DEVELOPMENT OF MOLECULAR MARKERS FOR MARKER ASSISTED SELECTION FOR SEED QUALITY TRAITS IN OILSEED RAPE

By

MD. MUKHLESUR RAHMAN

A Thesis
Submitted to the Faculty of Graduate studies
in Partial Fulfillment of the Requirements
of the degree of

DOCTOR OF PHILOSOPHY

Department of Plant Science
University of Manitoba
Winnipeg, Manitoba, Canada R3T 2N2

September 2007

TABLE OF CONTENTS

LISTS OF TABLES

LISTS OF FIGURES

ACKNOWLEDGMENTS

I am very much grateful to the Department of Plant Science, University of Manitoba, Winnipeg, Manitoba, Canada for providing all necessary support and help I received during my Ph.D. study. This research has been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), Bunge Canada and the Manitoba Canola Growers Association in an IRC program. Norddeutsche Pflanzenzucht (NPZ) Lembke, Hohenlieth, Germany is thanked for supplying proprietary yellow seeded lines. I gratefully acknowledge NSERC, the Faculty of Graduate Studies, University of Manitoba and the Government of Manitoba for their scholarship support. I wish to thank the University of Manitoba Students' Union, the Douglas L. Campbell Graduate Fellowship committee and the Clarence Bogardus Sharpe Memorial Scholarship committee for their financial support well as the Department of Plant Science and the Faculty of Graduate Studies, University of Manitoba for providing Travel Awards.

I would like to pay my deep sense of gratitude, heartfelt respect and immense indebtedness to my reverend supervisor Professor Dr. Peter B.E. McVetty, Department of Plant Science, University of Manitoba for his scholastic guidance, estimable criticism, valuable suggestion, corrections and numberless discussions in the whole doctoral period and all the manuscripts of my thesis throughout the study and research work. I appreciate his kindness, patience and encouragement, during my stay at University of Manitoba.

I want to extend my gratitude towards Dr. Genyi Li, Assistant Professor, Department of Plant Science, University of Manitoba, who is the co-supervisor of my Ph.D. advisory committee, for his support of my work in the molecular genetic lab, for

guidance to work in the development of different molecular markers, showing all molecular techniques more precisely, helpful comments and data analysis.

Thanks to Dr. Dana Schroeder, Assistant Professor, Department of Botany, University of Manitoba, for being a member of my Ph.D. advisory committee, and for great support and guidance to upgrade my knowledge on seed coat coloration mechanisms, fatty acids biosynthesis process, writing review papers, giving seminar presentation and also for her kindness.

I would like to record a special word of thanks to Zudong Sun, Jiefu Zhang, Zining Wang, Zhixia Niu at the molecular genetic lab for proper advice and cooperation during research work. I am grateful to all graduate students in the department of Plant Science for their encouragement during my study and research.

Special thanks to all greenhouse and field staff, especially Sandra Fuller, Judith Nugent-Rigby, Lynn Coleman, Bob Smith, Ian Brown and Cathy Bay. Thanks to Dave Audette and Pat Kenyon for fatty acids analysis of rapeseed samples. Thanks to those who contributed in any way but are not mentioned here.

I would like to expresses my boundless gratitude to Almighty Allah for enabling me for successful completion of this work.

Finally, I would like to pay my heartfelt thanks and gratitude to my family, my beloved mother who always pray for my success, my wife Sharmina Helal for her continuous encouragement, mental support, understanding and tolerance, my son Shadman Mansib, his innocent appearance and boundless smile encouraged me to work hard, my brother Dr. Habibur Rahman for his valuable suggestions during my research.

ABSTRACT

Molecular markers for seed quality traits including erucic acid content genes, seed coat color genes in *B. napus* and seed coat color genes in *B. rapa* were developed.

The inheritance of erucic acid content and the development of molecular markers for the erucic acid genes were studied in a mapping population of a cross between a high erucic acid content *B. napus* cultivar 'MillenniUM03'and a zero erucic acid content *B. napus* line 'SRYS-3'. Two genes (*Bn-FAE1.1* and *Bn-FAE1.2*) which control erucic acid content in *B. napus* were identified. Two BAC clones containing the *Bn-FAE1.1* and *Bn-FAE1.2* genes were identified from a *B. rapa* BAC library and a *B. oleracea* BAC library, respectively. The genes were sequenced and flanking sequences were obtained from the BAC clones. A single nucleotide polymorphism (SNP) in the *Bn-FAE1.1* gene was identified between the high and low erucic acid content alleles. Similarly, a two-base deletion was observed in the *Bn-FAE1.2* allele that produced nearly zero erucic acid in the canola lines compared to the high erucic acid content allele. Sequence dissimilarity of the flanking sequences of these two *FAE1* homologs allowed the design of gene/genome specific primers that amplified the respective *Bn-FAE1.1* and *Bn-FAE1.2* genes in the corresponding A-genome and C-genome in the allotetraploid *B. napus*. SNP detection and genotyping were performed for the *Bn-FAE1.1* gene with an ABI SNaPshot kit. Various lengths of polyT were added at the 5´end of the SNP primer to allow multiplexed detection of these SNP markers. A sequence characterized amplified region (SCAR) marker was developed from the two-base deletion in the *Bn-FAE1.2* gene and detected by an ABI 3100 genetic analyzer. For high throughput detection of the SCAR markers, a genome-specific reverse primer was fluorescently labeled with four different dyes and

combined with 20 different primers designed from the inside of the gene at different sequence positions to produce PCR products with different fragment sizes. The multiplexing of PCR products in both techniques significantly reduced the cost of marker detection for marker assisted selection (MAS) in plant breeding.

Inheritance of seed coat color and identification of molecular markers for individual seed coat color genes were studied in *Brassica napus* from crosses of five pure breeding black seeded cultivars/lines to three pure breeding yellow seeded lines. Trigenic inheritance was observed with the black color alleles dominant over the yellow color alleles at all three loci. Sequenced Related Amplified Polymorphism (SRAP) was employed to identify the molecular markers linked to the seed coat color genes using eleven hundred and sixty eight primer-pair combinations. Three SRAP markers very closely linked to the three different seed coat color genes were developed in *B. napus*. The first seed coat color gene (*Bn1*) marker was linked to the black/brown seed coat color trait of *B. napus*. This marker was converted into a SCAR marker using chromosome-walking technology. The second seed coat color gene (*Bn2*) marker was identified from the high density genetic map constructed in Dr. Li's lab using primer walking from the anchoring marker. This SRAP marker was also linked to the black/brown seed coat color gene located on the linkage group N13 of *B. napus*, and was converted into a co-dominant SNP marker using chromosome-walking technology. The third SRAP marker was closely linked to the third seed coat color gene (*Bn3*) in *B. napus* and co-segregated with the dark/light yellow seed coat color trait. These three markers linked to the three different seed coat color genes can be used in marker assisted selection in plant breeding and map based cloning of the yellow seed coat color genes in *B. napus*.

Digenic inheritance was observed in the segregating populations from a cross of a Canadian brown-seeded *B. rapa* cultivar 'SPAN' and a Bangladeshi yellow sarson *B. rapa* cultivar 'BARI-6'. A SRAP marker was identified as being tightly linked to the major seed coat color gene (*Br1*). This marker co-segregated with the brown seed color trait. Since the dominant SRAP marker was not suitable for MAS in plant breeding, the SRAP marker was sequenced and extended sequences were obtained using chromosome-walking technology. The flanking sequences of the SRAP marker contained 24 SNPs and a 12-bp deletion position that allowed the marker to be converted into a co-dominant SNP marker and a co-dominant SCAR marker, respectively. The SCAR marker was detected in the ABI 3100 genetic analyzer with four fluorescently labeled M13 primers integrated with different SCAR primers, which permitted pooling of PCR samples for high throughput detection. These multiplexed SCAR markers showed great potential in MAS in plant breeding.

FOREWARD

This thesis has been written in manuscript style. A general introduction and literature review precedes the three manuscripts that comprise the main part of the thesis. Each manuscript consists of an abstract, introduction, materials and methods, results, and discussion. All three manuscripts have been submitted to the Theoretical and Applied Genetics (TAG) journal for publication. One of the manuscripts has been accepted for publication in TAG journal. A general discussion, summary/conclusion, recommendations for future study, a list of references cited, and appendices follow the manuscripts.

1.0 GENERAL INTRODUCTION

Rapeseed (*Brassica napus* and *B. rapa*) is one of the most important oilseed crops cultivated in many parts of the world, used as a source of edible oil for human consumption and as a protein-rich meal for livestock feed. Recently, rapeseed oil is being used for the production of biodiesel in the transportation sector, which reduces the emissions of undesirable greenhouse gases in the atmosphere. The Brassica species naturally contain very long chain monounsaturated fatty acids in their seed oil including eicosanoic acid (C20:1), erucic acid (C22:1). Rapeseed cultivars contain high levels of erucic acid in their seed oil, which is used as industrial feedstock for the production of lubricants and other valuable products (Murphy and Sonntag 1991). In rapeseed, erucic acid is generated from oleic acid using an elongation process. Two alleles, E1 and E2 at two loci with additive gene action are responsible for the elongation process (Harvey and Downey 1964, Stefansson 1983). Identification of molecular markers for the erucic acid content genes will enhance breeding program efficiency for the development of high or low erucic acid content rapeseed/canola lines in *B. napus* when crosses of high and low erucic acid parents are used.

Development of double low 'canola' cultivars has created opportunities for low erucic acid rapeseed oil as a high quality vegetable oil in the world, which has perfect fatty acid composition including low saturated fatty acids, high oleic acid and moderate polyunsaturated fatty acids. However, further improvement of rapeseed/canola cultivars will be facilitated by the development of cultivars with higher oil, higher protein, and less fibre content in the seeds. These objectives can simultaneously be achieved by the

development of yellow seeded rapeseed/canola *B. napus* cultivars. It has been reported that the naturally occurring yellow seeded cultivars in *B. rapa, B. juncea* and *B. carinata* contain higher oil, higher protein and lower fibre than the black/brown seeded types (Downey *et al.* 1975). The yellow seeded Brassica genotypes have thinner and transparent seed coat resulting in a lower hull proportion with a bigger embryo, consequently, higher oil content in seeds (Stringam *et al.* 1974). Proanthocyanidins and tannins are the major compounds involved in seed coat pigmentation. These are deposited in the seed coat of black/brown seeded cultivars and reduce the digestibility of seed meal for poultry and livestock (Bell and Shires 1982, Slominski *et al.* 1994). The seed coat of the black/brown seeded cultivars contained more fibre and lower protein than those of the yellow seeded lines (Stringam *et al.* 1974). Although *B. napus* is by far the most important oilseed species among the plant family Brassicaceae, no yellow seeded types occur naturally. Therefore, yellow seeded *B. napus* lines have been developed from interspecific crosses with related species, namely, *B. rapa, B. oleracea* spp. *alboglabra, B. juncea* and *B. carinata* (Shirzadegan and Röbbelen 1985, Liu *et al.* 1991, Chen *et al.* 1988, Chen and Heneen 1992, Rashid *et al.* 1994, Qi *et al.* 1995, Tang *et al*. 1997, Meng *et al.* 1998, Rahman 2001b).

In recent years, molecular markers have been used widely in plant breeding for selecting a trait based on genotype rather than phenotype. This strategy is very effective for the traits which are difficult to assess, affected by environment and even controlled by multigene families. Identification of molecular markers tightly linked to the trait of interest could overcome the problems associated with the selection of desired traits at the very early stage of plant development. Molecular markers associated with the seed coat

color trait in *B. napus*, *B. juncea*, *B. carinata* and *B. oleracea* spp. *alboglabra* have been developed by various research groups (Van Deynze *et al.* 1995, Somers *et al*. 2001, Liu *et al.* 2005, Chen *et al.* 1997, Negi *et al.* 2000, Mahmood *et al.* 2005, Padmaja *et al.* 2005). Restriction Fragment Length Polymorphisms (RFLP), Random Amplification Polymorphism DNA (RAPD) and Amplified Fragment Length Polymorphism (AFLP) marker techniques have been used for mapping the seed coat color genes in *B. napus.* However, molecular markers tightly linked with individual genes in the yellow seed coat multigene family of *B. napus* are being considered for high efficiency selection for the trait and map-based gene cloning. No markers have yet been developed for the yellow seed coat color trait in *B. rapa.*

The objectives of this study were:

1. Development of gene and genome specific molecular markers for the erucic acid content genes in *B. napus.*
2. Study of the inheritance of the seed coat color genes in *B. napus*, development of SRAP molecular markers for individual genes from the multigene family for the seed coat color trait, and conversion of these markers into sequence-characterized amplification region (SCAR) markers and /or single nucleotide polymorphic (SNP) markers for high throughput detection for marker assisted selection in plant breeding.
3. Development of SRAP markers for the seed coat color genes in *B. rapa*, and conversion of the markers into SNP and/or SCAR molecular markers for marker assisted selection in plant breeding.

2.0 LITERATURE REVIEW

2.1.1. Brassica species

The Genus Brassica belongs to the flowering plant family Brassicaceae, formerly known as Cruciferae ("cross-bearing"), also known as the mustard family or cabbage family. Different species of Brassica have great economic importance worldwide by providing vegetable oil for human consumption, meal for livestock and vegetables as well. The harvested components include seeds (source of edible oil such as rapeseed/canola oil, mustard oil), roots (swedes, turnips), stems (kohlrabi), leaves (cabbage, brussels sprouts/miniature cabbages, bak choi, hakusai) and flowers (cauliflower, broccoli). Some Brassica species contain white or colored foliage or flowerheads and are being grown as ornamental plants (Wikipedia 2007) (http://en.wikipedia.org/wiki/Brassica).

2.1.2. Brassica U triangle

Among the different species of Brassica, three allotetraploids and three diploids species are particularly important for agricultural production. Interrelationship among these species is known as 'Triangle of U' as described by U (1935). Nagaharu U was a Korean botanist who was working in Japan, and his former Korean name Woo Jang-choon was transliterated into the Japanese alphabet. Woo resynthesized hybrids from diploid and tetraploid species and studied the chromosomal pairing behaviour in the resulted triploids. The 'Triangle of U' theory explained the contribution of the three diploids species for the development of allotetraploids species. According to the theory, the three allotetraploid Brassica species, namely, *B. carinata* A. Braun. (BBCC, 2n=34),

B. juncea L. Czern. (AABB, 2n=36) and *B. napus* L., (AACC, 2n=38) originated from the three diploid Brassica species *B. nigra* (L.) Koch. (BB, 2n=16), *B. oleracea* L. (CC, 2n=18), and *B. campestris* L. (syn. *B. rapa* L.; AA, 2n=20) (Figure 2.1).

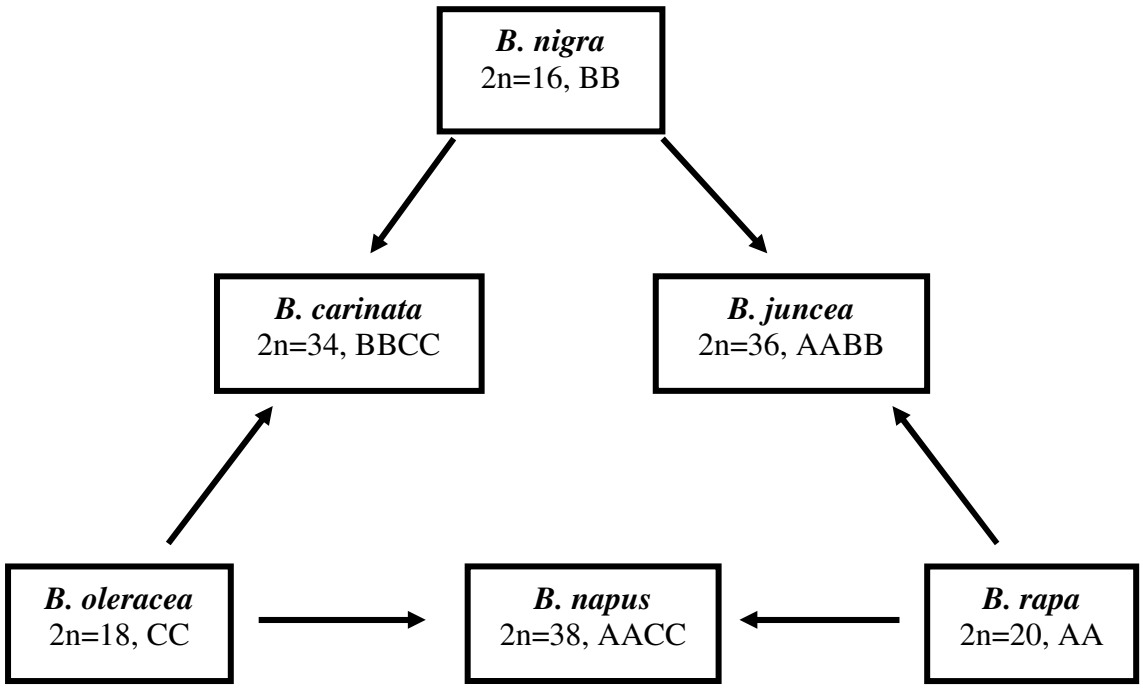


Figure 2.1: Brassica U-triangle

The amphidiploid *B. napus* (2n=38, AACC) originated from two diploid species, *B. rapa* (2n=20, AA) and *B. oleracea* (2n=18, CC) (U, 1935). Olsson (1960) developed artificial *Brassica napus* from *B. rapa* and *B. oleracea,* and showed the same agreement with U (1935). As *B. napus* is developed from the two diploid species, therefore, at least two sets of genes, one from *B. rapa* (A-genome) and the other from *B. oleracea* (C-genome) are responsible for most of the characteristics of *B. napus*.

2.2.1. Rapeseed and canola

Rapeseed also known as Rape, Oilseed Rape, Rapa and Rapaseed, consists of the species *B. napus* and *B. rapa*. Rapeseed cultivars which contain low erucic acid (less than 2%) in the seed oil and low glucosinolates in the meal (less than 30 μmol/g seed) are known as 'Canola'. The term 'Canola' is a registered trademark of the Canadian Canola Association derived from *Can*adian *O*il *L*ow *A*cid. Rapeseed has both spring and winter form cultivars; vernalization is required for the winter types. Spring type rapeseed varieties are mainly grown in Canada, Australia and the Indian subcontinent, whereas the winter type cultivars are grown mainly in Europe and in China.

2.2.2. History of Rapeseed

Rapeseed oil was first used as fuel in lamps by ancient civilizations in Asia and Europe. It was later used as cooking oil. Use of rapeseed oil in the Western World started when steam engines were developed and rapeseed oil was found to be a high quality lubricant. In Canada, rapeseed production was very limited until World War II. Rapeseed production in Canada increased significantly during the Second World War since it was required as a lubricant for the large number of steam engines used by naval and merchant ships during the war. However, the requirement for rapeseed oil suddenly declined at the end of the World War II. In 1954, the first registered high erucic acid, high glucosinolate rapeseed cultivar, 'Golden' was released in Canada (Stefansson 1983). Rapeseed oil was put on the market for edible purposes in 1956-57, but it was not well accepted because of its perceived low quality.

2.2.3. Development of 'Canola' in Canada

The first and only naturally occuring low erucic acid rapeseed line 'Liho' was discovered in 1963 at the University of Manitoba, Canada using a gas chromotography procedure. It had a mutation in the gene for erucic acid biosynthesis in *B. napus* (Stefanson and Hougen 1963). As a result the erucic acid biosynthetic process was genetically blocked by the mutant gene and thus reduced the amount of erucic acid in seed. Glucosinolate content in rapeseed meal adversely affects the nutritional value and palatability when fed to livestock and poultry (Hansen *et al.* 1997). In 1967, the low glucosinolate content variety Bronowski was identified in *B. napus.* The discovery of the erucic acid mutant gene and identification of a low glucosinolate content variety in *B. napus* enabled development of the double low (a low erucic acid content and a low level of glucosinolates) *B. napus* varieties. Finally in 1974, the first double low canola type rapeseed variety 'Tower' in *B. napus* was developed at the University of Manitoba and released in Canada (Stefansson and Kondra 1975).

2.2.4. Economic importance of Canola

The successful development of canola quality rapeseed significantly enhanced the status of canola in the world vegetable oil markets and animal feed industries, and now canola quality rapeseed ranks as the world's second largest oilseed producing crop next to soybean, providing about 11.8% of the world's edible oilseed supply (FAO 2006/2007) (http://www.fao.org/docrep/009/j8126e/j8126e05.htm). Rapeseed production has grown about 745% over the last 40 years worldwide, especially in the major rapeseed producing countries of China, Canada, Europe, India (Figure 2.2). The major field crops grown in

Canada are wheat, canola, barley, oats, maize, soybean, and flax. Canola is ranked as the second most economically valuable crop next to wheat, which contributed over 11 billion dollars of economic activity to Canada. In the year 2004, Canada earned over 2.0 billion dollars from canola exportation in the world (Canola Council of Canada, 2007a) (http://www.canola-council.org/industry.html). The area of cultivation and production of canola has sharply increased from 2759.7 thousand hectares to 5488.82 thousand hectares (increment is about 199%) and 3515.5 thousand tonnes to 9861.6 thousand tonnes (increment is about 280%), respectively, from the year 1990 to 2004 (FAO, 2005) (http://faostat.fao.org/site/336/DesktopDefault.aspx?PageID=336).

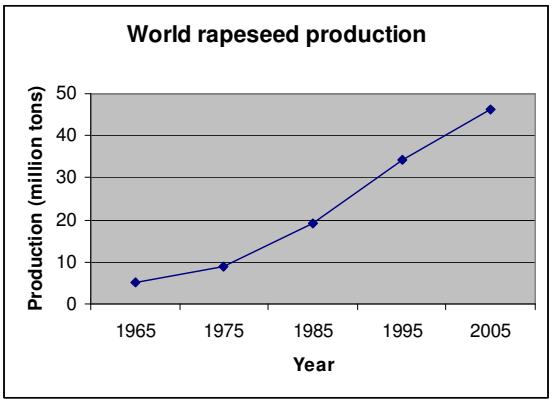


Figure 2.2: Rapeseed production increment in the world during 1965 - 2005 (Based on facts from FAO, 2005).

Recently, biodiesel has emerged as a new energy resource for transportation industries. It is mainly produced from vegetable oils including soybean oil in the USA, and rapeseed/canola oil in Canada and Europe. Traditionally 2% to 10% biodiesel is blended with petroleum diesel which generates fewer emissions, reduces CO_2, particulate matter, and volatile organic matter in the atmosphere and provides better lubrication and

low sulfur emissions in a large number of vehicles. The annual requirement of diesel fuel in Canada is about 60 billion liters, of which 26 billion liters is required for transportation industries. If 5% biodiesel is blended with diesel for transportation, a total of 1.3 billion liters of biodiesel is required in Canada. To obtain that much biodiesel from canola, on the basis of average oil content in seed (40%), about 3.25 million tonnes of additional canola need to be crushed, which is about 33% of total canola production (9.86 million tonnes) in Canada. As a result biodiesel opened a new opportunity for canola producers in Canada (Ag Canada 2007) (http://www.agcanada.com/ custompages/ stories_story .aspx? mid=46&id=1053).

2.3.1. Seed quality traits in Brassica

The major seed quality traits in Brassica are oil content, protein content and glucosinolate content. Lower dietary fibre content in the seed meal is also related to improved digestibility and nutritional value of seed meal (Slominski *et al.* 1994). Fatty acid composition of the oil is also a major determinant for seed quality in Brassica (Ohlrogge and Browse 1995). Erucic acid is one of the major fatty acids in rapeseed oil. Low erucic acid rapeseed oil is a very digestible healthy vegetable oil (Beare *et al.* 1963). In contrast, a high level of high erucic rapeseed oil intake in food is associated with fibriotic myocardium and increased blood cholesterol levels (Gopalan *et al.* 1974). However, high erucic acid rapeseed (HEAR) oil has several applications in the oleo-chemical industry for the production of high temperature lubricants, nylons, plastics, slip & coating agents, soaps, painting inks and surfactants (Topfer *et al.* 1995, Princen and Rothfus 1984). High oleic acid content (>70%) rapeseed seed oil is superior thermostable

vegetable cooking oil (Tanhuanpää *et al.* 1998). Linolenic acid is a poly unsaturated fatty acid, and is one of the essential fatty acids that reduces the plasma cholesterol levels in the human body (Eskin *et al*. 1996). However, linolenic acid contains three double bonds, and therefore is easily oxidized, causing rancidity and off-flavors, thus shortening the shelf life and fry life of the oil (Eskin *et al.* 1989, Prevot *et al.* 1990, Przybylski *et al.* 1993, Scarth *et al.* 1988). Oxidation of linoleic and linolenic acids is approximately 10 and 25 times higher, respectively, than that of oleic acid (Frankel 1991, Kinsella 1991, Carlson 1995, Horrobin 1995, Lands 1997). Partial hydrogenation of the double bond in cis configuration of unsaturated fatty acids (e.g. linolenic acid) generates trans configuration (trans fatty acids), raises serum low-density lipoprotein cholesterol (bad cholesterol) and lowers high-density lipoprotein cholesterol (good cholesterol) in humans and is directly associated with the increased risk of coronary heart disease (Zock *et al.* 1998). However, canola oil can be improved by the development of high oleic acid, low linolenic acid lines/cultivars. Scarth and McVetty (1999) reported that canola oil containing 75% oleic acid and 4% linolenic acid improved the edible oil quality while at the same time it can be used in the oleochemical industries.

Glucosinolates have anti-nutritional effects and adversely affect the nutritional value, taste and smell of food and animal feed stuffs (Hansen *et al.* 1997). Moreover, higher intake of glucosinolates can also have positive physiological effects for human being (Bjergegaard *et al.* 1994).

The seed coat of *B. napus* consists of three different cell layers including the palisade layer, several layers of crushed parenchyma cells, and a single aleurone cell layer (Van Caseele *et al.* 1982). The palisade layer is the major cell layer of the seed coat

containing higher fibre, which is reduced to two-thirds thickness in yellow seed coat in Brassica (Stringam *et al.* 1974). Yellow seeded Brassica contain higher oil, higher protein and lower fibre compared to that of black/brown seeded types (Downey *et al.* 1975). Due to reduced amount of lignins and polyphenols in the yellow seed coat the seed meal is well digestible for livestock (Rahman *et al.* 2001).

2.3.2. Fatty acid composition in the seed oil

The quality of canola oil is primarily determined by its constituent fatty acids. The major fatty-acid constituents in Brassica oil are palmitic acid (C16:0), stearic acid (C18:0), oleic (C18:1), linoleic acid (C18:2), linolenic acid (C18:3), arachidic acid (C20:0), eicosenoic acid (C20:1) and erucic acid (C22:1) (Ohlrogge and Browse 1995). In human diets, high levels of saturated fats (such as palmitic and stearic acid) have been positively correlated with increased blood cholesterol, arteriosclerosis and high a frequency of coronary heart disease (Eskin *et al.* 1996). Compared to the other vegetable

Figure 2.3: Comparison chart of different dietary fat in different oilseed crop oil (Canola Council of Canada 2007b) (http://www.canola-council.org/properties.html)

oils in the world, canola oil contains the lowest saturated fatty acids (5 to 7%), higher monounsaturated fatty acids (60 to 65% oleic acid), moderate levels of polyunsaturated fatty acids (about 20% linoleic acid and 10% linolenic acids) and trace amounts of erucic acid (Figure 2.3). The ratio of linoleic to linolenic acid in canola oil is 2:1 which is recognized as nutritionally favorable.

2.3.3. Fatty acid biosynthesis in seeds

In plant cells the fatty acid biosynthesis process is initiated in plastids and completed in the endoplasmic reticulum (Harwood 1988). Most of the fatty acids in plants have a chain length of 16 or 18 carbons and contain one to three double bonds all in a cis configuration. The fatty acid biosynthesis process shown in Figure 2.4 is as described by Murphy (1999) and Buchanan *et al.* (2000). In embryos, glucose-6-phosphate, phosphoenolpyruvate and pyruvate are produced through a glycolytic pathway in the cytosol from sucrose, and transferred into the plastid using different transporters. Acetyl-CoA, the source of the two-carbon units for the biosynthesis of fatty acids, is generated from pyruvate by the pyruvate dehydrogenase complex in plastids. Plant fatty acid biosynthesis is a cyclic process of a condensation reaction (1), reduction reaction (2), dehydration reaction (3) and finally another reduction reaction (4). Seven repetitions of this process produces a C_{16} fatty acid. The first turn condensation reaction is catalysed by ketoacyl-ACP synthetase III (KAS III) and the next six cycles are catalysed by the isoform of KAS I. Finally KAS II is used for the conversion of $C_{16:0}$ to $C_{18:0}$. The newly synthesized fatty acids release the acyl career protein (ACP) followed by attachment to acyl-CoA derivatives and export to the cytosol. In the endoplasmic reticulum, the acyl

chains are incorporated into glycerolipids through acylation of glycerol-3-phosphate by the action of the acyltransferases and produce phosphatidic acid. Diacylglycerol is formed from the phosphatidic acid by an enzymatic reaction with phosphatidic acid phosphatase.

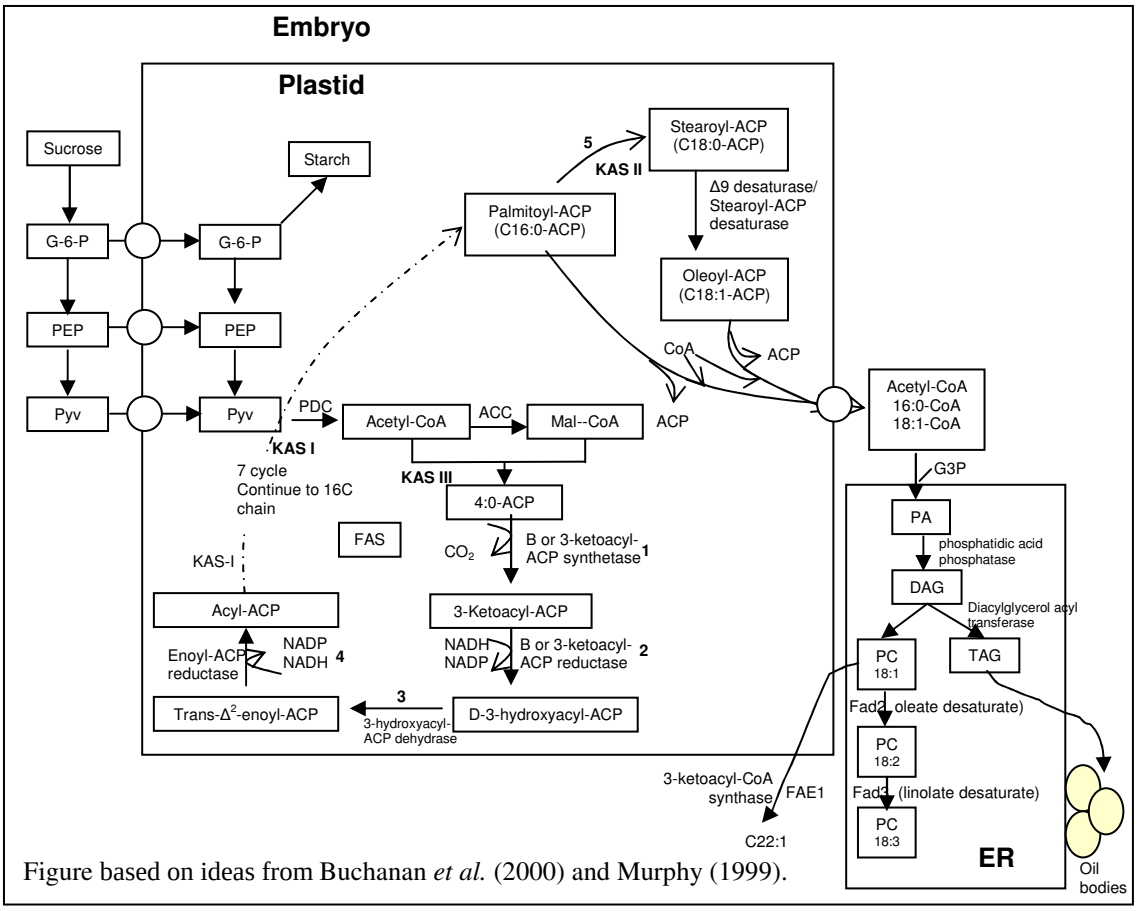


Figure based on ideas from Buchanan *et al.* (2000) and Murphy (1999).

Figure 2.4: Metabolic pathways of fatty acid and lipid biosynthesis in a seed. Mal-CoA: malonyl-CoA; PEP: phosphoenolpyruvate; PDC: pyruvate dehydrogenase complex; ACP: acyl carrier protein; PA: phosphatidic acid; DAG: diacylglycerol; TAG: triacylglycerol; PC: phosphatidyl choline; FAD: fatty acid desaturase; FAE: fatty acid elongase; KAS: ketoacyl-ACP synthetase.

Phosphatidyl choline (PC) and triacylglycerol are synthesized from phosphatidic acid . The enzymes involved in the desaturation of oleic acid to linoleic acid is Δ^{12} oleate

desaturase (FAD2), while the linoleic acid to linolenic acid desaturation enzyme is ω-3-fatty acid desaturase (FAD3). Erucic acid is synthesized from an elongation reaction from PC after a condensation reaction catalyzed by 3-ketoacyl-CoA synthase.

2.3.4. Fatty acid elongation

Oleic acid plays an important role for the biosynthesis of other fatty acids. It may undergo a desaturation process to produce linoleic acid and then linolenic acid or it can undergo a chain elongation process to form ecosenoic acid and subsequently erucic acid (Harwood 1996, Sharma *et al.* 2002). The elongation process has four different steps including the condensation of oleoyl-CoA and malonyl-CoA to form 3-ketoacyl-CoA in the first step, the reduction of the 3-ketoacyl-CoA to produce 3-hydroxyacyl-CoA in the second step, the dehydration of the 3-hydroxyacyl-CoA to form trans-(2,3)-enoyl-CoA in the third step and a further reduction of the trans-(2,3)-enoyl-CoA in the fourth and final step (Puyaubert *et al.* 2005). These reactions are catalyzed by four different enzymes 3-ketoacyl-CoA synthase, 3-ketoacyl-CoA reductase, 3-hydroxyacyl-CoA dehydratase and trans-2,3-enoyl-CoA reductase, respectively. The 3-ketoacyl-CoA synthase enzyme (KCS) is the major catalyzing enzyme utilized for the production of very long chain fatty acid (VLCFA) in seed oils (Puyaubert *et al.* 2005). The reactions involved in this elongation process have been characterized biochemically but gene expression and characterization of gene products still remains unknown. In *Arabidopsis thaliana* and jojoba, the enzyme 3-ketoacyl-CoA synthase was found to be encoded by an *FAE1* gene responsible for fatty acid elongation (Millar and Kunst 1997, Puyaubert *et al.* 2005). The two homologous sequences of the *FAE1* gene (*Bn-FAE1.1* and *Bn-FAE1.2*),

characterized in the embryos of *B. napus*, showed a similarity of 98.2% amino acid sequence of the encoded proteins (Barret *et al.* 1998a, Puyaubert *et al.* 2005). In *B. napus*, the gene encoding the second catalyzing enzyme 3-ketoacyl-CoA reductase for fatty acid elongation has been identified and named as *Bn-kcr* (*Bn-kcr1* and *Bn-kcr2*). It was observed that the *FAE1* and the *Bn-KCR* genes are expressed simultaneously in *B. napus* during the seed development, suggesting that expression of both genes is directly or indirectly linked (Puyaubert *et al.* 2005).

2.4.1. Yellow seeded Brassica

Seed oil in Brassica seed could be increased by the development of yellow-seeded lines/cultivars. It has been reported that the yellow-seeded Brassica produces higher oil, higher protein and lower fibre content over the black/brown seeded types (Downey *et al.* 1975). Even partially yellow-seeded types produced improved seed quality compared to the brown-seeded types in *B. napus* (Liu 1992). Rashid *et al.* (1995) observed 6% higher oil content, similar protein and 4% lower fibre contents in the yellow-seeded strains compared to the black-seeded strains of *B. napus*.

2.4.2. Yellow seeded Brassica: nutritive value of the seed meal

Brassica seed is mainly used for oil purposes, and the seed meal is used as feed stuff for poultry and livestock. Rapeseed meal contains about 40% protein with a well balanced amino acid composition. There are a few anti-nutritional compounds, including glucosinolates, sinapine, tannins, and phytic acid present in rapeseed meal. Glucosinolates have a negative effect on growth and health of animals. Sinapine produces

off-flavored trimethylamine in the seed meal that reduces the usability of the meal to the susceptible hens. The black/brown seed coat contains abundant tannins that reduce the digestibility of seed meal especially affecting the protein hydrolysis. Embryo-contained phytic acid affects phosphorus binding as well as other essential minerals (Uppström 1995, Griffiths *et al*. 1998, Matthäus 1998, Naczk *et al*. 1998, Velasco and Möllers 1998). The seed meal from yellow seeded lines has higher protein and less anti-nutritional compounds that make the meal favorable for use in the poultry or hog industries (Anjou *et al.* 1977, Slominski *et al.* 1994).

2.4.3. Yellow seeded Brassica: fibre content in the seed meal

Higher fibre content in the seed meal is less desirable for monogastric animals. The yellow seeded Brassica lines have less tannins and less fibre content in the seed meal which significantly improves the meal quality (Anjou *et al.* 1977, Slominski *et al.* 1994). Slominski *et al*. (1999) conducted a comparative study on nutritive value of seed meal derived from yellow-seeded *B. rapa, B. juncea* and *B. napus*, and reported that the meal from yellow-seeded *B. napus* had lowest dietary fibre content (271 g/kg dry matter) while it was highest in the black-seeded *B. napus* (352 g/kg dry matter). The yellow seeded *B. napus* meal generated the highest metabolizable energy when broilers were fed with this meal. These results indicated that meal from yellow seeded *B. napus* could be utilized successfully in the poultry industry.

2.5.1. Seed coat coloration

The pigmentation of the seed coat in Brassica occurs mainly due to deposition of polyphenols which are polymers of leucocyanidins (Leung *et al.* 1979, Hu 1988) that belong to the group of flavonoids (Nørbæk *et al.* 1999). According to Van Caseele *et al.* (1982), the seed coat of *B. napus* has three different cell layers including the palisade layer, several layers of crushed parenchyma cells, and a single aleurone cell layer. Chen and Meng (1984) reported that the palisade and crushed parenchyma layers belong to the outer integument of the ovary while the aleurone layer is the outermost layer of endosperm. Deposition of the seed coat pigment (flavonoids) occurs in the palisade and crushed parenchyma layers of the seed (Vaughan 1970, Stringam *et al.* 1974). Stringam *et al.* (1974) reported that the palisade layer is the cell layer of the seed coat that has higher fibre. It is about two-thirds thinner in yellow seed coat lines than in black/brown seed coat lines. The reduction of the palisade layer in yellow seeded lines reduces the proportions of polyphenol and lignin in seed (Anjou *et al*. 1977, Theander *et al*. 1977, Slominski *et al*. 1994).

2.5.2. Flavonoid biosynthesis

Flavonoids are the major compounds involved in the seed coat pigmentation process (Shirley 1998). Seed flavonoids have been classified into different groups that include flavonols, anthocyanins, phlobaphenes, isoflavones, and proanthocyanidins. Proanthocyanidin also known as condensed tannin is deposited only in the seed coat. It is synthesized through a common phenylpropanoid pathway in the flavonoid pathway (Lepiniec *et al.* 2006, Figure 2.5). The aromatic ring (naringenin) is generated by

chalcone synthase and chalcone isomerase. Oxidation of the aromatic ring yields a dihydrokaemperol (dihydroflavonol), which is hydroxylated at the 3´ or 5´ positions or on

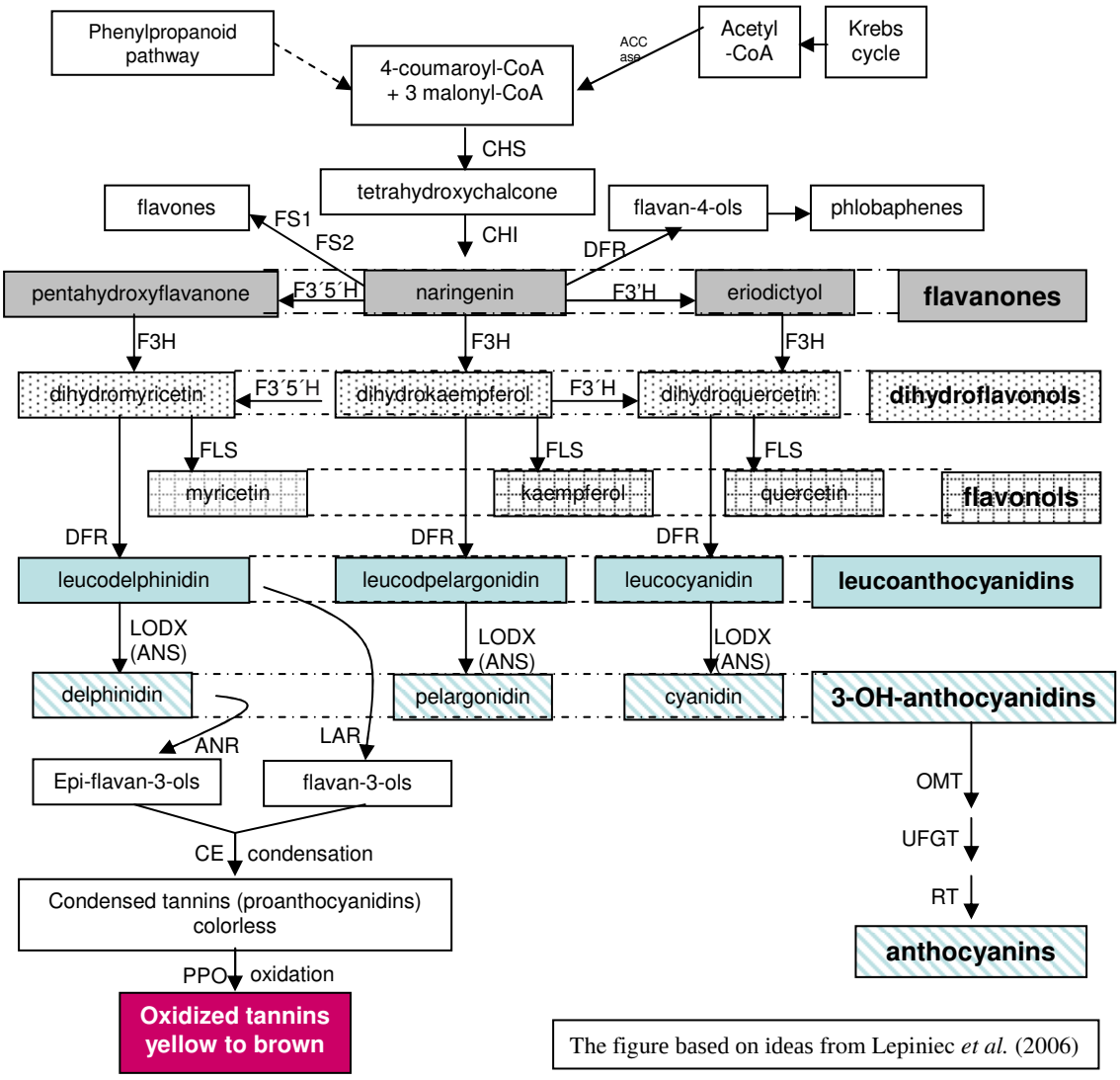


Figure 2.5: Flavonoid biosynthetic pathway. ACCase, acetyl CoA carboxylase; ANS, anthocyanidin synthase; ANR, anthocyanidin reductase; CE, condensing enzyme; CHS, chalcone synthase; CHI, chalcone isomerase; DFR, dihydroflavonol 4-reductase; F3H, flavanone 3-hydroxylase; F3´H, Flavonoid 3´ hydroxylase; F3´5´H, Flavonoid 3´5´ hydroxylase; FLS, flavonol synthase; LAR, leucoanthocyanidin reductase; LDOX, leucoanthocyanidin dioxygenase; OMT, *O*-methyltransferase; PPO, polyphenol oxidase; RT, rhamnosyl transferase; UFGT, UDP Flavonoid glucosyl transferase.

both positions by Flavonoid 3´ hydroxylase and Flavonoid 3´5´ hydroxylase to produce two other dihydroflavonols including dihydroquercetin and dihydromyricetin, respectively. These dihydroflavonols are converted by dihydroflavonol 4-reductase to produce different leucoanthocyanidins, and then by leucoanthocyanidin dioxygenase to yield delphinidin (purple color), pelargonidin (orange color) and cyanidin (red-magenta color), respectively, which can further be substituted by *O*-methyltransferase, UDP Flavonoid glucosyl transferase, and rhamnosyl transferase, resulting in decorated anthocyanins. In seeds, leucoanthocyanidins are substituted by leucoanthocyanidin reductase, generating flavan-3-ols that undergo a condensation reaction to produce proanthocyanidin. Leucoanthocyanidin reductase is absent in Arabidopsis seed and therefore, proanthocyanidin is formed by using anthocyanidin as a precursor through one or more enzyme-catalyzed reactions by anthocyanidin reductase (Abrahams *et al.* 2002, Achnine *et al.* 2004). Proanthocyanidins start to accumulate at the micropylar part of the embryo at the very early stage of fertilization (around 1 to 2 days after fertilization), followed by the endothelium and the chalaza end (about 5 to 6 days after fertilization). Oxidation of proanthocyanidin occurs during the seed desiccation period (30 to 40 days after fertilization) leading to brown pigment formation in the mature seed (Lepiniec *et al.* 2006). Yellow seed coat color occurs only when one or more gene(s) encoding different enzymes are mutated in the flavonoid biosynthetic pathway, so that proanthocyanidin fails to deposit in the seed and thus produces a transparent seed coat that makes that yellow embryo visible, resulting in yellow seed color.

2.6. Breeding for yellow seeded *B. napus*

In nature, the seed coat colors of all Brassica species are black/brown, and yellow-seeded mutants have been identified only in *B. rapa, B. juncea* and *B. carinata.* There is no naturally occurring yellow-seeded *B. napus* germplasm available. Therefore, development of yellow seeded *B. napus* has been a prime breeding objective for many years by various researchers around the world. The first yellow seeded *B. napus* cultivar HUA-yellow No.1 was registered in China in 1990 (Liu *et al.* 1991). This yellow seeded line was developed from a cross of yellow seeded *B. rapa* (ssp. *chinensis*) with black seeded *B. napus*, followed by backcross to *B. napus* and a pedigree selection (Liu 1983). However, the seed coat color was frequently affected by environmental influence resulting in black seed and/or black spot on seed coat. Shirzadegan and Röbbelen (1985) developed yellow seeded *B. napus* (winter type) from a cross of a black seeded winter type *B. napus* variety Quinta with a resynthesized yellow-brown seeded *B. napus* line derived from a cross of light yellow-brown seeded *B. oleracea* ssp. *alboglabra* and yellow seeded *B. rapa*. Zaman (1987) developed partially yellow-seeded *B. napus* from interspecific crosses of yellow seeded *B. carinata* with yellow seeded *B. rapa* followed by a backcross to *B. napus*, and from a cross of black seeded *B. napus* with yellow seeded *B. rapa*. He also attempted to develop a yellow seeded CC genome species from a cross of yellow seeded *B. carinata* (BBCC) and black seeded *B. alboglabra* (CC) which could be crossed with yellow seeded *B. rapa* to develop yellow seeded *B. napus.* However, this attempt did not generate a yellow seeded CC genome species. Chen *et al.* (1988) resynthesized *B. napus* from interspecific crosses between yellow-seeded *B. rapa* and yellow- or brown-seeded *B. alboglabra,* and the yellow seeded types were identified from

a segregating population. Chen and Heneen (1992) developed yellow-seeded *B. napus* lines from crosses of resynthesized *B. napus* lines containing the yellow seed coat color genes (Chen *et al.* 1988) with yellow-brown seeded *B. napus*, or with yellow-seeded *B. carinata*. However, the yellow seeded line did not reproduce truly in the following generations. Rashid *et al.* (1994) obtained yellow seeded *B. napus* through interspecific crosses of black seeded *B. napus* with two yellow seeded mustard species, *B. juncea* and *B. carinata.* To eliminate the B genome chromosomes from both hybrids, the F_1 generations were backcrossed to *B. napus.* Further crosses were made between backcross F_2 plants of the (*B. napus* ×*B. juncea*) × *B. napus* cross and backcross F_2 plants of the (*B. napus* ×*B. carinata*) × *B. napus* cross. The F_2 intercrosses population was grown in the field, and ninety-one yellow seeded plants were selected from 4858 plants. Qi *et al.* (1995) obtained yellow-seeded *B. napus* from a cross of yellow seeded *B. carinata* with a partially yellow-seeded *B. napus* line that carried the yellow seed coat color gene from *B. rapa.* Tang *et al*. (1997) developed 16 yellow seeded *B. napus* lines from various interspecific crosses using naturally occurring yellow seeded species, such as (*B. rapa* × *B. oleracea*) × *B. napus*, *B. napus* × *B. juncea*, and *B. napus* × *B. rapa*, intervarietal crosses of *B. napus*, and from irradiated progenies of *B. napus*. Meng *et al.* (1998) developed yellow seeded *B. napus* from interspecific cross of yellow seeded *B. rapa* and yellow seeded *B. carinata*. Trigenomic hexaploid Brassica (2n=54, AABBCC) was generated from the F_1 which was crossed with partial yellow or brown seeded varieties of *B. napus*. Most of the F_1 seeds were self-fertile and generated brown seed color. In the F_2 population, 73 out of 2590 open pollinated plants produced yellow seed coat color. Stable yellow seeded lines were developed from two successive self-pollinated generations of

the selected lines. Rahman (2001b) developed yellow seeded *B. napus* from interspecific crosses between yellow-seeded *B. rapa* var. 'yellow sarson' (AA), black-seeded *B. alboglabra* (CC), yellow-seeded *B. carinata* (BBCC) and resynthesized black-seeded *B. napus* (AACC) derived from a black-seeded *B. alboglabra* and yellow seeded *B. rapa* cross. Three approaches were taken to develop yellow seeded *B. napus* and yellow seeded CC genome species. In the first approach, yellow seeded *B. rapa* was crossed with yellow seeded *B. carinata*, and the trigenomic hybrid (ABC) was crossed with resynthesized black seeded *B. napus*. The hybrid was self-pollinated for number of generations and yielded yellow seeded *B. napus* lines. In the second approach, brown seeded trigenomic hexaploid Brassica (AABBCC) was developed from the trigenomic hybrid (ABC) of a cross of yellow seeded *B. rapa* and yellow seeded *B. carinata*. The hexaploid Brassica was crossed with the resynthesized brown seeded *B. napus.* The new hybrid was self-pollinated for several generations but failed to generate any yellow seeded *B. napus* lines. To develop yellow seeded CC genome species in the third approach, a cross was made between black seeded *B. alboglabra* and yellow seeded *B. carinata* and the F_1 was self-pollinated for several generations. A yellowish-brown seeded *B. alboglabra* (CC) line was developed which was crossed with yellow seeded *B. rapa*, but failed to generate any resynthesized yellow seeded *B. napus* lines. Burbulis and Kott (2005) identified canola type yellow seeded *B. napus* in the DH progenies of a cross between two black seeded and six yellow seeded lines.

2.7. Environmental influence in seed coat coloration

In general, warmer temperatures significantly reduce the color deposition in the seed coat of Brassica. Van Deynze *et al.* (1993) conducted systematic research on the

influence of temperature on seed coat color in the yellow seeded and black seeded lines of *B. napus.* The experiment was conducted in growth cabinets at 16°/12°C, 18°/14°C, 20°/16°C, 22°/18°C and 24º/20º C day/night temperatures with a 16hrs photoperiod (350 µmol $m^{-2}s^{-2}$). Yellow seeded lines had improved yellow seed color with increased temperatures and the reduction of color was linearly correlated to the increased temperatures. For example, two genotypes produced black or dark brown seed coat color at 16º C while the genotypes produced yellow seed coat color at 24º C. These results indicated that high temperature might block the function of genes or enzymes involved in the biosynthesis of pigments in seed. Marles *et al.* (2003) reported that the yellow seeded *B. carinata* produced yellow seed at 25°/20°C day/night temperature, but the seed coat color turned light brown at 18°/15°C. Burbulis and Kott (2005) conducted an experiment with 11 yellow to brown-yellow seeded lines grown at three different temperature conditions, 20°/16°C and 28°/24°C day/night temperatures in growth chambers, and an outside plot with day time temperatures above 30°C from bolting to seed maturity stage. The lines showed an increasing yellowness with higher temperatures, and were darker with cooler growing conditions.

Oil content in seed was not affected in warmer growing conditions, but oleic acid content was increased by 10% and linolenic acid content was decreased by 5% (Burbulis and Kott 2005), which is desired for improvement of seed quality. It has been cited earlier that proanthocyanidins are the final compound in the seed coat coloration process, and dihydroflavonol 4-reductase is one of the enzymes responsible for the biosynthesis of proanthocyanidins from dihydromyricetin (Figure 2.5). Dihydroflavonol 4-reductase transcripts were absent or less abundant in the yellow seeded lines of *B. carinata* when

they were grown at warm temperatures (25/20°C). Cooler (18/15°C) growing temperatures increased the dihydroflavonol 4-reductase expression resulting in deposition of pigments in the seed coat of yellow-seeded lines, indicating that temperature has a direct affect on biosynthesis of some transcripts responsible for seed coat pigmentation in plants (Marles *et al.* 2003).

2.8.1. Seed coat color inheritance in *B. napus*

Seed coat color inheritance in *B. napus* has been investigated by several researchers. Shirzadgen (1986) studied seed coat color inheritance and proposed a three-gene model for seed coat color in *B. napus.* According to his model, (i) yellow seed occurs only when all three loci are in the homozygous recessive condition, (ii) brown seed occurs when a dominant allele is present at either the Bl_2 or Bl_3 locus and the Bl_1 locus is either in homozygous recessive or in heterozygous conditions, and (3) black seed occurs when the Bl_1 is in homozygous dominant condition. Van Deynze and Pauls (1994) showed a similar agreement with a modification in the occurrence of black seed coat color as proposed by Shirzadgen (1986), where black seed occurs when the Bl_1 locus was in homozygous dominant and at least one dominant allele was present at any of the two loci. Henderson and Pauls (1992) confirmed a trigenic inheritance for seed coat color in the F_2 segregation populations of inbred lines (16:47:1) and the DH lines (4:3:1) in *B. napus*. Baetzel *et al.* (1999) observed a trigenic inheritance for seed coat color in the inbred (27:36:1) and DH (1:6:1) lines of *B. napus* and proposed that black seed coat occurred when all genes were in homozygous dominant conditions. Rahman *et al.* (2001) reported three or four gene loci were involved in the determination of seed color, and

yellow seeds were formed when all alleles were in homozygous recessive state. They classified the self-pollinated DH population progeny into five groups including black/dark brown, reddish-brown, partly-yellow, yellow-brown and yellow. Four gene segregation was reported when the first four seed color groups were pooled into one group and the yellow seeded lines were separated from the pool. In contrast, three gene segregation was observed with pooling of the black/dark brown, reddish brown and partly yellow seeds into one group, and those with yellow-brown and yellow seeds into a second group. Lühs *et al.* (2000) reported trigenic inheritance for seed coat color in *B. napus.* Apart from the dominant nature of black seed coat color of *B. napus*, Li *et al.* (2003) reported that different yellow seed coat color genes existed in different yellow seeded lines; and that these genes showed dominance, partial dominance and recessive allelic interactions. They also obtained a completely dominant yellow seeded line in a segregating population. Liu *et al*. (2005) investigated seed color inheritance in F_2, BC_1 and F_1-derived DH progenies and reported that seed coat color was controlled by maternal genotype and that a single gene locus was responsible for the yellow seed trait. This gene was partially dominant over the black seed trait. Liu *et al.* (2005) investigated the seed coat color inheritance in F_2, and BC_1 populations of the two crosses and also in a DH population and reported that three genes were responsible for seed coat color in *B. napus.* They identified a dominant yellow seed color gene that had epistatic effects on the other two independently segregating dominant black-seeded genes.

2.8.2. Seed coat color inheritance in *B. rapa*

Seed coat color inheritance in *B. rapa* has been reported by several researchers. In Brassica species the maternal genotype mainly controls the seed coat color, however, interplay between the maternal parent and the endosperm and/or embryonic genotype also affects the seed coat color (Chen and Heneen 1992, Rahman *et al.* 2001). Mohammad *et al.* (1942) and Jönsson (1975) reported that three independent genes were responsible for seed coat color inheritance where brown seed color was dominant over the yellow color type, and yellow seed occurred only when all three loci were in a homozygous recessive condition. Stringam (1980) observed a digenic inheritance of seed coat color in *B. rapa*, and proposed a model for seed color genes *Br1* and *Br3*. According to the model, presence of dominant alleles at both loci (*Br1-Br3-*) or presence of dominant alleles only at the first locus (*Br1-br3br3*) generated brown seed color; the yellow-brown seed occurred when the second locus contained at least one dominant allele and the first locus contained homozygous recessive alleles (*br1br1Br3-*); presence of homozygous recessive alleles at both loci produced yellow seeds (*br1br1br3br3*). Later, Schwetka (1982), Zaman (1989) and Rahman (2001a) supported this model for seed coat color inheritance in *B. rapa.* Ahmed and Zuberi (1971), Hawk (1982), Chen and Heneen (1992) observed a single gene inheritance pattern in *B. rapa* where the brown seed color was dominant over the yellow seed type. In contrast, Ahmed and Zuberi (1971) and Chen and Heneen (1992) observed single locus inheritance for seed coat color in *B. rapa*. Rahman (2001a) explained this apparent anomaly by suggesting that the second locus for yellow-brown seed color gene was in homozygous recessive (*br3br3*) condition and as a

result the first gene (*Br1*) was segregating as a single gene.

2.9.1. Use of molecular markers

The creation of genetic variation and subsequent selection of desirable genotypes is a key issue in crop improvement and plant breeding. Traditionally, selection of plant materials with desirable traits is carried out using phenotypic selection. Phenotypic traits are rather limited in number and only a limited number of qualitative and/or quantitative genes are responsible for most of the phenotypic variation. Furthermore the expression of genes is often influenced by the environment. In contrast, the development of a molecular marker linked to a trait of interest using different molecular marker techniques allows the monitoring of similarity/dissimilarity among different genotypes at the very early stages of plant development, independent of environmental effects. This knowledge can directly be used in a marker assisted selection program in plant breeding to identify a desirable genotype in a segregating population (F_2 or backcross). Marker assisted selection can speed breeding progress significantly. The use of molecular markers makes it possible to transfer a trait from one genotype to an other with high efficiency, even the traits which are highly influenced by the environment.

2.9.2. Molecular marker techniques

During 1950s, protein markers (biochemical markers) became popular for selection in plant breeding programs (Hunter and Markert 1957). Protein markers are of two types: isozymes and storage proteins, and separated mainly by SDS-PAGE (Sodium Dodecylsulfate-Polyacrylamide Gel Electrophoresis). However, these technologies were

not adequate for marker assisted selection in plant breeding or for construction of high-density genetic maps. More recently, different types of DNA molecular marker systems have been used for marker assisted selection in plant breeding, genetic map construction and map-based gene cloning. These DNA markers have been classified into two groups: (i) hybridization-based markers such as Restriction Fragment Length Polymorphisms (RFLP) (Weber and Helentjaris 1989), Expressed Sequence Tags (ESTs) (Qin *et al*. 2001), and Fluorescence In-situ Hybridization (FISH) (Pinkel *et al*. 1986), and (ii) PCR-based markers including Simple Sequence Repeats (SSR) (Herne *et al*. 1992), Random Amplification Polymorphic DNA (RAPD) (Williams *et al*. 1990), Amplified Fragment Length Polymorphism (AFLP) (Vos *et al*. 1995), Sequence Tagged Site (STS) (Fukuoka *et al*. 1994), Single nucleotide polymorphism (SNP) (Gabor *et al.* 1999), Sequence Related Amplified Polymorphism (SRAP) (Li and Quiros 2001) and others. The principles of these techniques differ and they generate varying amounts of data. RFLPs are the most reliable marker used for accurate scoring of genotypes, are co-dominant and can identify a unique locus, and are widely used for linkage map construction (Song *et al.* 1988). Genetic maps for various crop species including rice have been constructed using RFLP markers. However, RFLP analysis is labor intensive, time consuming, expensive. Analysis of ESTs is highly expensive and is not feasible for marker assisted selection in plant breeding. FISH is being used intensively in clinical practice, but hybridization efficiency is highly variable between laboratories (Babu and Wiktor 1991). Later, with the development of polymerase chain reaction (PCR) based techniques, most of the problems associated with hybridization-based techniques were overcome. SSR is also known as microsatellite, is highly variable to the repeat number, detects very high levels

of polymorphism, is co-dominant, and highly efficient for marker assisted selection in plant breeding. However, the development of SSR markers is expensive, time consuming and known DNA sequences are required. RAPD is very simple, relatively easy to perform and does not need DNA sequence information (Williams *et al.* 1990). However, the RAPD technique is non-reproducible in various laboratories as it is highly influenced by experimental conditions (Devos and Gale 1992, Staub *et al.* 1996). Paran and Michelmore (1993) developed the STS marker from RAPD marker. RAPD fragments are separated on an agarose gel, excise the bands and extract the DNA from gel, re-amplify the gel extracted DNA using the original RAPDs primers. The new amplified DNA is then cloned into a vector for DNA sequencing. New and comparatively longer PCR primers are designed from the sequenced data to amplify the genomic DNA, which showed the same profile to that of RAPD with better performance. Thus the problem associated with RAPD could be solved by using STS markers. However, problems are also associated with the STS markers. The RAPD fragments converted into STS markers may have completely different DNA fragments with similar fragment size in the agarose gel from various positions of the genomic DNA, which limits the efficient of the use of STS markers. With the advent of the AFLP technique, the problems encountered with RFLP, SSR and RAPD have been overcome. But the main disadvantages of this method are that it requires multiple steps, such as digestion of genomic DNA, ligation to the adaptor, amplification, and it is difficult to optimize the conditions for each step. For sequencing of an AFLP specific band, the fragment needs to be cloned into a vector, which is laborious (Li and Quiros 2001). SRAP is a PCR-based two primer-based marker system, a forward primer and a reverse primer with 17 or 18 nucleotides (Li and Quiros

2001). In the primer sequences, the first 10 or 11 bases are randomly selected nucleotides, followed by a CCGG sequence in the forward primer and an AATT sequence in the reverse primer, and three selective nucleotides at the 3´end. The reason of designing these special primers is to target the exons, promoters and introns sites of a gene (Lin *et al*. 1999). The annealing temperature for the first five cycles was set at 35°C that ensured the binding of both primers to many partially matched positions of the targeted DNA. The annealing temperature is then raised to 50°C for 35 cycles, which allowed the five cycled amplified DNA products to generate consistent bands in the rest of the cycles. Therefore, the SRAP marker technique generated a good number of monomorphic and polymorphic bands between the parents. Li and Quiros (2001) observed more than ten polymorphic loci with a single primer combination in two parental lines of *B. oleracea.* These observations allowed using SRAP technique with high efficiency in the development of molecular markers in marker assisted selection in plant breeding.

2.10.1. Molecular markers for seed color genes in *B. napus*

Only a few researchers have developed molecular markers for seed coat color genes in *B. napus*. Van Deynze *et al.* (1995) developed RFLP markers for two of the three seed coat color genes in *B. napus*. Somers *et al*. (2001) identified eight RAPD markers co-segregating with the major yellow seed coat color gene (*pigment1*) in *B. napus* that differentiated over 72% of individuals for seed coat color variation. Liu *et al.* (2005) reported that a single gene locus in the yellow seeded lines was partially dominant over black seeded lines, and two RAPD and eight AFLP markers were developed linked

to the seed coat color genes in *B. napus*. Two of the ten markers were very close (3.9 cM and 2.4 cM) to the seed coat color gene which allowed selection for yellow seeded individuals at an accuracy of 99.91%. However, these RAPD and AFLP markers were not adapted for larger scale marker assisted selection. Therefore, Liu *et al.* (2006) converted these RAPD and AFLP markers into reliable SCAR and cleaved amplified polymorphic sequence (CAPS) markers from the sequence information of the most closely linked 4 AFLP and 2 RAPD markers for seed coat color breeding in *B. napus.*

2.10.2. Molecular markers for seed color genes in other Brassica species

Chen *et al.* (1997) developed a RAPD marker linked to the seed coat color gene located at the terminal region of chromosome 1 of *B. alboglabra* using *B. campestris* - *alboglabra* addition lines. Heneen and Jørgensen (2001) used a brown seeded *B. campestris* - *alboglabra* monosomic additional line that carried chromosome 4 of *B. alboglabra* and reported that the embryo of a brown seeded additional line contained a gene for dark seed coat color. They developed a RAPD marker for the seed coat color trait linked to chromosome 4 of *B. alboglabra*. Negi *et al.* (2000) developed three AFLP markers tightly linked to the brown seed coat color trait in *B. juncea*. The AFLP marker was not suitable for large scale application in marker assisted selection in plant breeding. Therefore, PCR-walking technology was applied to convert the dominant AFLP marker into simple co-dominant SCAR markers, which distinguished the yellow and brown seeded *B. juncea* lines as well as the homozygous and heterozygous brown-seeded individuals. Mahmood *et al.* (2003) constructed an RFLP-based genetic map of *B. juncea* and described the linkage groups of A- and B-genome. On the basis of this genetic map,

later Mahmood *et al.* (2005) identified two QTL linked with two seed coat color genes located on both linkage group specific to A- and B-genome of *B. juncea*. Padmaja *et al.* (2005) developed three microsatellite markers (Ra2-A11, Na10-A08 and Ni4-F11) strongly associated with seed coat color genes. Two (Ra2-A11, Na10-A08) of the three markers were placed on linkage group-1 and the other marker (Ni4-F11) was placed on the linkage group-2. The markers Na10-A08 and Ni4-F11 co-segregated without any recombination with the seed coat color genes *BjSC1* and *BjSC2* in *B. juncea*, respectively.

3.0 HIGH THROUGHPUT GENOME SPECIFIC MOLECULAR MARKERS FOR ERUCIC ACID CONTENT GENES IN *Brassica napus* (L.)

Mukhlesur Rahman, Zudong Sun, Peter B.E. McVetty and Genyi Li

Department of Plant Science, University of Manitoba

Winnipeg, Manitoba, Canada R3T 2N2

Article has been submitted to Theoretical and Applied Genetics journal for publication

Key words: *B. napus*, *Bn-FAE1* gene, molecular marker, SNP, SCAR.

3.0 High throughput genome-specific molecular markers for erucic acid content genes in *Brassica napus* (L.)

3.1 Abstract

A single base change in the *Bn-FAE1.1* gene in the A genome and a two-base deletion in the *Bn-FAE1.2* gene in the C genome produce the nearly zero content of erucic acid observed in canola. A BAC clone anchoring *Bn-FAE1.1* from a *B. rapa* BAC library and a BAC clone anchoring *Bn-FAE1.2* from a *B. oleracea* BAC library were used in this research. After sequencing the gene flanking regions, it was found that the dissimilarity of the flanking sequences of these two *FAE1* homologs facilitated the design of genome specific primers that could amplify the corresponding genome in allotetraploid *B. napus*. The two-base deletion in the C genome gene was detected as a sequence characterized amplified region (SCAR) marker. To increase the throughput, one genome specific primer was labeled with four fluorescence dyes and combined with 20 different primers to produce PCR products with different fragment sizes. Eventually, a super pool of 80 samples was detected simultaneously. This dramatically reduces the cost of marker detection. The single base change in the *Bn-FAE1.1* gene was detected as single nucleotide polymorphic (SNP) marker with an ABI SNaPshot kit. A multiplexing primer set was designed by adding a polyT to the 5' primer end to increase SNP detection throughput through sample pooling. Furthermore, the *Bn-FAE1.1* and *Bn-FAE1.2* were integrated into the N8 and N13 linkage groups of our previously reported high density sequence related amplified polymorphism (SRAP) map, respectively. There were 124 SRAP markers in a N8 bin in which the *Bn-FAE1.1* gene specific SCAR marker was located and 46 SRAP markers in a N13 bin into which the *Bn-FAE1.2* SNP marker was

integrated. These three kinds of high throughput molecular markers have been successfully implemented in our canola/rapeseed breeding programs.

3.2 Introduction

Canola/rapeseed (*Brassica napus* L.) is a major oilseed crop in Canada, Europe, Australia, China and the Indian subcontinent. The quality of canola/rapeseed oil is determined primarily by its constituent fatty acids. The major fatty-acid constituents of Brassica oil are palmitic acid (C16:0), stearic acid (C18:0), oleic (C18:1), linoleic acid (C18:2), linolenic acid (C18:3), arachidic acid (C20:0), eicosenoic acid (C20:1), and erucic acid (C22:1). Erucic acid is one of the main fatty acids in rapeseed oil. Low erucic acid in rapeseed improves the quality of the oil because high erucic acid is relatively low in digestibility and has been associated with health problems (Beare *et al.* 1963). On the other hand, high erucic acid rapeseed (HEAR) has several potential applications in the oleo-chemical industry for the production of high temperature lubricants, nylon, plastics, slip and coating agents, soaps, painting inks, surfactants (Topfer *et al.* 1995).

In *B. napus,* additive alleles at two genes (*Bn-FAE1.1* and *Bn-FAE1.2*) control the erucic acid content in seeds (Harvey and Downey 1964), and these two genes are mapped in rapeseed (Ecke *et al.* 1995, Jourdren *et al.* 1996, Thormann *et al.* 1996). Development of low or high erucic acid content *B. napus* cultivars requires a long process of crossing, backcrossing and selfing of the segregating generations followed by identification of low or high erucic acid content lines from fatty acid profiles of seed lots from individual seeds by gas chromatography (GC). However, DNA molecular markers which are tightly linked to the erucic acid genes or inside the candidate genes can be applied with high

efficiency in marker assisted selection for rapid transfer of a character into an otherwise desirable genotype efficiently and effectively. For instance, 15 of 16 plants in the F_2 generation of canola/rapeseed crosses could be discarded at the seedling stage by selecting homozygous *Bn-FAE1.1* and *Bn-FAE1.2* genotypes with high or low erucic acid content instead of growing plants to full maturity and then doing seed quality analysis and selection.

The pathway for erucic acid biosynthesis and the major reactions involved in this pathway has been well characterized in Arabidopsis. Oleic acid is the main precursor for erucic acid biosynthesis via an elongation process in the developing embryos of *B. napus* (Xiaoming *et al.* 1998). In seeds of *Arabidopsis thaliana* it was reported that fatty acid elongation 1 (*FAE1*) is the candidate gene and that the gene product was essential for elongation of C18:1 (oleic acid) to C22:1 (erucic acid) (Lemieux *et al*. 1990, Kunst *et al*. 1992). The elongation process has four different steps. The first step is the condensation of oleoyl-CoA to malonyl-CoA to form a 3-ketoacyl-CoA. The second step is the reduction of the 3-ketoacyl-CoA to produce 3-hydroxyacyl-CoA. The third step is the dehydration of the 3-hydroxyacyl-CoA to form trans-(2,3)-enoyl-CoA. The final step it is the further reduction of the trans-(2,3)-enoyl-CoA. These reactions are catalyzed by four different enzymes 3-ketoacyl-CoA synthase, 3-ketoacyl-CoA reductase, 3-hydroxyacyl-CoA dehydratase and trans-(2,3)-enoyl-CoA reductase, respectively (Fehling and Mukherjee 1991). The role of the *FAE1* gene in producing erucic acid was genetically ascertained by genetic transformation of a low erucic acid content rapeseed (Lassner *et al.* 1996). In rapeseed, the two loci E1 and E2 of *FAE1* homologs encode the rapeseed 3-ketoacyl-CoA synthases for the elongation process to generate erucic acid from oleoyl-

CoA (Barret *et al*. 1998a, Fourmann *et al.* 1998). In *B. napus*, these two homologs of the *FAE1* gene (*Bn-FAE1.1* and *Bn-FAE1.2*) have been characterized. These two homologs show 99.4% nucleotide identity and a two-base deletion in the low erucic acid content line results in a functional loss of *Bn-FAE1.2* gene in the C genome (Fourmann *et al.* 1998). Katavic *et al*. (2002) reported that the single amino acid serine found at position 282 in the high erucic acid content line is substituted by phenyl-alanine in the low erucic acid content line due to a single base pair change in the *Bn-FAE1.1* gene in the A genome. In this report, BAC clones containing *Bn-FAE1.1* and *Bn-FAE1.2* genes from the A and C genome libraries were used to extend the sequence to the outside of these two genes to develop genome specific high throughput molecular markers. These markers will considerably facilitate the selection of the four different erucic acid content control alleles in canola/rapeseed breeding programs.

3.3 Materials and methods

Plant materials

The pure breeding high erucic acid content cultivar 'MillenniUM 03' (~55% erucic acid) was crossed with the pure breeding canola line 'SRYS-3' (0% erucic acid) and the F_1 was backcrossed to SRYS-3. The F_1, F_2, F_3 and BC_1 generations were grown in the greenhouse. Self-pollinated seeds of 279 F_2 and 159 BC_1 plants were examined for erucic acid content by gas chromatography (GC).

DNA extraction and sequencing of flanking sequence of the *FAE1* genes

DNA was extracted using a modified version of the CTAB method according to Li and Quiros (2001) from the flower buds of *B. rapa, B. oleracea, B. napus.* The

broccoli BAC library and library screening strategy, as described by Quiros *et al*. (2001), was used to pick up the BAC clone anchoring the *Bo-FAE1.2* gene. The *B. rapa* BAC library, constructed from a male sterile line (kindly provided by Dr. Quiros, University of California, Davis) was used to select the BAC clone containing the *Br-FAE1.1* gene. BAC DNA was prepared from the A-genome library and the C-genome library of pooled cultures of all clones from each 384-well plate. Targeted positive clones from positive plates were identified using a column and row (24x16) pooling strategy. The primers were designed according the *FAE1* sequence in the GenBank data (u50771, AF009563, Y14974, Y14975, Y14981). The flanking sequences of the *FAE1* gene in the A-genome and C-genome were sequenced by chromosome-walking with the selected BAC clones. With the extended flanking sequences, the polymorphic regions between the A-genome and C-genome were found and used to design genome specific primers.

Multiplexing SCAR markers by targeting the two-base deletion in the *Bn-FAE1.2* gene

The two-base deletion in the *Bn-FAE1.2* gene is located at the end of this gene. With the extended 3'-end flanking sequence of *Bn-FAE1.1* and *Bn-FAE1.2*, a polymorphic region was found and a genome specific primer FE42A targeting the *Bn-FAE1.2* in the C genome was designed and another 20 primers were located inside the gene to form 20 primer pairs for multiplexing detection (Table 3.1). The genome specific primer was labeled with one of the four fluorescent dyes (6-FAM, VIC, NED and PET, ABI, California) and the PCR products in a range of 200 to 500 bp were separated with an ABI 3100 DNA analyzer (ABI, California).

SNaPshot detection of the *Bn-FAE1.1* alleles

SNP markers were developed through targeting the single base change in the *Bn-FAE1.1* gene in the A genome. A genome specific primers, 'FEAG4 and FEAS1' were designed on the basis of the extended gene flanking sequence to amplify the *Bn-FAE1.1* gene in the A genome. For multiplexing, another 22 primers were designed to detect the single nucleotide polymorphism (SNP) in the *Bn-FAE1.1* by adding polyT at the 5' end of the primers (Table 3.1).

The first step for SNP detection was to produce genome specific PCR products containing the SNP position. PCR reactions were performed in a 10 µl volume containing 60 ng of genomic DNA, 100 µM of each dNTP, 0.15 µM of each primer, 1x PCR buffer, 1.5 mM $MgCl_2$ and 1 Unit of *Taq* polymerase. The PCR program was: 94°C for 3 min. followed by 35 cycles of 94°C for 1.0 min, 55°C for 1.0 min, 72°C for 1.0 min and final extension 72°C for 10 min.

SNP detection was performed using single nucleotide primer extension method in the ABI 3100 Genetic Analyzer using SNaPshot multiplex kit (ABI, California), following the instruction in the kit.

The SNaPshot products were pooled first and 2 µl pooled DNA was mixed with 8 µl formamide containing GeneScan^{TM} 120 LIZ^{TM} size standard (ABI, California). After the DNA was denatured for 5 minutes at 95°C, the DNA fragments were analyzed with an ABI 3100 Genetic Analyzer. Genotypes were scored manually, using peak color verification.

3.4 Results

Study on erucic acid content of *B. napus*

Self-pollinated seeds of 279 F_2 and 159 BC_1 plants from a cross of 'MillenniUM 03' and 'SRYS-3', were examined for erucic acid content. The parental lines, 'MillenniUM 03' and 'SRYS-3', have 56% and 0% erucic acid in seed, respectively. Twenty out of 279 plants produced low erucic content in F_2, and 39 out of 159 BC_1 plants showed low erucic acid content in seed. These results were consistent with a digenic (15:1 for F_2 and 3:1 for BC_1) segregation ratio indicating that two genes are responsible for erucic acid content (Table 3.2).

Gene specific SNP markers for *Bn-FAE1.1* in the A genome in *B. napus*

Gene specific markers for the A-genome (*Bn-FAE1.1* gene) were developed by targeting a transition type base substitution at position 846 (C to T) from the starting codon in low erucic acid content lines. Because of the high level of sequence conservation between *Bn-FAE1.1* and *Bn-FAE1.2*, some primers targeting the sequence dissimilarities between the *Bn-FAE1.1* and *Bn-FAE1.2* coding regions were found to be unreliable to produce genome specific PCR products. In order to design reliable primers for genome specific PCR products, chromosome-walking 2.5 kb upstream and 1.5 kb downstream of the flanking sequence of the *Bn-FAE1.2* gene from the broccoli BAC clone, and 750 bp upstream and 700 bp downstream flanking sequences of the *Bn-FAE1.1* gene from the *B. rapa* BAC clone was conducted. After Clustalw analysis (http://www.ebi.ac.uk/clustalw), dissimilarities between the upstream and downstream flanking sequences of *Bn-FAE1.1* and *Bn-FAE1.2* genes were found (Figure 3.1). On the basis of these dissimilar sequences, one primer FEAG4 in the upstream flanking region

and another primer FEAS1 inside the gene were designed to amplify a 1.2 kb A-genome specific fragment containing the single nucleotide polymorphic (SNP) position. The SNP was detected by using single base extension method with ABI SNaPshot Multiplex kit. The SNaPshot data were analyzed by GeneScan software and peak information was transformed manually for each loci. The polymorphic position containing homozygous 'C' for high erucic acid content exhibited a black peak, homozygous 'T' for low erucic acid content, a red peak, and 'C/T' for heterozygous plants, both a black peak and a red peak (Figure 3.2). Genotypes were assigned accordingly with E1E1 for the black peak, e1e1 for the red peak and E1e1 for both the black peak and the red peak. One hundred and seventy four F_2 plants were tested for the *Bn-FAE1.1* gene using SNaPshot. In this F_2 population, 43 plants exhibited the E1E1 genotype, 81 showed the E1e1 genotype and 50 showed the e1e1 genotype, a good fit to a 1:2:1 genotypic ratio for single locus segregation for the *Bn-FAE1.1* gene ($\chi^2 = 1.385$, P = 0.2 – 0.3). Similarly, in the BC_1 population 49 plants were found to be the E1e1 genotype and 59 were the e1e1 genotype, consistent with the 1:1 expected genotypic segregation for a single locus inheritance for *Bn-FAE1.1* gene ($\chi^2 = 0.926$, P = 0.3 - 0.5).

Gene specific SCAR markers for the *Bn-FAE1.2* gene in the C genome in *B. napus*

Gene specific markers for the *Bn-FAE1.2* gene in the C genome were developed after the genome specific primers were designed. The two-base deletion is located at 1425-1426 bp position for low erucic acid content lines which result in a truncated protein due to an early stop codon (Fourmann *et al.* 1998). Fortunately, this deletion was close to the downstream dissimilarity region where a C genome specific primer FE42A was designed. Twenty different primers within the gene coding region from the upstream

of the polymorphic region (CAAAT/C--AT) were also designed to amplify various fragments of varying from 270 bp to 484 bp in size. The DNA fragment specific for high erucic acid content lines consistently produced two-base larger fragments than those of low erucic acid content lines. This small difference was easily detected by the ABI 3100 Genetic Analyzer when the primer FE42A was labeled with a fluorescent dye. Therefore, the *Bn-FAE1.2* gene for homozygous high erucic acid content lines generated two-base larger fragments, the low erucic acid content lines had two-base smaller fragments and the heterozygous lines had both larger and smaller fragments (Figure 3.3). Since these twenty different primer combinations generated fragments that were different in size, it was possible to pool the PCR products from 20 different primer combinations to achieve multiplexed, high throughput markers. With one fluorescent dye color, 320 DNA samples (16 x 20) were separated in one run in 30 minutes with the ABI 3100 DNA analyzer (Figure 3.4).

A total of 174 F_2 plants were analyzed for the *Bn-FAE1.2* gene in the C genome. Of these, 46 were E2E2 genotype (single upper band for high ER), 45 were e2e2 genotype (single lower band for low ER) and 83 were E2e2 genotype (both upper and lower band for intermediate band). The F_2 population fit a 1:2:1 genotypic ratio for single locus inheritance for *Bn-FAE1.2* for high, intermediate and low erucic acid content alleles ($\chi^2 = 0.378$, P = 0.5 – 0.7). Similarly, 108 BC_1 plants segregated into two groups, 51 for the E2e2 genotype and 57 for the e2e2 genotype which also fit a 1:1 genotypic ratio for single locus segregation for *Bn-FAE1.2* ($\chi^2 = 0.334$, P = 0.5 – 0.7).

Association of *FAE1* molecular markers with erucic acid content

Erucic acid content in the F_2 and BC_1 populations were associated with the two

FAE1 homologs present in the A genome and the C genome. On the basis of SNP and SCAR markers, the F_2 plants were classified into nine different genotypes: E1E1E2E2, E1E1E2e2, E1E1e2e2, E1e1E2E2, E1e1E2e2, E1e1e2e2, e1e1E2E2, e1e1E2e2 and e1e1e2e2 for *Bn-FAE1.1* and *BN-FAE1.2* genes in the A genome and the C genome, respectively (Figures 3.2 and 3.4, and Table 3.3). A total of 174 F_2 plants were tested for both genes, of which 10 were E1E1E2E2, 23 E1E1E2e2, 11 E1E1e2e2, 24 E1e1E2E2, 37 E1e1E2e2, 19 E1e1e2e2, 11 e1e1E2E2, 24 e1e1E2e2 and 15 e1e1e2e2 (Table 3.3). The segregation ratio was consistent with a 1:2:1:2:4:2:1:2:1 genotypic segregation ratio for two genes ($\chi^2 = 2.091$, P = >0.95). All genotypic classes were tested for erucic acid content in seeds and it was found that all four dominant alleles contributed to the highest erucic acid content with an average of 41.2%, three dominant alleles contributed an average of 34.4% (average E1E1E2 = 34.5% and average E1E2E2 = 34.3%), two dominant alleles contributed an average of 23.7% (E1E1 = 24.2%, E1E2 = 24.1% and E2E2 = 22.9%), one dominant allele contributed an average of 12.5% (E1 = 12.0% and E2 = 13.0%) while e1e1e2e2 plants had an average 0.26% erucic acid content (Table 3.3). The results indicated that all four alleles had almost identical contribution to erucic acid content in seed with an average of 11.2% erucic acid per allele.

A further study was conducted to test the association of the two genes with erucic acid content in 108 BC_1 plants developed from a cross of 'MillenniUM 03' and 'SRYS-3', with 'SRYS-3' (low ER) as recurrent parent. The genotyping was done on the basis of SNP detection and two bases deletion detection in ABI 3100 genetic analyzer with four groups obtained from the BC_1 population as E1e1E2e2, E1e1e2e2, e1e1E2e2 and e1e1e2e2 (Figures 3.2 and 3.4, and Table 3.3). Of the 108 BC_1 plants, 26 were E1e1E2e2

with an average 19.8% erucic acid content, 23 E1e1e2e2 with and average 10.5% erucic acid content, 22 e1e1E2e2 with an average 10.4% erucic acid content and 35 e1e1e2e2 with an average 0.18% erucic acid content. The four different genotypic classes fit a 1:1:1:1 backcross segregation ratio for two genes ($\chi^2 = 3.373$, P = 0.3 – 0.5). The contributions of each allele for erucic acid content in seed in the BC_1 population were similar to each other at 10.2%.

3.5 Discussion

In rapeseed, the wild type genotypes produce high erucic acid content in the seed. Low erucic acid content genotypes possess mutations affecting both the E1 and E2 loci controlling erucic acid content in the seed (Harvey and Downey 1964). The fatty acid elongation1 (*FAE1*) gene encoding 3-ketoacyl-CoA synthase has been characterized in *Arabidopsis thaliana* (James *et al.* 1995) and jojoba (Lassner *et al.* 1996). Barret *et al.* (1998a) identified the *FAE1* gene in *Brassica napus* with two homologous sequences (*Bn-FAE1.1* and *Bn-FAE1.2*) from the embryo. The *Bn-FAE1.1* gene has been shown to be linked to the E1 locus for *B. rapa*, and *Bn-FAE1.2* was assigned to the E2 locus in *B. oleracea*. Only three nucleotide changes, a single base change in *Bn-FAE1.1* and a two-base deletion in *Bn-FAE1.2* account for the fatty acid difference of canola and rapeseed (Barret *et al.* 1998a, Fourmann *et al.* 1998, Katavic *et al.* 2002). By targeting these three base changes, high throughput gene-specific markers for both *BnFAE1.1* and *Bn-FAE1.2* were developed. These markers are perfect matches with erucic acid content and no recombination occurs between the markers and trait.

Sequence similarity among the homologous genes in the A and C genes in *B. napus* is very high for most genes even in intergenic regions, which is an obstacle for developing gene specific markers. The *Bn-FAE1.1* gene in the A genome and the *Bn-FAE1.2* gene in the C genome have similar sequences. Since the sequence similarity of these two genes is very high, it is difficult to amplify the SNP position for low ER in the A genome, because the same gene sequence is present in the C genome for high erucic acid content. Similarly, it is difficult to separately amplify the two-base deletion mutation for low erucic content genotypes in the C genome, because the same gene sequence is present in the A genome for high erucic acid which hinders the identification of genotypes with homozygous deletion alleles. To overcome these problems, the extended sequence from the flanking regions of *Bn-FAE1.1* in the A genome and *Bn-FAE1.2* in the C genome were determined. Sequence dissimilarities outside the genes regions in the A and C genomes were large enough for genome-specific primers to amplify the A genome specific *Bn-FAE1.1* containing a SNP position and the C genome specific *Bn-FAE1.2* containing a two-base deletion position. This is the first reported approach to amplify the A and C genome specific *FAE1* genes in *B. napus* for precise tagging of the high and low erucic acid content genotypes in *B. napus*.

Increasing throughput and reducing the cost for molecular marker detection is critical for large scale application of marker assisted selection in canola/rapeseed breeding programs since large number of plants have to be screened in a short time. In this report, multiplexed SNP detection and multiplexed SCAR markers were demonstrated using a series of SNP detection primers varying in size by adding polyT at the primer 5' end. Twenty two SNP detection primers allow a mixture of 22 reactions in

one well. In theory, 16896 samples (22 x 16 x 48) can be done with a medium throughput DNA analyzer, such as ABI 3100 DNA analyzer, within 24 hrs. Similarly with an ABI 3100 Genetic analyzer, over sixty thousand (4 x 20 x 16 x 48) samples can be detected with the SCAR markers for *Bn-FAE1.2* within 24 hrs since the genome specific primer labeled with four fluorescent dye colors can be used in our laboratory. Using this approach it is possible to pool 80 PCR products for detection in an ABI genetic analyzer which could reduce running time and cost by approximately 80 times.

The ABI 3100 genetic analyzer can be used with high efficiency to detect two-base difference markers. The two base deletion in the C genome could also be detected using SNaPshot, but using a SCAR marker is much cheaper than using SNaPshot since only a small amount of specific PCR product is required for detection, Additionally it is a one step PCR with a short PCR amplification program, about 15-20 cycles. Approximately 60,000 samples could be done with a medium throughput DNA analyzer and 360,000 samples could be done with a high throughput DNA analyzer, such as ABI 3700 genetic analyzer, within 24 hrs, with a per sample cost of just a few cents instead of more than one dollar.

Table 3.1 List of primers used for identification of erucic acid content genes in *B. napus*.

Primers amplifying the A-genome specific fragment containing SNP position.
Forward: FEAG4: CTCATTCCCGAGAAACACTGA
Reverse: FEAS1: CTACGATCTCCAGGCTTGTT

SNP detection primers:
FEAF1: TTTTTGCCGCTATTTTGCTCT
FEAF2: TTTTTTTTTTGCCGCTATTTTGCTCT
FEAF3: TTTTTTTTTTTTTTTGCCGCTATTTTGCTCT
FEAF4: TTTTTTTTTTTTTTTTTTTTGCCGCTATTTTGCTCT
FEAF5: TTTTTTTTTTTTTTTTTTTTTTTTTGCCGCTATTTTGCTCT
FEAF6: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCCGCTATTTTGCTCT
FEAF7: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCCGCTATTTTGCTCT
FEAF8: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCCGCTATTTTGCTCT
FEAF9: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCCGCTATTTTGCTCT
FEAF10: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCCGCTATTTTGCTCT
FEAF11: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCCGCTATTTTGCTCT
FEAF12: TTTTTTCGATCTCCAGGCTTGTTG
FEAF13: TTTTTTTTTTTCGATCTCCAGGCTTGTTG
FEAF14: TTTTTTTTTTTTTTTCGATCTCCAGGCTTGTTG
FEAF15: TTTTTTTTTTTTTTTTTTTCGATCTCCAGGCTTGTTG
FEAF16: TTTTTTTTTTTTTTTTTTTTTTTTCGATCTCCAGGCTTGTTG
FEAF17: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGATCTCCAGGCTTGTTG
FEAF18: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGATCTCCAGGCTTGTTG
FEAF19: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGATCTCCAGGCTTGTTG
FEAF20: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGATCTCCAGGCTTGTTG
FEAF21: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGATCTCCAGGCTTGTTG
FEAF22: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGATCTCCAGGCTTGTTG

List of SCAR primers specific to C-genome amplify the two bases deletion position:
Reverse: FE42A (labeled): GACCATCTTTAACCCTAAAACC
Forward: FE42D: CAATGTCAAAGCTTCAA
Forward: FE42D1: GGCTCTAAACAATGTCAAAGC
Forward: FE42E1: TGCAGTTTGGGTGGCTCT
Forward: FE42F: GTGTAACAGTGCAGTTTGGG
Forward: FE42F1: GGCTTTAAGTGTAACAGTGC
Forward: FE42G1: TTAGGGTCAGGCTTTAAG
Forward: FE42J: GGTAATAAAGTTTGGCAG
Forward: FE42K: GCAAAAGGAAGGATGAAG
Forward: FE41: GGCATACATAGAAGCAAAAG
Forward: FE42L: TGGTATGAGTTGGCATAC
Forward: FE42M1: CTAGCTCAATATGGTATGAG
Forward: FE42M: GGAAACACTTCATCTAGCTC
Forward: FE42N1: CATAGATTTGGAAACACTTC
Forward: FE42N2: CAACGTTACATAGATTTGG
Forward: FE42N: GGCATCAAGATCAACGTTAC
Forward: FE42O: TCGATGTAGAGGCATCAAG
Forward: FE42P: CCTAGCACCGATCGATGTAG
Forward: FE42Q: AGAACCTAGGCCTAGCACCG
Forward: FE42R: GTGCTAGAGAAGAACCTAGG
Forward: FE42S: AGCCGTGATTGATGTGCTAG

Table 3.2 Segregation for erucic acid in the F_2 [MillenniUM 03 x SRYS-3] and BC_1 [(MillenniUM 03 x SRYS-3) x SRYS-3] generations of *B. napus.*

Generation	Segregation ratio	Erucic acid content		χ^2	P
		7-45%	0-1%		
F_2	15 : 1	259	20	0.401	0.5 – 0.7
BC_1	3 : 1	120	39	0.019	0.7 – 0.9

Table 3.3 Erucic acid content in different genotypic class of the F_2 [MillenniUM 03 x SRYS-3] and BC_1 [(MillenniUM 03 x SRYS-3) x SRYS-3] populations. The genotyping were done on the basis of SNP and SCAR markers detected in ABI 3100 Genetic Analyzer as shown in Figure 3.2 and 3.3.

Generation / Genotypes	Expected genotypic ratio	Number of individuals	Mean Erucic acid (%) and SD*
F_2			
E1E1E2E2	1	10	41.2 ± 0.83
E1E1E2e2	2	22	34.5 ± 2.21
E1E1e2e2	1	11	24.2 ± 1.88
E1e1E2E2	2	23	34.3 ± 1.94
E1e1E2e2	4	37	24.1 ± 2.31
E1e1e2e2	2	19	12.8 ± 2.00
e1e1E2E2	1	11	22.9 ± 1.73
e1e1E2e2	2	24	12.0 ± 1.91
e1e1e2e2	1	13	0.26 ± 0.29
χ^2 (F_2) = 2.071 (P = > 0.95)			
BC_1			
E1e1E2e2	1	26	19.8 ± 1.90
E1e1e2e2	1	23	10.5 ± 1.61
e1e1E2e2	1	22	10.4 ± 1.11
e1e1e2e2	1	32	0.18 ± 0.22
χ^2 (BC_1) = 2.359 (P = 0.5 - 0.7)			

* Standard deviation

A. Upstream flanking region

```
FAEA    AAATTATCTTATTCCGGTCATGTTGGCCTTAACTAAAC TTCCACACATTTGTTTA- CTGA-
FAEC    ATGTGGTGACACGT-GGTTTGAAACCCACCAAATAATCGATCACAAAAAACCTAAGTTAA
          *   *   *     *     ** *           *     * *  * **  *     * * * *  *        ** * *  *

FAEA    TATTCG AGTATAAACTTTGCGGGAAAACTCATTCCCGAGAAACACTGATCCCATAATTAG
FAEC    GGATCGGTAATAACCTTT CTAATTAAT TTTGATTTAATT AAA-  TCACTCTTTT TA TTTA
            * **     * * * *  * * *         * *   *     *         * * *    *            * *  * * *

FAEA    TCAGAGTCTATGTCGGTTTAGCCTATC- ACTG- CTAAGTACAAAAATTCTCTCGAATTCA
FAEC    T- AAACCCCACTAAATTATGCGATATTGATTGTCTAAGTACAAAAATTCTCTCGAATTCA
          * *  *   *  *         * *     * * *   *  * *  ** ** * ** * * * * * * * * * * ** * * ** * * * * *
```

B. Downstream flanking region

```
FAEA    GTATGTATCTCTTA - - -  - - -  -TTTA - - CATGAAA - -  -TTTTTAAACGCCT - - -  -AACACT
FAEC    GTATGTATCTCTTAGACCATCTTTAACCCTAAAACCCTTTTTGAGATCCTTAAGAAAATT
          ** * * ** ** * * * *          * * * *   *  *  * * *    * * ***  *    * * *      * *  *  *

FAEA    ATGACGTCATATCCTAAATCAAACAA - -CTAAAACCCCGAGTT -  - - -  -CTATCATTTAT
FAEC    TTAATAATAAATTAAACCTTAAGTGGGGTTAAGGATCTC TGTTAAGAAACTTCCATTTTT
          *  *      *  * *    *    * * *          * * *      *     * **         * *   * * * * *  *
```

Figure 3.1 The upstream and downstream sequence comparison of the *Bn-FAE1.1* and the *Bn-FAE1.2* genes in the A-genome, *B. rapa,* and C-genome *B. oleracea*, respectively.

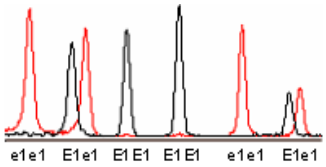


Figure 3.2 Showing multiplexed SNP detection by GeneScan software in ABI 3100 genetic analyzer to analyze the SNaPshot data. The peak information has been transformed manually for each locus [e.g. black for ‘C’ and the genotype is E1E1; red for ‘T’ and genotype is e1e1; and black/red for ‘C/T’ and genotype E1e1].

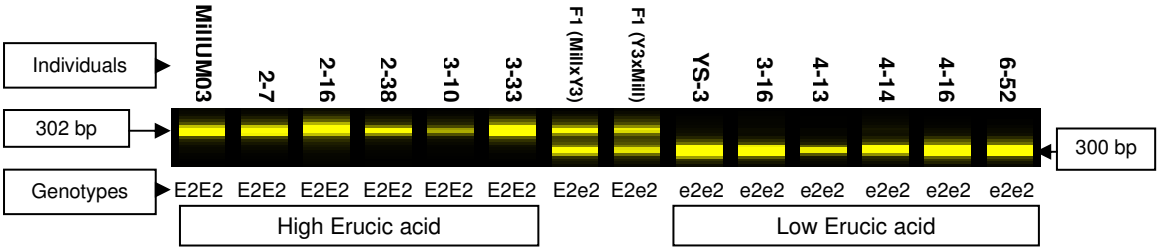


Figure 3.3 Detection of C-genome specific High ER, F_1, F_1´ and low ER lines using C-genome specific primers FE42A/FE42F. The upper band is responsible for high ER (E2E2), lower band is for low ER (e2e2) and double band is for intermediate ER (E2e2).

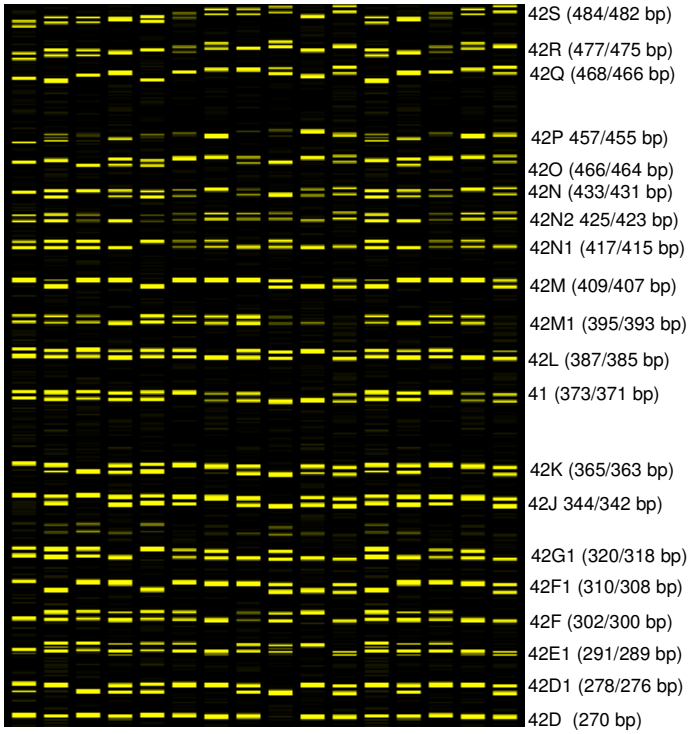


Figure 3.4 Identification of different erucic acid content segregating populations using labeled FE42A primer with twenty different upstream unlabelled primers to produce various DNA fragments from 270bp to 484bp. The upper band is two bases larger than the lower band for all pairs.

4.0 INHERITANCE OF SEED COAT COLOR GENES OF *Brassica napus* (L.) AND TAGGING THE GENES USING SRAP, SCAR AND SNP MOLECULAR MARKERS

Mukhlesur Rahman, Genyi Li and Peter B.E. McVetty

Department of Plant Science, University of Manitoba

Winnipeg, Manitoba, Canada R3T 2N2

Article has been submitted to Theoretical and Applied Genetics journal for publication

Key words: Seed coat color, molecular marker, SRAP, SCAR, SNP, *Brassica napus*

4.0 Inheritance of seed coat color genes of *Brassica napus* (L.) and tagging the genes using SRAP, SCAR and SNP molecular markers

4.1 Abstract

Inheritance of seed coat color in *Brassica napus* was studied in F_1, F_2, F_3 and backcross progenies from crosses of five black seeded varieties/lines to three pure breeding yellow seeded lines. Seed coat color in *B. napus* displayed primarily maternal inheritance, but a pollen effect was also found when yellow seeded lines were used as the female parent. Seed coat color varied from black to dark brown, brown, dark yellow, light yellow and bright yellow. Seed coat color was found to be controlled by three genes, with black color alleles dominant over yellow color alleles at all three loci. Sequenced Related Amplified Polymorphism (SRAP) was employed to identify molecular markers linked to the seed coat color genes. One tightly linked SRAP marker (SA12BG18388) linked to the black/brown seed coat color gene was identified in the F_3 and backcross populations. This SRAP marker was converted into a sequence-characterized amplification region (SCAR) markers using chromosome-walking technology. A second SRAP marker (SA7BG29245), very close to another black/brown seed coat color gene, was identified from a high density genetic map developed in our laboratory using primer walking from an anchoring marker. The marker was located on the linkage group N13 of the C-genome of *B. napus.* This marker also co-segregated with the black/brown seed coat color gene in *B. rapa.* Based on the sequence information of the flanking sequences, 24 single nucleotide polymorphisms (SNPs) were identified between the yellow seeded and black/brown seeded lines. SNP detection and genotyping clearly differentiated the black/brown seeded plants from dark/light/bright yellow-seeded plants and also

differentiated between homozygous (Y2Y2) and heterozygous (Y2y2) black/brown seeded plants. The third seed coat color gene was responsible for dark/light yellow seed in *B. napus*. A total of 768 SRAP primer pair combinations were screened in dark/light yellow seed coat color plants, and a close marker (DC1GA27197) linked to the dark/light seed coat color gene was developed. These three markers linked to the three different yellow seed coat color genes can be used to screen for yellow seeded lines in canola/rapeseed breeding programs. They will also be useful for map based gene cloning of the yellow seed color genes.

4.2 Introduction

Yellow seed coat color is desirable in *B. napus* because yellow-seeded lines have an advantage over the dark-seeded ones since they have a thinner seed coat resulting in 5 to 7% more oil content in the seed and increased value (Stringam *et al*. 1974, Shirazdegan and Robellen 1985, Liu *et al.* 1991). In addition, yellow seeded Brassica lines contain lower dietary fiber, higher protein and more digestible energy in the meal therefore they significantly increase the feed value for animal feed (Bell 1993). A feeding study conducted by Slominski *et al.* (1999) reported that the seed meal from yellow seeded Brassica lines provided higher metabolized energy compared to dark seeded lines for poultry. Therefore, development of yellow seeded lines/cultivars is one of the major breeding objectives in canola/rapeseed.

The inheritance of seed coat color in *B. napus* has been studied by several researchers. Due to allotetraploidy, multiple gene inheritance, maternal effects and environmental effects, the inheritance of seed coat color is quite complex (Van Deynze *et*

al. 1993, Rashid *et al.* 1994, Tang *et al.* 1997, Meng *et al.* 1998). In most cases, it has been reported that three genes are responsible for seed coat color in *B. napus.* Shirzadegan (1986) proposed a three-gene model in *B. napus* where yellow seed occurs only when all three loci are in a homozygous recessive condition; brown seed occurs when a dominant allele occurs at either the Bl_2 or Bl_3 locus and the Bl_1 locus is in either homozygous recessive or in heterozygous condition; and black seed occurs when the Bl_1 is in homozygous dominant condition. Trigenic inheritance for seed coat color in *B. napus* was also proposed by Van Deynze and Pauls (1994) who reported that black seed was formed when the 'A-locus' was homozygous dominant and at least one dominant gene was present at the 'B-locus'; brown seeds developed when one or more dominant genes were present at any of the three loci with at least one recessive gene present at the 'A-locus'; yellow seed coat color resulted when all three loci were in homozygous recessive condition. Henderson and Pauls (1992), Lühs *et al*. (2000), and Baetzel *et al*. (1999) also observed trigenic inheritance for seed coat color in *B. napus.* Rahman *et al.* (2001) reported that three to four gene loci were involved in the determination of seed color in DH populations of *B. napus.* Li *et al.* (2003) obtained completely different seed coat segregation in *B. napus* and reported that different yellow seed coat color genes existed in different yellow seeded lines and that these genes showed dominance, partial dominance and recessive allelic interactions. Liu *et al*. (2005) investigated seed color inheritance in the F_2, BC_1 and F_1-derived DH progenies and reported that seed coat color was controlled by maternal genotype and that a single gene locus, responsible for the yellow seed trait, was partially dominant over the black seed color trait.

Efficient selection of desirable traits is a key concern for crop improvement. Traditionally, phenotypic analysis of plants is used for the selection of desirable plants. However, phenotypic traits are rather limited in number and only a limited number of qualitative and/or quantitative genes are responsible for most of the phenotypic variation observed. Furthermore, expression of genes is often influenced by environment. In contrast, development of molecular markers linked to a trait of interest by use of different molecular-marker techniques allows monitoring similarity/dissimilarity between different genotypes at the very early stages of plant development. This knowledge can be used directly in marker assisted selection breeding to identify recurrent-parent-like genotypes from a segregating population (F_2 or backcross) and further backcrossing of these genotypes with the recurrent parent can speed up the breeding program significantly.

Different DNA molecular markers systems which generate varying amounts of information have been used for marker assisted selection in plant breeding. Some of these are Restriction Fragment Length Polymorphisms (RFLP) (Weber and Helentjaris 1989), Simple Sequence Repeats (SSR) (Herne *et al*. 1992), Random Amplification Polymorphism DNA (RAPD) (Williams *et al*. 1990), Amplified Fragment Length Polymorphism (AFLP) (Vos *et al*. 1995) and Sequence Related Amplified Polymorphism (SRAP) (Li and Quiros 2001). SRAP is simple, reliable and permits easy sequencing of the selected bands, targets coding sequences in the genome and results in a moderate to large number of dominant markers (Li and Quiros 2001). Therefore, the SRAP marker technique was used in this study for the development of DNA molecular markers linked to the seed coat color genes in *B. napus*.

Molecular markers for seed coat color genes in *B. napus* have been reported by only a few researchers. Van Deynze *et al.* (1995) identified two RFLP markers for two of three seed coat color genes using a double haploid (DH) population. Similarly, Somers *et al.* (2001) reported eight RAPD markers for single major gene (*pigment1*), and one and two RAPD markers for two additional genes for seed coat color, respectively. Liu *et al.* (2005) developed 2 RAPD and 8 AFLP markers for a single seed coat color gene linked to the yellow seed trait. Later, Liu *et al.* (2006) converted these RAPD and AFLP markers into reliable sequenced characterized amplified region (SCAR) and cleaved amplified polymorphic sequence (CAPS) markers for marker assisted selection in *B. napus* breeding.

The objectives of this study were to determine the inheritance of the yellow/black seed coat color genes in *B. napus* using five different cross combinations; to develop the SRAP molecular markers for individual seed coat color genes from this multi-gene family; and to convert the SRAP markers into SCAR markers or SNP markers for marker assisted selection in yellow seeded *B. napus* breeding.

4.3 Materials and methods

Plant materials

Three pure breeding spring type yellow seeded *B. napus* lines (SRYS-1, SRYS-2, SRYS-3) developed by Rahman (2001b) were obtained from Norddeutsche Pflanzenzucht (NPZ) Lembke, Hohenlieth, Germany, and four different pure breeding black-seeded spring type varieties, Sentry (canola type), MillenniUM 03 (high erucic acid rapeseed), Holly 276 (high oleic and low linolenic acid content), Allons (low linolenic

acid content) and one black seeded DH (double haploid) line DHBS126, all developed at the Department of Plant Science, University of Manitoba, were used in this research. The F_1, F_2 and F_3 plants were grown in a green house at the Department of Plant Science, University of Manitoba. Backcross seeds were produced by crossing the F_1 plant with the respective yellow seeded parental line. Inheritance of seed coat color was studied in the F_2 (i.e. in F_3 seeds) and BC_1 populations. Chi-square (χ^2) goodness of fit tests were used to check expected versus observed phenotypic segregation ratios for F_2 and BC_1 data. Prior to pooling data for the five crosses used in this study, χ^2 homogeneity tests were done to confirm that the data could be validly pooled (Strickberger 1976).

DNA extraction and SRAP molecular marker

DNA was extracted into 5ml Eppendorf tubes using a modified version of the CTAB method according to Li and Quiros (2001) from the flower buds of the parental cultivars or lines and their segregating populations. The quantity of the genomic DNA was assessed on ethidium bromide-stained agarose gels by visual comparison with a DNA standard. SRAP is a PCR-based marker system with two primers, a forward primer and a reverse primer each with 17 to 22 nucleotides. One primer is fluorescently labeled for detection in an ABI 3100 Genetic Analyzer (ABI, California). For multiplexing the PCR products, a four fluorescent dye set including, 6-FAM (blue), VIC (green), NET (yellow), PET (red) was used to label the SRAP primers, and LIZ (orange) color was used as the standard. All five fluorescent dyes were supplied by ABI (ABI, California). PCR amplification for SRAP was performed as described by Li and Quiros (2001).

Chromosome-walking and DNA sequencing

Siebert *et al.* (1995) have described a chromosome-walking method to generate flanking sequence adjacent to an original DNA sequence and the method was commercialized by Clontech Laboratories (California, USA). The GenomeWalker™ Universal Kit (Clontech, California) was used to obtain extended genomic sequences to the seed coat color marker sequence. Genomic libraries were constructed from DHBS126 (black seeded) parental line by individually digesting 25µg of genomic DNA by using the restriction enzymes *DraI, EcoRV, PvuII* and *StuI*. The digested DNA was purified and GenomeWalker Adaptor was ligated to the genomic DNA according to the protocol provided in the Clontech kit. Two rounds of PCR amplification are necessary in the Chromosome-walking steps. The second round PCR amplification generated sharp and strong bands. These bands were excised from an agarose gel and DNA was extracted using a Qiagen Gel Extraction kit according to the kit protocol. All the DNA fragments were sequenced using a BigDye®Terminator v1.1 Cycle Sequencing Kit (ABI, California).

PCR primer design

The SCAR, SNP and chromosome-walking primers were designed using the Primer3 (http://biotools.umassmed.edu/bioapps/ primer3_www.cgi; University of Massachusetts Medical School, Worcester, Massachusetts) software.

SNaPshot

SNP detection and genotyping were done using the single nucleotide primer extension method (SNaPshot) on the ABI 3100 Genetic Analyzer using a SNaPshot multiplex kit (ABI, Toronto). Genomic DNA containing the SNP position was amplified

with specific primers [MR54: GAGAATTGAGAGACAAAGC and MR13: TGCTCGTTCTTGACAACAC] targeting the corresponding SNP mutations. PCR amplification was performed in a 10 µl reaction mixture containing 60 ng of genomic DNA, 0.375 µM dNTP, 0.15 µM of each primer, 1x PCR buffer, 1.5 mM $MgCl_2$ and 1 unit *Taq* polymerase. PCR conditions for amplification were 94°C for 3 min, followed by 35 cycles of 94°C for 1.0 min, 55°C for 1.0 min, 72°C for 1.0 min and final extension 72°C for 10 min. SNP detection was performed using a SNP detection primer (MSNP1: GTGGTTGAGCGCTCAGTTGCA) and a SNaPshot multiplex kit according to the protocol in the ABI kit. All four ddNTPs were fluorescently labeled with a different color dye i.e. the nucleotide 'C' was black, 'T' was red, 'G' was blue and 'A' was green. The SNP products were separated in 4% ABI polyacrylamide gel with SNP36_POP4 run module and analyzed using ABI GeneScan software. The alleles of a single marker were identified and genotypes were scored manually by observing the different fluorescence color peaks.

SCAR detection

The SRAP marker was sequenced and the flanking sequences were obtained using Chromosome-walking technique. The forward SCAR primer MR47: TGAACTGTGGAAGCCAAGC and the reverse SCAR primer MR32: TCACCACTACGCGGTAACTG were designed from genomic DNA sequence, which amplified the black/brown seeded lines. PCR reactions were carried out as follows: each 10 µl reaction mixture contained 60 ng of genomic DNA, 0.375 µM dNTP, 0.15 µM of each primer, 1x of the PCR buffer, 1.5 mM $MgCl_2$ and 1 Unit *Taq* polymerase. The thermocycler program for amplification was: one cycle for 3 min at 94°C; 35 cycles of

1.0 min at 94°C, 1.0 min at 55°C, 1.0 min at 72°C; one cycle of 10 min at 72°C. The SCAR fragments were visualized in 1.5% agarose gel with ethidium bromide (10 µl EtBr/200ml 1x TAE buffer) for 20 min, and photographed under UV light. Marker segregation versus the seed coat color gene was tested for goodness of fit (χ^2 test) for selected Mendelian segregation ratios.

Genetic map construction

The polymorphic markers linked to the seed coat color trait were scored as one for presence and zero for absence and analyzed using MAPMAKER/EXP ver. 3.0 (Lander *et al.* 1987; Lincoln *et al.* 1992). The genetic map was constructed and map order was determined by three-point analysis with a minimum LOD threshold of 4.0. Map distance was converted to centiMorgans (cM) using the Kosambi function (Kosambi 1944).

4.4 Results

Seed coat color in F_1

Since the seed color in Brassica is determined primarily by the genotype of maternal tissue, i.e. the testa, all F_1 seeds exhibited the same color as the female parent when black seeded parents were used as maternal parent. However, when the yellow seeded lines were used as maternal parent and the black seeded parents were used as pollen source, the F_1 seed coat color turned into dull yellow or yellowish brown (Figure 4.1).

Seed coat color inheritance in *B. napus*

Inheritance of seed coat color was studied in the 1105 F_2 and 570 BC_1 plants using five different cross combinations (‘DHBS-126’ x ‘SRYS-1’; ‘Sentry’ x ‘SRYS-2’;

'Holly-276' x 'SRYS-2'; 'MillenniUM 03' x 'SRYS-3'; and 'Allons' x 'SRYS-3') of black seeded and yellow seeded lines. F_2 progenies varied from black to dark brown, brown, dark yellow, light yellow and bright yellow (Figure 4.2), indicating an incomplete dominance of the black color (Shirzadegan 1986). The yellow seeded F_2 progenies were pooled in one group and all other types of seeds were pooled in another group. The test of homogeneity for five crosses on F_2 data (homogeneity $\chi^2 = 3.358$, P = 0.3-0.5) indicated that the crosses could be pooled. The χ^2 goodness of fit test for the pooled data ($\chi^2 = 1.071$, P = 0.2-0.3) was consistent with a Mendelian 63:1 phenotypic segregation ratio for pooled seed coat color : yellow seed coat color. The homogeneity (homogeneity $\chi^2 = 3.049$, P = 0.5-0.7) for five backcrosses data also showed that the crosses could be pooled. The χ^2 goodness of fit test for the pooled data ($\chi^2 = 1.191$, P = 0.2-0.3) of the BC_1 progenies fit a 7:1 phenotypic segregation ratio for pooled seed color : yellow seed coat color (Table 4.1). These results indicated that three gene loci are responsible for yellow/black seed coat color of *B. napus.*

Molecular markers linked to black/dark brown/brown seed coat color genes

A total of four hundred different random primer pair combinations were used in *B. napus*, of which 26 primer pairs produced markers closely co-segregating with the black/brown seed coat color genes. A preliminary genetic map was constructed from 26 markers using MAPMAKER/EXP ver. 3.0 (Lander *et al.* 1987), where 2 markers were found to be linked in group-1 and 24 markers were linked in group-2.

One very closely linked marker (SA12BG18388) (Figure 4.3) was identified from group-1 and tested with 192 black-seeded and 256 yellow-seeded lines obtained from five different crosses. This confirmed that it was very close to the black/brown seed coat color

gene. This marker was also tested with 248 F_2 plants from the cross of MillenniUM 03 x SRYS-3, in which 194 were found to be linked to the black/brown seed coat color and 54 segregated with dark/light/bright yellow seed coat color, showing a single Mendelian gene segregation ratio 3:1 (χ^2 =1.376, P=0.2-0.3). Similarly, in 180 BC_1 plants, the marker segregated in a 1:1 segregation ratio for brown seed coat color and dark/light yellow/bright yellow seed coat color (χ^2 =0.554, P=0.3-0.5). A total of 248 F_2 plants contributed 496 alleles of which 9 recombinant alleles were identified producing a genetic distance of 1.81 cM from the molecular marker SA12BG18388 to the seed coat color gene. Similarly, in 180 BC_1 plants a total of 180 F_1-derived alleles generated 7 recombinant alleles producing a genetic distance of 3.88 cM between the molecular marker SA12BG18388 and the seed coat color gene. Sequence analysis of the SRAP marker generated a 334 bp fragment, which was extended to 1282 bp sequences (GenBank accession number EF608928) using the chromosome-walking technique (Figure 4.4). Different primers were designed from various positions in the flanking sequence. No primer pairs amplified the yellow-seeded lines. The SCAR marker was screened with 188 F_2 and 173 BC_1 plants of the MillenniUM 03 x SRYS-3 cross, and showed a similar banding pattern to that for the SRAP marker (Figure 4.5).

Twenty four markers were linked in group-2, of which one marker [212AB149 (ODD3PM17-149)] was anchored on the linkage group N13 in the high density genetic map developed in the Dr. Li's lab (Sun *et al.* 2007). A primer walking approach on the genetic map was taken to locate close markers linked to the seed coat color genes. Six markers were found to co-segregate with the seed coat color gene. From the anchoring marker, three markers [1204BB248 (ME2PM93-248)], [722CG245 (SA7BG29-245)],

[1115AB303 (ME2PM16-303)] were located towards the seed coat color gene and the other three markers [826BG304 (SA7PM37-304)], [819ER406 (PM88PM46-406)], [824EY277 (BG23PM120-277)] were present in the opposite direction of the gene (Figure 4.6). The entire chromosomal region encompassing the anchoring marker was 20.6 cM in length. The marker SA7BG29245 was very close to the seed coat color gene (Figure 4.3). Fortunately we found this SRAP marker (SA7BG29245) generated the same size fragment for the major seed coat color gene *Br1* in *B. rapa.* This marker tested with 192 black seeded and 256 yellow-seeded lines, derived from five different crosses of *B. napus*, showed a very close linkage with the black seed coat color trait. The marker was also tested with 217 F_2 and 175 BC_1 plants of the MillenniUM 03 x SRYS-3 cross. This marker also co-segregated with the black/brown seed coat color and was absent in the dark/light/bright yellow seed coat color lines, and it showed a single Mendelian gene inheritance pattern for F_2 3:1 (χ^2 =0.124, P=0.7-0.9) and BC_1 1:1 (χ^2 =0.463, P=0.3-0.5). Recombinants alleles were observed in 9 plants of 217 F_2 plants producing a genetic distance of 2.07 cM between the molecular marker SA7BG29245 and the seed coat color gene. On the other hand 8 recombinant alleles were identified in 175 BC_1 plants producing a genetic distance of 4.57 cM away from the seed coat color gene. This marker was sequenced in *B. napus* and generated the exactly same sequence (214 bp) to that of the marker in *B. rapa.* There was no marker sequence difference between the black and the yellow seeded lines in *B. napus*. However, we obtained a total of 1746 bp sequence in *B. napus* using various primers designed from the *B. rapa* sequences earlier submitted to the gene bank (GenBank Accession Number EF488953, EF488954). The new sequences in the black seeded parental cultivar ‘MillenniUM 03’ and the yellow seeded parental line

'SRYS-3' have been submitted in the GenBank (GenBank accession number EF608926 and EF608927). A total of 24 SNPs were identified from the flanking sequences for the black seeded lines and the yellow seeded lines in *B. napus*, and one SNP position was used to screen the F_2 and BC_1 plants from a cross of MillenniUM 03 x SRYS-3. The SNP position for homozygous black/brown seed color was 'C' and generated a black peak, heterozygous black/brown seed color had both 'C/T' and produced both a black and a red peak, and for the dark/light/bright yellow seed coat color was 'T' and produced a red peak (Figure 4.7). The SNP markers were tested with 176 F_2 plants, of which 40 were found to be homozygous dominant (Y2Y2), 95 were heterozygous (Y2y2) and 41 were homozygous recessive (y2y2), in accordance with a 1:2:1 genotypic segregation ratio for single Mendelian gene inheritance (χ^2 =1.125, P=0.2-0.3). Similarly, 168 BC_1 plants segregated in a single Mendelian gene 1:1 segregation ratio (χ^2 = 0.857, P=0.3-0.5). The genotyping profile for the SRAP marker showed the same profile as for the SNP marker.

Molecular marker linked to dark/light yellow seed coat color gene

Seven hundred and sixty eight SRAP primer pair combinations were used to screen three parental black seeded lines, eight dark/light yellow seeded lines (the marker SA12BG18388 and SA7BG29245 were absent in these lines), two brown seeded lines (the marker SA12BG18388 and SA7BG29245 were present in these lines) and three parental yellow seeded lines. The SRAP markers present in black seeded and dark/light yellow seeded lines, but absent in the two brown seeded and three yellow seeded parental lines (Figure 4.8) were tested with 48 black seeded and 48 yellow seeded lines, as well as used to screen 80 dark/light yellow seeded lines (marker SA12BG18388 and SA7BG29245 were absent in these lines). One close SRAP marker (DC1GA27197)

linked to the dark/light yellow seed coat color gene was identified from these populations (Figure 4.8). This marker was screened with 238 F_2 and 154 BC_1 plants from the [MillenniUM 03 x SRYS-3] and [(MillenniUM 03 x SRYS-3) x SRYS-3] crosses, respectively. The segregation pattern of the marker showed single Mendelian gene inheritance, i.e. a 3:1 ratio (χ^2 = 0.746, P = 0.3-0.5) in the F_2 and a 1:1 segregation ratio (χ^2 = 1.662, P = 0.1- 0.2) in the BC_1. Ten recombinant alleles were identified in the 238 F_2 plants producing a genetic distance of 2.10 cM, and 154 BC_1 plants showed 10 recombinant alleles producing a genetic distance of 6.5 cM between the molecular marker DC1GA27197 and the seed coat color gene.

4.5 Discussion

Seed coat color in Brassica species is determined by the genotypes of the maternal tissue when selfed or crossed with an other parent. However, a pollen effect was found when the yellow seeded *B. napus* lines were pollinated with the black seeded parental lines. A pollen effect in maternal yellow-seeded lines (a Xenia effect) is common in yellow sarson *B. rapa.* This effect could be used as an indicator for successful crosses. Yellow sarson was used as a parent line to create the yellow seeded materials used in this study, therefore, the gene(s) responsible for this Xenia effect may have been transferred into these yellow-seeded *B. napus* lines. Rahman *et al.* (2001) observed such a Xenia effect in the open pollinated yellow-seeded *B. napus* lines derived from yellow sarson. Liu *et al.* (2005) identified an intermediate seed coat color in the F_1 when yellow seeded *B. napus* line was crossed with black seeded lines. The trigenic inheritance for yellow/black seed coat color in *B. napus* observed in this study confirmed the reports of

Shirzadegen (1986), Van Deynze and Pauls (1994), Henderson and Pauls (1992), Lühs *et al.* (2000), and Baetzel *et al.* (1999).

SRAP molecular markers for seed coat color genes in *B. napus* were identified in F_2 and BC_1 plants using five different cross-derived populations. Using five different crosses permitted discovery of unique molecular markers linked to the seed coat color genes common to all crosses. Two SRAP markers very closely linked to the two major seed coat color genes responsible for black/brown seed coat color and one SRAP marker linked to the dark/light yellow seed coat color gene were identified. Analysis of SRAP markers with an ABI 3100 genetic analyzer using fluorescently labeled primers is a moderately expensive procedure. This may limit the use of these markers in less advanced laboratories, whereas simple PCR-based SCAR markers could be used with high efficiency using agarose gels in any laboratory. Therefore, a SCAR marker was developed from one of the SRAP markers using chromosome-walking technology. Conversion to SCAR markers from other types of markers has been reported by several researchers, for example SCAR markers from RAPD markers (Naqvi and Chattoo 1996, Lahogue *et al*. 1998, Barret *et al*. 1998b) and from AFLP markers (Negi *et al.* 2000, Adam-Blondon *et al.* 1998, Bradeen and Simon 1998). Chromosome-walking is one of the best methods to obtain flanking DNA sequence adjacent to the sequence of interest (Devic *et al.* 1997, Negi *et al.* 2000). Negi *et al.* (2000) applied the chromosome-walking method to convert AFLP markers to SCAR markers. The SRAP marker (SA12BG18388) found in this study was dominant in nature and 1282 bp of sequence was obtained from the flanking sequence of the black seeded parent, which was absent in the yellow seeded lines therefore, the SCAR marker developed from the flanking sequence remained in

dominant nature. The genotyping profile of the SCAR marker was same as the SRAP marker, evidence of the successful conversion of the SRAP marker into a SCAR marker.

We have developed an ultra-dense genetic map with over 13500 SRAP markers distributed into different linkage groups of *B. napus* (Sun *et al.* 2007). One random marker linked to the seed coat color gene was anchored on the genetic map which allowed us to find a closely linked seed coat color gene marker on linkage group N13 using primer walking from the anchoring marker. It was interesting to note that this SRAP marker (SA7BG29245) produced the same size DNA fragment and the same DNA sequence in the black seeded cultivar 'MillenniUM 03' of *B. napus* (AACC) and yellow sarson cultivar 'BARI-6' of *B. rapa* (AA). Therefore, we used over 2.0 kb of *B. rapa* sequence earlier submitted in GenBank to design different primers to generate extended sequence in *B. napus.* The flanking sequence of *B. napus* was very similar to the *B. rapa* sequence, which is an indication of the presence of the same gene in the A genome of *B. napus.* The resynthesized yellow seeded *B. napus* line 'SRYS-3' was developed using a yellow sarson line as a parental source where the yellow seed coat color gene from yellow sarson may have been introgressed into the C genome of yellow seeded *B. napus* line (Rahman 2001b). Our results confirm the insertion of the yellow seed coat color gene from the A genome of *B. rapa* to the C genome (linkage group N13) in *B. napus*. However, we were unable to identify the seed coat color gene location in the A genome linkage group of *B. napus.* The marker SA7BG29245 was dominant, less convenient for marker assisted selection in plant breeding. Therefore, this SRAP marker was converted into a co-dominant SNP marker. The ABI SNaPshot method used to detect the SNP was simple, required very little optimization and was high throughput in an ABI 3100 genetic

analyzer (Nirupma *et al*. 2004). Therefore, this method for genotyping the seed coat color gene was used. The SNP markers showed the same genotyping pattern as for the SRAP marker for the yellow/dark-yellow seeded plants in the F_2 and BC_1 generations, and also same phenotypic pattern for SNP markers and the SRAP marker, indicating that the SRAP marker was successfully converted into a SNP marker. The SNP marker approach is very efficient but rather expensive.

The segregation pattern for seed coat color in segregating generations showed that two genes were responsible for black/brown seed coat color trait and a third seed coat color gene was responsible for the dark/light yellow seed coat color gene. Two markers (SA12BG18388 and SA7BG29245) were linked to the black/brown seed coat color gene. A third seed color gene marker was tightly linked to the dark/light yellow seed coat color gene. This SRAP marker was distinct from the two other markers linked to the black/brown seed coat color genes.

Three molecular markers for the three genes controlling yellow/black seed color in *B. napu*s were successfully developed in this study. These included a SCAR, a SNP and a SRAP marker, all of which are simple and convenient for marker assisted selection in plant breeding. These markers are tightly linked to the three different seed coat color genes in *B. napus*, which will greatly facilitate the selection of yellow seeded lines in the seedling stage, from segregating populations. These markers will also be highly useful for map-based cloning for the three seed coat color genes in *B. napus*.

Table 4.1 Segregation of seed coat color in F_2 and BC_1 populations of black-seeded and yellow-seeded *Brassica napus* crosses.

Populations	Black/ brown /partially yellow	Pure Yellow seeds	Segregation ratio 63:1		
			χ^2	df	P
F_2 Populations					
DHBS126 x SRYS-1	209	1	3.307	1	0.05-0.1
Sentry x SRYS-2	194	3	0.500	1	0.3-0.5
Holly276 x SRYS-2	180	3	0.007	1	>0.95
MillenniUM 03 x SRYS-3	294	4	0.036	1	0.7-0.9
Allons x SRYS-3	215	2	0.579	1	0.3-0.5
Total			4.429	5	
Goodness of fit χ^2 (pooled data)	1092	13	1.071	1	0.2-0.3
Homogeneity χ^2 (Total – pooled data)			3.358	4	0.3-0.5
			7:1		
BC_1 Populations			χ^2	df	P
(DHBS126 x SRYS-1) x SRYS-1	43	6	0.003	1	>0.95
(Sentry x SRYS-2) X SRYS-2	107	17	0.166	1	0.5-0.7
(Holly-276 x SRYS-2) x SRYS-2	59	10	0.250	1	0.5-0.7
(MillenniUM03xSRYS-3)xSRYS-3	167	19	0.887	1	0.3-0.5
(Allons x SRYS-3) x SRYS-3	131	11	2.934	1	0.05-0.1
Total			4.240	5	
Goodness of fit χ^2 (pooled data)	507	63	1.191	1	0.2-0.3
Homogeneity χ^2 (Total – pooled data)			3.049	4	0.5-0.7

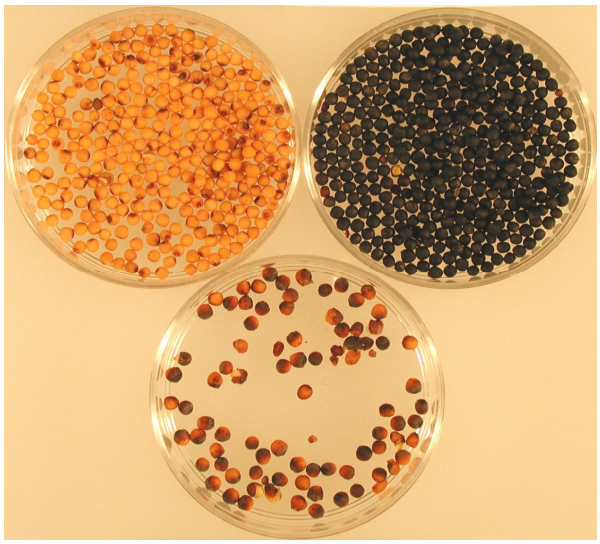


Figure 4.1 The F_1 seed of the cross between yellow-seeded (SRYS-3) and black-seeded (MillenniUM 03) lines of *B. napus.*

Figure 4.2 Seed coat color segregation in different F_2 [MillenniUM 03 x SRYS-3] and BC_1 [(MillenniUM 03 x SRYS-3) x SRYS-3] populations of *B. napus* crosses.

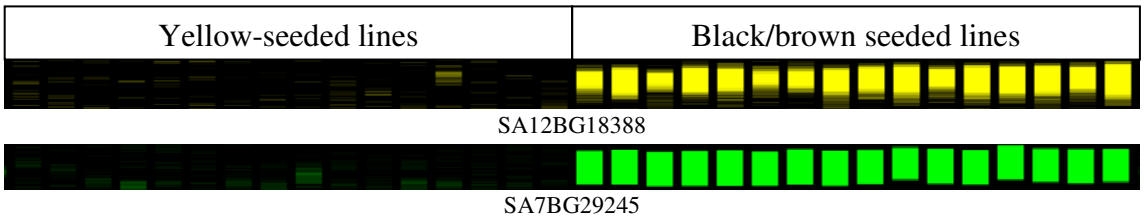


Figure 4.3 Two SRAP markers linked to the two different black/brown seed coat color genes in *B. napus*.

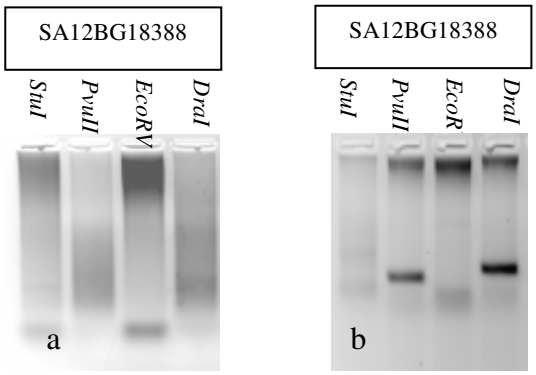
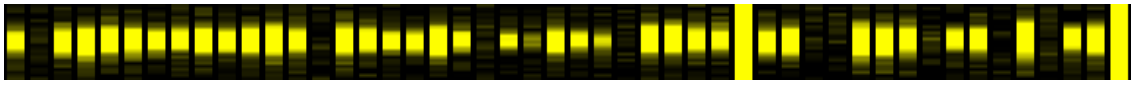
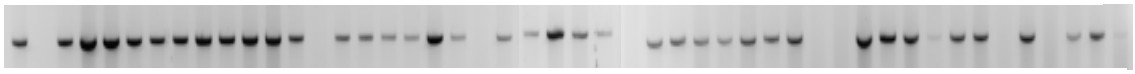


Figure 4.4 Chromosome-walking from the 5'-end of the marker SA12BG18388 sequence. Walking from the 3'-end of the marker SA12BG18388 sequence are not shown here. Two -step PCR using primer combination AP1/MWalk-20 and AP2/MWalk-21 from the marker SA12BG18388 were performed. The DNA was taken for first PCR and second PCR from four different genomic libraries constructed by *DraI, EcoRV, PvuII* and *StuI.* a. AP1+MWalk20, first round PCR; b. AP2+MWalk21, second round PCR.



4.5a. SRAP marker (SA12BG18388)



4.5b. SCAR marker converted from SRAP marker (SA12BG18388)

Figure 4.5 SRAP marker (SA12BG18388) segregation in some F_2 progenies of a cross of MillenniUM03 x SRYS-3 (4.5a). The same plant individuals were tested with the SCAR marker (4.5b).

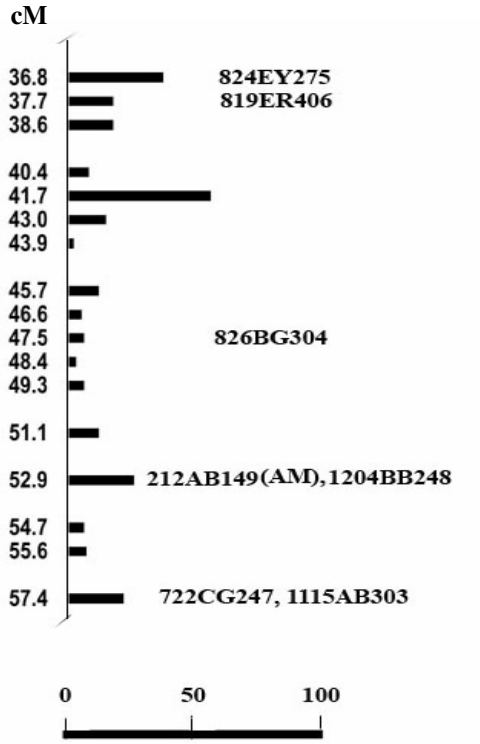


Figure 4.6 A partial genetic-linkage map of SRAP markers associated with the seed coat color gene locus (partial linkage group N13, Sun *et al.* 2007). Map distances are demonstrated in Kosambi units (centiMorgans, cM). AM: Anchoring marker.

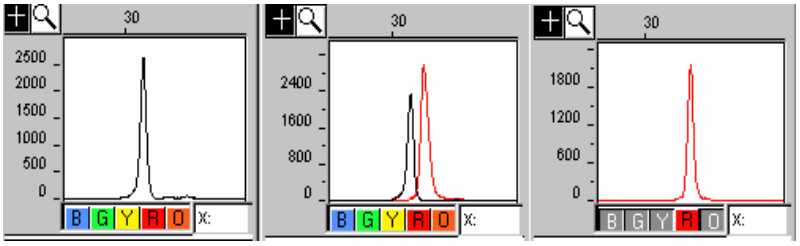


Figure 4.7 SNP detection in the using GeneScan software (ABI 3100 genetic analyzer). The peak information has been transformed manually for each locus [e.g. black for 'C' and the genotype is Y2Y2; red for 'T' and genotype is y2y2; and black/red for 'C/T' and genotype Y2y2].

Figure 4.8 SRAP marker (DC1GA27197) linked to the dark/light yellow seed coat color gene in *B. napus*.

5.0 DEVELOPMENT OF SRAP, SNP AND MULTIPLEXED SCAR MOLECULAR MARKERS FOR THE MAJOR SEED COAT COLOR GENE IN *Brassica rapa* L.

Mukhlesur Rahman, Peter B.E. McVetty and Genyi Li

Department of Plant Science, University of Manitoba

Winnipeg, Manitoba, Canada R3T 2N2

Article has been accepted for publication at Theoretical and Applied Genetics journal

Key words: *B. rapa*; seed coat color; molecular marker; SRAP; SNP; SCAR.

5.0 Development of SRAP, SNP and multiplexed SCAR molecular markers for the major seed coat color gene in *Brassica rapa* L.

5.1 Abstract

Seed coat color inheritance in *B. rapa* was studied in F_1, F_2, F_3, and BC_1 progenies from a cross of a Canadian brown-seeded variety 'SPAN' and a Bangladeshi yellow sarson variety 'BARI-6'. A pollen effect was found when the yellow sarson line was used as the maternal parent. Seed coat color segregated into brown, yellow-brown and bright yellow classes. Segregation was under digenic control where the brown or yellow-brown color was dominant over bright yellow seed coat color. A sequence related amplified polymorphism (SRAP) marker linked closely to a major seed coat color gene (*Br1/br1*) was developed. This dominant SRAP molecular marker was successfully converted into single nucleotide polymorphism (SNP) markers and sequence characterized amplification region (SCAR) markers after the extended flanking sequence of the SRAP was obtained using chromosome-walking. In total, 24 SNP's were identified from more than 2-kb sequence. A 12-bp deletion allowed the development of a SCAR marker linked closely to the *Br1* gene. Using the five fluorescence dye set supplied by ABI, four labeled M13 primers were integrated with different SCAR primers to increase the throughput of SCAR marker detection. Using multiplexed SCAR markers targeting insertions and deletions in a genome showed great potential for marker assisted selection in plant breeding.

5.2 Introduction

Brassica rapa is a major oilseed and vegetable species throughout the world as well as being one of the parent species of *B. napus*. Yellow seed coat color is desirable in any oilseed Brassica species because it has been reported that yellow-seeded varieties have a thinner seed coat than black seeded varieties, resulting in comparatively larger endosperm which contributes 5 to 7% more oil in the seed (Liu *et al.* 1991). The seed meal from yellow seeded varieties also contains higher protein and lower fibre content, which improves the meal quality for poultry and livestock (Shirzadegan and Röbellen 1985).

Early genetic studies by Mohammad *et al.* (1942) and Jönsson (1975) indicated that three genes are responsible for seed coat color segregation in *B. rapa.* Later, Stringam (1980) reported that two independent loci controlled seed color and proposed a model for seed coat color genes *Br1* and *Br3*. According to Stringam's model, presence of dominant alleles at both loci (*Br1* and *Br3*) or presence of dominant alleles only at the first locus (*Br1*) produce brown seed color, while presence of dominant alleles at a second locus (*Br3*) and homozygous recessive alleles at the first locus (*br1br1*) produce yellow-brown seeds. Yellow seeds are produced only when both loci present are in homozygous recessive condition (*br1br1br3br3*). Schwetka (1982), Zaman (1989) and Rahman (2001a) confirmed the seed coat color inheritance pattern in *B. rapa* as proposed by Stringam (1980).

Traditionally, selection of plant materials with desirable traits is carried out using phenotypic selection and the phenotypic variation is often inXuenced by environments. In contrast, the development of molecular markers linked to a trait of interest using different

molecular marker techniques, enables marker assisted selection (MAS) for the monitoring of similarly/dissimilarly among different genotypes at the very early stages of plant development, independent of environmental effects. This can significantly reduce the cost of producing breeding lines and can accelerate the breeding program dramatically. There are several molecular markers technologies available for marker assisted selection in plant breeding including restriction fragment length polymorphism (RFLP), simple sequence repeats (SSR), random amplification of polymorphic DNA (RAPD) (Williams *et al.* 1990, Karp *et al.* 1997), amplified fragment length polymorphism (AFLP) (Vos *et al.* 1995), and sequence related amplified polymorphism (SRAP) (Li and Quiros 2001). The principles of these marker techniques vary and they generate different amounts of information. The SRAP technique is simple and easy to perform, preferentially amplifies ORF or ORF related sequences and selected SRAP PCR products separated on a polyacrylamide gel are easy to sequence (Li and Quiros 2001). Therefore, the SRAP marker technique was used in this study for the identification of molecular markers linked to seed coat color genes in *B. rapa*.

Several molecular markers linked to seed coat color in Brassica species have been reported. Van Deynze *et al.* (1995) identified RFLP markers linked to a seed coat color gene in *B. napus*. Similarly, Somers *et al.* (2001) developed a RAPD marker for single major gene (*pigment1*) controlling seed coat color in *B. napus*. Liu *et al.* (2005) reported that yellow seed color was partially dominant over black seed color and developed 2 RAPD and 8 AFLP markers for the seed coat color gene in *B. napus*. The RAPD and AFLP markers developed by Liu *et al.* (2005) were not suitable for large scale marker assisted selection, therefore Liu *et al.* (2006) converted these markers into reliable

sequenced characterized amplified region (SCAR) and cleaved amplified polymorphic sequence (CAPS) markers for seed coat color breeding in *B. napus.* Negi *et al.* (2000) identified an AFLP marker for seed coat color gene in *B. juncea* and converted the marker into SCAR marker. In another study, SSR markers were developed for mapping and tagging the two independent loci controlling the seed coat color in *B. juncea* (Padmaja *et al.* 2005). Mahmood *et al.* (2005) identified QTLs associated with the seed coat color in *Brassica juncea* from an RFLP map using a doubled-haploid population. Chen *et al.* (1997) identified a RAPD marker linked to a seed coat color gene in a C genome chromosome of a *B. campestris – B. alboglabra* additional line. Heneen and Jørgensen (2001) identified a RAPD marker on chromosome 4 for brown seed color in *B. alboglabra* using *B. rapa – B. alboglabra* monosomic addition lines. To date, no seed coat color gene in *B. rapa* has been identified. In this study, the inheritance of seed coat color in *B. rapa* was analyzed using cross progeny from a cross of the self-incompatible variety 'SPAN' and the self-compatible yellow sarson variety 'BARI-6'. SRAP, SNP and multiplexed SCAR molecular markers closely linked to a seed coat color gene were developed. These molecular markers will be used for marker assisted selection in Brassica breeding and map-based cloning of this seed coat color gene.

5.3 Materials and methods

Plant materials

The pure breeding brown-seeded self-incompatible Canadian *B. rapa* variety 'SPAN' was crossed with the pure breeding yellow sarson self-compatible Bangladeshi *B. rapa* variety 'BARI-6' and the F_1 was backcrossed with 'BARI-6'. The F_1, F_2, F_3 and

BC_1 were grown in a greenhouse at the University of Manitoba. Segregation for seed coat color was studied in F_2 and backcross populations in the greenhouse. Plants from the segregating populations were grouped into brown, yellow-brown and yellow color seed produces. A χ^2 test was performed on the grouped data to check the goodness of fit of the segregating populations to the expected Mendelian phenotypic segregation ratio. A total of 224 F_2 and 197 BC_1 plants were used for seed coat color segregation analysis and molecular marker development for the seed coat color trait.

DNA extraction and SRAP molecular marker development

DNA was extracted using a modified CTAB method according to Li and Quiros (2001) from the flower buds of parental lines and their segregating populations. SRAP PCR amplification was the same as that of Li and Quiros (2001). Instead of autoradiography for signal detection, a five fluorescent dye set including, 6-FAM (blue), VIC (green), NET (yellow), PET (red), and LIZ (orange) supplied by Applied Biosystems (ABI), was used to separate SRAP PCR products with an ABI 3100 Genetic Analyzer (ABI, California).

Chromosome-walking and sequencing

The chromosome-walking method is commonly used to determine genomic sequence flanking the known sequence of molecular markers. Siebert *et al.* (1995) have described a chromosome-walking method on uncloned human genomic DNA, which was commercialized by Clontech Laboratories (Clontech Laboratories Inc, Mountain View, California). The Genome Walker™ Universal Kit was used to obtain flanking chromosome sequence of the molecular marker linked to seed coat color. The procedure was performed according to the protocol provided in the Clontech kit. Genomic DNA of

'SPAN' (brown seeded parent) was digested with restriction enzymes *DraI, EcoRV, PvuII* and *StuI*. Sharp and strong bands were obtained after a second PCR amplification. These bands were excised from an agarose gel and DNA was extracted using a Qiagen Gel Extraction kit. All the DNA fragments were sequenced using a BigDye®Terminator v1.1 Cycle Sequencing Kit.

SNP detection with SNaPshot

SNP primer (GTGGTTGAGCGCTCAGTTGCA) and SCAR primers (MR13, TGCTCGTTCTTGACAACAC; MR54, GAGAATTGAGAGACAAAGC) used in this study were designed using the Primer3 (http://biotools.umassmed.edu/bioapps/ primer3_www.cgi) software. SNPs were detected with an ABI SNaPshot kit (ABI, Toronto). Genomic DNA was amplified first with specific primers targeting the corresponding SNP mutations. The PCR reaction was set up in 10 µl of reaction mix containing 60 ng of genomic DNA, 0.375 µM dNTP, 0.15 µM of each primer, 1x PCR buffer, 1.5 mM MgCl2 and 1 unit *Taq* polymerase. The PCR running program was 94°C for 3 min, followed by 35 cycles of 94°C for 1.0 min, 55°C for 1.0 min, 72°C for 1.0 min and final extension at 72°C for 10 min. The amplified fragments were further analyzed with SNP detection primers and SNaPshot was performed according to the protocol in the ABI kit. The final products were separated with an ABI 3100 Genetic Analyzer. All four ddNTPs were fluorescently labeled with a different color dye i.e. the nucleotide 'C' was black, 'T' was red, 'G' was blue and 'A' was green. The alleles of a single marker were identified by different fluorescence color peaks after the data was analyzed with ABI GeneScan software.

SCAR detection

The forward primer MR13 (TGCTCGTTCTTGACAACAC) and the reverse primer MR54 (GAGAATTGAGAGACAAAGC) were designed to target a deletion mutation that occurred in the black-seeded lines. To detect this deletion with the ABI 3100 Genetic Analyzer, an M13-tailed primer method (Boutin-Ganache *et al.* 2001) was applied where the M13 primer sequence (CACGACGTTGTAAAACGAC) was added to the 5' primer end of MR13 to create a primer MR1313 (CACGACGTTGTAAAACGACTGCTCGTTC TTGACAACAC). The M13 primer was labeled with four fluorescence dyes, 6-FAM, VIC, NED, and PET supplied by the ABI Company. In the PCR amplification, four different PCR reactions were set by four fluorescently labeled primers with separately unlabeled MR1313 and MR54 primers. The PCR reactions were mixed together in a 10 µl volume containing 60 ng of genomic DNA, 0.375 µM dNTP, 0.10 µM of M13 primer, 0.05 µM of MR1313 primer, 0.10 µM of MR54 primer, 1x PCR buffer, 1.5 mM MgCl2 and 1 Unit *Taq* polymerase. PCR was performed at 94°C for 3 min, six cycles at 94°C for 50 sec, 60°C for 1.0 min with a 0.7°C decrease of annealing temperature at each cycle, 72°C for 1.0 min, and then twenty cycles at 94°C for 30 sec, 56°C for 30 sec, 72°C for 1.0 min for denaturing, annealing and extension, respectively. The PCR amplification products from different dye colors were pooled together so that each well contained four different fluorescently labeled DNA fragments which were detected in an ABI 3100 Genetic Analyzer.

5.4 Results

Seed coat color inheritance in *B. rapa*

Inheritance of seed coat color was analyzed using 224 F_2 individuals from a cross of 'SPAN x BARI-6'. It was found that seed coat color was mainly controlled by the genotypes of plants bearing seeds; therefore brown F_1 seeds were produced when a brown-seeded variety was used as the maternal parent. A pollen effect was observed when yellow sarson was used as the female parent so the F_1 seed coat color was dark yellow instead of bright yellow. Seeds on F_2 plants segregated into brown, yellow-brown and bright yellow color (Figure 5.1), indicating incomplete dominance of the brown color as described by Shirzadegan (1986). Of 224 F_2 plants, 164 had brown seed color, 48 had yellow-brown seed color and 12 had bright yellow seed color. The Chi-square test showed that seed coat color segregated in a ratio of 12:3:1 ($\chi^2 = 1.238$, P = 0.5-0.7), confirming digenic inheritance of the trait. However, when the 164 brown seeded plants were placed in one group and the 60 yellow-brown and bright yellow seeded plants were placed in another group, seed coat color segregated in a monogenic inheritance pattern ($\chi^2 = 0.381$, P = 0.5-0.7).

Self pollinated seeds of 197 BC_1 plants from the [(SPAN x BARI-6) x BARI-6] cross were also used for seed coat color segregation analysis. The seed coat colors in BC_1 also segregated into brown, yellow-brown and bright yellow classes. Of the 197 BC_1 plants, 95 had brown seed color, 55 had yellow-brown seed color and 47 had bright yellow seed color plants. The progenies fit a digenic (2:1:1, $\chi^2 = 0.898$, P = 0.5-0.7) segregation ratio for seed coat color. However, when the 95 brown seeded plants were placed in one group and the 102 brown yellow and bright yellow seeded plants were

placed in another group, seed coat color segregated in a monogenic manner (1:1, χ^2 = 0.248, P = 0.5-0.7).

SRAP molecular markers for seed coat color

Forty eight different SRAP primer pairs were used for the development of molecular markers for the seed coat color trait in *B. rapa.* Initially, sixteen brown-seeded lines and sixteen bright yellow-seeded lines from the BC_1 population were used for the identification of molecular markers using all 48 primer combinations. The markers SA7BG29-245, ME2FC1-266, FC1BG69-530, PM88PM78-435, SA12BG18-244 and SA12BG38-306 were found to be linked to the seed coat color with few recombinants. After testing these markers using the F_2 and BC_1 generations, the marker SA7BG29-245 was found to be closely linked to seed coat color. This marker was screened in the F_2 and BC_1 generations and showed a 3:1 segregation ratio (χ^2 = 0.215, P = 0.5-0.7) in the F_2 and a 1:1 ratio in the BC_1 (χ^2 = 0.408, P = 0.5-0.7). These results indicated that this SRAP marker was tightly linked to the brown (*Br1*) seed coat color gene, the major seed coat color gene in *B. rapa.* There were two recombinant alleles in a total of 448 alleles in the 224 F_2 plants equal to a genetic distance of 0.45 cM between the molecular marker SA7BG29-245 and the seed coat color gene. Similarly, in the BC_1 generation, there were two recombinant alleles in a total of 197 F_1-derived alleles used to produce this generation, equal to a genetic distance of 1.02 cM between the molecular marker SA7BG29-245 and the seed coat color gene.

Chromosome-walking and SNP development

The SRAP molecular marker SA7BG29-245 was sequenced and its flanking sequences were obtained by chromosome-walking. Two-step PCR reactions were

performed. The first PCR amplification using the left side marker specific primer MWalk27 and adaptor specific primer AP1 produced a smear in all lanes (Figure 5.2a). The second PCR amplification using the adaptor specific primer AP2 and marker specific primer MWalk28 produced a single strong band with *EcoRV* and *PvuII* (Figure 5.2b). Similarly, the first PCR amplification using the adaptor specific primer AP1 and marker specific primer MWalk24 from the right end generated a smear in all lanes (Figure 5.2c); and the second PCR amplification using the adaptor specific primer AP2 and marker specific primer MWalk25 generated two strong bands with *DraI* and *StuI* (Figure 5.2d). A total of 529bp was extended from left end and 427bp from right end and in total an 1170bp fragment was obtained from brown seeded variety 'SPAN' (GenBank Accession Number EF488953, EF488954). Unfortunately, the sequence did not match any gene in Arabidopsis after BLAST analysis against the Arabidopsis database (http//www.arabidopsis.org). After sequencing the corresponding region in the yellow-seeded parent, 24 SNPs were found between the brown-seeded and yellow-seeded parent lines after the sequences were aligned with CLUSTALW software (http://www.ebi.ac.uk/clustalw) (appendix viii). The SNPs were detected with an ABI SNaPshot Multiplex kit. For example, one SNP position (at 1041 bp position of 'SPAN') for homozygous brown seed color was 'C' and generated a black peak, heterozygous plants, 'C/T', generated both a black peak and a red peak, and homozygous yellow-brown or bright yellow seed coat color, 'T', generated a red peak (Figure 5.3). Since the marker was closely linked to the major seed coat color gene *Br1*/*br1*, the black peak identified homozygous brown seed color Br1Br1 genotypes; the dual black and red peaks identified heterozygous brown seed color Br1br1 genotypes; while the red peak identified

homozygous bright yellow or yellow-brown seed color br1br1 genotypes. The SNP markers were tested using both the F_2 and BC_1 generations, and were found to be at the same genetic distance from the seed coat color gene as the SRAP molecular marker SA7BG29-245.

Development of multiplexed SCAR markers

On the basis of 1,170 bp for the SRAP marker and its flanking sequences, no deletion or insertion polymorphic region was found between brown and yellow seeded lines. Therefore, chromosome-walking was performed again to obtain additional extended flanking sequence from the left side. With the new chromosome-walking sequence, a 12-bp deletion in the brown seeded lines or a 12-bp insertion in the yellow-brown or bright yellow seeded lines was identified (appendix viii), which were used for the development of multiplexed SCAR markers. Primers MR1313 and MR54 were designed to target the 12-bp deletion. Together with the 19-bp M13 sequence, a 388-bp fragment for brown seeded lines and a 400-bp fragment for yellow-brown or bright yellow seeded lines were produced, respectively (Figure 5.4). Since the SCAR marker was not far from the SRAP marker and SNPs mentioned previously, the genotyping of the SCAR marker in 224 F_2 plants and 197 BC_1 plants were exactly the same as that of the SRAP and SNP markers.

5.5 Discussion

The *B. rapa* yellow sarson parent line variety 'BARI-6' was taxonomically different from the Canadian *B. rapa* parent line variety 'SPAN'. Yellow sarson belongs to ssp. *trilocularis* and is self-compatible, while 'SPAN' belongs to ssp. *oleifera* and is

self-incompatible. Using a self-compatible parent in the cross made it easier to self plants in the greenhouse. A pollen effect was observed when yellow sarson was used as the female parent, resulting in dark yellow F_1 seeds instead of bright yellow F_1 seeds. This is known as a Xenia effect in yellow sarson and could be used as an indicator for successful crosses. This phenomenon was also observed by Rahman *et al.* (2001) who used an open pollinated yellow-seeded *B. napus* line that was derived from yellow sarson, suggesting that yellow sarson contains the gene(s) for Xenia effect. Stefansson and Hougen (1963) reported such effect of pollen (Xenia effect) on the fatty acid profile of the seeds of self-pollinated and reciprocal cross-pollinated *B. napus* plants.

Digenic inheritance with dominant epistasis was observed for seed coat color segregation in *B. rapa.* The dominant epistatic gene was responsible for brown color and the hypostatic gene was responsible for yellow-brown seed color, and yellow seed color was observed when both the genes were in homozygous recessive condition. These results confirm the seed coat color segregation results reported by Stringam (1980) and by Rahman (2001a).

A dominant SRAP marker is less convenient than a co-dominant marker for large scale marker assisted selection in plant breeding. Consequently, the dominant SRAP marker developed in this study was converted to co-dominant SNP and SCAR markers, following the lead of several researchers who converted their dominant markers into co-dominant markers, such as SCAR marker from RAPD markers (Naqvi and Chattoo 1996, Lahogue *et al.* 1998, Barret *et al*. 1998b) and AFLP markers (Negi *et al.* 2000, Adam-Blondon *et al*. 1998, Bradeen and Simon 1998), and SCAR and CAPS markers from RAPD and AFLP markers (Liu *et al.* 2006). There was no difference between brown-

seeded and yellow-seeded lines in the 214 bp sequence of the SRAP marker. A single nucleotide polymorphic position is required for the development of co-dominant SNP markers. Co-dominant SCAR markers are developed from the insertion or deletion fragments position in any of the two sequences. Even, development of CAPS markers required the DNA fragments size range of 500 to1500 bp (Barret *et al*. 1998b). Therefore, a 214 bp SRAP sequence limits the development of any co-dominant SNP, SCAR or CAPS markers. However, the extended flanking sequence from the SRAP marker allowed the development of SCAR or SNP co-dominant markers. A chromosome-walking approach was used to obtain the flanking sequence adjacent to the SRAP marker. It has been proven that chromosome-walking is one of the best methods for determining the flanking sequence adjacent to a sequence of interest (Devic *et al.* 1997, Negi *et al.* 2000). Negi *et al.* (2000) successfully converted the AFLP markers to the SCAR markers using chromosome-walking method and isolated the large-sized fragments adjacent to the AFLP markers which did not require any optimization for different walking. We obtained more than 2.0 kb flanking sequences from the SRAP markers that showed 24 SNPs and a 12 bp deletion or a 12 bp insertion site which allowed developing SNP markers and SCAR markers, respectively.

The SNaPshot method used in this study is simple, requires very little optimization and is high throughput using an ABI 3100 genetic analyzer (Nirupma *et al*. 2004). SNP markers are co-dominant, and have been found to be abundant in genomic sequences and therefore can potentially be used for marker assisted selection. The SNP markers developed in this study used to screen the F_2 and BC_1 generations showed the same pattern as the SRAP marker, indicating that the SRAP marker was successfully

converted into SNP markers that were closely linked to the *Br1* seed coat color gene. The major shortcoming of the SNP marker approach is cost.

A cost effective alternative to SNP markers are SCAR markers, most especially multiplexed SCAR markers. In this study, a 12-bp deletion in the brown seeded lines allowed the development of multiplexed co-dominant SCAR markers. Here we used four fluorescently labeled M13 primers with single unlabeled primer that allowed pooling four PCR products for the detection in an ABI 3100 genetic analyzer (four fluorescently labeled M13 primers were universally used to combine with any co-dominant multiplexing SCAR markers in our laboratory). However, in principle, any primers covering this 12 bp deletion region would produce two bands with a 12 bp sequence difference. Using the M13 primer labeled with four fluorescent dye colors and a series of primers that produced fragments with 12 bp differences in length permitted the pooling of several hundred amplified DNA samples for signal detection using the ABI Genetic Analyzer. Multiplexed SCAR markers can reduce the running cost of the ABI DNA Genetic Analyzer dramatically and significantly increase the efficiency of marker assisted selection in a breeding program compared to the high cost of SNP detection. For example, we designed 20 unlabeled primers to target a two-base deletion position in the *Bn-FAE1-2* gene of the C genome of *B. napus* and combined with a genome specific primer that was labeled with four fluorescent colors to form 80 primer pairs in total, and each primer pair was used to amplify different DNA samples. After PCR, 80 samples were pooled and 1280 (16 x 80) samples were analyzed with an ABI 3100 Genetic Analyzer in 40 minutes (unpublished data). The running cost was reduced by 80 times compared with that of SNP detection with the ABI SNaPShot detection kit. Actually

more unlabeled primers could be designed to increase the pooled samples to reduce the cost further. Therefore, multiplexing any co-dominant SCAR markers targeting deletions or insertions (INDELs) has great potential for marker assisted selection in plant breeding if a sample pooling strategy as described in this report is implemented.

Figure 5.1 Seed coat color segregation in the progenies of a cross of yellow-seeded 'BARI-6' and 'brown-seeded 'SPAN' in *B. rapa*. Seeds in the left Petri dish representing 'SPAN' and brown-seeded progeny; in the middle Petri dish, yellow-brown progeny; in the right Petri dish, 'BARI-6' and bright-yellow progeny

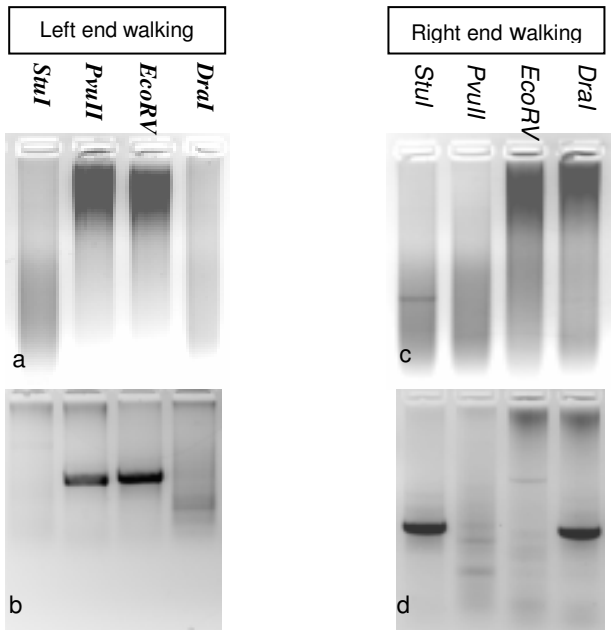


Figure 5.2 PCR walking from left end and right end of the marker (SA7BG29-245) sequence. Two-step PCR using primer combination AP1/MWalk27 and AP2/MWalk28 from the left end; and another two-step PCR from the right border with the primer combinations AP1/MWalk24 and AP2/MWalk25 were performed. The DNA were taken for first PCR and second PCR from four different genomic libraries constructed by *DraI, EcoRV, PvuII* and *StuI.* a. AP1+MWalk27, first round PCR; b. AP2+MWalk28, second round PCR; c. AP1+MWalk24, first round PCR; d. AP2+MWalk25, second round PCR.

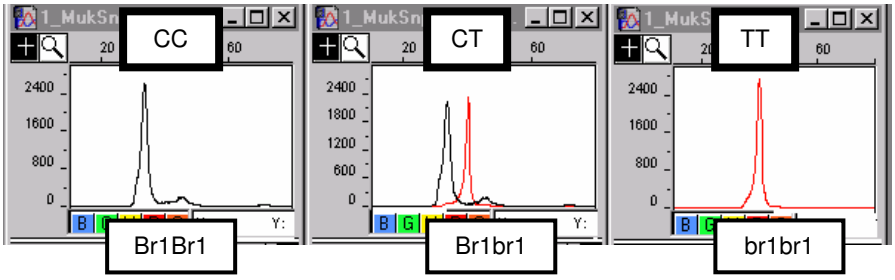


Figure 5.3 Figure showing SNP detection by GeneScan software (ABI 3100 genetic analyzer) to analyze the SNaPshot Multiplex kit data. The peak information was transformed manually for each locus [e.g. black for 'C' and the genotype *Br1Br1*; red for 'T' and genotype *br1br1*; and black/red for 'C/T' and genotype *Br1br1*].

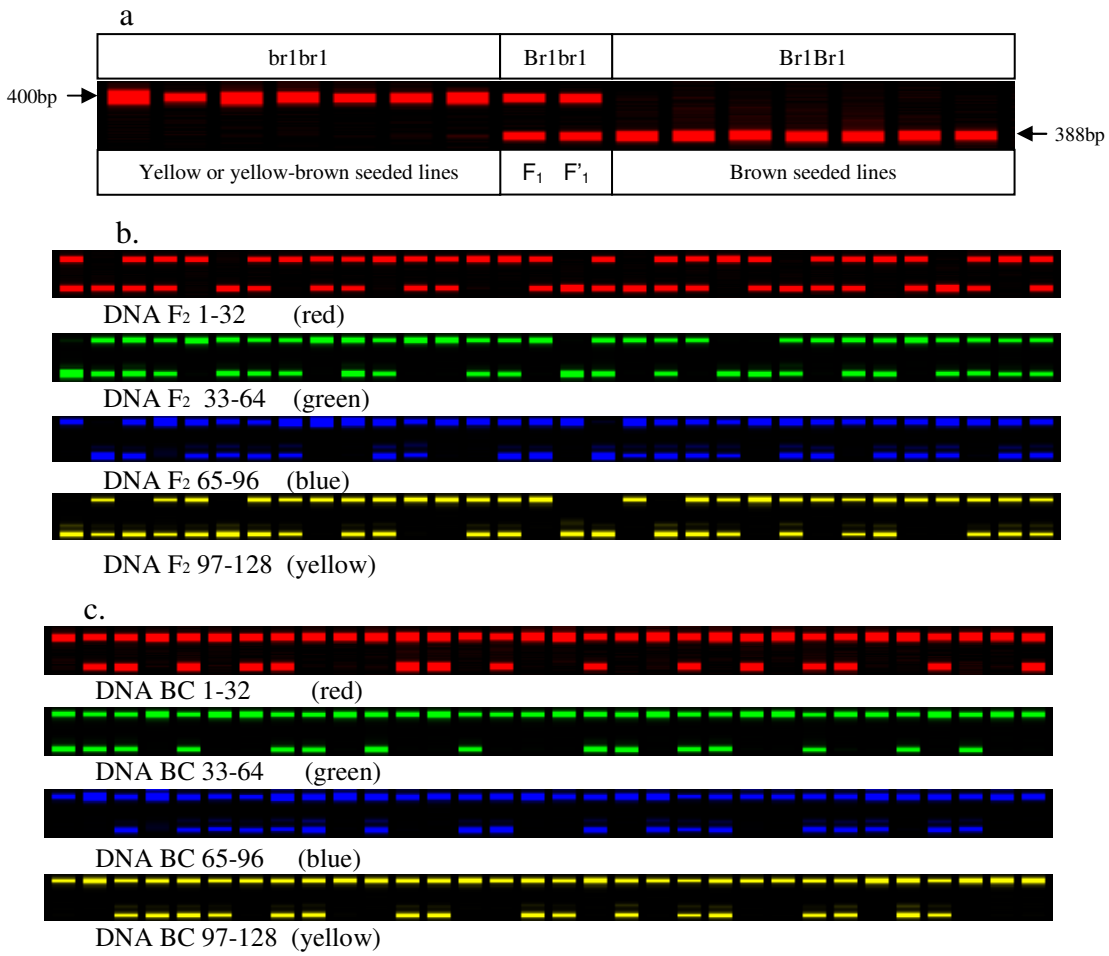


Figure 5.4 Multiplexed SCAR marker linked to seed coat color was detected in ABI 3100 genetic analyzer using four different fluorescently labeled primers M13 with unlabeled MR1313 and MR54. The marker linked to the brown seed gene (*Br1Br1*) produced 388 bp, the yellow or yellow-brown (*br1br1*) gene generated 400 bp and the heterozygotes (*Br1br1*) produced both 388 bp and 400 bp fragments. a. SCAR marker segregation in the brown-seeded, yellow-seeded and F_1 genotypes; b. SCAR marker segregation in the F_2; c.SCAR marker segregation in the BC_1.

6.0 GENERAL DISCUSSION

In this research, molecular markers for seed quality traits including erucic acid content, yellow seed coat color of *Brassica napus* and *B. rapa* have been developed for marker assisted plant selection in the breeding program.

A mapping population was developed from a cross of high erucic acid content variety 'MillenniUM 03' and zero erucic acid content line 'SRYS-3' for the identification of molecular markers for erucic acid content genes in *B. napus.* Digenic inheritance for erucic acid content was identified from the F_2 and backcross progenies in agreement with earlier research reported by Stefansson (1983). The fatty acid elongation1 (*FAE1*) gene is responsible for the biosynthesis of erucic acid had been characterized in Arabidopsis. The *FAE1* gene has two copies in *B. napus*, *Bn-FAE1.1* and *Bn-FAE1.2*, located in the A genome and C genome of *B. napus*, respectively. A single base substitution mutation and a two base deletion mutation have been identified in the *Bn-FAE1.1* and *Bn-FAE1.2* genes, respectively, for the low erucic acid and high erucic acid content lines in *B. napus* (Katavic *et al*. 2002, Fourmann *et al.* 1998). In this research, two BAC clones containing the *Bn-FAE1.1* and the *Bn-FAE1.2* genes identified from a *B. rapa* BAC library and a *B. oleracea* BAC library displayed the same single base substitution mutation and two base deletion mutation, respectively, as described by Katavic *et al*. (2002) and Fourmann *et al.* (1998). Sequence analysis of the two genes in this study showed a very high sequence similarity, a confirmation of the report of Fourmann *et al.* (1998). The single base polymorphism position present in the *Bn-FAE1.1* gene was absent at the same position in the *Bn-FAE1.2* gene, similarly, a two base deletion mutation in the *Bn-FAE1.2* gene was not present in the *Bn-FAE1.1* gene. The high similarity between the two *Bn-FAE* genes

provided obstacles for the development of molecular markers from inside the genes. However, this problem was solved by identifying the sequence dissimilarities between the flanking sequences of the two genes in the A and C genomes of *B. napus.* The polymorphism in the flanking sequences allowed genome-specific and gene-specific primers to be developed resulting in molecular markers for the erucic acid content genes. The single base polymorphic position in the *Bn-FAE1.1* gene was detected using the ABI SNaPshot multiplex kit, and the two base deletion in the *Bn-FAE1.2* gene for low erucic acid content was detected using fluorescently labeled SCAR markers in an ABI 3100 genetic analyzer. Although the two mutation positions in the *FAE1* gene have been characterized before, this is the first report of molecular markers developed inside the two genes and specific to the two genomes in *B. napus*. These molecular markers have been used with an accuracy of 100% for selecting plant for their genotypes for the *Bn-FAE1.1* and the *Bn-FAE1.2* genes and the erucic acid content of their seeds. Multiplexing of the PCR products for both markers were used to reduce the cost and time of analysis for erucic acid genotyping in segregating generations in the plant breeding program.

Yellow seeded *B. napus* genotypes have been developed by several researchers from different interspecific crosses of Brassica using naturally occurring yellow seeded species including *B. rapa, B. carinata* and *B. juncea* as a source of yellow seed coat color genes (Liu *et al.* 1991, Shirzadegan and Röbbelen 1985, Chen *et al.* 1988, Chen and Heneen 1992, Rashid *et al.* 1994, Qi *et al.* 1995, Tang *et al*. 1997, Meng *et al.* 1998, Rahman 2001b). Most of the artificially developed yellow seeded lines were affected by temperature. However, the yellow seeded lines used in this research obtained from NPZ, Lembke, Germany, developed by Rahman (2001b), were stable in true yellow color both

in the green house and in the field. This stability was useful for studying the seed coat color inheritance and for the development of tightly linked molecular markers co-segregated with the seed coat genes in *B. napus*. Three gene segregation for seed coat color in *B. napus* was observed in this study, where the black/brown color was dominant over the yellow seed coat color trait. These results agree with the seed coat color inheritance reports of most researchers who have developed the yellow seeded *B. napus* lines from interspecific crosses. The F_1 seed color of the crosses of yellow seeded lines as maternal parent and the brown seeded lines as donor parents produced brown-yellow color instead of bright yellow color, which is known as 'Xenia effects' (Rahman *et al.* 2001). This brown-yellow color occurred because the yellow seeded *B. napus* lines were developed from an interspecific cross used yellow sarson *B. rapa* as a source of yellow seed coat color genes. Yellow sarson expresses a naturally occurring Xenia effect in the F_1 seeds and the genes responsible for this effect might have been transferred from yellow sarson into the yellow seeded lines of *B. napus*.

Molecular markers linked to the seed coat color genes in *B. napus* have been developed by RFLP, RAPD, AFLP markers techniques (Van Deynze *et al.* 1995, Somers *et al*. 2001, Liu *et al.* 2005). A SRAP marker technique was used to develop the molecular markers linked to the seed coat color genes in *B. napus.* SRAP is a PCR-based two primer marker system, a forward primer and a reverse primer with 17 or 18 nucleotides (Li and Quiros 2001). The thermocycler program for SRAP amplification is set an annealing temperature at 35°C for first five-cycle that ensured the binding of both primers at many partially matched positions of the target DNA. The temperature is then raised to 50°C for the 35 cycles that permits the five-cycle amplified DNA products to

generate consistent bands in the rest of the cycles. As a result, the SRAP marker technique generated a good number of monomorphic and polymorphic loci between the parents. Li and Quiros (2001) observed more than ten polymorphic loci with a single primer combination in the two parental lines of *B. oleracea.* This observation was very useful for detecting any SRAP molecular markers linked to a trait of interest in *B. napus.* Initially 8 brown seeded parental and segregating lines, and 8 yellow seeded parental and segregating lines were used to screen 400 primer pairs in 384 PCR plates. Since most SRAP markers are dominant, candidate markers linked to the black/brown seeded coat color genes were developed instead of the markers linked to the yellow seed coat color genes. This is because the black/brown seed coat color gene markers must be absent in the yellow seeded lines, whereas the yellow seed coat color gene markers might be present in the brown seed coat color trait when the trait is controlled by the gene in heterozygous condition.

Three different SRAP markers were developed for tagging the three different seed coat color genes in the gene family of *B. napus*. The first SRAP marker was tightly linked to the black/brown seed coat color trait of *B. napus.* This marker was sequenced and extended sequences were obtained by chromosome-walking as described by Siebert *et al.* (1995). Conversion of seed coat color AFLP markers into SCAR markers have been successfully conducted by Liu *et al.* (2006) in *B. napus*, and Negi *et al.* (2000) in *B. juncea.* The second seed coat color gene SRAP marker was identified from a high density genetic map of *B. napus* (Sun *et al.* 2007). This marker also associated with the black/brown seed coat color trait of *B. napus* present on the linkage group N13 of C genome of *B. napus*. This marker generated the same size DNA fragment and the same

DNA sequence in the black seeded cultivar 'MillenniUM 03' of *B. napus* (AACC) and yellow sarson cultivar 'BARI-6' of *B. rapa* (AA). We have deposited over 2.0 kb *B. rapa* sequences for brown seeded and yellow seeded lines into GenBank database. These sequences allowed designing different primers to generate about 1.7 kb of extended sequence in *B. napus.* The flanking sequences between black/brown seeded lines and the yellow seeded lines generated 24 SNPs which were used to convert the SRAP marker into a co-dominant SNP marker. The first two SRAP markers were linked to the black/brown seed coat color genes, and it was assumed that the third seed coat color gene was responsible for the dark/light yellow seed coat color traits in *B. napus*. Therefore, the third SRAP seed coat color gene marker linked to the dark/light seed coat color trait was developed using the black seeded parental lines, dark/light yellow seeded lines, and yellow seeded parental lines. Development of the three different SRAP markers linked to the three different seed coat color genes were highly applicable for screening 100% yellow seeded lines from the segregating populations in the marker assisted selection in plant breeding.

Seed coat color in *B. rapa* was found to be controlled by two genes, one was for brown seed coat color trait and the other was for dark/light yellow seed coat color trait. This result showed an agreement with the seed coat color inheritance model proposed by Stringam (1980) and later supported by Schwetka (1982), Zaman (1989), Rahman (2001a). A Xenia effect was observed at the F_1 seed color when the yellow sarson *B. rapa* was hybridized with the brown seeded pollen donor parent. The Xenia effect is naturally occurring in yellow sarson *B. rapa*, which could be used as an indicator for successful crosses. Yellow sarson genotypes belong to the *B. rapa* ssp. *trilocularis*, are special

ecotypes that originated in India, and are self-compatible (Rahman *et al.* 2001). Molecular markers linked to the seed coat color traits in *B. rapa* have not been developed before. However, *B. rapa* was used for the development of *B. rapa- alboglabra* monosomic additional lines with *B. oleracea* ssp. *alboglabra*, and a RAPD marker linked to the seed coat color gene in the *B. oleracea* ssp. *alboglabra* was developed from the additional chromosome line (Chen *et al.* 1997, Heneen and Jørgensen 2001). In this research, a dominant SRAP marker linked to the major seed coat color gene *Br1* in *B. rapa* for brown seed color has been developed, and the marker was successfully converted into co-dominant SCAR and SNP markers for the marker assisted selection in plant breeding. Sequence analysis of the SRAP marker showed the same sequence as the second seed coat color gene marker in *B. napus*, indicating the presence of the same gene in *B. rapa* and in *B. napus.*

7.0 SUMMARY/CONCLUSION

Canola oil has excellent fatty acid composition and low saturated fat levels. Canola meal has protein with excellent amino acid composition. Canola seed quality can be further improved by higher oil content, higher protein and lower fibre content in the seed through the development of yellow seeded lines. Erucic acid is one of the most important fatty acids in rapeseed oil, where low erucic acid rapeseed oil is used for edible purposes, and high erucic acid rapeseed oil is used for industrial purposes. The breeding of high/low erucic acid content rapeseed as well as yellow seeded *B. napus* and *B. rapa* could be enhanced in efficiency by the development of molecular markers linked to the desired traits. Therefore, the objectives of the current studies were development of molecular markers for erucic acid content genes, seed coat color genes in *B. napus* and seed coat color genes in *B. rapa.*

High throughput genome-specific and gene-specific molecular markers for erucic acid content genes in *B. napus* were developed using a *B. rapa* (A-genome) BAC library and a *B. oleracea* (C-genome) BAC library. The fatty acid elongation 1 (*FAE1*) gene sequence was taken from the Arabidopsis gene bank sequence data and subsequently BAC clones containing the *Bn-FAE1.1* (E1 locus) and the *Bn-FAE1.2* (E2 locus) were identified from the *B. rapa* BAC library and the *B. oleracea* BAC library, respectively. The full length gene sequences of the *Bn-FAE1.1* gene and the *Bn-FAE1.2* gene in *B. napus* were identified by sequencing of the respective BAC clones. Sequence alignment between the *Bn-FAE1.1* genes for high and low erucic acid content lines in *B. napus* showed a single nucleotide polymorphism (SNP). Similarly a two base deletion mutation was observed in the *Bn-FAE1.2* gene for low erucic acid lines in *B. napus.* The sequence

similarity between the *Bn-FAE1.1* gene and the *Bn-FAE1.2* gene was also very high. The SNP position in the *Bn-FAE1.1* gene (in A-genome) for high and low erucic acid content does not have polymorphism at the same position in the *Bn-FAE1.2* gene (in C-genome). Similarly, the two base deletion in the *Bn-FAE1.2* gene for low erucic acid content does not have polymorphism at the same position in the *Bn-FAE1.1* gene, which created challenges for developing gene-specific markers. Therefore, genome-specific primers were designed from the outside of the gene sequences to amplify the A-genome (E1 locus) and C-genome (E2 locus) specific fragments containing the two polymorphic positions. The SNPs were detected with an ABI SNaPshot kit and the two base deletion SCAR markers were detected with an ABI 3100 Genetic Analyzer. Association of the two genes markers with the erucic acid content were tested in the F_2 and backcross progeny of a cross of a high and a low erucic acid content lines. There was a 100% match with the erucic acid profiles of individual progeny plants. For example, E1E1E2E2 genotypes had the highest erucic acid content (>42%), e1e1e2e2 genotypes had low erucic acid content (>1%), while E1e1E2e2 genotypes had about 20-22% erucic acid content, with all four erucic acid conditioning alleles having an almost identical contribution to erucic acid content in the seed with an average of 10.2% erucic acid per allele. High throughput SCAR marker detection was carried out with four fluorescently labeled reverse primers combined with 20 different primers to produce PCR products with different fragment sizes. Multiplexing of the PCR products allowed analyzing over sixty thousand of samples per day using a medium capacity ABI 3100 Genetic Analyzer. This dramatically reduces the cost of marker detection. Similarly, a multiplexing primer set for the *Bn-FAE1.1* gene was designed by adding a polyT to the 5' primer end to

increase SNP detection throughput through sample pooling. These multiplexed SCAR and SNP markers have great potential for marker assisted selection in plant breeding and have been successfully implemented in canola/rapeseed breeding programs.

Five different mapping populations were developed between crosses of five pure breeding black seeded varieties/line and three pure breeding yellow seeded lines in *B. napus*. Seed coat color were scored from black, dark brown, brown, dark yellow, light yellow and bright yellow, which were controlled by three genes. A pollen effect was found when yellow seeded lines were used as a female parent that produced brown yellow color F_1 seeds. SRAP methodology was applied to develop molecular markers linked to the individual seed coat color genes from the multi-gene family in *B. napus*. In total, four hundred SRAP primer pairs were tested to the DNA of 8 black seeded and 8 yellow seeded lines, and then 32 black seeded and 32 yellow seeded lines. Twenty-six markers which tightly co-segregated with the black/brown seed coat color genes were found. A genetic map was constructed using these 26 SRAP markers. Two markers were found to belong to one group and 24 markers were found to belong to a second group. One of the two SRAP markers in group one was very close to a black/brown seed coat color gene of *B. napus*. This marker was sequenced and extended sequences were obtained using chromosome-walking technology. A SCAR marker was developed from this SRAP marker. A second seed coat color gene marker was developed from the high density consensus map developed in Dr. Li's lab using primer walking from the anchoring marker for the second group. This marker was tightly linked to the black/brown seed coat color gene present on linkage group N13 of the C genome of *B. napus.* This marker generated the same SRAP fragments and same DNA sequences in *B.*

napus and in *B. rapa.* Based on about 2.0 kb sequence information of *B. rapa* earlier submitted into GenBank database, the flanking sequences of *B. napus* were obtained by designing different primers using *B. rapa* sequence. The second SRAP marker was converted into a simple co-dominant SNP marker. The third SRAP marker was developed using 768 SRAP primer pair combinations, and was associated with dark/light yellow seed coat color gene in *B. napus.* These three seed coat color genes markers can be used to screen for yellow seeded lines in the segregating generations of canola/rapeseed breeding programs.

A seed coat color gene in *B. rapa* was mapped in the F_2 and backcross populations of the cross of a brown seeded variety 'SPAN' and a yellow sarson *B. rapa* variety 'BARI-6'. Two genes were identified for inheritance of seed coat color in *B. rapa.* A pollen effect on maternal parents (a Xenia effect) was observed in yellow sarson *B. rapa.* This effect was used as an indicator for successful crosses. The SRAP technique was employed to develop the molecular markers for seed coat color genes. A tightly linked dominant SRAP marker was developed which was associated with the brown seed coat color gene in *B. rapa*. Since the SRAP marker was dominant in nature, conversion of the marker into co-dominant SNP and SCAR markers was done using the chromosome-walking technique. In total, 24 SNP's and a 12 bp deletion positions were identified from the 2.0 kb extended DNA sequences between the brown seeded and the yellow seeded lines, which were used to develop SNP and SCAR markers, respectively. These two markers were screened in 224 F_2 plants and 197 BC_1 plants, and demonstrated exactly the same phenotype and genotyping profile as that of the SRAP marker, indicating that the dominant SRAP marker was converted successfully into the co-dominant SNP and SCAR

markers. Four fluorescently labeled M13 primers were integrated with the forward SCAR primers to amplify the genomic DNA with unlabeled reverse primer permitted the pooling of amplified DNA samples for signal detection using the ABI Genetic Analyzer. This multiplexed SCAR markers can reduce the cost of marker analysis dramatically and significantly increase the efficiency of marker assisted selection in a breeding program.

8.0 RECOMMENDATIONS FOR FURTHER STUDY

Molecular markers for erucic content genes were developed from inside the genes specific to the A-genome and the C-genome, which were used with 100% accuracy to screen the genotypes and approximate erucic acid content profiles in F_2 and backcross segregating populations. The erucic acid content gene sequences in *B. napus* are universal for all rapeseed/canola lines/cultivars. These markers should be used to screen segregating progeny from canola x rapeseed crosses for their erucic content profile. This will greatly enhance the efficiency of canola and HEAR cultivar development.

Three different seed coat color SRAP markers in *B. napus* were developed in this study. All three markers were very closely linked to the seed coat color genes in *B. napus* and therefore, it should be possible to clone these seed coat color genes using positional cloning method. Successful cloning of these genes will open a new era for the development of yellow seeded *B. napus* using plant transformation. Precise location and sequencing of the yellow seed genes should be used for the development of molecular markers from inside the genes for screening seed coat color segregating plants at the very early stages of crop development with 100% accuracy. The first SRAP marker was sequenced and about 1.2 kb DNA additional sequence was obtained by chromosome-walking. Unfortunately, the 1.2 kb sequence was present only in the black seeded parents and was absent in the yellow seeded lines. Therefore, the SCAR marker developed from the SRAP marker remained a dominant marker. Further chromosome-walking might be useful for the development of a co-dominant SCAR or SNP markers for marker assisted selection in plant breeding. The third seed coat color SRAP marker in *B. napus* was not

sequenced and therefore, it is recommended that this marker be sequenced for the development of co-dominant SCAR or SNP markers.

Two genes were identified for the seed coat color trait in *B. rapa.* SRAP, SCAR and SNP markers were developed for the major seed coat color gene (*Br1*) responsible for brown coat color in *B. rapa.* Since the markers were very close to the seed coat color gene it should be possible to clone this gene by positional cloning. Molecular markers for the second seed coat color gene (*Br3*) linked to the brown-yellow seed coat color trait in *B. rapa* has not been developed yet for marker assisted selection in plant breeding. Therefore, it is recommended that molecular markers linked to the dark/light yellow seed coat color gene in *B. rapa* be developed.

9.0 REFERENCES

Abrahams, S., Tanner, G.J., Larkin, P.J. and Ashton, A.R. 2002. Identification and biochemical characterization of mutants in the proanthocyanidin pathway in Arabidopsis. Plant Physiol. 130:561–76.

Achnine, L., Blancaflor, E.B., Rasmussen, S. and Dixon, R.A. 2004. Colocalization of lphenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis. Plant Cell 16:3098–109.

Adam-Blondon, A.F., Sevignac, M., Bannerot, H. and Dron, M. 1998. SCAR, RAPD and RFLP markers linked to a dominant gene (*Are*) conferring resistance to anthracnose in common bean. Theor. Appl. Genet. 88:865–870.

Ag Canada (2007). Farm Business Communications. [Online Database) URL. http://www.agcanada.com/custompages/stories_story.aspx? mid=46&id=1053 (25 February 2007).

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

Anjou, K., Lönnerdal, B., Uppström, B. and Åman, P. 1977. Composition of seeds from some Brassica cultivars. Swedish J. Agric. Res. 7:169-178.

Babu, V.R. and Wiktor, A. 1991. A fluorescence in-situ hybridization technique for retrospective cytogenetic analysis. Cytogenet. Cell Genet. 57(1):16-7.

Baetzel, R., Wolfgang, F., Axel, V. and Lühs, W. 1999. Development of yellow-seeded high-erucic acid rapeseed (*Brassica napus* L.) Proceeding of the 10th International Conference of Rapeseed Congress, Canberra, Australia.

Barret, P., Delourme, R., Renard, M., Domergue, F., Lessire, R., Delseny, M. and Roscoe, T.J. 1998a. A rapeseed *FAE1* gene is linked to the E1 locus associated with variation in the content of erucic acid. Theor. Appl. Genet. 96:177-186.

Barret, P., Delourme, R., Foisset, N. and Renard, M. 1998b. Development of a SCAR (sequence characterized amplified region) marker for molecular tagging of the dwarf BREIZH (*Bzh*) gene in *Brassica napus* L. Theor. Appl. Genet. 97: 828–833.

Beare, J.I., Campbell, J.A., Youngs, C.G. and Craig, B.M. 1963. Effects of saturated fat in rats fed rapeseed oil. Can. J. Biochem. Physiol. 41:605-612.

Bell, J.M. 1993. Factors affecting the nutritional value of canola meal: A review. Can. J. Anim. Sci. 73:679-697.

Bell, J.M. and Shires, A. 1982. Composition and digestibility by pigs of hull fractions from rapeseed cultivars with yellow or brown seed coats. Can. J. Anim. Sci. 62:557-565.

Bjergegaard, C., Li, P.W., Michaelsen, S., Møller, P., Otte, J. and Sørensen, H. 1994. Glucosinolates and their transformation products – compounds with a broad biological activity. Bioactive. Subst. Food Plant Origin 1: 1–15.

Boutin-Ganache, I., Raposo, M., Raymond, M. and Deschepper, C.F. 2001. M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. Biotechniques 31(1):24–28.

Bradeen, J.M. and Simon, P.W. 1998. Conversion of an AFLP fragment linked to the carrot Y2 locus to a simple, co-dominant, PCR based marker form. Theor. Appl. Genet. 97: 960–967.

Buchanan, B., Gruissem, W. and Jones, R. 2000. Biochemistry and molecular biology of plants. P. 475. eds. American Society of Plant Physiologists.

Burbulis, N. and Kott, L.S. 2005. A new yellow-seeded canola genotype originating from double low black-seeded *Brassica napus* cultivars. Can. J. Plant Sci. 85: 109–114.

Canola Council of Canada 2007a. Canadian Canola Industry. [Online Database] URL. http://www.canola-council.org/industry.html (06 May 2007).

Canola Council of Canada 2007b. Canola Oil Properties and Uses. [Online Database] URL. http://www.canola-council.org/properties.html (06 May 2007).

Carlson, S.E. 1995. The role of PUFA in infant nutrition. INFORM. 6: 940-946.

Chen, B.Y., Heneen, S.K. and Josson, R. 1988. Resynthesis of *Brassica napus* L. through interspecific hybridization between *B. alboglabra* Bailey and *B. campestris* L. with special emphasis on seed color. Plant Breed. 101:52–59.

Chen, B.Y. and Heneen, W.K. 1992. Inheritance of seed color in *Brassica campestris* L. and breeding for yellow-seeded *B. napus* L. Euphytica 59:157-163.

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

Chen, B.Y. and Meng, J.L. 1984. Preliminary observation on testa development of rape (*Brassica napus* L.) seeds (in Chinese with English summary). J. Central China Agric. Univ. 3:5-8.

Devic, M., Albert, S., Delseny, M. and Roscoe, T.J. 1997. Efficient PCR walking on plant

genomic DNA. Plant Physiol. Biochem. 35: 331-339.

Devos, K.M. and Gale, M.D. 1992. The use of random amplified polymorphic DNA in wheat. Theor. Appl. Genet. 84:567-572.

Downey, R.K., Stephensson, B.R. and Stringam, D.I. 1975. Breeding rapeseed and mustard crops. P. 157-178. In: Oilseed and Pulse Crops in Western Canada. Harapiak J.T. eds.

Ecke, W., Uzunova, M. and Weileder, K. 1995. Mapping the genome of rapeseed (*Brassica napus* L.). II. Localization of genes controlling erucic acid synthesis and seed oil content. Theor. Appl. Genet. 91:972-977.

Eskin, N.A.M., Vaisey-Genser, M., Durance-Todd, S. and Przybylski, R. 1989. Stability of low linolenic acid canola oil to frying temperatures. J. American Oil Chemists Soc. 66:1081-1084.

Eskin, N.A.M., Macdonald, B.E., Przybylski, R., Malcomson, L.J., Scarth, R., Mag, T., Ward, K. and Adolphe, D. 1996. Canola Oil. P. 1-95. Chapter 1 in Bailey's Industrial Oil and Fat Products. Fifth edition. Vol. 2. Edible oil and Fat Products: Oils and Oilseeds. eds. John Wiley & Sons, Inc. New York.

FAO 2005. FAOSTAT. [Online Database] URL. http://faostat.fao.org/site/336/default.aspx (25 February 2007).

FAO 2006-2007. FOOD OUTLOOK, Global Market Analysis. (Online Database] Oilseed, oil and oilmeals. World oilseeds and products markets at a glance. URL. http://www.fao.org/docrep/009/j8126e/j8126e05.htm (06 May 2007).

Fehling, E. and Mukherjee, K.D. 1991. Acyl-CoA elongase from higher plant (*Lunaria annua*): metabolic intermediates of very-longchain acyl-CoA products and substrate specificity. Biochim. Biophys. Acta 1082: 239–246.

Fourmann, M., Barret, P., Renard, M., Pelletier, G., Delourme, R. and Brunel, D. 1998. The two genes homologous to Arabidopsis FAE1 co-segregate with the two loci governing erucic acid content in *Brassica napus*. Theor. Appl. Genet. 96:852-858.

Frankel, E.N. 1991. Review: Recent advances in lipid oxidation. J. Sci. Food Agric. 54: 495-511.

Fukuoka, S., Inoue, T., Miyao, A., Monna, L., Zhong, H.S., Sasaki, T. and Minobe, Y. 1994. Mapping of sequence-tagged sites in rice by single strand conformation polymorphism. DNA Res. 1: 271-277.

Gabor, T. M., Korf, I., Yandell, M.D., Yeh, R.T., Gu, Z., Zakeri, H., Stitziel, N.O., Hillier L., Kwok, P., Warren, R. and Gish, A. 1999. General approach to single-nucleotide

polymorphism discovery. Genetics 23:452 – 456.

Griffiths, D.W., Birch, A.N.E. and Hillman, J.R. 1998. Antinutritional compounds in the Brassicaceae: Analysis, biosynthesis, shemistry and dietrary effects. J. Hort. Sci. Biotechnol. 73:1-18.

Gopalan, G.D., Krisnamurthy, D., Shenolikar, I.S. and Krisnamurthy, K.A.V.R. 1974. Myocardial changes in Monkeys fed on mustard oil. Nutr. Metab. 16:352-365.

Harvey, B.L. and Downey, R.K. 1964. The inheritance of erucic acid content in rapeseed (*Brassica napus* L.). Can. J. Plant Sci. 44:104-111.

Harwood, J.L. 1988. Fatty acid metabolism. Annu. Rev. Plant Physiol. 39:101-138.

Harwood, J.L. 1996. Recent advances in the biosynthesis of plant fatty acids. Biochem. Biophys. Acta 1301:7-56.

Hawk, J.A. 1982. Single gene control of seed color and hypocotyls color in turnip rape. Can. J. Plant Sci. 62:331–334.

Heneen, W.K. and Jørgensen, R.B. 2001. Cytology, RAPD, and seed color of progeny plants from *Brassica rapa-alboglabra* aneuploids and development of monosomic addition lines. Genome 44:1007–1021.

Hansen, M., Laustsen, A.M., Olsen, C.E., Poll, L. and Sørensen, H. 1997. Chemical and sensory quality of broccoli (*Brassica oleracea* L. var. *italica*). J. Food Quality 20: 441–459.

Henderson, C.A.P. and Pauls, K.P. 1992. The use of haploidy to develop plants that express several recessive traits using light-seeded canola (*B. napus*) as an example. Theor. Appl. Genet. 83:476-479.

Herne, C.M., Ghosh, S. and Todd, J.A. 1992. Microsatellite for linkage analysis of genetic traits. Trends Genet. 8:288-294.

Horrobin, D.F. 1995. Medical roles of metabolites of precursor EFA.INFORM.6:428-435.

Hu, X.J. 1988. Studies on the relationship between seed coat color and pigment content in different types of rapeseed cultivars. EUCARPIA Cruciferae News L.13:44-45.

Hunter, R.L. and Markert, C.L. 1957. Histochemical demonstration of enzymes separated by zone electrophoresis in starch gels. Science 125: 1294-1295.

James, D.W.J., Lim, E., Keller, J., Plooy, I., Ralston, E. and Dooner, H.K. 1995. Directed tagging of the Arabidopsis Fatty Acid Elongation 1 (*FAE1*) gene with the maize transposon Activator. Plant Cell 7:309–319.

Jourdren, C., Barret, P., Horvais, R., Foisset, N., Delourme, R., Renard, M. 1996. Identification of RAPD markers linked to the loci controlling erucic acid level in rapeseed. Mol. Breed. 2:61-71.

Jönsson, R. 1975. Yellow-seeded rape and turnip rape. II. Breeding for improved quality of oil and meal in yellow-seeded materials (in Swedish with English summary). Sveriges Utsadesförenings Tidsk rift 85:271-278.

Karp, A., Kresovich, S., Bhat, K.V., Ayad, W.G. and Hodgkin, T. 1997. Molecular tools in plant genetic resources conservation: a guide to the technologies; in IPGRI Technical Bulletin No. 2. International Plant Genetic Resources Institute, Rome, Italy.

Katavic, V., Mietkiewska, E., Barton, D.L., Giblin, E.M., Reed, D.W. and Taylor, D.C. 2002. Restoring enzyme activity in nonfunctional low erucic acid *Brassica napus* fatty acid elongation 1 by a single amino acid substitution. Eur. J. Biochem. 269:5625–5631.

Kinsella, J.E. 1991. α-linolenic acid: Functions and effects on linoleic acid metabolism and eicosanoid-mediated reactions. Adv. Food Nutr. Res. 35:1-184.

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

Kunst, L., Taylor, D.C. and Underhill, E.W. 1992. Fatty acid elongation 1 in developing seeds of *Arabidopsis thaliana*. Plant Physiol. Biochem. 30 : 425-434.

Lahogue, F., This, P. and Bonquet, A. 1998. Identification of a co-dominant SCAR marker linked to the seedlessness character in grapevine. Theor. Appl. Genet. 97: 950–959.

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

Lands, W.E.M. 1997. The two faces of essential fatty acids. INFORM 8:1141-1147.

Lassner, M.W., Lardizabal, K. and Metz, J.G. 1996. A jojoba b-ketoacyl-CoA synthase cDNA complements the canola fatty acid elongation 1 mutation in transgenic plants. Plant Cell 8:281–292.

Lemieux, B., Miquel, M., Somerville, C. and Browse, J. 1990. Mutants of Arabidopsis with alterations in seed lipid fatty acid composition. Theor.Appl.Genet. 80:234-240.

Lepiniec, L., Debeaujon, I., Routaboul, J., Baudry, A., Pourcel, L., Nesi, N. and Caboche, M. 2006. Genetics and Biochemistry of Seed Flavonoids. Annu. Rev. Plant Biol. 57:405–30.

Leung, J., Fenton, T.W., Mueller, M.M. and Clandinin, D.R. 1979. Condensed tannins of rapeseed meal. J. Food Sci. 44:1313-1316.

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.

Li, J., Chen, L., Liang, Y., Ye, X. and Liu, L. 2003. Research and commercial application of the complete dominance yellow-seeded gene in *Brassica napus* L. Proceedings of the 11th International Rapeseed Congress, Copenhagen, Denmark. 1:202-204.

Lin, X., Kaul, S., Rounsley, S., Shea, T. and Benito, M.I. 1999. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. Nature 402:761–767.

Lincoln, S., Daly, M. and Lander, E. 1992. Constructing genetic maps with MAPMAKER/EXP 3.0. Whitehead Institute Technical Report, 3rd edn. Whitehead Technical Institute, Cambridge, Mass.

Liu, H.L. 1983. Studies on the breeding of yellow seeded *Brassica napus* L.. Proceedings of the 6th International Rapeseed Congress, Paris, France. 1:637-641.

Liu, H.L., Han, J.X. and Hu, X.J. 1991. Studies on the inheritance of seed coat color and other related characteristics of yellow seeded *Brassica napus*. Proceedings of the 8th International Rapeseed Congress, Saskatoon, Canada. 5:1438-1444.

Liu, H.L. 1992. Studies on the inheritance of yellow-seeded *Brassica napus* L.. Acta Agronomica Sinica 18: 241–248.

Liu, X.P., Tu, J.X., Chen, B.Y. and Fu, T.D. 2005. Identification and inheritance of a partially dominant gene for yellow seed color in *Brassica napus.* Plant Breed. 124 (1):9–12.

Liu, Z., Fu, T.D., Tu, J. and 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(2):303-10.

Liu, Z., Fu, T.D., Wang, Y., Tu, J., Chen, B.Y., Zhou, Y., Ma, C. and Shan, L. 2006. Development of SCAR and CAPS markers for a partially dominant yellow seed coat gene in *Brassica napus* L.. Euphytica 149:381-385.

Lühs, W., Baetzel, R. and Friedt, W. 2000. Genetic analysis of seed color in rapeseed (*Brassica napus* L.). J. Genet. Plant Breed. 36:111-115.

Mahmood, T., Ekuere, U., Yeh, F., Good, A.G. and Stringam, G.R. 2003. RFLP linkage analysis and mapping genes controlling the fatty acid profile of *Brassica juncea* using reciprocal DH populations. Theor. Appl. Genet. 107:283-290.

Mahmood, T., Rahman, M.H., Stringam, G.R., Raney, P.J. and Good, A.G. 2005. Molecular markers for seed color in *Brassica juncea*. Genome 48:755-760.

Marles, S.M.A., Gruber, M.Y., Scoles, G.J. and Muir, A.D. 2003. Pigmentation in the developing seed coat and seedling leaves of *Brassica carinata* is controlled at the dihydroflavonol reductase locus. Phytochemistry 62:663–672.

Matthäus, B. 1998. Effect of dehulling on the composition of antinutritive compounds in various of rapeseed. Fett/Lipid 100:295-301.

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

Millar, A.A. and Kunst, L. 1997. Very-long-chain fatty acid biosynthesis is controlled through the expression and specificity of the condensing enzyme. Plant J. 12 :121-131.

Mohammad, A., Sikka, S.M. and Aziz, M.A. 1942. Inheritance of seed color in some oleiferous *Brassiceae*. Indian J. Genet. 2:112-127.

Murphy, I.S. and Sonntag, J.G. 1991. Erucic, behenic: feedstocks of the 21st century. Inform 2:449-463.

Murphy, D.J. 1999. Production of novel oils in plants. Curr. Opin. Biotect. 10:175-180.

Naczk, M., Amarrowicz, R., Sullivan, A. and Shahidi, A. 1998. Current research developments on polyphenolics of rapeseed/canola: a review. Food Chemis. 62:489-502.

Naqvi, N.I. and Chattoo, B.B. 1996. Development of a sequence characterized amplified region (SCAR) based indirect selection method for a dominant blast-resistance gene in rice. Genome 39:26–30.

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

Nirupma, P., Schowinsky, V., Kokanovic, O., Magnuson, V. and Ghosh, S. 2004. A comparison between SNaPshot, pyrosequencing, and biplex invader SNP genotyping methods: accuracy, cost, and throughput. J. Biochem. Biophys. Methods 60:1-12.

Nørbæk, R., Nielsen, J.K. and Kondo, T. 1999. Flavonoids from owers of two *Crocus chrysanthus-biflorus* cultivars: `Eye-Catcher' `Spring Pearl' (Iridaceae). Phytochemistry 51:1139-1146.

Ohlrogge, J. and Browse, J. 1995. Lipid biosynthesis. The Plant Cell 7:957-970.

Olsson, G. 1960. Species crossing within the genus Brassica. II. Artificial *Brassica napus* L.. Hereditas 46:357-396.

Padmaja, K.L., Arumugam, N., Gupta, V., Mukhopadhyay, A., Sodhi, Y.S., Pental, D. and Pradhan, A.K. 2005. Mapping and tagging of seed coat color and the identification of markers for marker-assisted manipulation of the trait in *Brassica juncea* microsatellite. Theor. Appl. Genet. 111(1):8-14.

Paran, I. and Michelmore, R.W. 1993. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor. Appl. Genet. 85:985-993.

Pinkel, D., Strume, T. and Gray, G.W. 1986. Cytogenetic analysis using quantitative, high-sensitively fluorescence hybridization. Proc. Natl. Acad. Sci. USA 83:2934.

Prevot, A., Perrin, J.L., Laclaverie, G., Auge, P. and Coustille, J.L. 1990. A new variety of low linolenic rapeseed oil; characteristics and room-odor tests. J. Am. Oil Chem. Soc. 67:161-164.

Princen, L.H. and Rothfus, J. 1984 Development of new crops for industrial raw materials. J. Am. Oil Chem. Soc. 61:281-289.

Przybylski, R., Malcolmson, L.J., Eskin, N.A.M., Durance-Todd, S., Mickle, J. and Carr, R. 1993. Stability of low linolenic acid canola oil to accelerated storage at 60ºC. Lebensmittel-Wissenschaft und -Technologie. Food Science and Technology 26:205-209.

Puyaubert, J., Wilfrid, D., Patricia, C., Sylvette, C., Annick, B. and Rene, L. 2005. Temporal gene expression of 3-ketoacyl-CoA reductase is different in high and low erucic acid *Brassica napus* cultivars during seed development. Biochimica et Biophysica Acta 1687:152-163.

Qi, C.K., Fu, S.Z. and Pu, H.M. 1995. A successful transfer of yellow-seeded trait from *Brassica carinata* to *B. napus*. Proc. 9th Int. Rapeseed Congress Cambridge 4:1137-1140.

Qin, L., Prins, P., Jones, J.T. and Popeijus, H. 2001. Genest, a powerful bi-directional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP. P62 The international Conference on Plant Animal Genome IX January. San Diego, CA.

Quiros, C.F., Grellet, F., Sadowski, J., Suzuki, T., Li, G. and Wroblewski, T. 2001. Arabidopsis and Brassica comparative genomics: sequence, structure and gene content in the ABI-Rps2-Ck1 chromosomal segment and related regions. Genetics 157(3):1321-30.

Rahman, M.H. 2001a. Inheritance of petal color and its independent segregation from seed color in *Brassica rapa*. Plant Breed. 120:197-200.

Rahman, M.H. 2001b. Production of yellow-seeded *Brassica napus* through interspecific crosses. Plant Breed. 120: 463-472.

Rahman, M.H., Joersbo, M. and Poulsen, M.H. 2001. Production of yellow-seeded *Brassica napus* of double low quality. Plant Breed. 120:473-478.

Rashid, A., Rakow, G. and Downey, R.K. 1994. Development of yellow seeded *Brassica napus* through interspecific crosses. Plant Breed. 112:127-134.

Rashid, A., Rakow, G. and Downey, R.K. 1995. Agronomic performance and seed quality of black seeded cultivars and two sources of yellow seeded *Brassica napus*. Proceedings of 9th International Rapeseed Congress, Cambridge, U.K. p. 1141–1146.

Scarth, R., McVetty, P.B.E., Rimmer, S.R. and Stefansson, B.R. 1988. Stellar low linolenic-high linoleic acid summer rape. Can. J. Plant Sci. 68:509-511.

Scarth, R. and McVetty, P.B.E. 1999. Designer oil canola-A review of new food-grade Brassica oils with focus on high oleic, low linolenic types. Proceeding of the 10th International Conference of Rapeseed Congress, Canberra, Australia.

Schwetka, A. 1982. Inheritance of seed color in turnip rape (*Brassica campestris* L.). Theor. Appl. Genet. 62:161-169.

Sharma, R., Aggarwal, R.A.K., Kumar, R., Mohapatra, T. and Sharma, R.P. 2002. Construction of recombinant inbreds in mustard (*Brassica juncea*). Genome 45:467-472.

Shirley, B.W. 1998. Flavonoids in seeds and grains: physiological function, agronomic importance and the genetics of biosynthesis. Seed Sci. Res. 8:415–22.

Shirzadegan, M. 1986. Inheritance of seed coat color in *Brassica napus* L.. Z Pflanzenzuecht 96:140-146.

Shirzadegan, M. and Röbbelen, G. 1985. Influence of seed color and hull proportions on quality properties of seeds in *Brassica napus* L. Fette. Seifen. Anstrichm. 87:235-237.

Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A. and Lukyanov, S.A. 1995. An improved method for walking in uncloned genomic DNA. Nuc. Acids Res. 23:1087–1088.

Slominski, B.A., Campbell, L.D. and Guenter, W. 1994. Carbohy-drates and dietary fiber components of yellow-seeded and brown-seeded canola. J. Agric. Food Chem. 42:704-707.

Slominski, B.A., Simbaya, J., Campbell, L.D., Rakow, G. and Guenter, W. 1999. Nutritive value for broilers of meals derived from newly developed varieties of yellow-seeded canola. Anim. Feed Sci. Technol. 78:249-262.

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

Song, K.M., Obsorn, T.C. and Williams, P.H. 1988. Brassica taxonomy based on nuclear restriction fragment length polymorphism (RFLPs). 1. Genome evolution of diploid and amphidiploid species. Theor. Appl. Genet. 75:784–794.

Staub, J., Bacher, J. and Peotter, K. 1996. Sources of potential errors in the application of Random Amplified Polymorphic DNAs in cucumber. Hort. Sci. 31:262–266.

Stefansson, B.R. and Hougen, F.W. 1963. Selection of rape plants (*Brassica napus*) with seed oil practically free from erucic acid. Can. J. Plant Sci. 44: 359-364.

Stefansson, B.R. 1983. The development of improved rapeseed cultivars. p. 143-159. In Kramer JKG, Sauer FD, Pigden WJ (eds.) High and low erucic acid rapeseed oils. Academic Press, New York.

Stefansson, B.R. and Kondra, Z.P. 1975. Tower summer rape. Can. J. Plant Sci. 55:343-344.

Strickberger, M.W. 1976. Genetics, 2nd edn. Macmillan Publ. Co., Inc., New York.

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

Stringam, G.R., McGregor, D.I. and Pawlowski, S.H. 1974. Chemical and morphological characteristics associated with seed-coat color in rapeseed. Proceeding of the 4th International Rapeseed Conference, Giessen p. 99-108.

Sun, Z., Wang, Z., Tu, J., Zhang, J., Yu, F., McVetty, P.B.E. and Li, G. 2007. An ultra dense genetic recombination map for *Brassica napus*, consisting of 13551 SRAP markers. Theor. Appl. Genet. 114(8):1305-1317.

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

Tanhuanpää, P., Vilkki, J. and Vihinen, M. 1998. Mapping and cloning of *FAD2* gene to develop allele-specific PCR for oleic acid in spring turnip rape (*Brassica rapa* ssp.*oleifera*). Mol. Breed. 4:543–550.

Theander, O., Åman, P., Miksche, G.E. and Yasuda, S. 1977. Carbohydrates, polyphenol, and lignin in seed hulls of different colors from turnip rapeseed. J. Agric. Food Chem. 25:270-273.

Thormann, C.E., Romero, J., Mantet, J., Osborn, T.C. 1996. Mapping loci controlling the concentrations of erucic and linolenic acids in seed oil of *Brassica napus* L.. Theor. Appl. Genet. 93:282-286.

Topfer, R., Nartini, N. and Schell, J. 1995. Modification of plant lipid synthesis. Science 268:681-686.

Uppström, B. 1995. Seed chemistry. In: Kimber, and D.I. McGregor (Eds.), Brassica Oilseeds – Production and Utilization. P. 217-242. CAB, Wallingford, UK.

U, N. 1935. Genome-analysis in Brassica with special reference to the experimental formation of *B. napus* and its peculiar mode of fertilization. Japan J. Bot. 7:389–452.

Van Caseele, L., Mills, J.T., Sumner, M. and Gillespie, R. 1982. Cytological study of the palisade development in the seed coat of Candle canola. Can. J. Bot. 60:2469-2475.

Van Deynze, A.E., Beversdorf, W.D. and Pauls, K.P. 1993. Temperature effects on seed color in black- and yellow-seeded rapeseed. Can. J. Plant Sci. 73:383-387.

Van Deynze, A.E., Pauls, K.P. 1994. The inheritance of seed coat color and vernalization requirement in *Brassica napus* using double haploid population. Euphytica 74:77-83.

Van Deynze, A.E., Landry, B.S. and Pauls, K.P. 1995. The identification of restriction fragments length polymorphisms linked to seed color genes in *Brassica napus*. Genome 38:534–542.

Vaughan, J.G. 1970. Cruciferae. P. 49-62. In: J. G. Vaugan (ed.), The Structure and Utilization of Oilseeds. Chapman and Hall, London.

Velasco, L. and Möllers, C. 1998. Non-destructive assessment of sinapic acid esters in Brassica species: II. Evaluation of germplasms and identification of phenotypes with reduced levels. Crop Sci. 38:1650-1654.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Freijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. 1995. AFLP a new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407-4414.

Weber, D. and Helentjaris, T. 1989. Mapping RFLP loci in maize using B-A translocations. Genetics 121:583–590.

Williams, G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18:6531–6535.

Xiaoming, B., Pollard, M. and Ohlrogge, J. 1998. The Biosynthesis of Erucic Acid in Developing Embryos of *Brassica rapa*. Plant Physiol. 118:183–190.

Zaman, M.W. 1987. Potential for species introgression in *Brassica napus* with special reference to earliness and seed color. Ph.D. thesis, Dept. of Plant Breeding, Institution for Vaxtforadeling, Swedish Univ. Agric. Sci., Uppsala, Sweden.

Zaman, M.W. 1989. Inheritance of seed color in *Brassica campestris*. Sveriges Utsadesförenings Tidskrift 99:205-207.

Zock, P.L., Urgert, R., Hulshof, P.J. and Katan, M.B. 1998. Dietary trans-fatty acids: a risk factor for coronary disease. Ned. Tijdschr Geneeskd. 142(30):1701-4.

APPENDICES

Appendix i

SRAP-1, SCAR, SRAP-2, SNP and SRAP-3 markers segregation in the F_2 populations of a 'MillenniUM 03 x SRYS-3' cross and in the backcross populations of a (MillenniUM 03 x SRYS-3) x SRYS-3' cross of *Brassica napus*.

Segregation of SRAP-1 (SA12BG18388) marker in F_2 families and backcross

Populations	Total	Marker association		Segregation ratio	χ^2	P
		+ ve	- ve			
F_2	248	194	54	3 : 1	1.376	0.2-0.3
Backcross	180	95	85	1 : 1	0.554	0.3-0.5

Segregation of SCAR (converted from the SRAP-1) marker in F_2 families and backcross

Populations	Total	Marker association		Segregation ratio	χ^2	P
		+ ve	- ve			
F_2	188	137	51	3 : 1	0.453	0.5-0.7
Backcross	173	90	83	1 : 1	0.283	0.5-0.7

Segregation of SRAP-2 (SA7BG29245) marker in F_2 families and backcross

Populations	Total	Marker association		Segregation ratio	χ^2	P
		+ ve	- ve			
F_2	217	165	52	3 : 1	0.124	0.7-0.9
Backcross	175	92	83	1 : 1	0.463	0.3-0.5

Segregation of SNP (converted from the SRAP-2) marker in F_2 families and backcross

Populations	Total	Genotypes			Segregation ratio	χ^2	P
		Y2Y2	Y2y2	y2y2			
F_2	176	40	95	41	1 : 2 : 1	1.125	0.2-0.3
Backcross	168		90	78	1 : 1	0.857	0.3-0.5

Segregation of SRAP-3 (DC1GA27197) marker in F_2 families and backcross

Populations	Total	Marker association		Segregation ratio	χ^2	P
		+ ve	- ve			
F_2	238	172	66	3 : 1	0.746	0.3-0.5
Backcross	154	85	69	1 : 1	1.662	0.1-0.2

Appendix ii

Flanking sequences of the seed coat color SRAP marker (SA12BG18388) in the black seeded lines of *B. napus.*

Black seeded *B. napus* (1282 bp)

TAAACGTCCATATTTGTTCTTTTACGACTATTCCTCTTCGTCTTCACCACTACG
CGGTAACTGTAATCCACCCTTATGAATCTGTAATGCTCTTAAATCCTCTTTTA
ATCCGACGTCCCAGCTCCCCCTCTTTCTCCTTCTATATAAAGCCCTAATCTGAT
GTTTAATCCAAACCCTTATTCATAACCTCCACTGAAGAAGGTACAGCGACCG
CCATGAAATCCACAGAAAAGTCTCCAGTCCCCAGCCGCAAGCCGAAATTGAT
CTCCTTTTTCGGTTAATCTATTTCTCAGAGGCAAGAAACATTGCAAAAGGCAG
GACCTTTATTGGAACATTGCAAAAGGCAGGACCTTTATTGGGTTGGAGCTCCT
CCAATACGACGAACAGGTACGAGTGCTTATATCCAACTTTTTTCCATCGTAGT
TCATTGAGTTTTGGGTTTGATACAAAATGAGATTAAACAAGGCCGTGAATTTT
GGGTTTAATATCTACTCTGTCAGGGGACTGTGATGCATGGGTTCATCTCCTCT
ACTCGTGCTCCACGATGCTCGCCTCATCCTAAAGCAGAAACAACTTACAGCC
TTCTGAAATTTATTGAGATGCAAAAAGATTTATTGAGTTGCTGATCAGAGTTT
AACAATTTCTTTTTTTGAATGAAAAGTTTAACAATATCCCTCTTGCACAACTTT
GTTCTCTTCGTGCTTGACGGCATCAGCCTCTCTATTGACGCAGACAGGTTCAG
GTTCCGCTCCAACAACAGCATCATAATTCTCTCTCCCTTTCTCTCATGATGACT
GGCCGTATAACATATACCATCATTGACCCCATTGCTTGGCTTCCACAGTTCAG
GTTTGCCATAGATTTTTATAGTAGTAGTAGAACTATAAAAGCAATGTTAATGC
TTAAATCTCACTATAATTATTATACGTTACAGCACCAAAGGCATGATACTTAT
CAGTTTTTCTTCAACCACCAGGCGCACCAGATCTAATGGGAAACACGGTCTAT
TACAAATAGTCTGTTGATACATTTTATTTGATTTCAGTAGTTCACGTATTAGG
ACGAAACATGCCTCACATTAACATGTCTATGTGACATGATCTTTGTGGAACAG
CCTGAGAGACTTTTCCAAAGGAATTAGTAAGATTTCTGCAGTGAAGCTTTGGC
TACTACTAGGCCTCGTGCTACTCCAACGGCTCGTACTATGGAGATAGAGCAA
GACTTACTAACCAGGTCATTTAATCTACTCTTTTACCAGCCCGGGCCGTCGAC
CACGCGTCCC

Appendix iii

Flanking sequences of the seed coat color SRAP marker (SA7BG29245) in the black seeded and the yellow seeded lines of *B. napus.*

Black seeded *B. napus* (1746 bp)

TAAAGTTGTGGTCTCTGATCGATCCTCATGTCGTGAGGGGTTTGCCTTCTCGG
AGAGGTAGGAAGAGAAAAATGAGGTCTCCCAGTGGCGATATGAGAGTTGCC
GACGATAATGTTCAGATGGTTCCGCATCTTAGTTATGTTGCTCGTGCTGACGC
GCCCGCTACATATTTGCCTAAAATCCCTCATGTTGGTCATGCGCTTCATAATG
CTTCCAAATCGGGAAATCTCCTGCTGGACCGATTATATTTGTTCGGGTATCTT
TGTTTTCTCCAGATCGCTCTGATCTTCCTTTTCGTCATCCACCAGTGCTTAGTG
GGTCTCCGCATGCATACATGCCTTTTACAGATGAGGCGAAGAAGTTGAGCAT
GCTCGCTTGGTTGATTCTTTGGGTTACTCGCATGGCTTCATTTTTAGGTATAAT
GGCAAGGATGGCTTTATTTCTGATTGATAGGCGTGCAATCGCTTATGTCATAG
ATTATTGAAGTATGTAGGTGCTTCTGAGAAATCTTCGTTTATTTTCAAGAATT
CATATGGAGATCTCGTGATTGCTACTTTTCGTGTAAGCTCTTTGTTTTCATATG
CTATTTTATGACGCATTGATGTCCTGGTTTCTTACTATTTGTTTTCCTCCTCTCC
CCCTCTTTTTAGTTAGAGGGTCGGGAGATTCTGCTTATGACGGGCAAGGATAA
GAGAGTGAAGAACTTGGATTGATCATTGGAACGTATGACGGAAGAGGCCGA
GCAAGATTGATTGGATGCTTGGTTGTCGCAATAGAAGGCTGATCCTCTCAAG
GCTCACATTGTCGATATGCCTCAAGCTCTTGAGAATTGAGAGACAAAGCTTC
GTGCAGTGGATAGCTCTCGGCAGGACATCCAAGCCGCTCAAGTGGAGCTTCA
AGAGATTCAGGCCGAGGTGGTTTGATTGTGCTCCGAGATTGATCTGGTGAAG
GCCGCTAATAAGGCCCTCGAGGAAGAGAAAGAGAGGCTAAATTCGCGGAGC
TAAAGGAGTCAAATAATCGTTGCGTTGTGGTTGAGCGCTCAGTTGCACGGAA
GTTGGTCAAATCAGTTTATGAGCCTCGCTTTGGTTGGGTTCGCCGATATGCCG
CCGATCAGGAACAGGTTTGGGTGTCGCATGTTGAATTTCACCGTATGTCCGGT
TCTCTTGGCTACCTTGGTGTTGTCAAGAACGAGCATGTTTTTCCGGTCCATTC
AGATTTTGTGGCGGGTCTTGAAATCAAAGAAAGCTCGTATCTGGTGAAGGTG
AATTCCATGGATGAGCCCACCCTTGACGTTGAGGATTTAGAAGTTTCTCACGA
TGCGGTGCTCAGATCATCTAGCGATGTGCCATCAGTTTCTTCTCGTTCTTGAG
TGTATTGTATATGTTGTCTCTTTTTTTGTATCATTCTTTTCTATATAAATTGTTC
ATTATTTTTATGCTTCGTTCGTTCTTCATAGTTTTTTAGTGTGTATGTACGCTTT
GTAAGCAACTCTTTTGTTTGTGGATTCAATTATGCCACTTCGCTTCAAGGTTC
GTTTACGACTTTGCCTTGGCTAAAATAAATCGAGCGGGTGATCAGTCCGCTAT
GCTCATAGGGAGCGTGGTTTCGCGGATTAAATTGGGCCACTGTCTATGTGTTC
GTTCAACGTTGACAATTAATTTTTATTCTTGAGACAGTTCTCGTTATGGCGAT
CTATCCCTCGTTCCTCGAGGAACTTCAGCGTAAGGGAGGCATTCCTAATTAAG
A

Yellow seeded *B. napus* (1746 bp)

TAAAGTTGTGGTCTCTGATCGATCCTCATGTCGTGAGGGGTTTGCCTGCTCGG
AGAGGTAGGAAGAGAAAAATGAGGTCTCCCAGTGGCGATATGAGAGTTGTC
GACGATAATGTTCAGATGGTTCCGCATCTTAGTTATGTTGCTTGTGCTGACGC
GCCCGCTAAATATTTGCCTAAAATCCCTCATGTTGGTCATGCGCATCATAATG
CTTCCAAATCGGCAAATCTCCTGCTGGACCGATGATATTTGTTCGGGTATCTT
TGTTTTCTCCTGATCGCTCTGATCTTCCTTTTCGTCATCCACCAGTGCTTAGTG
GGTCTCCGCATGCATACATGCCTTTTACAGCTGAGGCGAAGAAGTTGAGCAT
GCTCGCTTGGTTGATTCTTTGGGTTACTCGCATGGCTTCATTTTTAGGCATAAT
GGCAAGGATGGCTTTATTTCTGATTGATAGGCGTGCAATCGCTTATGTCATAG
ATTATTGAAGTATGTAGGTGCTTCTGAGAAATCTTCGTTTATTTTTAAGAATT
CATATGGAGATCTCGTGATTGCTACTTTTCGTGTAAGCTCTTTGTTTTCATATG
CTATTTTATGACGCATTGATGTCCTGGTTTCTTACTATTTGTTTTCCTCCTCTCC
CCCTCTTTTTAGTTAGAGGGTCGGGAGATTCTGCTTATGACGGGCAAGGATAA
GAGAGTGAAGAACTTGGATTGATCATTGGAACGTATGACGGAGGAGGCCGA
GCAAGATTGATTGGATGCTTGGTTGTCGCAATAGAAGGCTGATCCTCTCAAG
GCTCACATTGTCGATATGCCTCAAGCTCTTGAGAATTTAGAGACAAAGCTTCG
TGCAGTGTATAGCTCTCGGCAGGACATCCAAGCCGCTCAAGTGGAGCTTCAA
GAGATTCGGGCCGAGGTGGTTTGATTGTGCTCCGAGATTGATCTGGTGAAGG
CCGCTAATAAGGTCCTCGAGGAAGAGAGAGAGAGGCTAAATTCGCGGAGCT
AAAGGAGTCAAATAATCGTTGCGTTGTGGTTGAGCGCTCAGTTGCATGGAAG
TTGGTCAAATCAGTTTATGAGCCTCGCTTTGGTTGGGTTCGCCGATATGCCGC
CGATCAGGAACAGGTTCGGGTGTCGCATGTTGAATTTCACCATATGTCCGGCT
CTCTTGGCTACCTTGGTGTTGTCAAGAACGAGCATGTTTTGCCGGTCCATTCA
GATTTTGTGGCGGGTCTTGAAATCAAAGAGAGCTCGTATCTGGTGAAGGTGA
ATTCCATGGATGAGCCCACCCTTGACGTTGAGGATTTAGAAGTTTCTCACGAT
GCGGTGCTCAGATCATCTAGCGATGTGCCATCAGTTTCTTCTCGTTCTTGAGT
GTATTGTATATGTTGTCTCTTTTTTTGtATCATTCTTTTCTATATAAATTGTTCA
TTATTTTTATGCTTCGTTCGTTCTTCATAGTTTTTTAGTGTGTATGTACGCTTTG
TAAGCAACTCTTTTGTTTGTGGATTCAATTATGCCACTTCGCTTCAAGGTTCGT
TTACGACTTTGCTTTGGCTAAAATAAATCGAGCGGGTGATCAGTCCGCTATGC
TCATAGGGAGCGTGGTTTCGCGGATTAAATTGGGCCACTGTCTATGTGTTCGT
TCAACGTTGACAATTAATTTTTATTCTTGAGACAGTTCACGTTATGGCGATCT
ATCCCTCGTTCCTCGAGGAACTTCAGCGTAAGGGAGGCATTCCTAATTAAGA

Appendix iv

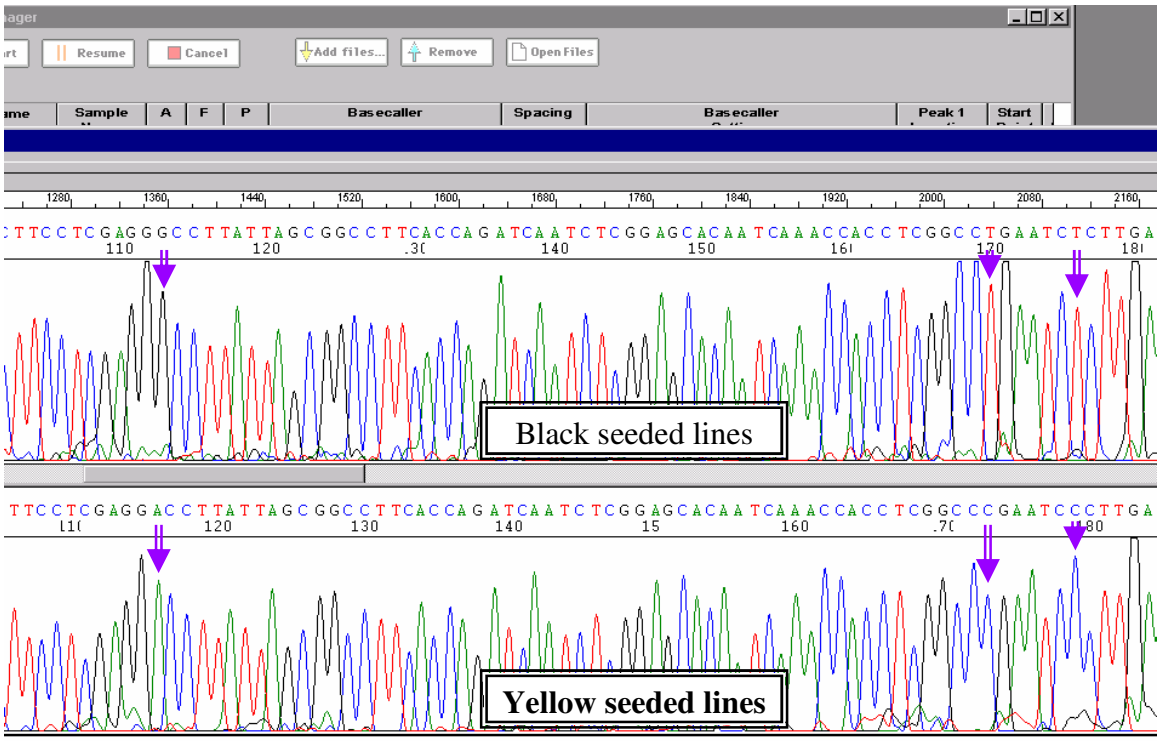
Sequence alignment of the flanking sequences from the SRAP marker SA7BG29245 between the black seeded and the yellow seeded *B. napus* lines.

```
BSnapus         TAAAGTTGTGGTCTCTGATCGATCCTCATGTCGTGAGGGGTTTGCCTTCTCGGAGAGGTA 60
YSnapus         TAAAGTTGTGGTCTCTGATCGATCCTCATGTCGTGAGGGGTTTGCCTGCTCGGAGAGGTA 60
                ********************************************** ************

BSnapus         GGAAGAGAAAAATGAGGTCTCCCAGTGGCGATATGAGAGTTGCCGACGATAATGTTCAGA 120
YSnapus         GGAAGAGAAAAATGAGGTCTCCCAGTGGCGATATGAGAGTTGTCGACGATAATGTTCAGA 120
                ****************************************** ****************

BSnapus         TGGTTCCGCATCTTAGTTATGTTGCTCGTGCTGACGCGCCCGCTACATATTTGCCTAAAA 180
YSnapus         TGGTTCCGCATCTTAGTTATGTTGCTTGTGCTGACGCGCCCGCTAAATATTTGCCTAAAA 180
                ************************** ****************** **************

BSnapus         TCCCTCATGTTGGTCATGCGCTTCATAATGCTTCCAAATCGGGAAATCTCCTGCTGGACC 240
YSnapus         TCCCTCATGTTGGTCATGCGCATCATAATGCTTCCAAATCGGCAAATCTCCTGCTGGACC 240
                ********************* ******************* ******************

BSnapus         GATTATATTTGTTCGGGTATCTTTGTTTTCTCCAGATCGCTCTGATCTTCCTTTTCGTCA 300
YSnapus         GATGATATTTGTTCGGGTATCTTTGTTTTCTCCTGATCGCTCTGATCTTCCTTTTCGTCA 300
                *** ***************************** **************************

BSnapus         TCCACCAGTGCTTAGTGGGTCTCCGCATGCATACATGCCTTTTACAGATGAGGCGAAGAA 360
YSnapus         TCCACCAGTGCTTAGTGGGTCTCCGCATGCATACATGCCTTTTACAGCTGAGGCGAAGAA 360
                ********************************************** ************

BSnapus         GTTGAGCATGCTCGCTTGGTTGATTCTTTGGGTTACTCGCATGGCTTCATTTTTAGGTAT 420
YSnapus         GTTGAGCATGCTCGCTTGGTTGATTCTTTGGGTTACTCGCATGGCTTCATTTTTAGGCAT 420
                ************************************************************ **

BSnapus         AATGGCAAGGATGGCTTTATTTCTGATTGATAGGCGTGCAATCGCTTATGTCATAGATTA 480
YSnapus         AATGGCAAGGATGGCTTTATTTCTGATTGATAGGCGTGCAATCGCTTATGTCATAGATTA 480
                ************************************************************

BSnapus         TTGAAGTATGTAGGTGCTTCTGAGAAATCTTCGTTTATTTTCAAGAATTCATATGGAGAT 540
YSnapus         TTGAAGTATGTAGGTGCTTCTGAGAAATCTTCGTTTATTTTTAAGAATTCATATGGAGAT 540
                ****************************************** ******************

BSnapus         CTCGTGATTGCTACTTTTCGTGTAAGCTCTTTGTTTTCATATGCTATTTTATGACGCATT 600
YSnapus         CTCGTGATTGCTACTTTTCGTGTAAGCTCTTTGTTTTCATATGCTATTTTATGACGCATT 600
                ************************************************************

BSnapus         GATGTCCTGGTTTCTTACTATTTGTTTTCCTCCTCTCCCCCTCTTTTTAGTTAGAGGGTC 660
YSnapus         GATGTCCTGGTTTCTTACTATTTGTTTTCCTCCTCTCCCCCTCTTTTTAGTTAGAGGGTC 660
                ************************************************************

BSnapus         GGGAGATTCTGCTTATGACGGGCAAGGATAAGAGAGTGAAGAACTTGGATTGATCATTGG 720
YSnapus         GGGAGATTCTGCTTATGACGGGCAAGGATAAGAGAGTGAAGAACTTGGATTGATCATTGG 720
                ************************************************************

BSnapus         AACGTATGACGGAAGAGGCCGAGCAAGATTGATTGGATGCTTGGTTGTCGCAATAGAAGG 780
YSnapus         AACGTATGACGGAGGAGGCCGAGCAAGATTGATTGGATGCTTGGTTGTCGCAATAGAAGG 780
                ************* ************************************************

BSnapus         CTGATCCTCTCAAGGCTCACATTGTCGATATGCCTCAAGCTCTTGAGAATTGAGAGACAA 840
YSnapus         CTGATCCTCTCAAGGCTCACATTGTCGATATGCCTCAAGCTCTTGAGAATTTAGAGACAA 840
                ************************************************** ********

BSnapus         AGCTTCGTGCAGTGGATAGCTCTCGGCAGGACATCCAAGCCGCTCAAGTGGAGCTTCAAG 900
YSnapus         AGCTTCGTGCAGTGTATAGCTCTCGGCAGGACATCCAAGCCGCTCAAGTGGAGCTTCAAG 900
                ************** ************************************************

BSnapus         AGATTCAGGCCGAGGTGGTTTGATTGTGCTCCGAGATTGATCTGGTGAAGGCCGCTAATA 960
YSnapus         AGATTCGGGCCGAGGTGGTTTGATTGTGCTCCGAGATTGATCTGGTGAAGGCCGCTAATA 960
                ****** ********************************************************
```

```
BSnapus         AGGCCCTCGAGGAAGAGAAAGAGAGGCTAAATTCGCGGAGCTAAAGGAGTCAAATAATCG 1020
YSnapus         AGGTCCTCGAGGAAGAGAGAGAGAGGCTAAATTCGCGGAGCTAAAGGAGTCAAATAATCG 1020
                *** ************** ****************************************

BSnapus         TTGCGTTGTGGTTGAGCGCTCAGTTGCACGGAAGTTGGTCAAATCAGTTTATGAGCCTCG 1080
YSnapus         TTGCGTTGTGGTTGAGCGCTCAGTTGCATGGAAGTTGGTCAAATCAGTTTATGAGCCTCG 1080
                **************************** ******************************

BSnapus         CTTTGGTTGGGTTCGCCGATATGCCGCCGATCAGGAACAGGTTTGGGTGTCGCATGTTGA 1140
YSnapus         CTTTGGTTGGGTTCGCCGATATGCCGCCGATCAGGAACAGGTTCGGGTGTCGCATGTTGA 1140
                ******************************************* ****************

BSnapus         ATTTCACCGTATGTCCGGTTCTCTTGGCTACCTTGGTGTTGTCAAGAACGAGCATGTTTT 1200
YSnapus         ATTTCACCATATGTCCGGCTCTCTTGGCTACCTTGGTGTTGTCAAGAACGAGCATGTTTT 1200
                ******** ********* ****************************************

BSnapus         TCCGGTCCATTCAGATTTTGTGGCGGGTCTTGAAATCAAAGAAAGCTCGTATCTGGTGAA 1260
YSnapus         GCCGGTCCATTCAGATTTTGTGGCGGGTCTTGAAATCAAAGAGAGCTCGTATCTGGTGAA 1260
                 ****************************************** *****************

BSnapus         GGTGAATTCCATGGATGAGCCCACCCTTGACGTTGAGGATTTAGAAGTTTCTCACGATGC 1320
YSnapus         GGTGAATTCCATGGATGAGCCCACCCTTGACGTTGAGGATTTAGAAGTTTCTCACGATGC 1320
                ************************************************************

BSnapus         GGTGCTCAGATCATCTAGCGATGTGCCATCAGTTTCTTCTCGTTCTTGAGTGTATTGTAT 1380
YSnapus         GGTGCTCAGATCATCTAGCGATGTGCCATCAGTTTCTTCTCGTTCTTGAGTGTATTGTAT 1380
                ************************************************************

BSnapus         ATGTTGTCTCTTTTTTTGTATCATTCTTTTCTATATAAATTGTTCATTATTTTTATGCTT 1440
YSnapus         ATGTTGTCTCTTTTTTTGTATCATTCTTTTCTATATAAATTGTTCATTATTTTTATGCTT 1440
                ************************************************************

BSnapus         CGTTCGTTCTTCATAGTTTTTTAGTGTGTATGTACGCTTTGTAAGCAACTCTTTTGTTTG 1500
YSnapus         CGTTCGTTCTTCATAGTTTTTTAGTGTGTATGTACGCTTTGTAAGCAACTCTTTTGTTTG 1500
                ************************************************************

BSnapus         TGGATTCAATTATGCCACTTCGCTTCAAGGTTCGTTTACGACTTTGCCTTGGCTAAAATA 1560
YSnapus         TGGATTCAATTATGCCACTTCGCTTCAAGGTTCGTTTACGACTTTGCTTTGGCTAAAATA 1560
                ********************************************** ************

BSnapus         AATCGAGCGGGTGATCAGTCCGCTATGCTCATAGGGAGCGTGGTTTCGCGGATTAAATTG 1620
YSnapus         AATCGAGCGGGTGATCAGTCCGCTATGCTCATAGGGAGCGTGGTTTCGCGGATTAAATTG 1620
                ************************************************************

BSnapus         GGCCACTGTCTATGTGTTCGTTCAACGTTGACAATTAATTTTTATTCTTGAGACAGTTCT 1680
YSnapus         GGCCACTGTCTATGTGTTCGTTCAACGTTGACAATTAATTTTTATTCTTGAGACAGTTCA 1680
                ************************************************************

BSnapus         CGTTATGGCGATCTATCCCTCGTTCCTCGAGGAACTTCAGCGTAAGGGAGGCATTCCTAA 1740
YSnapus         CGTTATGGCGATCTATCCCTCGTTCCTCGAGGAACTTCAGCGTAAGGGAGGCATTCCTAA 1740
                ************************************************************

BSnapus         TTAAGA 1746
YSnapus         TTAAGA 1746
                ******
```

Appendix v

SNPs positions between the black seeded and the yellow seeded lines in *B. napus.*



Appendix vi

Seed coat color, SRAP, SNP and SCAR markers segregation in the F_2 populations of a 'SPANxBARI-6' cross and in the backcross populations of a (SPANxBARI-6)xBARI-6' cross of *B. rapa*.

Seed coat color segregation							
Population	Total (421)	Brown	Dark/light yellow	Bright yellow	Expected ratio	χ^2	P
F_2	224	164	48	12	12 : 3 : 1	1.238	0.5-0.7
Backcross	197	95	55	47	2 : 1 : 1	0.898	0.5-0.7

SRAP markers						
Population	Total (423)	+ve	- ve	Expected ratio	χ^2	P
F_2	224	165	59	3 : 1	0.215	0.5-0.7
Backcross	199	95	104	1 : 1	0.408	0.5-0.7

SNP markers							
Population	Total (423)	Black	Black/Red	Red	Expected ratio	χ^2	P
F_2	224	51	115	58	1 : 2 : 1	0.597	0.3-0.5
Backcross	199		94	105	1 : 1	0.608	0.3-0.5

SCAR markers							
Population	Total (423)	B	AB	A	Expected ratio	χ^2	P
F_2	224	51	115	58	1 : 2 : 1	0.597	0.3-0.5
Backcross	199		94	105	1 : 1	0.608	0.3-0.5

Appendix vii

Flanking sequences of the seed coat color SRAP marker (SA7BG29245) in the black seeded and the yellow seeded lines of *B. rapa.*

Black seeded *B. rapa* (2052 bp)

TGGTCTCTGATCGATCCTCATGTCGTGAGGGGTTTGCCTTCTCGGAGAGGTAG
GAAGAGAAAAATGAGGTCTCCCAGTGGCGATATGAGAGTTGCCGACGATAAT
GTTCAGATGGTTCCGCATCTTAGTTATGTTGCTCGTGCTGACGCGCCCGCTAC
ATATTTGCCTAAAATCCCTCATGTTGGTCATGCGCTTCATAATGCTTCCAAAT
CGGGAAATCTCCTGCTGGACCGATTATATTTGTTCGGGTATCTTTGTTTTCTCC
AGATCGCTCTGATCTTCCTTTTCGTCATCCACCAGTGCTTAGTGGGTCTCCGC
ATGCATACATGCCTTTTACAGATGAGGCGAAGAAGTTGAGCATGCTCGCTTG
GTTGATTCTTTGGGTTACTCGCATGGCTTCATTTTTAGGTATAATGGCAAGGA
TGGCTTTATTTCTGATTGATAGGCGTGCAATCGCTTATGTCATAGATTATTGA
AGTATGTAGGTGCTTCTGAGAAATCTTCGTTTATTTTCAAGAATTCATATGGA
GATCTCGTGATTGCTACTTTTCGTGTAAGCTCTTTGTTTTCATATGCTATTTTA
TGACGCATTGATGTCCTGGTTTCTTACTATTTGTTTTCCTCCTCTCCCCCTCTTT
TTAGTTAGAGGGTCGGGAGATTCTGCTTATGACGGGCAAGGATAAGAGAGTG
AAGAACTTGGATTGATCATTGGAACGTATGACGGAAGAGGCCGAGCAAGATT
GATTGGATGCTTGGTTGTCGCAATAGAAGGCTGATCCTCTCAAGGCTCACATT
GTCGATATGCCTCAAGCTCTTGAGAATTGAGAGACAAAGCTTCGTGCAGTGG
ATAGCTCTCGGCAGGACATCCAAGCCGCTCAAGTGGAGCTTCAAGAGATTCA
GGCCGAGGTGGTTTGATTGTGCTCCGAGATTGATCTGGTGAAGGCCGCTAAT
AAGGCCCTCGAGGAAGAGAAAGAGAGGCTAAATTCGCGGAGCTAAAGGAGT
CAAATAATCGTTGCGTTGTGGTTGAGCGCTCAGTTGCA**C**GGAAGTTGGTCAA
ATCAGTTTATGAGCCTCGCTTTGGTTGGGTTCGCCGATATGCCGCCGATCAGG
AACAGGTTTGGGTGTCGCATGTTGAATTTCACCGTATGTCCGGTTCTCTTGGC
TACCTTGGTGTTGTCAAGAACGAGCATGTTTTTCCGGTCCATTCAGATTTTGT
GGCGGGTCTTGAAATCAAAGAAAGCTCGTATCTGGTGAAGGTGAATTCCATG
GATGAGCCCACCCTTGACGTTGAGGCTTTAGAAGTTTCTCACGATCCGGTGCT
CAGATCATCTAGCGATGTGCCATCAGTTTCTTCTCGTTCTTGAGTGTATTGTAT
ATGTTGTCTCTTTTTTTGTATGATTCTTTTCTATATAAATTGTTCATTATTTTTA
TGCTTCGTTCGTTCTTCATAGTTTTTTAGTGTGTATGTACGCTTTGTAAGCAAC
TCTTTTGTTTGTGGATTCAATTATGCCACTTCGCTTCAAGGTTCGTTTACGACT
TTGCCTTGGCTAAAATAAATCGAGCGGGTGATCAGTCCGCTATGCTCATAGG
GAGCGTGGTTTCGCGGATTAAATTGGGCCACTGTCTATGTGTTCGTTCAACGT
TGACAATTAATTTTTATTCTTGAGACAGTTCTCGTTATGGCGATCTATCCCTCG
TTCCTCGAGGAACTTCAGCGTAAGGGAGGCATTCCTAATTAAGACGGTTTATC
AAAAAAAAAATTAGAAACAACTTAAATGATGTTGATTAAGAGGCCTTGAAGC
ACAACGCAAAAGTGAAAAGAACTTAGATAGGTAAATCAAATTGAAACACTA
ATAATAAACAGTTTGTCGACGGACTTTTACGGCATTCTCGGGCTTCTACGACT
TGGGCTTTTTTTGGTGGCATCTAAGGTTCTCTGTGGAACTCGTTACTGTTTGGG

CTCTTGCAAGACCTGTTGCGACCATGGCGTGCTGTGGCATTCGTGGTTTGGTA
GTCAGGTTTACTTGCCTCGTAGTTTGTAGTGTGGGCTCCACG

Yellow sarson *B. rapa* (2064 bp)

TGGTCTCTGATCGATCCTCATGTCGTGAGGGGTTTGCCTGCTCGGAGAGGTAG
GAAGAGAAAAATGAGGTCTCCCAGTGGCGATATGAGAGTTGTCGACGATAAT
GTTCAGATGGTTCCGCATCTTAGTTATGTTGCTTGTGCTGACGCGCCCGCTAA
ATATTTGCCTAAAATCCCTCATGTTGGTCATGCGCATCATAATGCTTCCAAAT
CGGCAAATCTCCTGCTGGACCGATGATATTTGTTCGGGTATCTTTGTTTTCTCC
TGATCGCTCTGATCTTCCTTTTCGTCATCCACCAGTGCTTAGTGGGTCTCCGCA
TGCATACATGCCTTTTACAGATGAGGCGAAGAAGTTGAGCATGCTCGCTTGG
TTGATTCTTTGGGTTACTCGCATGGCTTCATTTTTAGGCATAATGGCAAGGAT
GGCTTTATTTCTGATTGATAGGCGTGCAATCGCTTATGTCATAGATTATTGAA
GTATGTAGGTGCTTCTGAGAAATCTTCGTTTATTTTTAAGAATTCATATGGAG
ATCTCGTGATTGCTACTTTTCGTGTAAGCTCTTTGTTTTCATATGCTATTTTAT
GACGCATTGATGTCCTGGTTTCTTACTATTTGTTTTCCTCCTCTCCCCCTCTTTT
TAGTTAGAGGGTCGGGAGATTCTGCTTATGACGGGCAAGGATAAGAGAGTGA
AGAACTTGGATTGATCATTGGAACGTATGACGGAGGAGGCCGAGCAAGATTG
ATTGGATGCTTGGTTGTCGCAATAGAAGGCTGATCCTCTCAAGGCTCACATTG
TCGATATGCCTCAAGCTCTTGAGAATTTAGAGACAAAGCTTCGTGCAGTGTAT
AGCTCTCGGCAGGACAGTGTTAATCGTATCCAAGCCGCTCAAGTGGAGCTTC
AAGAGATTCGGGCCGAGGTGGTTTGATTGTGCTCCGAGATTGATCTGGTGAA
GGCCGCTAATAAGGTCCTCGAGGAAGAGAGAGAGAGGCTAAATTCGCGGAG
CTAAAGGAGTCAAATAATCGTTGCGTTGTGGTTGAGCGCTCAGTTGCA**T**GGA
AGTTGGTCAAATCAGTTTATGAGCCTCGCTTTGGTTGGGTTCGCCGATATGCC
GCCGATCAGGAACAGGTTCGGGTGTCGCATGTTGAATTTCACCATATGTCCG
GCTCTCTTGGCTACCTTGGTGTTGTCAAGAACGAGCATGTTTTGCCGGTCCAT
TCAGATTTTGTGGCGGGTCTTGAAATCAAAGAGAGCTCGTATCTGGTGAAGG
TGAATTCCATGGATGAGCCCACCCTTGACGTTGAGGCTTTAGAAGTTTCTCAC
GATCCGGTGCTCAGATCATCTAGCGATGTGCCATCAGTTTCTTCTCGTTCTTG
AGTGTATTGTATATGTTGTCTCTTTTTTTGTATGATTCTTTTCTATATAAATTGT
TCATTATTTTTATGCTTCGTTCGTTCTTCATAGTTTTTTAGTGTGTATGTACGCT
TTGTAAGCAACTCTTTTGTTTGTGGATTCAATTATGCCACTTCGCTTCAAGGTT
CGTTTACGACTTTGCTTTGGCTAAAATAAATCGAGCGGGTGATCAGTCCGCTA
TGCTCATAGGGAGCGTGGTTTCGCGGATTAAATTGGGCCACTGTCTATGTGTT
CGTTCAACGTTGACAATTAATTTTTATTCTTGAGACAGTTCTCGTTATGGCGA
TCTATCCCTCGTTCCTCGAGGAACTTCAGCGTAAGGGAGGCATTCCTAATTAA
GACGGTTTATCAAAAAAAAAAATTAGAAACAACTTAAATGATGTTGATTAAGA
GGCCTTGAAGCACAACGCAAAAGTGAAAAGAACCTAGATAGGTAAATCAAA
TTGAAACACTAATAATAAACAGTTTGTCGACGGACTTTTACGGCATTCTCGGG
CTTCTACGACTTGGGCTTTTTTTGGTGGCATCTAAGGTTCTCTGTGGAACTCGT
TACTGTTTGGGCTCTTGCAAGACCTGTTGCGACCATGGCGTGCTGTGGCATTC
GTGGTTTGGTAGTCAGGTTTACTTGCCTCGTAGTTTGTAGTGTGGGCTCCACG

Appendix viii

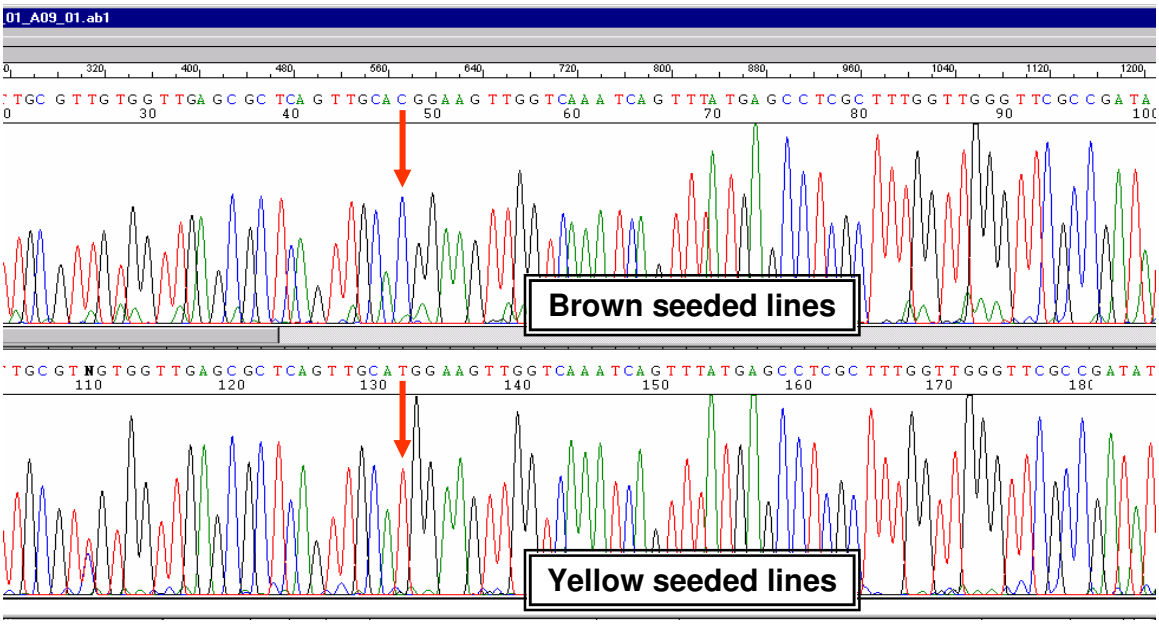
Sequence alignment of the flanking sequences from the SRAP marker SA7BG29245 between the brown seeded and the yellow seeded *B. rapa* lines.

```
BSrapa          TGGTCTCTGATCGATCCTCATGTCGTGAGGGGTTTGCCTTCTCGGAGAGGTAGGAAGAGA 60
YSrapa          TGGTCTCTGATCGATCCTCATGTCGTGAGGGGTTTGCCTGCTCGGAGAGGTAGGAAGAGA 60
                ************************************ ********************

BSrapa          AAAATGAGGTCTCCCAGTGGCGATATGAGAGTTGCCGACGATAATGTTCAGATGGTTCCG 120
YSrapa          AAAATGAGGTCTCCCAGTGGCGATATGAGAGTTGTCGACGATAATGTTCAGATGGTTCCG 120
                ********************************** ************************

BSrapa          CATCTTAGTTATGTTGCTCGTGCTGACGCGCCCGCTACATATTTGCCTAAAATCCCTCAT 180
YSrapa          CATCTTAGTTATGTTGCTTGTGCTGACGCGCCCGCTAAATATTTGCCTAAAATCCCTCAT 180
                ****************** ****************** ***********************

BSrapa          GTTGGTCATGCGCTTCATAATGCTTCCAAATCGGGAAATCTCCTGCTGGACCGATTATAT 240
YSrapa          GTTGGTCATGCGCATCATAATGCTTCCAAATCGGCAAATCTCCTGCTGGACCGATGATAT 240
                ************* ******************** ******************** ****

BSrapa          TTGTTCGGGTATCTTTGTTTTCTCCAGATCGCTCTGATCTTCCTTTTCGTCATCCACCAG 300
YSrapa          TTGTTCGGGTATCTTTGTTTTCTCCTGATCGCTCTGATCTTCCTTTTCGTCATCCACCAG 300
                ************************* ************************************

BSrapa          TGCTTAGTGGGTCTCCGCATGCATACATGCCTTTTACAGATGAGGCGAAGAAGTTGAGCA 360
YSrapa          TGCTTAGTGGGTCTCCGCATGCATACATGCCTTTTACAGATGAGGCGAAGAAGTTGAGCA 360
                ************************************************************

BSrapa          TGCTCGCTTGGTTGATTCTTTGGGTTACTCGCATGGCTTCATTTTTAGGTATAATGGCAA 420
YSrapa          TGCTCGCTTGGTTGATTCTTTGGGTTACTCGCATGGCTTCATTTTTAGGCATAATGGCAA 420
                ************************************************ ***********

BSrapa          GGATGGCTTTATTTCTGATTGATAGGCGTGCAATCGCTTATGTCATAGATTATTGAAGTA 480
YSrapa          GGATGGCTTTATTTCTGATTGATAGGCGTGCAATCGCTTATGTCATAGATTATTGAAGTA 480
                ************************************************************

BSrapa          TGTAGGTGCTTCTGAGAAATCTTCGTTTATTTTCAAGAATTCATATGGAGATCTCGTGAT 540
YSrapa          TGTAGGTGCTTCTGAGAAATCTTCGTTTATTTTTAAGAATTCATATGGAGATCTCGTGAT 540
                ********************************** ***************************

BSrapa          TGCTACTTTTCGTGTAAGCTCTTTGTTTTCATATGCTATTTTATGACGCATTGATGTCCT 600
YSrapa          TGCTACTTTTCGTGTAAGCTCTTTGTTTTCATATGCTATTTTATGACGCATTGATGTCCT 600
                ************************************************************

BSrapa          GGTTTCTTACTATTTGTTTTCCTCCTCTCCCCCTCTTTTTAGTTAGAGGGTCGGGAGATT 660
YSrapa          GGTTTCTTACTATTTGTTTTCCTCCTCTCCCCCTCTTTTTAGTTAGAGGGTCGGGAGATT 660
                ************************************************************

BSrapa          CTGCTTATGACGGGCAAGGATAAGAGAGTGAAGAACTTGGATTGATCATTGGAACGTATG 720
YSrapa          CTGCTTATGACGGGCAAGGATAAGAGAGTGAAGAACTTGGATTGATCATTGGAACGTATG 720
                ************************************************************

BSrapa          ACGGAAGAGGCCGAGCAAGATTGATTGGATGCTTGGTTGTCGCAATAGAAGGCTGATCCT 780
YSrapa          ACGGAGGAGGCCGAGCAAGATTGATTGGATGCTTGGTTGTCGCAATAGAAGGCTGATCCT 780
                ***** ******************************************************

BSrapa          CTCAAGGCTCACATTGTCGATATGCCTCAAGCTCTTGAGAATTGAGAGACAAAGCTTCGT 840
YSrapa          CTCAAGGCTCACATTGTCGATATGCCTCAAGCTCTTGAGAATTTAGAGACAAAGCTTCGT 840
                ****************************************** ****************

BSrapa          GCAGTGGATAGCTCTCGGCAGGACA------------TCCAAGCCGCTCAAGTGGAGCTT 888
YSrapa          GCAGTGTATAGCTCTCGGCAGGACAGTGTTAATCGTATCCAAGCCGCTCAAGTGGAGCTT 900
                ****** ******************            ***********************

BSrapa          CAAGAGATTCAGGCCGAGGTGGTTTGATTGTGCTCCGAGATTGATCTGGTGAAGGCCGCT 948
YSrapa          CAAGAGATTCGGGCCGAGGTGGTTTGATTGTGCTCCGAGATTGATCTGGTGAAGGCCGCT 960
                ********** ***************************************************

BSrapa          AATAAGGCCCTCGAGGAAGAGAAAGAGAGGCTAAATTCGCGGAGCTAAAGGAGTCAAATA 1008
YSrapa          AATAAGGTCCTCGAGGAAGAGAGAGAGGCTAAATTCGCGGAGCTAAAGGAGTCAAATA 1020
                ******* ************** ***************************************


BSrapa          ATCGTTGCGTTGTGGTTGAGCGCTCAGTTGCACGGAAGTTGGTCAAATCAGTTTATGAGC 1068
YSrapa          ATCGTTGCGTTGTGGTTGAGCGCTCAGTTGCATGGAAGTTGGTCAAATCAGTTTATGAGC 1080
                ******************************** ***************************

BSrapa          CTCGCTTTGGTTGGGTTCGCCGATATGCCGCCGATCAGGAACAGGTTTGGGTGTCGCATG 1128
YSrapa          CTCGCTTTGGTTGGGTTCGCCGATATGCCGCCGATCAGGAACAGGTTCGGGTGTCGCATG 1140
                ********************************************* ************
```

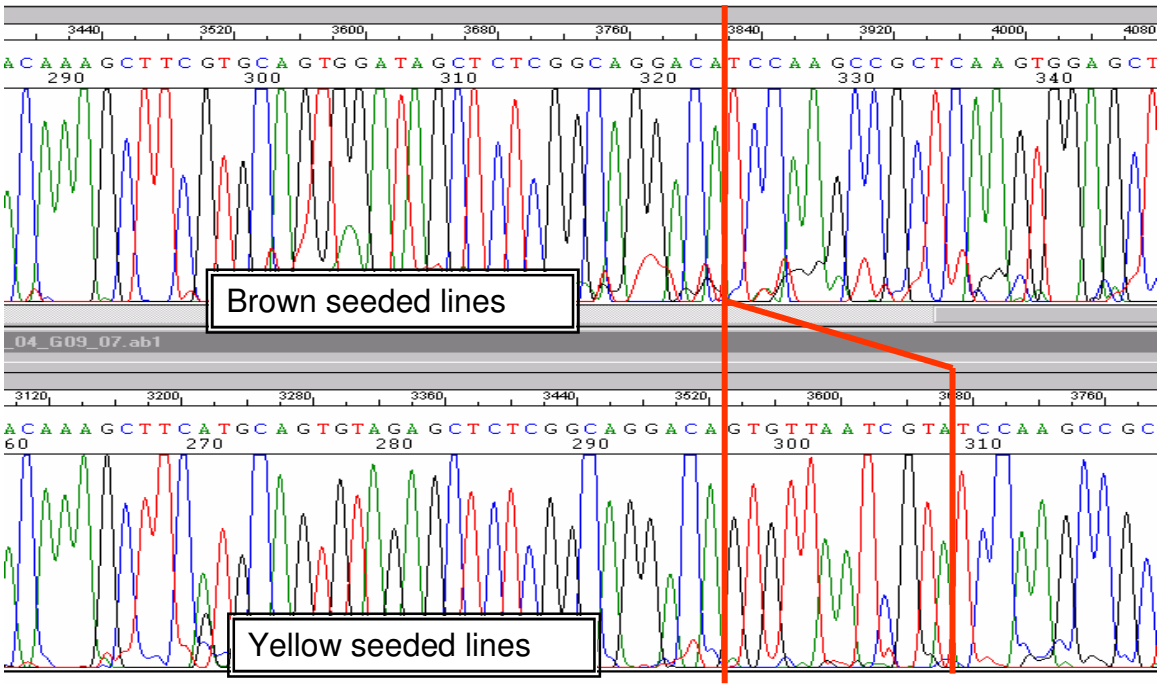
```
BSrapa          TTGAATTTCACCGTATGTCCGGTTCTCTTGGCTACCTTGGTGTTGTCAAGAACGAGCATG 1188
YSrapa          TTGAATTTCACCATATGTCCGGCTCTCTTGGCTACCTTGGTGTTGTCAAGAACGAGCATG 1200
                ************ ********* ************************************

BSrapa          TTTTTCCGGTCCATTCAGATTTTGTGGCGGGTCTTGAAATCAAAGAAAGCTCGTATCTGG 1248
YSrapa          TTTTGCCGGTCCATTCAGATTTTGTGGCGGGTCTTGAAATCAAAGAGAGCTCGTATCTGG 1260
                **** ****************************************** **************

BSrapa          TGAAGGTGAATTCCATGGATGAGCCCACCCTTGACGTTGAGGCTTTAGAAGTTTCTCACG 1308
YSrapa          TGAAGGTGAATTCCATGGATGAGCCCACCCTTGACGTTGAGGCTTTAGAAGTTTCTCACG 1320
                ************************************************************

BSrapa          ATCCGGTGCTCAGATCATCTAGCGATGTGCCATCAGTTTCTTCTCGTTCTTGAGTGTATT 1368
YSrapa          ATCCGGTGCTCAGATCATCTAGCGATGTGCCATCAGTTTCTTCTCGTTCTTGAGTGTATT 1380
                ************************************************************

BSrapa          GTATATGTTGTCTCTTTTTTTGTATGATTCTTTTCTATATAAATTGTTCATTATTTTTAT 1428
YSrapa          GTATATGTTGTCTCTTTTTTTGTATGATTCTTTTCTATATAAATTGTTCATTATTTTTAT 1440
                ************************************************************

BSrapa          GCTTCGTTCGTTCTTCATAGTTTTTTAGTGTGTATGTACGCTTTGTAAGCAACTCTTTTG 1488
YSrapa          GCTTCGTTCGTTCTTCATAGTTTTTTAGTGTGTATGTACGCTTTGTAAGCAACTCTTTTG 1500
                ************************************************************

BSrapa          TTTGTGGATTCAATTATGCCACTTCGCTTCAAGGTTCGTTTACGACTTTGCCTTGGCTAA 1548
YSrapa          TTTGTGGATTCAATTATGCCACTTCGCTTCAAGGTTCGTTTACGACTTTGCTTTGGCTAA 1560
                ************************************************** ********

BSrapa          AATAAATCGAGCGGGTGATCAGTCCGCTATGCTCATAGGGAGCGTGGTTTCGCGGATTAA 1608
YSrapa          AATAAATCGAGCGGGTGATCAGTCCGCTATGCTCATAGGGAGCGTGGTTTCGCGGATTAA 1620
                ************************************************************

BSrapa          ATTGGGCCACTGTCTATGTGTTCGTTCAACGTTGACAATTAATTTTTATTCTTGAGACAG 1668
YSrapa          ATTGGGCCACTGTCTATGTGTTCGTTCAACGTTGACAATTAATTTTTATTCTTGAGACAG 1680
                ************************************************************

BSrapa          TTCTCGTTATGGCGATCTATCCCTCGTTCCTCGAGGAACTTCAGCGTAAGGGAGGCATTC 1728
YSrapa          TTCTCGTTATGGCGATCTATCCCTCGTTCCTCGAGGAACTTCAGCGTAAGGGAGGCATTC 1740
                ************************************************************

BSrapa          CTAATTAAGACGGTTTATCAAAAAAAAAATTAGAAACAACTTAAATGATGTTGATTAAGA 1788
YSrapa          CTAATTAAGACGGTTTATCAAAAAAAAAATTAGAAACAACTTAAATGATGTTGATTAAGA 1800
                ************************************************************

BSrapa          GGCCTTGAAGCACAACGCAAAAGTGAAAAGAACTTAGATAGGTAAATCAAATTGAAACAC 1848
YSrapa          GGCCTTGAAGCACAACGCAAAAGTGAAAAGAACCTAGATAGGTAAATCAAATTGAAACAC 1860
                ********************************* ****************************

BSrapa          TAATAATAAACAGTTTGTCGACGGACTTTTACGGCATTCTCGGGCTTCTACGACTTGGGC 1908
YSrapa          TAATAATAAACAGTTTGTCGACGGACTTTTACGGCATTCTCGGGCTTCTACGACTTGGGC 1920
                ************************************************************

BSrapa          TTTTTTTGGTGGCATCTAAGGTTCTCTGTGGAACTCGTTACTGTTTGGGCTCTTGCAAGA 1968
YSrapa          TTTTTTTGGTGGCATCTAAGGTTCTCTGTGGAACTCGTTACTGTTTGGGCTCTTGCAAGA 1980
                ************************************************************

BSrapa          CCTGTTGCGACCATGGCGTGCTGTGGCATTCGTGGTTTGGTAGTCAGGTTTACTTGCCTC 2028
YSrapa          CCTGTTGCGACCATGGCGTGCTGTGGCATTCGTGGTTTGGTAGTCAGGTTTACTTGCCTC 2040
                ************************************************************

BSrapa          GTAGTTTGTAGTGTGGGCTCCACG 2052
YSrapa          GTAGTTTGTAGTGTGGGCTCCACG 2064
                ************************
```

Appendix ix

SNP and a 12bp deletion positions between the brown seeded and the yellow seeded lines of *B. rapa.*



Sequence alignment showing a SNP position between the brown-seeded lines and the yellow seeded lines of *B. rapa.*



Sequence alignment showing a 12 bp deletion position between the brown-seeded lines and the yellow seeded lines of *B. rapa.*