of the Indian 'Toria' lines over the yellow-seeded 'Yellow Sarson' lines. But Stringam (1980) found that brown seed color trait was determined by two independent dominant genes in *B. rapa*. There are three or four independent recessive genes conditioning yellow seed color trait in *B. napus* (Shirzadegan, 1986; Liu 1992, Rahman et al., 2001). However, Liu et al. (2005) reported that one partially dominant gene controled the yellow seed-coat color trait in *B. napus*. These traditional analyses on the seed coat color trait suggest that the inherence of this trait is complex.

Molecular markers, such as restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and simple sequence repeats (SSR) are used to map the genes controlling seed coat color trait in different *Brassica* species (Teutonico and Osborn, 1994; Chen et al., 1997; Somers et al., 2001; Liu et al., 2005). In *B. rapa*, Teutonico and Osborn (1994) mapped a locus controlling seed coat color on linkage group 5. The bulked segregant analysis (BSA) with AFLPs and SSRs are used to identify markers linked closely to seed coat color trait in *B. juncea*, and one AFLP marker was converted to a sequence characterized amplified region (SCAR) marker (Negi et al., 2000).

Currently, very little is known about the genes controlling seed coat color trait in *Brassica* crops although there are 19 transparent tasta (tt) genes and a few others, such as transparent tasta glabra (ttg1 and ttg2) that are cloned and analyzed functionally in Arabidopsis (Baudry et al., 2004, 2006; Debeaujon et al., 2003; Shirley et al., 1995). To better understand the genes controlling the seed coat color and to manipulate this trait in *Brassica* crops, a Mendelian gene controlling seed coat color trait in *B. rapa* was targeted through map-based gene cloning in this report. Sequence related amplified polymorphism (SRAP) was used to find some molecular markers that were linked to the seed coat color trait and then chromosomal walking was performed using the Arabidopsis genome sequence as a reference.

## Materials and methods

There were 559 microspore-derived doubled haploid (DH) lines of *B. rapa* that were used for gene tagging. These DH lines were developed from a cross of a black-seeded Chinese cabbage DH line, 'Y195-93′, and a yellow-seeded Chinese

cabbage DH line, 'Y177-12'.

Genomic DNA was extracted using a modified version of the 2xCTAB method as described by Li and Quiros (2001). SRAP PCR products were separated with an ABI 3100 Genetic Analyzer and a five-color fluorescent dye set, including 'FAM' (blue), 'VIC' (green), 'NET' (yellow) and 'PET' (red), and 'LIZ' (orange as the standard) was used. All PCR reactions were set up using the same components and amplification program reported by Li and Quiros (2001). Samples from four different color labeled primers were pooled together after running PCR reactions and 2.5 μl of the pooled samples was added to a 5.5 μl mixture of formamide and 500-LIZ size standard (ABI, California), and then denatured at 95 °C for three minutes. The plates containing the samples were then loaded into the auto sampler of the ABI 3100 Genetic Analyzer (ABI, California).

Bulk segregant analysis (BSA) was used to tag the seed coat color gene. Equal quantities of DNA from yellow-seeded lines and black-seeded lines selected from the mapping DH population were pooled together to create the yellow-seeded and black-seeded bulks, respectively. The bulks were subjected to SRAP analysis to identify putative markers linked to the seed coat color gene. Then the candidate SRAP markers were used to run the whole population.

SRAP markers that were linked to the seed coat color trait were sequenced via the following protocol. Denatured polyacrylamide gels were used to separate SRAP PCR products. After electrophoresis, the DNA in the gels was colored with a silver staining kit (Promega, Madison, Wisconsin). The gel pieces containing the polymorphic bands were cut and put into a 1.5 ml tube and 550μl DNA elution buffer (500 mM NH4oAc, 10 mM Mg(oAc)2, 1 mM EDTA, 0.1% SDS) was added. After incubation at 37°C and shaking at 200 rpm for 24 hours, the eluted DNA was precipitated with ethanol and used as template for checking the band size with an ABI 3100 Genetic Analyzer. The PCR products with the same size as the SRAP markers were sequenced directly with ABI 3100 DNA analyzer.

A BAC library was constructed following the protocol available at http://www.genome.arizona.edu/agi/Final. A *B. rapa* male sterile line was used and a BAC cloning vector, pIndigoBAC, was purchased from Epicentre (Madison, Wisconsin). After transformation into *E. coli* ElectroMAX DH10B (Invitrogen, Toronto), colonies were selected and put into 384-well plates with a QBot robotic system (Genetix, New Milton, U.K.). PCR-based screening of the BAC library was performed with plate pools, column and row pools using a robotic liquid handling system (Tecan, Toronto).

When SRAP markers were sequenced, new primers based on the sequence were designed to amplify 4 yellow-seeded and 4 black-seeded lines. When there were more than two bases that were different in yellow-seeded and black-seeded DH lines, specific primers were designed on the basis of these sequence differences to amplify the DNA from yellow-seeded and black-seeded DH lines. Thus the SRAP markers were converted to sequence characterized amplified region (SCAR) markers. Each polymorphic locus was scored as a dominant marker. Linkage analysis was performed on the segregation data of all markers and the seed coat color trait in the 559 DH lines using the software package Mapmaker version 2.0 for Macintosh (Lander et al., 1987).

## Results

The F1 seeds of the reciprocal crosses between the black-seeded parent, 'Y195-93′, and the yellow-seeded parent, 'Y177-12', were identical in color to that of the self-pollinated seeds borne on the maternal parents of the cross, indicating that seed coat color trait was the same as the maternal phenotype and that the black-seeded trait was dominant over the yellow-seeded one. Of the 559 DH lines used for the analysis, 254 DH lines were black-seeded and 305 DH lines were yellow-seeded. The segregation ratio of yellow-seeded versus black-seeded lines was 1:1 ($X^2$, p=0.086), indicating that one Mendelian gene controlled the seed coat color trait in this population.

Using a BSA strategy, 1100 SRAP primer combinations were used to amplify four DNA bulks from 16 (4 × 4) yellow-seeded DH lines and 4 others from 16 (4x4) black-seeded. After checking for the polymorphism, 48 out of the SRAP 1100 primer combinations were selected to amplify 16 yellow-seeded and 16 black-seeded DH lines. Finally, 14 primer pairs were used to analyze the whole mapping population to develop 14 SRAP molecular markers that were linked to seed coat color gene. These 14 SRAP molecular markers were used to generate a genetic map which included the yellow-seeded gene in Chinese cabbage.

The bands for seven SRAP markers, YG338, YB512, YR431, YYb197, YY396, YB458 and YB308, were cut from polyacrylamide gel and DNA was recovered and sequenced. After BLAST analysis with TAIR Arabidopsis database (http://www.arabidopsis.org), four of the marker sequences matched with the annotated genes in Arabidopsis. The sequences of YG338, YB512, YR431 and YYb197 corresponded to the Arabidopsis genes, AT5G26680, AT3G62850, AT5G63330 and AT2G19110, respectively.

YB512 was the closest SRAP molecular marker on the map to the seed coat color gene, and the sequence of this marker matched with a gene, AT3G62850, on Arabidopsis chromosome 3. Using the gene sequence from the flanking region of At3g62850 in Arabidopsis, several primer pairs were designed to amplify the DNA from 4 yellow-seeded and 4 black-seeded DH lines. After sequencing, SNPs were identified, but unfortunately these SNPs did not co-segregate with the seed coat color trait in the yellow-seeded and black-seeded DH line population. Thus the gene order around the SRAP marker YB512 in Chinese cabbage was not conserved compared to the corresponding Arabidopsis gene order in this region. Consequently the chromosome working with At3g62850 could not be continued.

The primers designed with the sequence of the closest SRAP marker YB512 only amplified the DNA from yellow-seeded DH lines. Since the *B. rapa* material used for BAC library construction was black-seeded, all these primers designed with the sequence of the marker YB512 were not usable for BAC library screening. To continue the chromosome walking with this closely linked marker, a genome walking kit (Clontech, Toronto) was used to extend the sequence of the

SRAP marker YB512 to its flanking regions and 1 kb extra sequence outside the marker in Chinese cabbage was obtained. With the primer walking sequence, new primers were designed to amplify the DNA from yellow-seeded and black-seeded DH lines, and to discover SNPs which co-segregated with the seed coat color trait. Meanwhile, these new primers allowed the selection of a BAC clone, 'A6L12', from the *B. rapa* BAC library. The BAC ends of 'A6L12' clone was sequenced, and BLAST analysis with TAIR database showed that the end sequences of the BAC 'A6L12' matched with AT5G24690 and AT5G24650 on Arabidopsis chromosome 5, respectively. Since the sequence from another closely linked SRAP marker YG338 matched with AT5G26680 on Arabidopsis chromosome 5, the chromosomal region containing the marker YB512 and YG338 in *B. rapa* should correspond to the syntenic region on chromosome 5 rather than that on chromosome 3 in Arabidopsis eventually.

The further chromosomal walking focused on the syntenic region on Arabidopsis chromosome 5. SNPs were discovered in the *B. rapa* homologs of AT5G27840, AT5G27410, AT5G27220, AT5G26680, AT5G26160, AT5G25510, AT5G25040 and AT5G24650 on Arabidopsis chromosome 5, and a few SNPs were also converted to SCAR markers. After testing the DH line mapping population, these SNPs in *B. rapa* showed the same gene order as that in Arabidopsis, and the SNP for the homolog of At5g24650 was the closest marker with a genetic distance of 0.20 cM. Consequently, the seed coat color gene was presumed to be very close to this SNP marker.

## Discussion

The goal of this research was to clone the gene controlling the seed coat color trait through a map-based gene cloning strategy. Since the whole genome sequence in Arabidopsis is available, the close relation of *Brassicas* to Arabidopsis offers a powerful tool to the *Brassica* community (Paterson et al., 2001). Since it is easy to sequence SRAP molecular markers and approximately 50% of SRAPs target the gene regions (Li and Quiros, 2001), SRAP molecular markers allows the identification of the corresponding region in Arabidopsis. However the dissimilarity between the *Brassica* and Arabidopsis genomes may result in misleading information and comparative genomics between *Brassicas* and Arabidopsis should be performed cautiously. For instance, the sequence of the SRAP molecular marker YB512 in this report matched with a gene, At3g62850, on chromosome 3 in Arabidopsis. Actually the flanking genes of At3g62850 in Arabidopsis are different from the genes surrounding the At3g62850 homolog in *B. rapa*. Therefore, chromosome walking in *B. rapa* with the sequence of the flanking genes of At3g62850 in Arabidopsis is impossible. In this case, a BAC library and more SRAP molecular markers helped solving this difficulty and eventually a closely linked SNP molecular marker was developed that allows the continuation of the chromosome walking to finally clone the gene that controls the seed coat color trait in *B. rapa* in the near future.

## References


Ahmed S.U., Zuberi M.I. (1971) Inheritance of seed coat color in *Brassica campestris L.* variety Toria. Crop Sci 11:309–310.

Baudry A., Heim M.A., Dubreucq B., Caboche M., Weisshaar B., Lepiniec L. (2004) TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in Arabidopsis thaliana. Plant J. 39(3):366-80.

Baudry A, Caboche M, Lepiniec L. (2006) TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell-specific accumulation of flavonoids in *Arabidopsis thaliana*. Plant J. 2006 46(5):768-79.

Chen B.Y., Jorgensen R.B., Cheng B.F., Heneen W.K. (1997) Identification and chromosomal assignment of RAPD marker linked with a gene for seed coat color in a *Brassica campestris-alboglabra* addition line. Hereditas 126:133–138.

Debeaujon I., Nesi N., Perez P., Devic M., Grandjean O., Caboche M., Lepiniec L. (2003) Proanthocyanidin-accumulating cells in Arabidopsis testa: regulation of differentiation and role in seed development. Plant Cell 15(11):2514-31

Lander E., Green P., Abrahamson J., Barlow A., Daley M., Lincoln S., Newburg L. (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181.

Li G. and Quiros C.F. (2001) Sequence related amplified polymorphism (SRAP) a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in Brassica. Theor Appl Genet 103:455-461.

Liu H.L. (1992) Studies on inheritance of yellow-seeded *Brassica napus* L. Acta Agron Sin 18:241–249.

Liu Z.W., Fu T.D., Tu J.X., Chen B.Y. (2005) Inheritance of seed color and identification of RAPD and AFLP markers linked to the seed color gene in rapeseed (*Brassica napus* L.). Theor Appl Genet 110:303–310.

Meng J..L, Shi S.W., Gan L., Li Z.Y., Qu X.S. (1998) The production of yellow-seeded Brassica napus (AACC) through crossing interspecific hybrids of B. campestris (AA) and B. carinata (BBCC) with B. napus. Euphytica 103:329–333.

Negi M.S., Devic M., Delseny M., Lakshmikumaran M. (2000) Identification of AFLP fragments linked to seed coat color in *Brassica juncea* and conversion to SCAR marker for rapid selection. Theor Appl Genet 101:146–152.

Paterson A.H., Lan T.H., Amasino R., Osborn T.C., Quiros C. (2001) *Brassica* genomics: a complement to, and early beneficiary of, the Arabidopsis sequence. Genome Biol. 2: REVIEWS1011.1.

Rahman M.H. (2001) Production of yellow-seeded through interspecific crosses. Plant Breed 120:463–472.

Shirley B.W., Kubasek W.L., Storz G., Bruggemann E., Koornneef M., Ausubel F.M., Goodman H.M. (1995) Analysis of Arabidopsis mutants deficient in flavonoid biosynthesis. Plant J. 8(5):659-71.

Somers D.J., Rakow G.., Prabhu V.K., Friesen K.R.D. (2001) Identification of a major gene and RAPD markers for yellow seed coat color in *B. napus*. Genome 44:1077–1082.

Stringam G.R. (1980) Inheritance of seed color in turnip rape. Can J Plant Sci 60:331–335.

Stringam G.R., McGregor D.I., Pawlowski S.H. (1974) Chemical and morphological characteristics associated with seed coat color in rapeseed. In: Proc 4th Int Rapeseed Conf. Giessen, pp 99–108.

Tang Z.L., Li J.N., Zhang X.K., Chen L., Wang R. (1997) Genetic variation of yellow-seeded rapeseed lines (*Brassica napus* L.) from different genetic sources. Plant Breed 116:471–474.

Teutonico R.A., Osborn T.C. (1994) Mapping of RFLP and quantitative 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.