Development of Linkage Map of *Brassica juncea* using molecular markers and detection of Quantitative Trait Loci for oil content, seed protein and fatty acids

BY

ROGER N. W. WATTS

A Thesis

Submitted to the Faculty of Graduate Studies

in Partial Fulfillment of Requirements

for the Degree of

MASTER OF SCIENCE

Department of Plant Science

University of Manitoba

Winnipeg, Manitoba

Table of contents

Abstract

A genetic linkage map of mustard (*Brassica juncea*) was developed using two double haploid populations produced from crosses between a low erucic cultivar "ZEM1" and two moderate erucic acid lines "Vniimk351" and "Vniimk405" with the use of SSR and SRAP markers. The linkage map of the ZEM1xVniimk351 population included 13 linkage groups with an overall length of 791 cM with an average marker interval of 5.7 cM. The linkage map of the ZEM1xVniimk405 population also contained 13 linkage groups with a distance of 623 cM and an average marker interval of 4.6 cM. Using the linkage maps for the two populations, QTLs were detected for seed oil, protein and fatty acids. QTL analysis for fatty acids indentified QTLs on LG1, 7 and 12 for the ZEM1xVniimk351 population and LG1, 3 and 4 for the ZEM1xVniimk405 population. Analysis for the seed oil and protein content in the ZEM1xVniimk351 population identified 2 QTLs on LG1 and LG4 and 1 QTL on LG1 respectively. The QTL analysis ZEM1xVniimk405 of oil and protein content identified 1 QTL for oil and protein on LG1. The variation of fatty acids was shown to be the result of monogenic inheritance of the *FAE1* gene in both populations.

Acknowledgements

I would like to thank my supervisor Dr. Muhammad Tahir and my committee members Dr. Genyi Li and Dr. Dana Schroeder for their help and direction throughout my research. This project was a collaborative effort with the University of Manitoba and DOW Agrosciences Ltd. through the funding of the NSERC Industrial Fellowship. I would like to thanks both Gary Turnbull and Greg Gingera of DOW for their support in making this research project happen.

I would like to thank the Plant Science staff that helped with the completion of this project. Our lab technician Ralph Kowatsch offered his experience and guidance which both help and encouraged me throughout my research. I would also like to thank the field and support staff Lynn Coleman, Judith Nugent-Rigby and Robert Smith for their help with the execution of the field trial and Ian Brown and Cathy Bay in the greenhouse for their assistance throughout my research. Additionally I would like to thank Debbie Witko and Dave Audette in the seed quality lab for the help they provided. Thank you to my fellow graduate students, Sean Asselin, Tyler Guerrieri, and Nasir Javed for their help and support through my research. A special thanks to Dr. Peter McVetty, Bev Godard and Martha Blouw for their guidance and help in the completion of this project.

Finally I would like to thank my wife Debbie and our families for their ongoing support throughout my research. Without them none of this would be possible.

List of Figures

List of Tables

List of Abbreviations

ABI – Applied Biosystems Incorporated

AFLP – Amplified Fragment Length Polymorphism

bp – Base pair

cM – centiMorgan

DH – Double Haploid

DNA - Deoxyribonucleic acid

FA – Fatty Acid

FAD – Fatty Acid Desaturase

FAE – Fatty Acid Elongase

GC – Gas Chromatography

ISSR – Inter-Simple Sequence Repeats

LG – Linkage Group

LDL - Low Density Lipoprotein

MUFA – Monounsaturated Fatty Acid

NIR – Near Infrared Spectroscopy

PUFA – Polyunsaturated Fatty Acid

QTL – Quantitative Trait Loci

RAPD – Random Amplified Polymorphism

RFLP – Restriction Fragment Length Polymorphism

SCAR – Sequence Characterized Amplified Region

SFA – Saturated Fatty Acid

SNP – Single Nucleotide Polymorphism

SRAP – Sequence Related Amplified Polymorphism

SSR – Simple Sequence Repeat

STS – Sequence Tagged

TAG – Triacylglycerol

ZEM – Zero Erucic Mustard

1.0 Introduction

Brassica production in Western Canada is a significant part of farm gate sale for producer. *Brassica napus* production accounted for 7.5 million hectares of the cultivated land in Canada and is responsible for contributing $15.4 billion to the Canadian economy (Canola Council of Canada, 2011). The demand for canola oil has increased steadily over the years due to the beneficial characteristics of the oil. Recently there has been an increase in the requirement of the use of biodiesel as a supplement to the fossil fuel derived conventional diesel. This growing demand for biodiesel as an alternative to fossil fuels results in a growing demand for oilseed production in Canada (Canola Council of Canada, 2007).

The growing demand for Canola oil has initiated the investigation of options to increase the Canola production in Western Canada. The increase of production in Canada is limited to the amount of land that is suitable for *B. napus* production and the yield potential of the current hybrid cultivar. As the demand is growing greater other strategies for increasing oil production are being studied. Some options include 1) increasing the yield of conventionally grown canola (*B. napus* and *B. rapa*); 2) increase of the oil content produced as a percentage of the seed and 3) production of other Brassica species such as *B. juncea* which can be grown in marginal or uncultivated cropping regions.

The accumulation of fatty acids in *Brassica* seeds is a complicated process controlled by an undetermined number of genes. As with other complicated traits such yield and maturity, oil content is quantitatively inherited and improvement through conventional breeding requires intense selection for the improvement of the trait. The ability to use molecular breeding techniques to identify quantitative trait loci (QTL) for the increase of oil content and the

selection of oil quality characteristic would allow for the selection of desired genotypes for the improvement of these traits. The identification of QTL would allow for an increase in the efficiency of breeding programs for the improvement of these traits.

Brassica juncea offers an opportunity for the increase in oilseed production in Western Canada. *Brassica juncea* is grown world wide as an oilseed and for condiment mustard production. *B. juncea* has shown greater heat and drought tolerance than is observed in *Brassica napus* and has similar oil content and oil profile to *B. napus*. These characteristics make *B. juncea* a candidate to increase oilseed production by introducing oilseed production in regions of Canada where Canola production has been limited.

1.1 Research objectives

The long term goal of *Brassica juncea* research is the improvement of seed oil content and to improve the fatty acid profile to be similar to that of current *B. napus* cultivars. This study was completed to improve the knowledge in *B. juncea* for improving the economically significant traits of seed oil and protein content and the fatty acid profile. The studies involved the two steps:

1. Development of linkage map using SSR and SRAP molecular markers.
2. QTL analysis of two DH populations for oil and protein content and fatty acid profile.

The use of molecular breeding using linkage maps and QTL can be used to accelerate the breeding efforts for the future improvement of these traits to meet this demand in the future.

2.0 Literature review

2.1 History of *Brassica juncea* cultivation

Brassica juncea is one of the earliest cultivated species of the *Brassica* family in agricultural history. It was described in Sumerian and Sanskrit text dating back to 3000 BC (Hemingway, 1995) as a commonly grown condiment. It is suggested (Gomez-Campo, 1999) that China and the Middle East most likely were the center of origin for *B. juncea*. In general, cultivation of *Brassica* as an oilseed crop dates back to 2000 BC in India (Pradash, 1980). In Europe cultivation of *Brassica* became more prominent in the 13th century where the seed oil was primarily use as lamp oil (Snowdon et al., 2007). The cultivation and production of *Brassica* also known as rapeseed was introduced in Canada during 1942 as a source of lubricant for war ships in World War II (Downey, 1989).

2.2 Biology of *Brassica* species

The cultivated oilseed *Brassica* family consists of three diploid species and where polyploidy species resulted from natural hybridization among the diploid species. The relationship among the six *Brassica* species is commonly described as Triangle of U (U, 1935) as shown in Figure 2.1. The *B. juncea* emerged (Axelsson et al., 2000) with the spontaneous hybridization between *Brassica rapa* ("AA" genome) and *Brassica nigra* ("BB" genome). Therefore, *B. juncea* can be considered as an amphidiploid species containing of two complete genomes (AABB). The *B. napus* is also an amphidiploids species ("AACC") sharing the "A" genome from *B. rapa* and containing the "C" genome from *Brassica oleracea*. The development of synthetic *B. napus* and *B. juncea* further supports the 3 genome model of *Brassica* crop species (Becker et al., 1995; Bhat and Sarla, 2004; Song et al., 1995; Srivastava et al., 2001)

Triangle of U

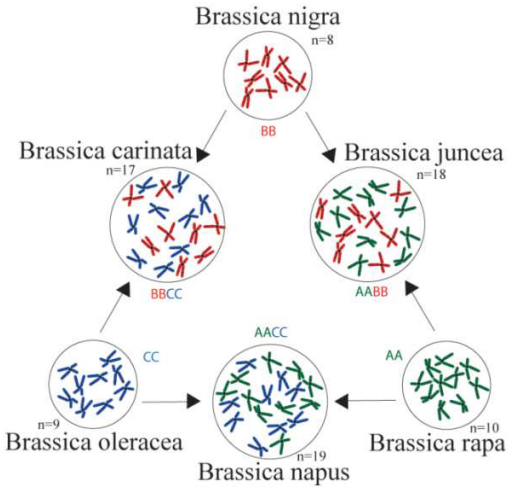


Figure 2.1: Representation of the genome organization between the *Brassica* sp. This shows the genome relationship between the amphidiploid species within the *Brassica* sp. (Adapted from the Wikimedia Commons file "Image:Triangle of U.png" http://en.wikipedia.org/w/index.php?title=File%3ATriangle_of_U_Simple1.PNG)

Studies of the common origins of the *Brassica* species have revealed that there is significant homology between the "A", "B" and "C" genomes. There is evidence that *B. rapa* and *B. oleracea* come from a similar linage whereas *B. nigra* originated differently (Warwick and Black, 1991). Further studies of the homologous regions within *Brassica* species suggest that "A", "B" and "C" genomes originated from ancestral species containing 6 (n=6) chromosomes (Truco et al., 1996). Mapping studies of the homology of *Arabidopsis* and *Brassica napus* genome displayed a number of homologous regions between the two species (Parkin et al., 2005). In addition to the homology observed with between *Arabidopsis* and *B. napus,* similarities have been observed in the *B. nigra* and *B. juncea* genomes with respect to *Arabidopsis* (Lagercrantz, 1998; Panjabi et al., 2008b).

There are similarities in the plant architecture of the *Brassica* species (Rakow and Woods., 1987). Both *B. juncea* and *B. napus* initiate growth in a rosette from which a vertical stem

elongates containing the racemous floral structure. Both species display showy yellow flower consisting of four petals and four alternating sepals. The stigma is receptive for approximately 3 days prior and after flower opening and this contributes to the high rate of self fertilization within each of the species. The ovary of both species consists of two carpals that develop into an elongated bivalent silique prior to fertilization.

The production of *B. napus* occurs in mild climates of Europe and Asia as a winter annual whereas spring annuals are more predominantly grown in the more moderate regions such as Northern Europe, Australia and Canada. In contrast production of (*B. juncea*) for mustard oil as annual crop is dominated in drier regions of Southern Asian continent and Australia (Edwards et al., 2007). *Brassica juncea* has a greater tolerance for heat and drought conditions and the mature siliques are less prone to shattering, reducing yield losses, making it well adapted for cultivation in these environments (Hemingway, 1995).

2.3 Rapeseed oil profile and development

The oil content of the seed is of significant value in *Brassica* oilseeds. Mature seeds contain approximately 40% oil with triglycerides making up close to 90% of the oil in the seed (Uppstrom, 1995). The oil profile has a high concentration of monounsaturated fatty acids and a low level of saturated fats (Uppstrom, 1995). It has been shown that monounsaturated fatty acids decrease the low density lipoprotein (LDL) cholesterol which is the major contributor to coronary heart disease (Lecerf, 2009). The byproduct of the oil extraction process produces a meal which has protein levels between 20-30% of the seed weight. The meal has significant commercial value and marketed as a supplement in animal feed.

Erucic acid is a significant component to the fatty acid profile of *Brassica*, with levels as high as 40% to 50% observed in high erucic lines (Edwards et al., 2007; Snowdon et al., 2007). Erucic acid produces a bitter taste to the oil and has been shown to have adverse health effects. Studies with rats have shown that the increased consumption of oils high in erucic acid resulted in increases in occurrence of cardiovascular disease (Beare, 1963; Gopalan, 1974; Hulan, 1977). The identification of germplasm with a zero erucic acid level (>2% of oil profile) in *B. napus* was identified in the late 1960 due to increased advances in high throughput screening of fatty acid profiles using gas chromatography (Downey and Harvey, 1963). A similar discovery in *B. juncea* identified two zero erucic acid mustard cultivars (ZEM1 and ZEM2) in 1991(Kirk and Oram, 1981). These lines were made openly available to breeders around the world.

Brassica species have naturally high levels of glucosinolates which produce the sharp flavours in the end products such as meal for animal feed. Glucosinolates are found in the meal byproducts of the oil extraction process which have anti-nutritional characteristics which reduce the palatability of the meal and the release of toxic by-products in the digestion of monogastrics which inhibits the use of the meal in animal feed (Bell, 1982; Bille et al., 1983). High levels of glucosinolates reduce the ability to market the feed meal as a supplement in animal feed rations.

The first low glucosinolate *B. napus* lines were identified in a Polish spring rape Bronowski in 1968 (Kondra and Stefansson, 1970) which was later shown to have recessive alleles at three loci contributing to the low aliphatic glucosinolate (>3 mg/g seed) level. This led to the development of the first double low spring rapeseed cultivar 'Tower' with zero erucic acid and low glucosinolate levels in 1974 (Stefansson and Kondra, 1975) . This "double low" rapeseed was given the name Canadian oil low acid (Canola) which helped develop the edible oil market in

Canada. Canola has now become the most economically significant crop grown in the western Canadian prairies.

The first low glucosinolate mustard (*B. juncea*) line was identified in 1990 allowing for the development of double low or Canola quality mustard oil (Love et al., 1991). This line was stable for low glucosinolate levels but was later shown to be nullisomic (2n=34) (Cheng et al., 2001). Ongoing breeding efforts in Australia and Canada resulted in reduce glucosinolate levels to develop Canola quality breeding material of *B. juncea* (Burton et al., 2003). This allowed for the production of Canola quality mustard oilseeds in regions where *B. napus* production was limited due to environmental factors such as reduced rainfall and high temperatures. In Canada the first double low *B. juncea* cultivars, Arid and Amulet, were released in 2002. In Australia there is active breeding for canola quality *B. juncea* to improve the agronomic characteristics of the crop (Burton et al., 2008).

2.4 *Brassica* production

2.4.1 Economic significant of *Brassica* oilseed production

Rapeseed oil production made up 13.6% of the total oilseed production in the world in 2010 second to soybean production which made up approximately 59% of overall production (FAO, 2010). Demand for canola rapeseed oil in Canada has grown from 3.8 million tonnes in 2000 to 7.3 million tonnes in 2010 fueled mainly by increased exports to Japan, China, Mexico and the United Arab Emirates (AAFC, 2010). Canada is the second largest Canola/Rapeseed producing country contributing 18.4% of the world's production in 2010 (AAFC, 2010). Canola acres have increased from 3.7 million acres in 2001 to 6.8 million ha in 2010 as the demand has increased

(StatCan, 2011). Canadian oilseed production added $15.4 billion to the Canadian GDP in 2010 making it a significant part of western Canadian economy (Canola Council of Canada, 2011).

Alternative uses for oilseeds as feedstock for fuel production thus reducing the reliance on nonrenewable fossil fuels, have resulted in an increase in demand for seed oil production. Canola oil can be used as a high quality source of oil for biodiesel production. The Canadian government has put forward a mandate that will requires a minimum 5% biodiesel blend by 2015 with a 2% biodiesel blend implemented by 2012 (Canola Council of Canada, 2007). The 5% biodiesel blend demands 1.5 billion litres of biodiesel which will require an increase of 2.9 million tonnes in oilseed production per year to meet current export and food uses (Canola Council of Canada, 2007).

2.4.2 *Brassica juncea* production

Brassica juncea cultivation in Western Canada is primarily for the purpose of condiment mustard production (oriental and brown mustard), and it goes along with the production of *Sinapsis alba* L. another species which produces the common yellow mustard (Skrypetz, 2003). Canada is the largest exporter and second largest producer of condiment mustard seed in the world, exporting nearly half of the global supply (AAFC, 2009). The production fluctuates between *B. juncea* and *Sinapsis alba* depending on the market price but *B. juncea* production typically makes up half of all mustard production in western Canada (Skrypetz, 2003). Mustard (*B. juncea*) production in 2009 totaled 515 thousand hectares in western Canada, with the majority of production located in Saskatchewan (Saskatchwan Agriculture, 2011). Oilseed mustard production in Canada is not significant as it competes directly with current canola production. *B. napus* cultivars have greater yield potential and agronomic performance than current oilseed mustard varieties due to increased breeding efforts.

B. juncea is a significant proportion of the overall oilseed production in India, Bangladesh and Pakistan. However, production data from these countries does not separate rapeseed production from mustard oil production (Edwards et al., 2007). Due to the drier conditions and cultural harvesting practices in India, *B. juncea* generally makes up 80% of the oilseed production (Edwards et al., 2007). Oilseed mustard production in Australia occurs in the drier regions of the country where *B. napus* production is limited (Burton et al., 2003). Similarly oilseed mustard production has been successfully introduced into semi-arid regions of western Canada where rapeseed production is limited.

2.5 Breeding objectives in *Brassica*

Breeding in *Brassica* has been used to improve agronomic and seed quality traits in the development of commercial cultivars. The manipulation of fatty acids in oilseed *Brassica* has been the focus of breeding efforts in rapeseed since the early 1950's with the goal of improving the oil profile (Uppstrom, 1995). This initiated the identification of germplasm producing seed oil with low erucic acid and reduced alpha glucosinolate levels to develop the first double low rapeseed varieties (Stefansson and Kondra, 1975). Further breeding efforts in *Brassica napus* were focused on the modification of oil quality in canola quality rapeseed. This included the reduction of linolenic acid which reduces shelf life of oil due to oxidation, and lead to unwanted flavours (Uppstrom, 1995). Linolenic acid made up between 8-12% of the oil profile of conventional canola lines. The identification of mutants producing lower linolenic levels reduced seed concentrations to less than 3% of the overall oil content (Scarth et al., 1988).

Manipulation of the linolenic acid levels in *B. napus* has helped the development of designer oils for industrial purposes. Low linolenic acid levels have been shown to increase the stability of the oil under heat and reduce the oxidation (Warner et al., 1994). Mid oleic oil is currently

marketed in the food industry where the oleic acid content is 67-75% of the oil profile. High oleic acid (>75%) is targeted for blending with other oils to increase oleic acid content or for industrial frying purposes (Scarth and McVetty, 1999).

The increased demand for vegetable oil will require greater production of oilseeds in Canada. This can be achieved by increasing the overall hectares devoted to oilseed production, increasing seed yield or increasing total seed oil. Focus on *Brassica juncea* has been on the condiment mustard production however there is promise that oilseed mustard can contribute to increase in overall production and help increase oil production within western Canada (Woods et al., 1991).

The identification of genetic sources of low erucic and low glucosinolate mustard lines has allowed for the development of canola quality *B. juncea* cultivars (Kirk and Oram, 1981; Love et al., 1991). Currently much of the breeding efforts for canola quality mustard oil have been focused in India and Australia where mustard oil production has already been established. *B. juncea* breeding in Canada has produced canola quality mustard cultivars including Arid and Amulet released in 2002. This has allowed for canola oil production within regions that currently cannot sustain *B. napus* cultivation This includes the drier regions of the Canadian prairies (Woods et al., 1991) and has shown promise in similar regions of Australia (Burton et al., 2003).

2.6 Fatty acid biosynthesis

2.6.1 Fatty acid biosynthesis and storage in *Brassica*

The process of fatty acid (FA) biosynthesis in plants has been studied extensively. FA accumulation has been shown to be initiated in the chloroplast of leaves and protoplast of developing seeds (Baud et al., 2008; Ohlrogge and Browse, 1995). The synthesis of methyl-CoA

by the carboxylation of acetyl-CoA provides the base from which the fatty acid chain is elongated. Acetyl groups are added to the fatty acid chain to elongate the fatty acid to palmitic acid C16:0 (Konishi et al.,1996). Palmitic acid undergoes further elongation and desaturation to produce stearic acid and oleic acid through the enzymatic action of Ketoacyl-ACP synthase (KAS) and stearoyl-ACP desaturase respectively (Ohlrogge and Jaworski, 1997; Pidkowich et al., 2007). These end products are exported to the cytoplasm where further modifications of the fatty acids are completed. This includes the further desaturation of oleic acid to produce linoleic and linolenic acid by the fatty acid desaturase enzymes FAD2 and FAD3 respectively (Ohlrogge and Browse, 1995).

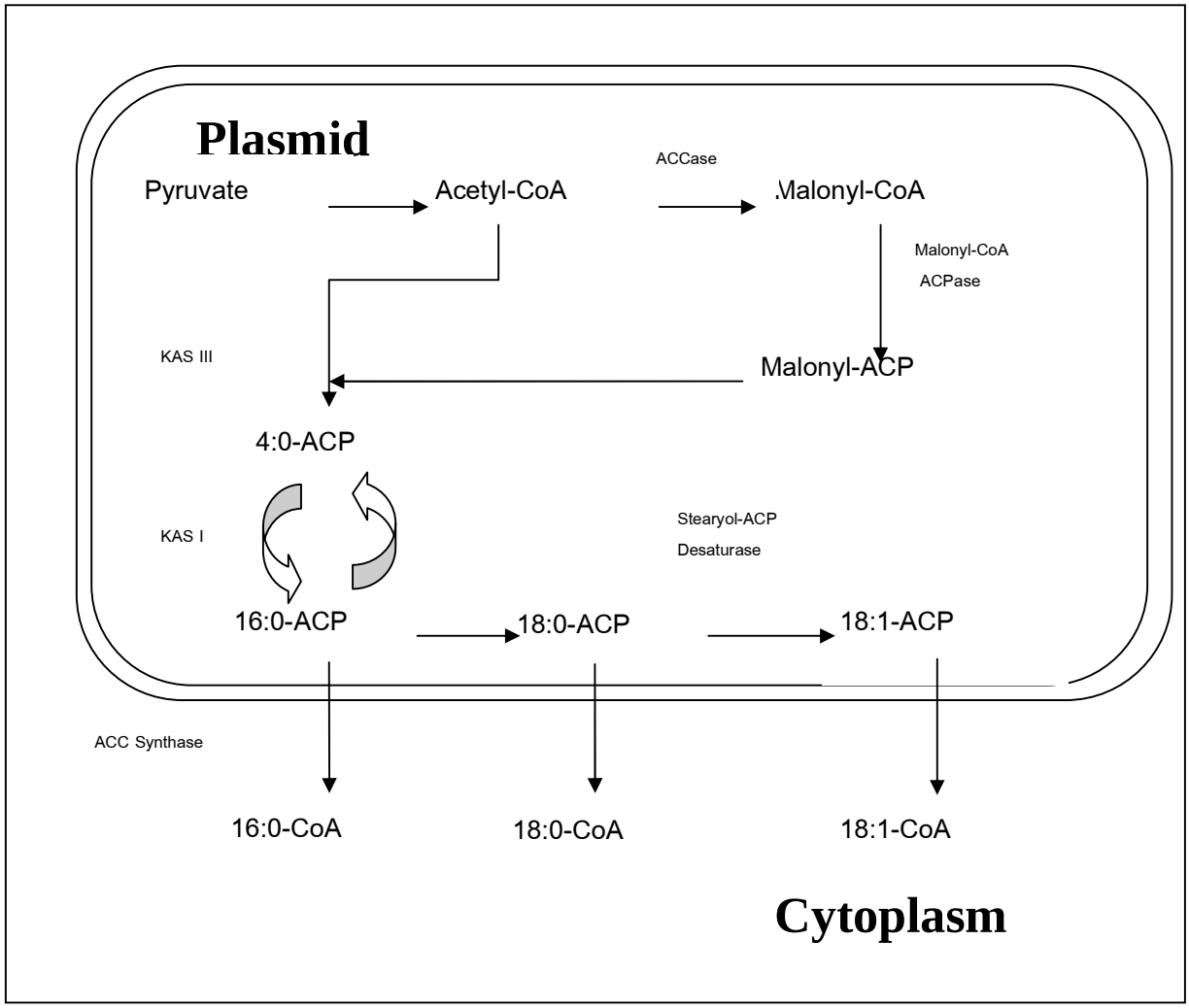


Figure 2.2: Overview of the pathway for the synthesis of fatty acids in the plasmid of the plant cell. Pyruvate produced during glycolysis becomes the substrate for Acetyl-CoA which is used in the formation of fatty acids. Acetyl CoA is converted to Malonyl-CoA by the enzyme ACCase. Malonyl-CoA is transformed into Malonyl-ACP by Malonyl-CoA ACPase. Acetyl CoA and Malonyl-ACP are elongated to produce 4:0-ACP by the enzyme KASIII. The carbon chain undergoes multiple elongation reactions using the enzyme KASI enzyme which adds two carbon chains to the molecule. This reaction is repeated to produce 16:0-ACP. A final elongation of 16:0-ACP is completed with KASII to produce 18:0-ACP. Further manipulation of the 18:0-ACP is desataturated to produce 18:1-ACP with the enzyme Stearoyl-ACP Desaturase. The elongated carbon chains are converted into fatty acid chains 16:0-CoA, 18:0-CoA and 18:1-CoA with the enzyme ACCsynthase and can then be transported into the cytoplasm. Adapted from Baud et al. (2008), Weselake et al. (2009) and Li-Beisson et al. (2010).

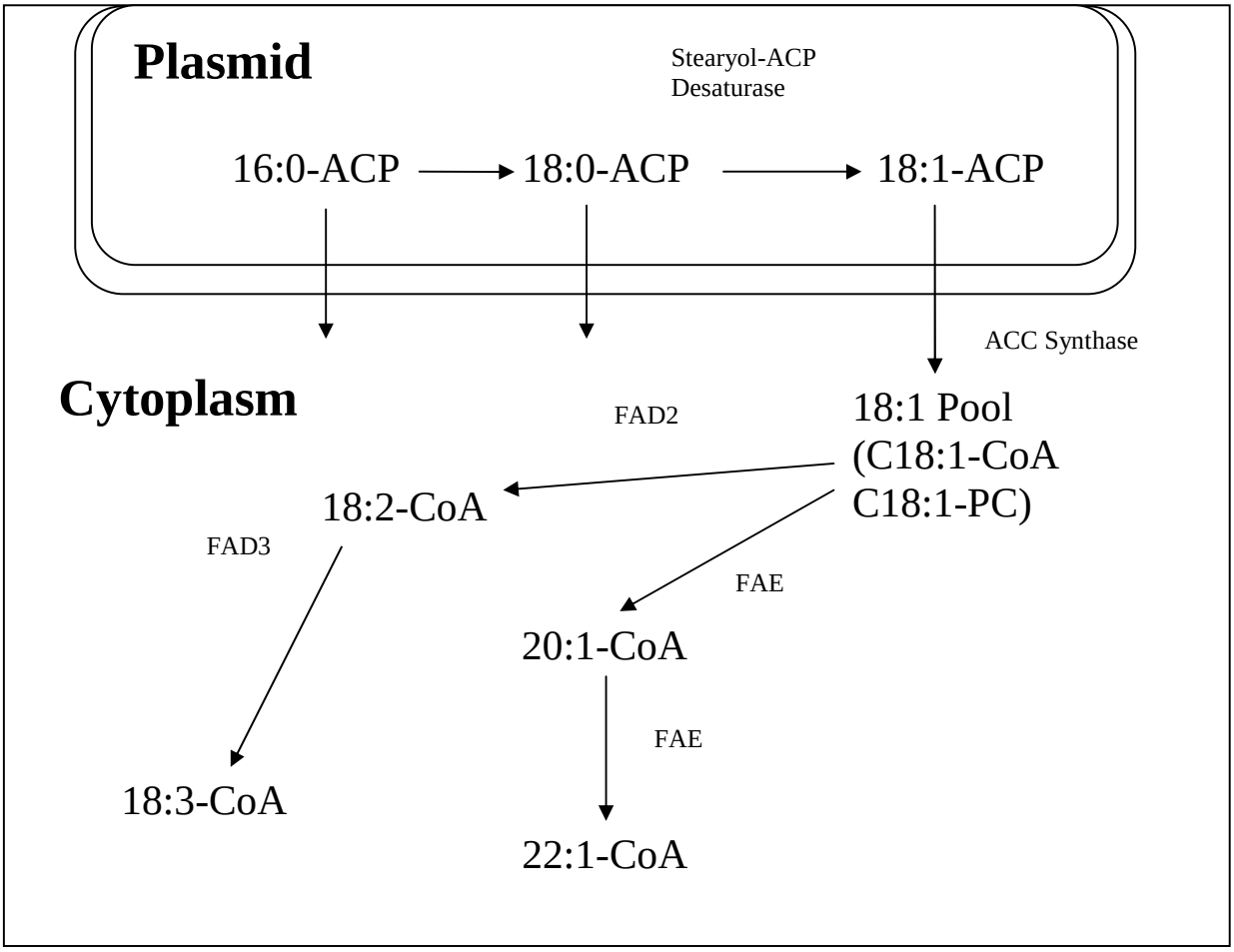


Figure 2.3: Simplified scheme of fatty acid elongation and fatty acid desaturation of 18:1 fatty acids produced in plastids. These fatty acids are elongated with the enzyme FAE to produce 20:1 and 20:2 fatty acid chains. Desaturation to produce 18:2 is catalyzed with the desaturase enzyme FAD2 and further desaturation to produce 18:3 is completed with FAD3. Adapted from Baud et al. (2008)

Fatty acids are stored in the plant cell as triacylglycerol (TAG) structure which contains 3 fatty acid chains that are bound to a glycerol backbone (Stymne and Stobart, 1993). The structure is synthesized in the endoplasmic reticulum where the fatty acids are bound to a glycerol-3-phosphate (Gly-3-P) molecule. The TAGs produced are stored in oil bodies within a single phospholipid layer. These bodies are developed in the ER however the precise method is not fully understood. It has been suggested that the TAG accumulate in the ER where they are

budded off surrounded by a mono-phospholipid layer that is thought to be derived from the ER phospholipid bilayer (Robenek et al., 2006). These bodies are released into the cytosol where, during seed maturation, the water content of the cells decrease and result in the oil bodies converging together bound by oleosins which bind the oil bodies together during cytoplasmic compression (Murphy and Vance, 1999).

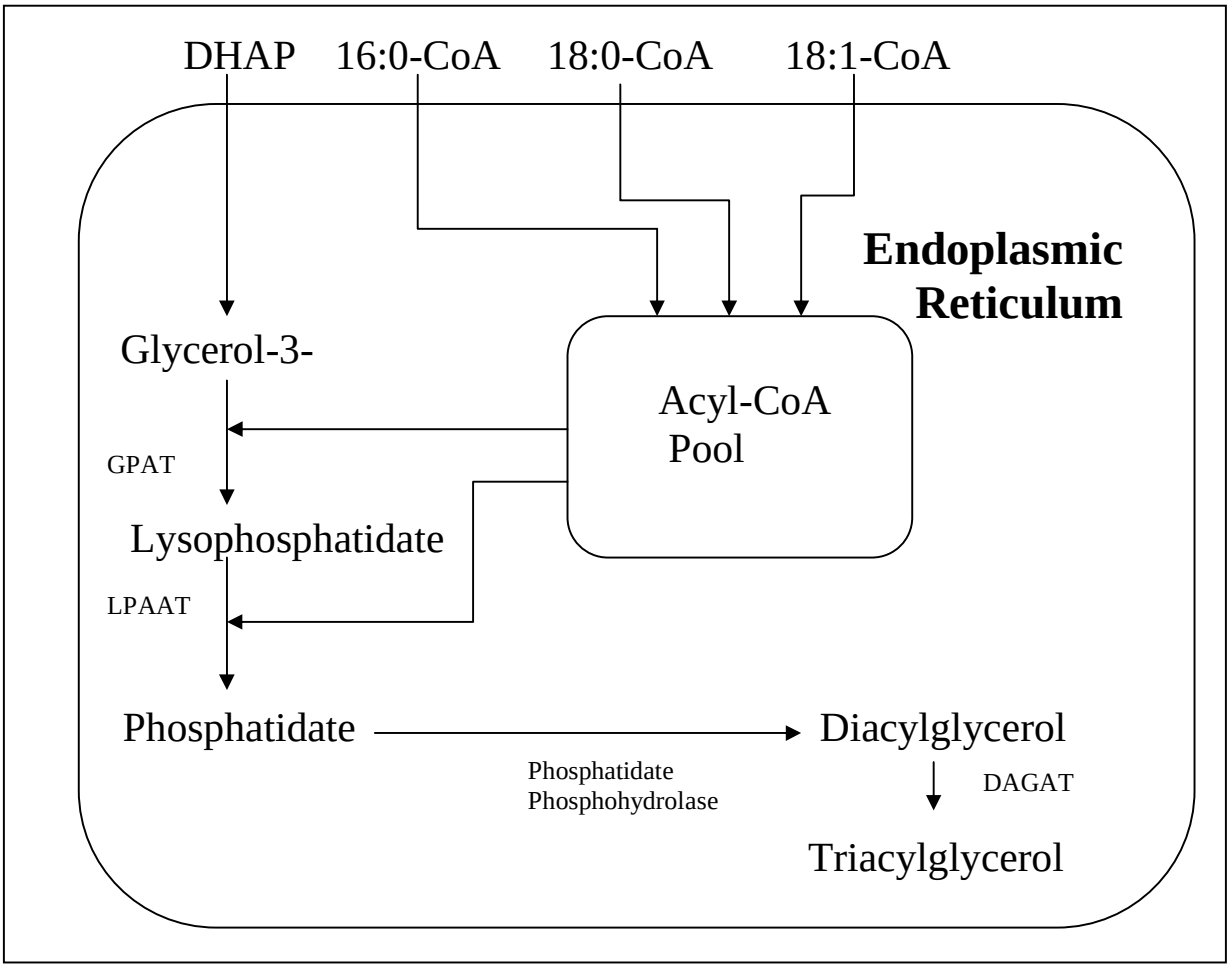


Figure 2.4: Simplified diagram of triacylglycerol synthesis in plant cells. Fatty acids synthesized in plasmids and modified in the cytoplasm are transported to the endoplasmic reticulum. Hydroxyacetone phosphate (DHAP) originating in the cytoplasm is converted into the glycerol backbone for triacylglycerol molecule with the enzyme glycerol-3-phosphate dehydrogenase. Kennedy pathway catalyzes the reactions which binds Acyl-CoA molecules to the Gly-3-P molecule to produce triacylglycerol molecule. Adapted from Baud (2008).

2.6.2 Genetic control of erucic acid accumulation

Erucic acid (C22:1) is one on the products of very long fatty acid chain production in the endoplasmic reticulum. Early studies in *Arabidopsis* identified FAE1 as being responsible for the accumulation of C20:1 and C22:1fatty acids (James et al., 1995). In *B. juncea* a study identified FAE candidate genes similar to those found in *B. napus* to control the expression of erucic acid (Gupta et al., 2004*)*. The *FAE1.1* and *FAE1.2* loci were mapped to linkage groups LG3 and LG17 respectively. Similar to *B. napus* the two loci are located on each of the diploid genomes AA and BB and control erucic acid in an additive manner. Analysis of low erucic and high erucic lines in *B. juncea* uncovered 4 SNP in *FAE1.1* and 3 SNP in *FAE1.2* that contribute to the variation between high and low erucic acid (Gupta et al., 2004).

The identification of the genes contributing to erucic acid accumulation has determined that the trait is controlled by 2 loci that contribute additively to the overall expression of erucic acid in the seed. The loci of that control erucic acid have been identified as E1 on the A genome and E2 on the B and C genome for both *B. napus* and *B. juncea*. High erucic acid lines having erucic acid in the seed greater than 40% are homozygous dominant for all 4 alleles (E1E1E2E2). Conversely low erucic (>2%) are homozygous recessive for all four alleles (e1e1e2e2) (Barret et al., 1998; Gupta et al., 2004; Rahman et al., 2008). There are intermediate levels of erucic acid which are determined by the relative contribution of dominant and recessive alleles.

2.7 Genetic studies in *Brassica*

2.7.1 Molecular biology techniques for genetic studies

The detection and manipulation of genetic polymorphisms has played a key role in breeding for economically important crop traits. For decades plant breeders has used polymorphism for

morphological traits as markers to trace the desired traits by their linkage with detectable morphological traits or markers. However, the rapid development in the field of molecular biology has given plant breeders the power to use molecular markers (protein, DNA) for manipulation of crop traits. Restriction fragment length polymorphisms (RFLP) were the first DNA- probe based genetic markers widely used to identify polymorphisms. RFLPs had a high rate of reproducibility allowing them to be used to develop some of the first molecular linkage maps. Additionally RFLPs produced co-dominant markers allowing for the identification of heterozygote genotypes. RFLPs however are very labour intensive, required some sequence knowledge to develop probes and required large amounts of high quality DNA for analysis. This limited the ability for RFLPs to be used in high throughput screening.

With the development of polymerase chain reaction (PCR) new marker techniques were developed which overcame the issue of sequence knowledge required by RFLPs. Random amplified polymorphism (RAPD) and amplified fragment length polymorphism (AFLP) were two of the first marker types developed using PCR technology to screen for polymorphisms. RAPD markers used a randomly developed oligonucliotide primer to amplify random fragments within the genome. These markers were highly effective in identifying polymorphisms but due to the random binding of the single primer reproducibility was inconsistent. This restricts the use of RAPD for marker assisted selection. AFLP were utilized to produce molecular linkage maps due to the high number of polymorphisms identified however this method required a number of steps to complete which reduced efficiency. Both RAPD and AFLP produced dominant markers which do not allow for the identification of heterozygote genotypes from homozygous dominant genotypes restricting the populations that could be analyzed. Other marker techniques that have

been developed that do not require sequence knowledge include sequence related amplified polymorphism (SRAP), and inter-simple sequence repeats (ISSR) markers.

Marker technology has improved to develop a number of other marker systems which require some sequence knowledge such as single nucleotide polymorphism (SNP) (Gupta et al., 2004), simple sequence repeats (SSR) (Uzunova and Ecke, 1999), sequence tagged site (STS) (Panjabi et al., 2008b) and sequence characterized amplified region (SCAR) markers (Rahman et al., 2007). These markers are more difficult to develop but produce more reproducible results due to the specificity of the primers that are developed.

2.7.2 SSR markers

Simple sequence repeat (SSR) markers identify difference in microsatellite regions within the genome. Microsatellites are regions of tandem repeats of di, tri and tetra nucleotide regions that occur repeatedly throughout the genome (Hamada et al., 1982; Vergnaud et al., 1991). These regions are non-coding and have been shown to occur frequently throughout the genomes in plants and animals (Innan et al., 1997; Morgante and Olivieri, 1993; Tautz, 1994; Weber and Wong, 1993). Their function has not been clearly identified however variation in canine species and diseases in humans have been linked to microsatellite regions (Fondon and Garner, 2004; Pearson et al., 2005). Microsatellites are the result of replication slippage during meiosis which causes mismatches between the DNA strands (Tautz, 1994). SSR are identified by probing BAC libraries (Lowe et al., 2004) or sequence data (Tang et al., 2008) to identify the microsatellite regions in DNA. Unique oligonucliotide primers are designed flanking the tandem repeats and PCR is used to amplify this region. Polymorphisms between the genotypes are reflected by the number of repeats that occur within the tandem repeat region. These are identified as DNA band size differences of the PCR products which can be as small as 2 base pairs (bp). The small

differences in band size requires a high fidelity method of visualizing the band such as polyacrylamide gels (Morin and Smith, 1995) or tagged florescence capillary fragments analysis using instruments such as Applied Biosystems ABI PRISM® 3100 Genetic Analyzer (Oetting et al., 1995).

2.7.3 SRAP markers

Sequence related amplified polymorphisms (SRAP) are markers which have been used to identify high rates of polymorphism with no sequence information available. The method uses two primers which have a specific motif. The forward primer is 18 bp long consisting of an 11 bp region of which the sequence of no specific consequence (Li and Quiros, 2001). This region is followed by an AATT region which is then followed by a random trinucleotide. The reverse primer follows a similar structure to the forward primer however it is 17 bp long with a 10bp region followed by GGCC and another random trinucleotide region. It has been identified that coding regions of genes in plants have a higher proportion of A and T and a high proportion of C and G nucleotides in the exon regions (Lin et al., 1999). This favors the forward primer to anneal in coding regions of the genome and the reverse primer to favor annealing to non-coding region. Initial studies showed a high proportion of markers that are amplifying in regions with a high concentration of coding genes (Li and Quiros, 2001). Initial five cycles of the PCR have a lower annealing temperature which will allow for nonspecific binding of the primers. The annealing temp is then raised for the remainder of the cycles to amplify specific regions amplified in the first five cycles.

2.7.4 Genetic Maps using Molecular Markers

Molecular markers have been used to produce genetic maps that represent the genome based on the recombination frequency of the polymorphic markers within a mapping population. A large

number of genetic maps have been developed for oilseed rape and mustard using F_2:F_3, recombinant inbred, backcross and double haploid populations (Table 2.2). Polymorphism identified between the parents using molecular markers are indications of genetic variation found within the genome. These maps can be used to locate the loci influencing or controlling the quantitative traits such as yield, seed quality, and disease resistance by the association of such quantitative trait loci (QTL) with molecular markers placed on a genetic linkage map. Early genetic maps were constructed using primarily RFLP markers however more recently maps have been published using SSR, SRAP, STS, AFLP, SCAR and RAPD PCR based markers (Edwards et al., 2007; Snowdon et al., 2007).

2.7.5 Computer software for developing genetic maps

To develop high density linkage or genetic maps, computer programs are used to compute the distance among the markers and determine the order in which they are placed on genetic maps. Mapmaker was one of the first widely used programs used to develop high density linkage maps which operated on a MSDOS platform (Lincoln, 1993). This program required a greater understanding of programming and was time consuming to develop maps with a greater number of markers without previous linked markers identified. JoinMap is a Windows based program which identified linkage groups based on the logarithm of odds (LOD) to identify linkage groups (Van Ooijen, 2001). This overcomes the need to identify anchored markers in the development of high density linkage maps.

2.7.6 Molecular mapping in *Brassica*

Linkage maps have been developed using RAPD, RFLP and AFLP markers for both *B. napus* (Cheung et al., 1997a; Ferreira et al., 1994; Foisset et al., 1996; Landry et al., 1991; Lombard and Delourme, 2001; Sharpe et al., 1995; Uzunova et al., 1995) and *B. juncea* (Aggarwal et al.,

2003; Cheung et al., 1997b; Lionneton et al., 2002; Mahmood et al., 2003; Sharma et al., 2002a). More recently, maps using SSR markers have been developed in *B. napus* (Cheng et al., 2009; Lowe et al., 2004; Piquemal et al., 2005; Uzunova and Ecke, 1999) and *B. juncea* (Ramchiary et al., 2007b). Mapping in *B. juncea* was completed using markers developed from EST sequences of *Arabidopsis* to develop oligonucliotide primers (Panjabi et al., 2008a). These primers were designed to identify polymorphism within intron regions of single copy genes which were used to develop a linkage map of the *B. juncea* genome. SNP and STS markers have been used in the development of linkage maps of *B. napus* (Qiu et al., 2006; Suwabe et al., 2008). A public source (www.*Brassica*.info) of SSR markers have been made available (Agarwal et al., 2008). These markers have been contributed from a number of sources and made available for public use in *Brassica* genomic studies.

Table 2.1: Overview of the linkage maps and marker technologies developed in amphidiploids species of *Brassica*.

Species	Markers used	Map length	Number of Linkage groups	Reference
B. juncea	**RFLP**	**173.9 cM**	**6**	**Sharma et al., 1994**
B. juncea	**RFLP**	**243.3 cM**	**9**	**Upadhyay et al., 1996**
B. juncea	**RFLP**	**2073 cM**	**18**	**Cheung et al., 1997a**
B. juncea	**RFLP**	**1266 cM**	**18**	**Axelsson et al., 2000**
B. juncea	**RAPD**	**790.4 cM**	**21**	**Sharma et al., 2002a**
B. juncea	**AFLP, RAPD**	**1641 cM**	**18**	**Lionneton et al., 2002**
B. juncea	**RFLP**	**1564 cM**	**18**	**Mahmood et al., 2003**
B. juncea	**AFLP, RFLP**	**1629 cM**	**18**	**Pradhan et al., 2003**
B. juncea	**RFLP, AFLP, SSR**	**1840 cM**	**18**	**Ramchiary et al., 2007b**
B. juncea	**STS**	**1922 cM**	**18**	**Panjabi et al., 2008b**
B. napus	**AFLP, RAPD, RFLP, SSR, SCAR**	**1574 cM**	**19**	**Lombard and Delourme, 2001**
B. napus	**AFLP, RAPD, RFLP, SCAR, SSR**	**2023 cM**	**19**	**Delourme et al., 2006**
B. napus	**RFLP, RAPD, SSR**	**1441 cM**	**19**	**Uzunova et al., 1995**
B. napus	**RFLP**	**1016 cM**	**22**	**Ferreira et al., 1994**
B. napus	**RFLP**	**1656 cM**	**19**	**Sharpe et al., 1995**
B. napus	**SSR, SCAR**	**2619 cM**	**19**	**Piquemal et al., 2005**
B. napus	**RFLP**	**1238 cM**	**18**	**Howell et al., 1996**
B. napus	**RFLP, AFLP, SNP, SSR**	**2031 cM**	**19**	**Cryer et al., 2005**
B. napus	**RFLP, AFLP, RAPD, SNP, SSR**	**1792 cM**	**19**	**Suwabe et al., 2008**
B. napus	**AFLP, RFLP, SSR, STS**	**1685 cM**	**19**	**Qiu et al., 2006**
B. napus	**RFLP**	**1968 cM**	**19**	**Parkin et al., 2005**
B. napus	**SSR**	**1957 cM**	**19**	**Lowe et al., 2004**
B. napus	**RFLP**	**1656 cM**	**19**	**Parkin et al., 1995**
B. napus	**AFLP, RAPD, RFLP, SSR**	**1912 cM**	**19**	**Lombard and Delourme, 2001**
B. napus	**AFLP, RAPD, RFLP, SSR**	**2023 cM**	**19**	**Lombard and Delourme, 2001**

2.8 QTL analysis in Brassica

2.8.1 QTL analysis

The identification of QTL based on a linkage map requires the analysis of the genotypic and phenotypic variation within the population that is being studied. Phenotypic data collected from field trials is used to determine if there is a significant correlation between the genetic region determined by the marker segregation and the variation observed within the quantitative trait. Multiple techniques have been developed to predict these regions and to minimize statistical errors in the identification of QTL and their description follows.

2.8.2 Single marker analysis

Single marker analysis is the simplest method of detecting a QTL. Statistical analysis such as Analysis Of Variance (ANOVA), t-test or linear regression can be used to determine the segregation of quantitative trait with the single marker. Linear regression is commonly used for single marker analysis as the coefficient of determination (R^2) can be used to determine phenotypic variation that can be explained by the marker. There are however a number of downfalls to the use of single marker analysis for the identification of QTL. The further the marker is away from the actual QTL reduces the ability to identify the correlation between the marker and the QTL. This is due to recombination that can occur between the marker and the QTL which effects the R^2 estimation of the overall effects of the QTL. This can result in the miss calculation of the QTL, resulting in the underestimation of the overall effect (Tanksley, 1993). This can be overcome by using a larger number of markers covering the entire genome (Tanksley, 1993).

2.8.3 Interval mapping

Simple interval mapping (SIM) was developed to identify QTL by regression analysis of adjacent markers along the linkage map. This uses a linkage map covering the whole genome as the basis of the analysis tests for QTLs located in the intervals between markers on identified linkage groups. This overcomes the limitations of single marker analysis by providing a larger number of markers that covers the whole genome. This reduces the effects of recombination of the marker and the QTL which results in a more effective estimation of the overall effect of the QTL on the trait making it a more powerful tool for marker assisted selection (Lander and Botstein, 1989; Liu, 1998).

2.8.4 Composite Interval mapping

To overcome the limitation of interval mapping composite interval mapping (CIM) was developed. Using SIM analysis to identify QTL results in bias based on the markers due to their position on the linkage map and is effected by other QTL located on the same linkage group (Jansen, 1993; Zeng, 1993). CIM overcomes this by combining SIM with linear regression in addition to the incorporation of additional markers in a statistical model (Jansen, 1993; Jansen, 1994; Zeng, 1993; Zeng, 1994). This method allows for more accurate identification of QTL and compared to the other methods described.

2.9 Tools for QTL analysis

2.9.1 Software for identification of QTL

The identification of QTL requires the correlation of phenotypic data with the inheritance of genetic regions within a segregating population. Single marker analysis of a population can be completed using ANOVA, t-test or linear regression to identify the significant correlation

between single marker and the trait(s) being studied. This analysis can be completed using simple statistical software to identify QTL. QTL analysis using SIM can be completed using the software programs Mapmaker/QTL (Lincoln, 1993) and QGene (Nelson, 1997). Programs such as Win QTL Cartographer (Basten et al., 2001) and MapQTL (Van Ooien et al., 2002) use CIM to identify QTL.

2.10 QTL in *Brassica*

2.10.1 QTL analysis in *Brassica*

The identification of significant QTL can be used as a tool to improve selection in breeding programs and the development of markers for marker assisted selection. In oilseed rape (*B. napus*) there has been a significant amount of research in the area of QTL identification. Studies to identify QTL in *B. napus* has been completed in economically important traits such as, oil content (Burns et al., 2003; Delourme et al., 2006; Ecke et al., 1995; Qiu et al., 2006; Zhao et al., 2005), fatty acid composition (Barker et al., 2007; Burns et al., 2003; Ecke et al., 1995; Qiu et al., 2006; Rajcan et al., 1999; Zhao et al., 2008) and yield (Chen et al., 2007; Gul, 2003; Quijada et al., 2006; Udall et al., 2006; Zhao et al., 2006). As the demands on oilseeds increase there has been more research for improvement in *Brassica* species. In responses there has been a number of studies completed looking to identify QTL to help improve the efficiency of breeding of successful cultivars.

2.10.2 QTL for oil content in *B. juncea*

There are limited reports on QTL analysis for oil content in *B. juncea*. Analysis of a *B. juncea* population comprising recombinant inbred lines (RIL) and characterized for RAPD markers identified three QTL for oil content, each accounting for approximately a 6% increase in oil

accumulation (Sharma et al., 1999). In another study (Lionneton et al., 2002), an AFLP based genetic map developed from 131 DH lines was used to identify two QTL for increases in oil content, which were located on linkage groups LG 11 and LG 18 accounting for 17.1% and 9.2% of the variation observed, respectively. These traits had opposite effects with the QTL on LG 11 corresponding to an increase in oil content whereas the QTL on LG 18 resulted in a decrease in oil content. Unlike other studies, no correlation between fatty acid QTL and oil content was observed (Mahmood et al., 2006b). A double haploid *B. juncea* population containing 112 lines were used to develop a RFLP linkage map consisting of 300 linked loci and the map was used to detect QTL for oil content. The analysis identified six QTL for oil content accounting for 65.3% of the observed variation in oil content in the population. Using a DH population of *B. juncea* containing 123 lines a linkage map was developed a map using AFLP, RFLP and SSR markers (Ramchiary et al., 2007). From this map seven QTL for oil content were identified on the A genome and showed strong correlation between oil content and erucic acid levels.

2.10.3 QTL analysis for Fatty Acids in *B. juncea*

Oil quality is a major focus for breeders in *Brassica* breeding. In the development of canola quality *B. juncea* it is necessary to reduce erucic acid (22:1) levels. Additionally there is a focus on developing specialty oils in *Brassica* including high oleic and high erucic acid lines for specific markets. All of the breeding objectives can benefit from marker assisted selection as they would allow for selection of alleles that produce the desired oil profile prior to field screening.

In *B. juncea* there has been a number of studies identifying QTL for the major fatty acids. QTL analysis for fatty acid in a DH population of *B. juncea* identified 2 QTL for erucic acid (Mahmood et al., 2003). This was confirmed in another experiment looking at another DH

population of *B. juncea* (Lionneton et al., 2004). In both of these studies there was a correlation between QTL identified for erucic acid and oleic, linoleic and linolenic acid levels. This corresponds to another work (Gupta et al., 2004) that has identified 2 *FAE1* genes shown to act additively in determining erucic acid levels in *B. juncea*. From these studies the negative relationship between oleic and erucic acid levels was confirmed. This was reinforced in a study that modified the expression of the *FAE* and *FatB* genes through transformation with showed a decrease in erucic acid and an increase in oleic, linoleic and linolenic fatty acid levels (Sinha et al., 2007). Studies to identify QTL for oleic acid have repeatedly identified 2 QTL which generally correspond to QTL identified in erucic acid (Lionneton et al., 2002; Lionneton et al., 2004; Mahmood et al., 2003; Sharma et al., 2002a). Similarly, common QTL are shared between oleic acid and linoleic and linolenic as oleic acid is the primary substrate for these fatty acids (Lionneton et al., 2002; Lionneton et al., 2004; Mahmood et al., 2003).

2.10.4 Agronomic improvement of *Brassica juncea*

The development of crop cultivars requires the improvement of agronomic traits such as maturity and seed yield. The inheritance of these agronomic traits is complex because they are highly influenced by environmental factors which make the selection of such traits more difficult for selection. This requires multi-location testing to account for variation due to environmental factors and then isolate the genetic factors involved in the performance of these traits.

The genetic control of yield requires multiple genes responding to sink and sources signals as well as environmental effects that contribute to the overall yield of a crop. In *B. juncea* the investigation of QTL for yield contributing characteristics and overall yield was studied looking at a DH population containing 112 individuals (Mahmood et al., 2005b). In this study 32 QTL for five yield components in *B. juncea* including number of siliques, seeds per silique, silique

length, test weight and silique weight were identified (Ramchiary et al., 2007b). Further investigation analyzed overall yield related to yield characteristics including pods per plant, silique length and seeds per silique. It was revealed that there was no significant yield QTL that could be identified and used as reliable method of selecting high yield varieties (Mahmood et al., 2005b).

Days to maturity in *B. juncea* is another important factor used for the selection of cultivars. QTL analysis for maturity in *B. juncea* has not been heavily studied as the phenotypic evaluation for this trait is routinely collected during the field testing. The analysis of a DH *B. juncea* population identified two QTL tightly linked to maturity which also correlated with QTL identified for days to flower and end of flowering. This shows a strong correlation between flowering and the determination of maturity in *B. juncea* (Mahmood et al., 2006a).

2.10.5 QTL for days to flower in *B. juncea*

The genetic factors that contribute to the days to flowering are also a quantitatively inherited. Days to flowering, days to end of flowering and flowering period are all components that are related to the maturity of oilseed mustard lines. These traits are easily measured by observations made during the field trial season and are used in the selection of cultivars. The investigation of F2 *Brassica* sp. identified three QTL in the diploid species *B. rapa*, *B. nigra* and *B. oleracea* which corresponded to six QTL identified *B. juncea* suggesting a common ancestrally inherited loci that related to flowering time (Axelsson et al., 2001). A study using a RFLP map developed from a DH *B. juncea* population showed a four times greater genetic advance by the use of phenotypic and genotypic evaluation of lines which could be a strong tool for breeders for increasing the efficiency of the breeding program (Mahmood et al., 2006a). Another study

identified three QTL for days to flower which also showed a significant negative correlation between flowering and oil content in *B. juncea* (Lionneton et al., 2004).

3.0 Linkage map development in *B. juncea* using SSR and SRAP Markers

3.1 Abstract

A genetic linkage map of mustard (*Brassica juncea*) was developed using molecular markers. The map is based on two genetically diverse double haploid (DH) populations produced from crosses between a yellow seeded mustard cultivar "ZEM1" and two mustard lines "Vniimk351" and "Vniimk405". The microspores from F1 hybrids of the crosses Zem1 X Vniimk351 and Zem1 X Vniimk405 were used to develop 106 and 57 DH lines, respectively. A parental screening for 363 SSR (simple sequence repeats) and SRAP (sequence related amplified polymorphisms) revealed that only 32% of the SSR markers were polymorphic. Both DH populations were then genotyped for the polymorphic SSR and SRAP markers to construct the linkage maps. The linkage map based on the ZEM1xVniimk351 population included 13 linkage groups housing 79 SSR and 61 SRAP markers. The map covered an overall length of 791 cM with an average interval of 5.7 cM between the markers. The linkage map developed from the ZEM1xVniimk405 population also contained 13 linkage groups populated by 76 SSR and 62 SRAP markers. The second linkage map spanned a distance of 623 cM overall with an average interval of 4.6 cM between the markers. The maps developed for the two populations share 47 common markers which are distributed throughout the linkage maps. The development of the molecular maps for these two DH populations through the use of SSR and SRAP markers can be used for further studies of the *B. juncea* genome.

3.2 Introduction

As described in the Triangle of U (U, 1935) *Brassica juncea* is an amphidiploid species which originated by natural hybridization between *B. rapa* (A genome) and *B. nigra* (B genome). Therefore, *B. juncea* is designated as having AB genome consisting of 18 chromosomes where 8 chromosomes originated from *B. nigra* and the 10 chromosomes from the *B. rapa.* The approximate size of the *B. juncea* genome is estimated at 1105 Mbp and the 2C genome mass of 2.29 pg/2C which is similar to the size of the *B. napus* genome which ranges from 1129-1235 Mbp and a mass of 2.34-2.56 pg/2C (He et al., 2003).

Several researchers in the recent past have developed the genetic linkage maps of *B. juncea* using various types of molecular markers such as RFLP, RAPD (Sharma et al., 2002a), AFLP (Lionneton et al., 2002; Pradhan et al., 2003; Ramchiary et al., 2007a) and STS (Panjabi et al., 2008a). The SSR markers have been used in the mapping of *B. juncea* for further saturation of an existing molecular map consisting of RFLP and AFLP markers (Ramchiary et al., 2007b). SRAP markers have been used in the comparative studies of the genetic diversity of *B. juncea* in the Asian and European continent (Wu et al., 2009). The linkage maps published by these researchers describe 18 linkage groups representing each of the 18 chromosomes of *B. juncea.* The coverage in terms of genetic distance ranges between 1266 cM and 1933 cM (Axelsson et al., 2000; Lionneton et al., 2002; Mahmood et al., 2003; Panjabi et al., 2008b; Pradhan et al., 2003; Ramchiary et al., 2007b) with the exception of a RAPD based map which has been described containing 21 linkage groups covering 790 cM distance. These linkage maps have been used for the identification of QTL for oil (Cheung et al., 1998; Mahmood et al., 2006b), fatty acid (Gupta et al., 2004; Jagannath et al., 2010; Lionneton et al., 2002; Mahmood et al.,

2003), protein content (Mahmood et al., 2006b) and agronomic traits such as yield (Ramchiary et al., 2007b) and maturity (Mahmood et al., 2006a).

The choice for the type of molecular markers to be used for linkage mapping studies depends on many factors including their ability to detect polymorphism, ease of use, reproducibility and most importantly how effectively they can be used in high throughput screening for genotypic selection (Agarwal et al., 2008; Varshney et al., 2005). RFLP markers have been shown to be highly reproducible however they require prior sequence information, require the use of radioactive isotopes, and protocols to use them are generally laborious. Similarly, AFLP require repeated steps which are time consuming for high throughput screening. Microsatellite markers have been shown to have a high degree of polymorphism and are reproducible for the purpose of screening (Tautz, 1994). The SRAP marker can be used to identify genomic regions of interest as the amplified regions can be used for obtaining sequence data (Li and Quiros, 2001). The use of microsatellite and SRAP markers for the development of genetic maps in *B. juncea* has been limited. Currently there are number of publicly available libraries of SSR and SRAP markers developed in *B. napus*, *B. rapa* and *B. oleracea* which can be investigated for use in the development of a linkage map of the *B. juncea* genome.

The goal of this study was to develop linkage maps in two DH *B. juncea* populations developed from the cross of ZEM1/Vniimk351 and ZEM1/Vniimk405 using SSR and SRAP markers that have been effectively used in the development of linkage maps in *B. napus.* These genetic maps would be valuable to understand the genome organization of *B. juncea* and development of markers linked to economically important traits such as oil and yield.

3.3 Materials and methods

3.3.1 Plant materials and DNA isolation

The parental lines for the development of the two mapping populations used in this study originated from differing genetic backgrounds. ZEM1 was identified in Australia where it was identified as being a natural mutation that had very low erucic acid expression (Kirk and Oram, 1981). Later studies identified polymorphisms in both the *FAE1.1* and *FAE1.2* genes that resulted in the zero erucic acid designation of the line (Gupta et al., 2004). These lines were made publicly available for breeding and development of canola quality *B. juncea* (Burton et al., 2003). The Vniimk lines were obtained from public breeding programs in Russia. Vniimk 351 has been studied in plant regeneration studies as negative control similar to ZEM (Bonfils et al., 1992). Vniimk 405 has been used for the study of *Albugo candiada* (White rust) resistance and virulence of *Albugo candiada* races (Prabhu et al., 1998). Both Vniink351 and 405 are moderate erucic acid lines similar to native *B. juncea* species and have been developed for the condiment mustard market in Russia.

Two DH populations of *B. juncea* were developed using microspore culture developed by Dow AgroSciences (DAS) Inc, Saskatoon, Saskatchewan. The first population comprising 106 DH lines was developed from the F_1 hybrid of the pure line of Zem1 and Vniimk 351. The second population was developed from the F_1 of the cross of the pure lines of ZEM1 and Vniimk 405. This population included 56 DH lines. Both populations were developed using DH protocols proprietary to DAS. The DH populations were genotyped for molecular markers at the department of Plant Science at the University of Manitoba, Winnipeg. Genomic DNA was isolated from young leaves at the two to three leaf stage for each of the 172 DH lines and the three parental lines using a modified 2X CTAB method. Approximately 2 g of plant tissue was

macerated in liquid nitrogen for genomic DNA extraction. The ground tissue was mixed with 20 ml of 2x CTAB buffer (2% CTAB, 20mM EDTA, 100 mM Tris, 1.4 M NaCl, pH-7.5) into 50 ml tubes and incubated at 65 ºC for 90 min. After incubation 20 ml of chloroform was added and vigorously mixed followed by separation in a centrifuge at 3200 rpm for 15 min to remove plant material and proteins. The supernatant was extracted and placed in a new 50 ml tube. Another chloroform treatment was repeated as described above using equal volume to supernatant. Following the centrifugation the supernatant was removed and placed in new 50 ml tube. The genomic DNA was precipitated from 15 ml of supernatant using 0.5 volume of isopropanol, mixed and centrifuged at 3000 rpm for 2 min. The supernatant was discarded and the DNA was washed with 70% ethanol and the pellet was air dried and dissolved in 2 ml of distilled water.

3.3.2 Identification of polymorphisms and molecular mapping

For polymorphism detection, the DNA extracted from the parental lines was used to amplify the fragments by PCR for 384 SSR markers. The SSR primer sequences were obtained from the *Brassica* Microsatellite Exchange (*Brassica*.info). All PCR reactions for SSR markers were performed in 10 µl volumes containing 4.6 µl of dH_2O, 1 µl 10X PCR buffer (100mM Tris-HCl, 500mM KCl, 0.1% Triton X-100, pH-9.3), 0.8 µl 25mM $MgCl_2$, 0.04 µl 25mM dNTP, 0.3 µl (0.5U) Taq polymerase, 0.05 µl 10 pM forward primer, 0.15 µl 10 pM reverse primer, 0.05 µl 20 pM M13 labeled primer, and 3 µl of template DNA standardized to 30 ug/ul. The PCR program was as follows: 94 ºC for 5 min; 5 cycles 94ºC for 50 sec, 56 ºC for 50 sec, 72 ºC for 1 min, with a 0.8 ºC decrease in annealing temperature at each cycle; then 25 cycles with 94 ºC for 50 sec, 51 ºC for 50 sec, 72 ºC for 1 min followed by a final elongation at 72 ºC for 5 min. For each of the forward primers a 25 bp M13 sequence (5'-CACGACGTTGTAAAACGAC-3') was added to the 5' end to allow for separation of PCR products for fragment analysis using ABI 3100 Genetic

Analyzer (Applied Biosystem, CA, USA). M13 primers labeled with four colors (VIC-green, NED-yellow, FAM-blue, PET-red) were added to the PCR products. For the parental screening the VIC M13 primer was used to label all the PCR reactions. Data collected from the ABI was processed using GeneScan (Applied Biosystems, CA, USA) software and visualized using Genographer2.0 http://hordeum.oscs.montana.edu/software/genographer/ index.html) to detect polymorphism between the parental lines. Marker combinations that showed polymorphic band patterns were selected for characterization or genotyping of each of the two DH populations. For population screening the selected markers were amplified following the same PCR protocol as used for parental screening (described above). Each PCR reaction was labeled with one of the four colored M13 markers to allow for multiplex screening of the DH population. The PCR products from each marker were pooled in equal volumes for fragment analysis. In a 384 well plate 2.5 µl of pooled produces are mixed with 7.5 µl of a 40ul 400-LIZ standard to 2mL formamide solution. The mixture is denatured at 95 ºC for 5 min and immediately placed on ice. The denatured products were separated in the ABI 3100 Genetic Analyzer.

The SRAP screening used 16 primer pairs which contained 4 different florescent labeled forward primers (VIC-green, NED-yellow, FAM-blue, PET-red) and an unlabeled reverse primers described in Sun et al. (2007). All PCR reactions for SRAP markers were performed in 10 µl volumes containing 4.6 µl of dH_2O, 1 µl 10X PCR buffer (100mM Tris-HCl, 500mM KCl, 0.1% Triton X-100, pH-9.3), 0.8 µl 25mM $MgCl_2$, 0.04 µl 25mM dNTP, 0.3 µl (0.5U) Taq polymerase, 0.15 µl 10pM labeled forward primer, 0.15 µl 10pM reverse primer, and 3 µl of template DNA standardized to 30 µg/µl. The PCR program was as follows: 94 ºC for 5 min; 5 cycles 94 ºC for 50 sec, 35 ºC for 50 sec, 72 ºC for 1 min; then 25 cycles with 94 ºC for 50 sec, 51 ºC for 50 sec, 72 ºC for 1 min followed by a final elongation at 72 ºC for 5 min. The PCR

reaction followed the same protocol as described for the SSR marker screening. Analysis of the amplified fragments was completed using 2.5 µl of multiplexed PCR products. The fragments were denatured by adding 7.5 µl of formamide mixed with 500-LIZ standard and analyzed in ABI 3100 Genetic Analyzer as described in SSR procedure.

The SSR and SRAP marker data collected from the ABI 3100 genetic analyzer were processed using GenScan (Applied Biosystems, CA, USA). The ABI files were converted to a virtual gel-like image and scored for all polymorphisms using the Genographer 2.0 software (Genographer, http://hordeum.oscs.montana.edu/software/genographer/index.html). The SSR markers were scored as co-dominant and dominant markers and SRAP markers were scored as dominant markers. The marker scoring data for both populations was recorded in a spreadsheet independently and the linkage analysis was completed using JoinMap® 3.0 mapping software (Van Ooijen, 2001).

3.4 Results

3.4.1 Linkage map for ZEM1 x Vniimk351 DH population

The genetic linkage map for the ZEM1 x Vniimk351 DH populations was constructed based on a total of 160 polymorphic loci as revealed by 89 SSR and 71 SRAP markers. A total of 10 SSR and 10 SRAP markers were not associated with any of the linkage groups in our constructed map. Logarithm of odds (LOD) values ranging from 3 to 10 was used to select and identify 13 linkage groups populated with 140 markers. The total map coverage in length was 791 cM with an average of 5.6 cM distance between the markers. This distance roughly translates to 1.39 Mbp/cM based on the estimated 1105 Mbp that makes up the B. juncea genome (He et al., 2003).

The SSR marker used in this study has also been reported in other *Brassica* mapping studies http://www.*Brassica*.info/resource/markers/ssr-exchange.php). However, the SSR markers selected for such mapping studies have had limited use in the development of genetic maps of *Brassica juncea* inhibiting the ability to make comparisons with existing linkage maps (Koundal et al., 2008; Ramchiary et al., 2007b). The map developed for the ZEM1 x Vniimk351 population shared 11 markers with a recently published map of *B. juncea* (Ramchiary et al., 2007b). The SRAP markers selected for this study were used in genetic studies in *Brassica rapa* (Rahman et al., 2007) and were used to develop an ultra dense linkage map of *B. napus* (Sun et al., 2007). There has been no current publications (to our knowledge) have reported the utilization of SRAP markers used in this map for molecular screening of *B. juncea*.

Table 3.1: Outline of the marker distribution and map distance of the linkage groups identified for the ZEM1xVniimk351 population.

LG	Number of Molecular Markers	Length	Average interval	Number SSR	Number SRAP
1	37	119.9 cM	3.3 cM	21	16
2	10	70.5 cM	7.8 cM	6	4
3	9	68.3 cM	8.5 cM	6	3
4	7	50.0 cM	8.3 cM	2	5
5	12	61.8 cM	5.6 cM	6	6
6	9	53.5 cM	6.7 cM	7	2
7	7	58.1 cM	9.7 cM	4	3
8	6	75.0 cM	15.0 cM	3	3
9	5	60.9 cM	15.2 cM	0	5
10	4	28.8 cM	9.6 cM	3	1
11	17	62.4 cM	3.9 cM	10	7
12	14	47.7 cM	3.7 cM	10	4
13	3	34.2 cM	17.1 cM	1	2
Total	140	791.1 cM	5.7 cM	79	61

Segregation distortion was calculated for each of the markers tested used on the DH population. The expected 1:1 segregation ratio was tested and of the 140 tested 9 (6 SSR and 3 SRAP) were not found to follow the expected segregation. Of those markers 4 showed significant distortion and the other 5 were missing loci due to poor image resolution. The markers showing significant distortion were found on linkage groups (LG) 1, 4, 5, 10 and 13. The missing loci were associated with markers found on LG 1, 4, 9, 10, and 12.

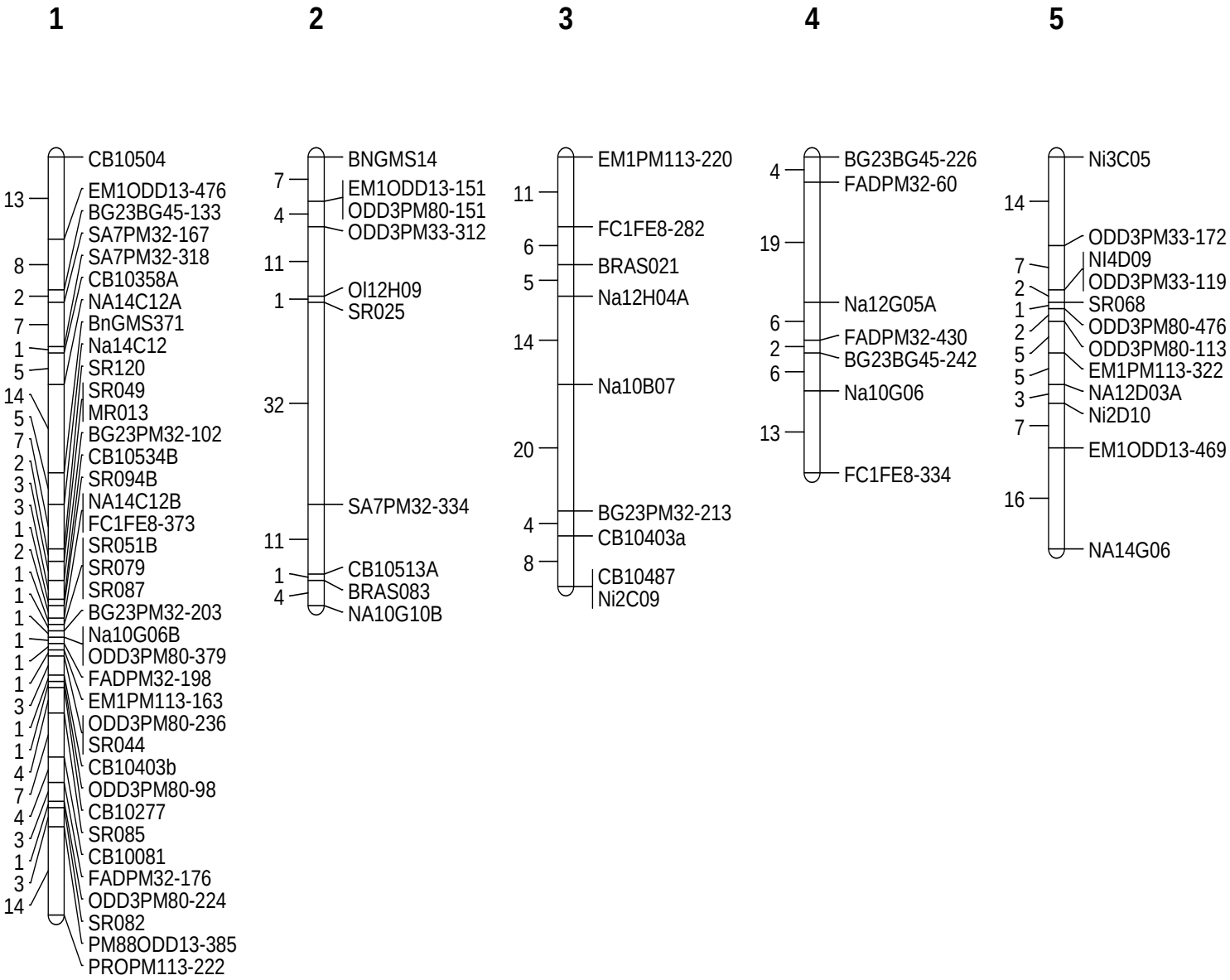


Figure 3.1: 13 Linkage groups identified for the ZEM1xVniimk351 population. Markers starting with Na, Ni, Ra, BNGMS, BRMS, CB, SR, MP, MD and BRAS identify the SSR markers. All other notations indicate the SRAP markers associated with the linkage groups.

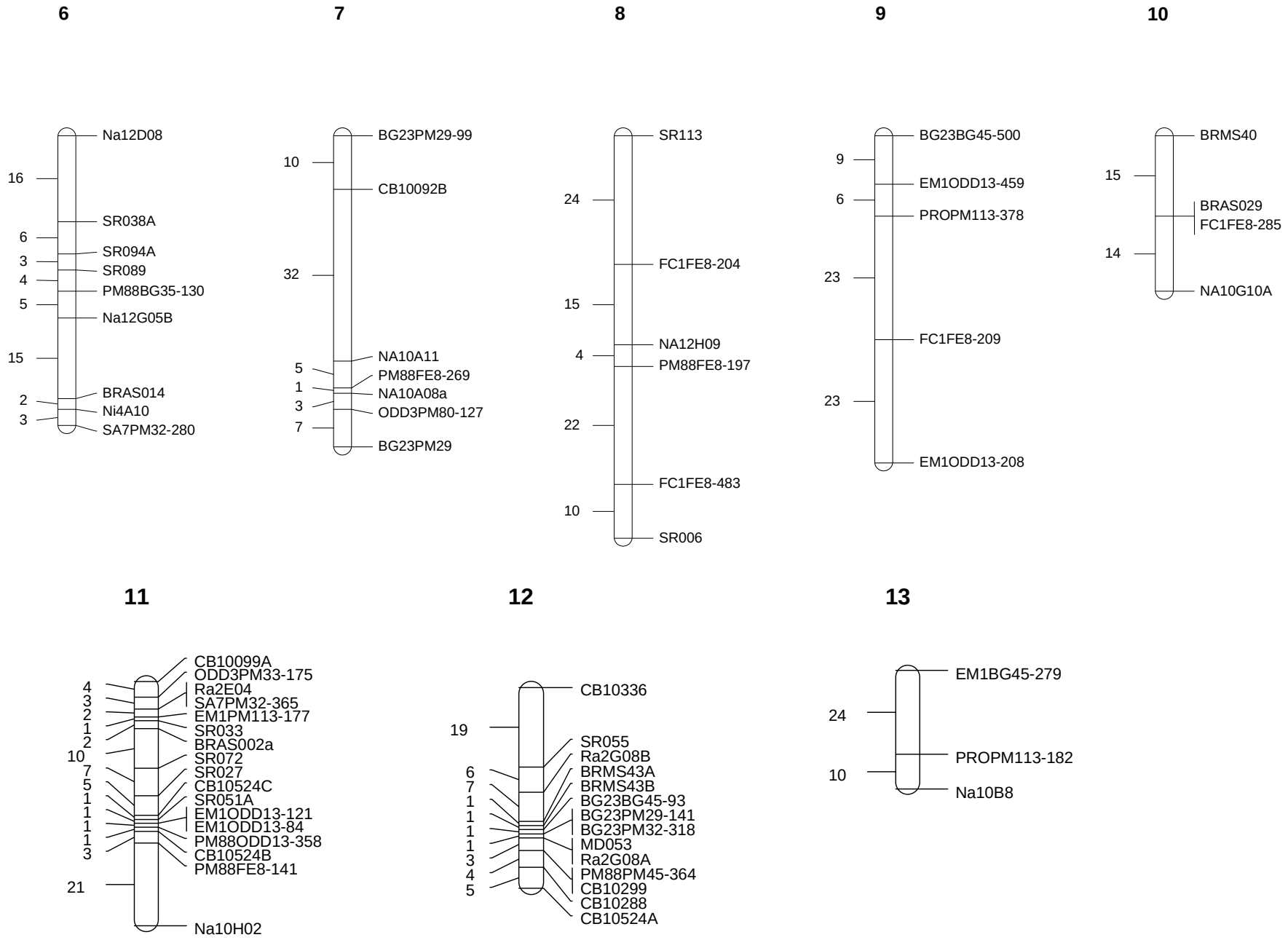


Figure 3.1 continued

3.4.2 Genetic map of ZEM1 x Vniimk405 DH Population

Using the same group of markers used in the previous linkage map described in section 3.4.1, 162 polymorphic markers were used to develop a linkage map consisting of 94 SSR and 68 SRAP markers for the ZEM1 x Vniimk405 population. A total of 19 SSR and 6 SRAP markers were not associated with any of the identified linkage groups. The final map contained 137 markers distributed over 13 linkage groups based on remaining markers. The map covers an overall length of 623 cM with an average distance between markers of 4.6 cM. This distance roughly translates to 1.77 Mbp/cM based on the estimated 1105 Mbp that makes up the B. juncea genome (He et al., 2003). Similar to the observations made with the genetic map developed for the ZEM1 x Vniimk351 population, there was no correlation of markers observed with other published linkage maps of *B. juncea*.

Table 3.2: Outline of the marker distribution and map distance of the linkage groups identified for the ZEM1xVniimk405 population.

LG	# of Molecular Markers	Length	Average interval	Number SSR	Number SRAP
1	24	82.2 cM	3.6 cM	14	10
2	18	67.9 cM	4.0 cM	7	11
3	16	73.0 cM	4.9 cM	10	6
4	14	60.0 cM	4.6 cM	10	4
5	12	34.5 cM	3.1 cM	5	7
6	11	81.0 cM	8.1 cM	6	5
7	9	30.0 cM	3.8 cM	7	2
8	8	28.6 cM	4.1 cM	5	3
9	6	35.8 cM	7.2 cM	2	4
10	6	48.8 cM	9.8 cM	2	4
11	5	36.3 cM	9.1 cM	2	3
12	4	21.9 cM	7.3 cM	2	2
13	4	23.6 cM	7.8 cM	3	1
Total	137	623.6 cM	4.6 cM	75	62

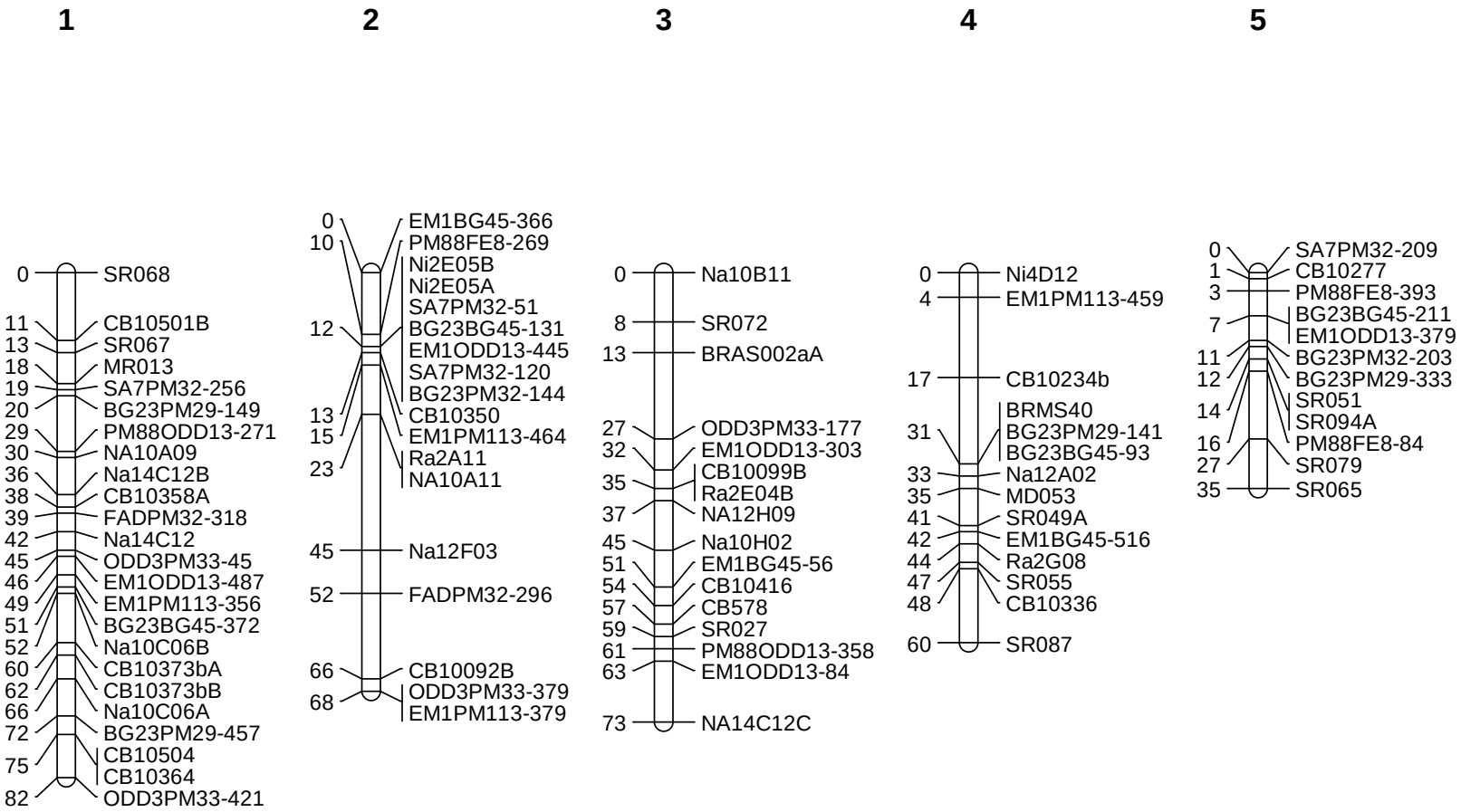


Figure 3.2: 13 Linkage groups identified for the ZEM1xVniimk351 population. Markers starting with Na, Ni, Ra, BNGMS, BRMS, CB, SR, MP, MD and BRAS identify the SSR markers. All other notations indicate the SRAP markers associated with the linkage groups

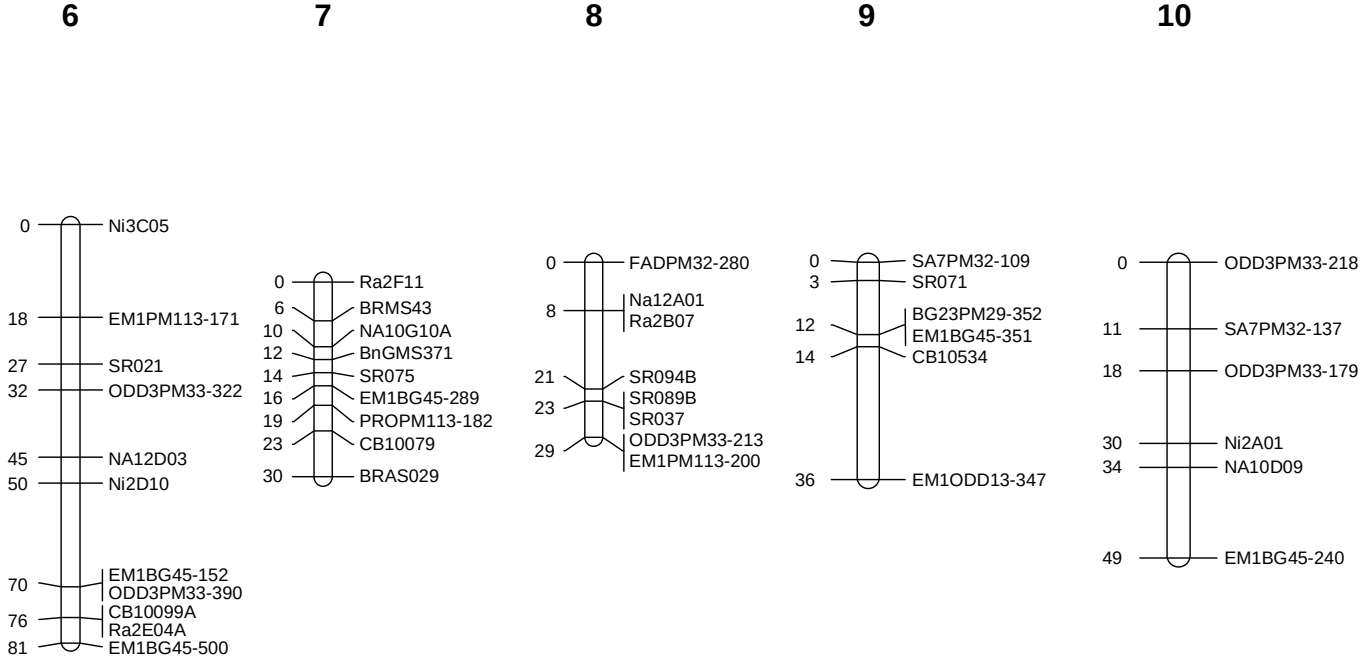


Figure 3.2 continued

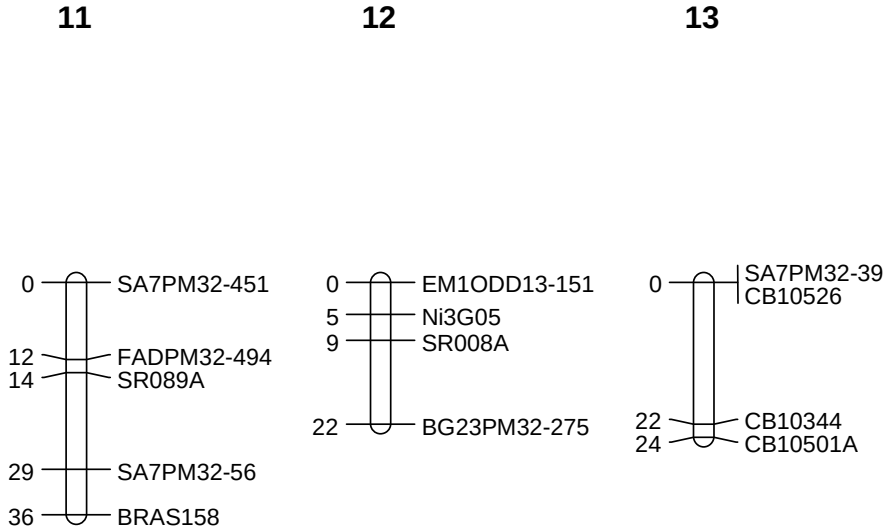


Figure 3.2 continued

3.4.3 Segregation distortion in linkage maps

Segregation distortion was calculated for each of the markers tested used on the DH population. The expected 1:1 segregation ratio was tested and of the 137 tested 13 (9 SSR and 4 SRAP) markers showed significant distortion. Of the 13 markers with significant distortion 1 marker on LG 7 was due to missing loci due to poor resolution during screening. There were 6 distorted markers on LG 8 with the segregation towards ZEM1. On LG 7 there were 2 distorted markers with segregation towards Vniimk405. Additionally 3 markers on LG 4 showed distortion, 2 towards Vniimk405 and 1 towards ZEM1. One marker on LG 6 had significant segregation distortion towards ZEM1.

3.4.4 Marker consensus between linkage maps

An analysis of the consensus between the two maps for both populations was conducted using JoinMap 3.0 to identify common markers. Little correlation was observed between the markers evaluated for the two populations. The linkage maps developed for the two populations totaled 230 markers of which 47 (8 SRAP and 39 SSR) markers were shared between both. There is an observed correlation between a few markers on some of the linkage groups as seen in Table 3.3. Similarities were found between LG1 in the ZEM1/Vniimk351 map and LG1 in the ZEM1/Vniimk405 map with 6 markers common between the two maps. However 6 markers were also shared with LG1 of ZEM1/Vniimk351 map and LG5 of ZEM1/Vniimk405 map. Other similarities included LG 13 and LG 14 in the ZEM1/Vniimk351 map with LG 3 and LG 4 in the ZEM1/Vniimk405 map respectively. Both shared 6 common markers between the two linkage groups.

The marker distribution varied between both genetic maps. The genetic map for the ZEM1xVniimk351 has an average marker interval of 5.7 cM however this ranged between 3.6

and 15.1 cM in specific linkage groups. This was reflective of the marker density with the genetic map. In this map markers are most dense on LG1, LG11 and LG13. The remainder of the linkage groups on this map had a wider distribution of markers and fewer markers per linkage groups with produced much of the variance in marker distribution. There was also some clustering of markers observed in LG 2 and LG 7 with a wide gap between markers at both ends of the linkage group.

In the map developed for the ZEM1/Vniimk405 population there is a total of 137 markers which was similar to the 140 markers found on the map for ZEM1/Vniimk351 however the overall distance of the map was almost 100 cM less. This reflects the evenly distributed markers within the ZEM1xVniimk405 map within each of the linkage groups. The average distance between markers in this map was 4.6 cM and a range between 3.1cM and 9.8 cM in the individual linkage groups. Clustering of markers was observed in LG2 in ZEM1/Vniimk405 map.

Table 3.3: List of common markers and their associated linkage groups identified in the linkage maps developed for the two DH B. juncea populations. POP1 refers to the ZEM1xVniimk351 population and POP2 refers the ZEM1xVniimk405 population.

Marker	LG in POP1	LG in POP2	Marker	LG in POP1	LG in POP2	Marker	LG in POP1	LG in POP2
CB10358A	1	1	EM1ODD13-151	2	12	BRAS002aA	13	3
CB10504	1	1	SR068	5	1	EM1ODD13-84	13	3
NA14C12C	1	1	Ni3C05	5	6	PM88ODD13-358	13	3
Na14C12	1	1	NA12D03	5	6	Na10H02	13	3
Na14C12B	1	1	Ni2D10	5	6	SR072	13	3
MR013	1	1	SR089B	6	8	Ra2E04B	13	3
SR049A	1	4	SR094B	6	8	SR051A	13	5
SR087	1	4	SR089A	6	11	CB10099A	13	6
CB10277	1	5	CB10092B	7	2	Ra2E04A	13	6
BG23PM32-203	1	5	NA10A11	7	2	CB10336	14	4
SR079	1	5	PM88FE8-269	7	2	BG23BG45-93	14	4
SR051	1	5	NA12H09	8	3	BG23PM29-141	14	4
SR094A	1	5	BRMS40	10	4	MD053	14	4
SR051B	1	5	BRAS029	10	7	SR055	14	4
BnGMS371	1	7	NA10G10A	10	7	Ra2G08	14	4
CB10534	1	9				PROPM113-182	15	7

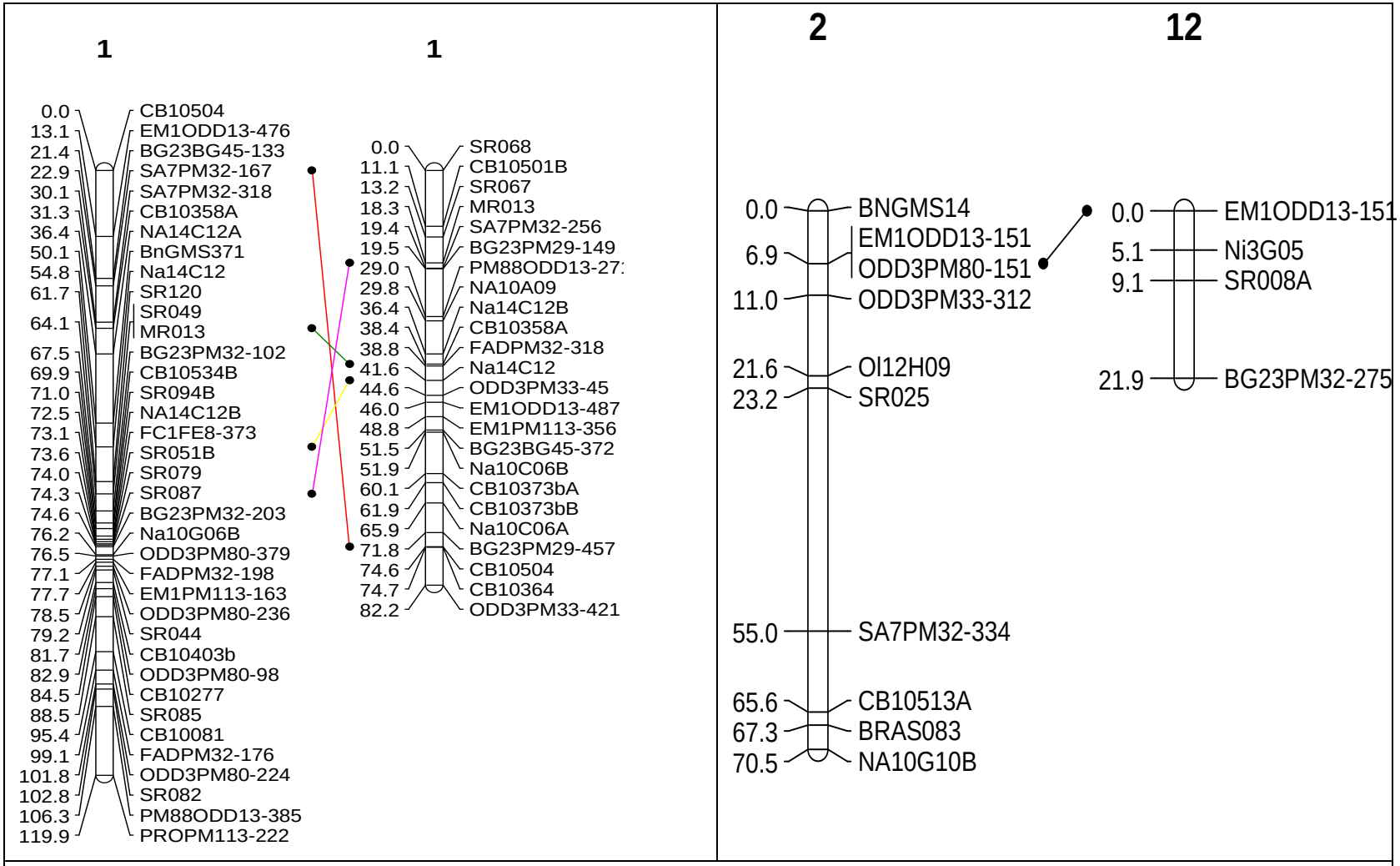


Figure 3.3: Location of the regions containing common markers identified between the two maps developed for the ZEM1xVniimk351 and Zem1x Vniimk405 populations. The ZEM1xVniimk351 linkage groups are located on the left and the common markers are connected between the two linkage groups

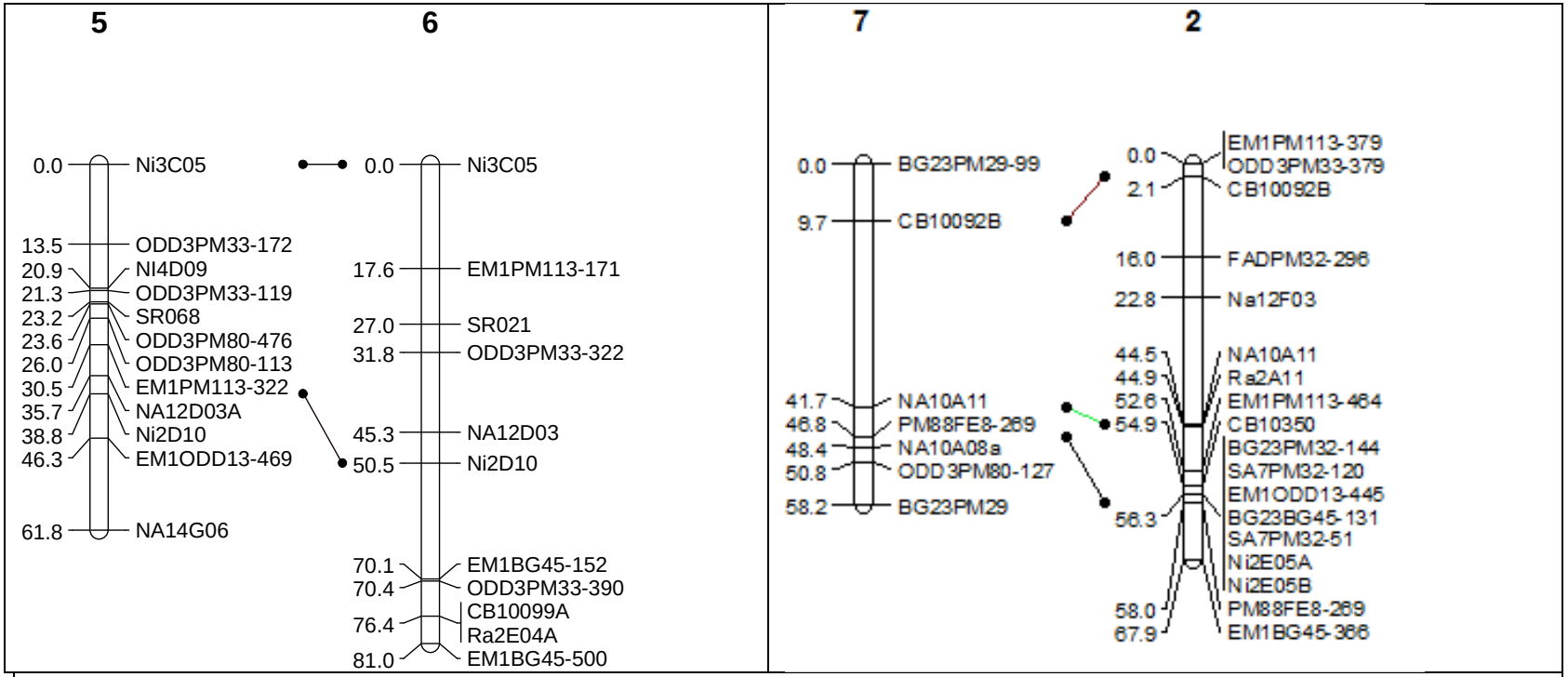


Figure 3.3 continued

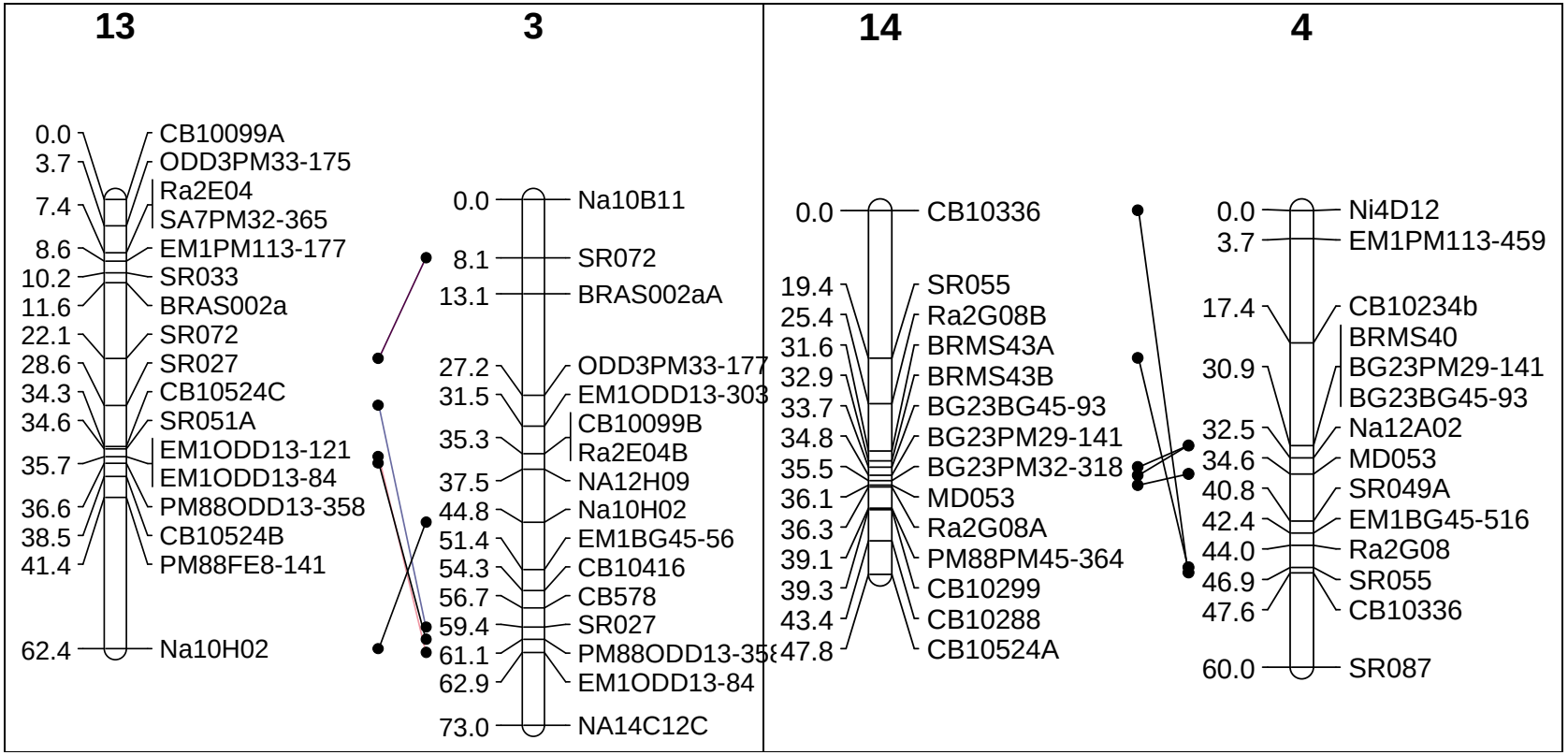


Figure 3.3 continued

3.5 Discussion

This study developed partial linkage maps of two DH *B. juncea* populations using SSR and SRAP molecular markers. The DH populations used were developed using a cross of yellow seeded lines of *B. juncea* ZEM1 and Vniimk351 (POP1) and ZEM1 and Vniimk405 (POP2). From the ZEM1/Vniimk351 population, 107 DH lines were developed and used for the development of a linkage map. A population of 57 DH lines was also developed from the cross between ZEM1/Vniimk405 and analyzed for linkage map development.

Current molecular maps of the *B. juncea* genome have relied heavily on RFLP and AFLP markers for the development of linkage maps. The only published map that has utilized SSR markers (Ramchiary et al., 2007b) contains 63 SSR markers within a dense map including 1297 AFLP and 72 RFLP markers. Similarly the use of SRAP markers have been limited to the study of genetic diversity of *B. juncea* lines (Wu et al., 2009). The SRAP markers used originated from early *Brassica napus* studies and they showed good polymorphisms in *B. juncea*. The SRAP markers selected were not the same as in those used in previous mapping studies for the development of the two maps due to marker availability. Both SSR markers and SRAP markers have not been used to develop linkage maps for *B. juncea*.

The genetic map developed from the ZEM1/Vniimk351 DH population was comprised of 13 linkage groups with a total length of 791 cM and an average marker interval of 5.7 cM. This map does not span the entire 18 linkage groups observed in other currently available maps. The published linkage maps of *B. juncea* that cover the entire genome (Axelsson et al., 2001; Cheung et al., 1997b; Lionneton et al., 2002; Mahmood et al., 2003; Panjabi et al., 2008b; Pradhan et al., 2003; Ramchiary et al., 2007b) have a length ranging from 1266 cM to 2073 cM with an average length of 1705 cM. Similar results were found in the linkage map developed from the

ZEM1/Vniimk405 DH population. This map covered 13 linkage groups with a total length of 623 cM and an average marker interval of 4.6 cM. As the maps for both populations covered only 13 of the 18 known linkage groups in *B. juncea*, this accounts for the reduced map distance in comparison to currently published linkage maps.

A comparison of the linkage map distance as compared to the physical distance of the genome can show the relative differences between the currently published maps and the maps presented in this study. As the B. juncea genome is estimated to contain approximately 1105 Mbp the distance represented in the ZEM1 x Vniimk351 and ZEM1 x Vniimk405 population covers a distance of 1.39 Mbp/cM and 1.77 Mbp/cM respectively. The distance between each cM unit in the linkage maps developed in this study is significantly wider when compared to the relative distance of current linkage maps that cover a range from 1266 cM (0.87 Mbp/cM) to 2073 cM (0.57 Mbp/cM) with an average distance of 1705 cM (0.64Mbp/cM). This suggests that the linkage maps developed in this study are covering approximately half of the physical distance of the genome as compare to other complete linkage maps for B. juncea as the number of base pairs per centiMorgan are approximately half of those observed in these other studies.

As mentioned earlier the use of SSR markers in published *B. juncea* linkage maps has been limited. In the screening of SSR markers used in the construction of the two maps 341 SSR markers were screened for the ZEM1/Vniimk351 and ZEM1/Vniimk405 populations and 79 and 75 markers were identified as being polymorphic respectively. A recently published map of *B. juncea* identified 63 polymorphic markers from 363 SSR markers that were obtained from the same public source used in this study (Ramchiary et al., 2007b). These markers were developed from different *Brassica* species including *B. oleracea*, *B. rapa* and *B. nigra* and *B. napus*. They were primarily used for the development of linkage maps for *B. napus* and *B. rapa* however they

had limited use in the other *Brassica* species mentioned. This low rate of polymorphisms in the *B. juncea* populations suggests that these markers are preferentially weighted to the A genome and C genome that is common between *B. napus*, *B. oleracea* and *B. rapa*. Studies suggest that *Brassica* genomes are derived from a hexaploid ancestry that resulted from the triplification of an unknown ancient diploid species (Truco et al., 1996). This common ancestry results in a number of homologous regions identified within the A, B and C genome of *Brassica* sp. However studies looking at the inherent similarities of the three *Brassica* genomes shows that the B genome of *B. nigra* comes from a more distant descendant than that observed between the A and C genomes of *B. rapa* and *B. oleracea* respectively (Attia and Röbbelen, 1986; Palmer et al., 1983; Quiros et al., 1991; Song et al., 1988). These studies suggest a high degree of homology with the C genome of *B. oleracea* and the A genome of *B. rapa* in comparison to the B genome of *B. nigra*. Similar studies have shown a wide genetic distance between the A and C genome to that of the B genome and suggests that the A and C genome has a common ancestral lineage (Lagercrantz, 1998; Lagercrantz and Lydiate, 1996). This suggests that the coverage of the B genome in both of the linkage maps developed for the ZEM1 x Vniimk351 and ZEM1 x Vniimk405 populations may be limited as the markers used were developed for screening the *Brassica* A and C genome This genetic distance will result in fewer markers amplified within the B genome which could explain the limited amount of polymorphic markers identified by screening the *B. rapa* and *B. napus* derived SSR marker sequences in *B. juncea*. Also the rate of polymorphisms identified in the amplified regions may be less in the B genome than is present in the A and C genome resulting in a lower rate of polymorphisms identified within other *Brassica* species.

This divergence has been illustrated in the development of euploid inter-genetic translocation lines developed between *B. napus* and *B. nigra* (Struss et al., 1995) where B genome traits such as erucic acid were observed but there was no inter-genomic chromosome pairing observed (Struss et al., 1991) such as observed in wheat rye introgression lines (Ren et al., 1990). However the 2 genes identified for erucic acid in *B. carinata* and *B. juncea* have been identified on the B genome of the amphidiploid species suggesting that there was an introgression of the erucic acid genes from the A and C genome of these species to result in the additional erucic acid gene appearing in the B genome of these two *Brassica* species (Struss et al., 1996). The genetic variation between B genome and the A and C genome was illustrated using repetitive sequence probes which showed a wide variation between the sequences found between these three genomes (Somers and Demmon, 2002). Similar patterns of genome duplication were observed between the A and C genome in the *B. napus* genome which leads to the similarity of replication of markers within the napus genome (Parkin et al., 2003).

The limited number of polymorphisms observed in the microsatellite screening was further observed with the relatively high degree of polymorphisms observed in the SRAP markers screened within the two populations. A total of 15 marker pairs were screened for both populations. This resulted in the identification of 71 polymorphic markers in the ZEM1 x Vnnimk351 and 68 polymorphic markers in the ZEM1xVniimk405. Due to the random binding of the SRAP markers there is an equal opportunity of amplification throughout the genome allowing for more amplification within the B specific region where the SSR markers may have been limited due to their focus on A and C specific genome amplification.

Both maps contained markers that had shown significant marker segregation distortion. Homology between 2 linkage groups showed evidence of marker segregation which could be

explained by amplification of multiple linkage groups causing such distortion. Studies have shown homologous regions between the A and B genome (Lagercrantz and Lydiate, 1996). In the ZEM1/Vniimk405 linkage map significant marker segregation was observed in 3 linkage groups. These linkage groups could share homology between the A and B genome which can result in a deviation within these regions due to amplification of multiple regions within the genome.

Currently there are a limited number of published linkage maps for the *Brassica juncea* genome. As mentioned earlier, the available linkage maps are primarily developed using RFLP and AFLP marker technology. A recent map (Ramchiary et al., 2007b) introduced 69 SSR marker to an existing map developed using RFLP and AFLP markers. The maps developed using the population derived from ZEM1xVniimk351 and ZEM1xVniimk405 shared 11 SSR markers present in the linkage map developed by Ramchiary et al. (2005). There was a correlation observed between LG 5 of ZEM1xVniimk351 and LG 6 of ZEM1XVniimk405 with 2 markers located on LG15 of the Ramchiary map. Outside of this there were common markers found on LG 3, LG 6and LG 13 of the ZEM1xVniimk351 population which correlated with LG 17, LG 16 and LG1 of the Ramchairy's map respectively. For the linkage maps for the ZEM1xVniimk405 population there were correlations between markers on LG 2, LG7, LG8, and LG10 which coincided with LG 9, LG 3, LG 12, and LG11 respectively. Markers on LG4 and LG12 of the ZEM1xVniimk405 shared markers located on LG 18 of the Ramchiary maps. As the LG12 of the map is a small linkage group it is possible that this could be an unlinked portion of the LG4. As use of SSR markers has been limited to one currently published *B. juncea* linkage map there was no way to confirm a significant correlation with other linkage maps of *B. juncea* populations with the maps developed for the ZEM1xVniimk351 and ZEM1xVniimk405 populations.

Additionally there has been no published maps developed using the SRAP primers utilized in the development of the linkage maps. However, as these markers rely on a random binding for the primers to anneal therefore the reproducibility of these markers beyond the populations being studied are not a reliable source for determining correlations between existing linkage maps.

Attempts have been made to correlate the A, B and C genomes with other linkage maps developed for *B. rapa* and *B. napus*. Studies have repeatedly shown the similar lineage between A and C genomes in comparison to B genome using molecular markers to determine the degree of relativeness (Johnston et al., 2005; La Mura et al., 2010; Liu and Wang, 2006; Lysak, 2005; Lysak and Lexer, 2006). These studies have shown that there is greater genetic distance between the B genome that the A and C genome. This variation may relate to the increased selection for markers that amplify in A genome regions with markers that have been studied exclusively in *B. napus*. This has been homology identified between B genome and *B. napus* has been displayed using STS markers (Plieske and Struss, 2001) however as recent research has shown there is still a significant degree of variation between these genomes which could cause the limited coverage of the B genome when markers are amplified.

3.6 Future Research

The further development of the molecular maps using existing SSR and new molecular markers can help improve the understanding of the genomics of the *B. juncea* species. Further screening using SSR and SCAR markers can be used to correlate the genetic regions of *B. juncea* with other highly studied species such as *B. napus* and *B. rapa*. Additionally, anchoring the maps developed in this study with existing maps using RFLP and AFLP markers will aid in correlation within the currently published maps. These linkage maps offer another tool to breeders and molecular biologists to further study the *Brassica* sp. and aid in the improvement of oil quality and agronomic performance to help improve oilseed production within Canada and elsewhere.

4.0 QTL Analysis for oil, protein and fatty acid profile of two DH *Brassica juncea* populations

4.1 Abstract

Quantitative trait loci were detected for seed oil, protein and fatty acids using molecular markers in two double haploid (DH) populations derived from the crosses between a low erucic variety ZEM1 and two moderate erucic acid linesVniink351 and Vniimk405. In both populations the erucic and oleic acid content ranged between 0.17% to 31.0% and 15% to 49% respectively of the overall fatty acids contained in the seed oil. The differences observed in the erucic and oleic acid content corresponded to variations observed within other fatty acids including linoleic, eicosenoic and arachidic fatty acid. A QTL analysis for fatty acids indentified QTLs on linkage groups 1, 7 and 12 for the ZEM1xVniimk351 population and on linkage groups 1, 3 and 4 for the ZEM1xVniimk405 population. The deviation of fatty acids in both population show significant evidence for the monogenic inheritance of the *FAE1* gene which results in the variation observed in the fatty acid profile of both populations. The QTL analysis for the seed oil and protein content in the ZEM1xVniimk351 population identified 2 QTL on LG1 and 4 and 1 QTL on LG1 respectively. The QTL analysis ZEM1xVniimk405 of oil and protein content identified 1 QTL for oil and protein on LG1. A significant positive correlation between erucic acid and seed oil content suggests that the variation in erucic acid content accounted for much of the differences observed in the seed oil content in both populations.

4.2 Introduction

The primary use for oilseed *Brassica* crops in Canada is for the extraction of oil for use in the food industry while the meal obtained as the byproduct of the oil extraction process is used as a protein supplement in animal feed (StatCan, 2012). A mandate by the Canadian government to increase the use of biodiesel will result in an increase demand for oilseeds of which canola in the major source in Canada (Canola Council of Canada, 2007). This increase requires research to focus on the increase of oil production from current cultivated acres and the increase of overall oilseed production within Canada. *Brassica juncea* offers the introduction of oilseed production in regions that are not suitable for to *Brassica napus* production (Woods et al., 1991).

Brassica juncea has limited production within western Canada and is mainly localized to the southern regions of Saskatchewan and Alberta due to the hot and dry condition within these regions (Woods et al., 1991). Compared to the traditionally grown *B. napus, B. juncea* has improved drought and heat tolerance allowing it to grow in regions not suited to current canola production (Burton et al., 2008; Kumar et al., 1984; Kumar et al., 1987; Kumar et al., 1994; Sakova et al., 1995; Wright et al., 1995). The majority of the *B. juncea* currently grown in western Canada has been used for condiment mustard production which requires low seed oil content for processing coupled with moderate erucic acid and high glucosinolate levels which are undesirable traits for canola quality oil (Hemingway, 1995; Skrypetz, 2003). High oil and low erucic acid lines of *B. juncea* have been used for the development of canola quality mustard which would allow the canola quality oil production to be introduced in regions which are more suitable for *B. juncea* production (Burton et al., 2004).

The improvement of oil content and oil quality is a significant component of breeding oilseed *Brassica juncea*. Selection requires the evaluation of phenotypic data collected through testing

in the field at a variety of locations to identify the desired qualities. QTL analysis can been used to understand the genetic regions that significantly contribute to the inheritance of oil content and help identify genetic regions that control desired oil quality traits. Using molecular maps developed for *B. juncea* from RFLP, AFLP and SSR markers have been used for the investigation of quantitative traits such as oil, yield and oil quality traits (Aggarwal et al., 2003; Cheung et al., 1998; Lionneton et al., 2002; Mahmood et al., 2005a; Mahmood et al., 2006b; Sharma et al., 1999; Sharma et al., 2002a). QTL analysis has helped identify the number of genes involved and aided in the isolation of genes involved in the inheritance of traits such as erucic, linolenic, oleic fatty acid expression in *Brassica* sp (Fourmann et al., 1998; Gupta et al., 2004; Hu et al., 2006; Jourdren et al., 1996b; Lionneton et al., 2002).

The goal of this study is to identify QTL for seed oil, protein and the fatty acid content of two DH *Brassica juncea* populations. Using linkage maps developed for both populations, a QTL analysis was completed using data collected from individual lines to identify genetic regions involved in the inheritance of seed oil, protein and specific fatty acids within the oil. Genotypic screening of QTL can allow breeders to screen populations for desired phenotypes prior to field testing improving the efficiency of breeding programs.

4.3 Material and methods

Two DH populations of *B. juncea* were developed by DOW AgroSciences Inc. using their proprietary microspore culture protocol. The first population was developed using a yellow seeded, high glucosinolate, low erucic acid pure line of *B. juncea* (ZEM1) crossed to a yellow seeded high glucosinolate, moderate erucic acid *B. juncea* pure line (Vniimk351). The F_1 of this cross was used as a source of microspores to develop 106 DH lines. Similarly the second population was developed using the same yellow seeded, high glucosinolate, low erucic acid pure line of *B. juncea* (ZEM1) crossed to another yellow seeded high glucosinolate, moderate erucic acid *B. juncea* pure line (Vniimk405). The F_1 of this cross was used as a source of microspores to develop 56 DH lines. The DH lines were grown in the greenhouse and self pollinated to produce seed for field evaluation of DH lines.

The DH populations were grown under field conditions in a replicated nursery format in Chile during the winter 2009. The same populations were also grown in the summer of 2010 at the University of Manitoba, Winnipeg and Carman, Manitoba. The individual DH lines and parental lines were planted in single 2m rows and replicated 4 times at each location. The Chile trials were completed by Dow AgroSciences at their winter nursery site. Seed harvested at all three trial locations was analyzed at the University of Manitoba seed quality lab to obtain the data on oil and protein content and fatty acid profile.

For each of the DH lines in both populations the fatty acid profile was measured using Gas Chromatography in the quality lab at the University of Manitoba. Samples of approximately 30mg were collected from each line and placed into eppendorf vials where 0.5 ml of heptane was added. The samples were crushed and let stand for approximately 12 hrs to extract the oil. The transesterification of the triglycerides into fatty acid methyl esters was completed by adding 100

µl solution of 0.5M sodium methoxide reagent (15.5g sodium methylate in 500 ml anhydrous methanol) to the sample (F.W. Hougen and Bodo, 1973; Liu, 1984). The vial was shaken vigorously for 20 seconds and left to stand for 20 minutes and repeated once again after the time had elapsed. The solution then had 100 µl of acidified water (0.3% acetic acid) added to the vial and mixed gently to combine the two solutions. The combined solution was left to stand at 4 degrees C for 1-2 hours allowing the mixture to settle (Christie, 1989). From the reaction mixture 250 µl was extracted and transferred to a 2 ml autosampler vile fitted with a 250 µl insert. The samples were analyzed using a Varian model 3900 Gas Chromatographer fitted with a CP-Wax52 CB capillary column and flame ionization detector. The column was 15 m x 0.32 mm fused with silica coated with a 0.025 micron polyethylene glycol phase (Varian, Walnut Creek, USA). Ultra high purity (UHP) helium was used as the inert carrier gas was set at a flow rate of 2.0 ml/min. Varian model CP-1177 injectors were used in the Varian model CP-8400 autosampler with an inlet liner which was a cup splitter with 10% OV-1 on Chromsorb-W HP. The injector operated at 250 degrees C with the column oven programmed between 190 to 240 degrees C and the detector temperature at 280 degree C. The peaks were detected using Varian Star Workstation software. The unit was calibrated using a GLC #421 reference standard (Nu-Check Prep, Elysain, Minnesota) to ensure proper GC operation.

The total seed oil and protein content were measured using near infrared spectroscopy based on protocol developed at the University of Manitoba based on the algorithms developed from validated samples (Shenk and Westerhaus, 1991; Windham et al., 1989). Samples were evaluated on a NIR system model 6500 Near Infrared Spectroscope (FOSS, North America). Results were collected using standard protocols in the quality lab at the University of Manitoba (Kim et al., 2007). Analysis was conducted on 4 gram samples and the data was collected on

WinISI v.1.04a software. Oil and protein levels were expressed as the total of the overall seed content.

Data was analyzed using Statistical Analysis System 9.2 (SAS Institute Inc. NC). The analysis of variance for oil and protein content as well as the fatty acid profile were performed using PROC MIXED. The variance between each line was analyzed as a complete randomized block design experiment with 4 replicates of each line. The analysis was complete using the model $Yijk = \mu + G_i + E_j + GE_{ij} + B_{jk} + \varepsilon_{ijk}$ where μ is the mean, G_i is the effect of the i^{th} line, E_j is the effect of the j^{th} environment, GE_{ij} is the interaction of the i^{th} line with the j^{th} environment, B_{jk} is the effect of the k^{th} replicate in the j^{th} environment, and ε_{ijk} is the random error.

QTL analysis for both populations was competed using WinQTL Cartographer v. 2.5 (Basten et al., 2001) using composite interval mapping. The detection of the QTL was performed at 1 cM intervals using a 5 cM window size using backward/forward regression analysis. The LOD values were determined by threshold values estimated by 1000 permutations of trait data across all genetic intervals. The QTL analysis was done on oil content, protein content, palmitic, stearic, oleic linoleic, linolenic and eicosenoic and erucic acid levels.

4.4 Results

4.4.1 Oil and protein content variation

The oil content of the ZEM1xVniimk351 DH population ranged from 24-45% while it was 33-45% for the ZEM1xVniimk405 DH population. The overall distribution of the oil percentage in the ZEM1xVniimk 351 population had a mean value of 38% and followed a normal distribution skewed towards the higher oil content. The ZEM1 parental line had an oil content of 39% as compared to the 46% observed in the Vniimk351 parental line. The distribution of the ZEM1xVniimk405 population also showed a normal distribution across the population with a mean of 39% and is also skewed toward the higher oil content. The overall oil content of the ZEM1 and Vniimk405 parental lines were observed as 39% and 46% respectively.

This distribution of the protein levels in both populations followed the similar trends as observed for the oil content. The protein content in the ZEM1xVniimk351 DH population followed normal distribution within a 21-39% range and a mean of 27%. The distribution of the protein content was skewed toward a lower protein level contrasting the distribution observed for the oil content observed within this population. The protein content in the ZEM1xVniimk405 DH population followed a normal distribution within a 21-33% range and a mean of 26%. The parental lines contained a protein content of 26% in ZEM1, 22% in Vniimk351 and 23% in Vniimk405was observed in the ZEM1, Vniimk351 and Vniimk405 with values of 26%, 22% and 23% respectively.

The sum total of the oil and protein content within the seed within the ZEM1 X Vniimk351 population followed a normal distribution with the data ranging from 62 to 73% with a mean of 66%. In the ZEM1xVniimk 405 the population followed a similar distribution within the range

of 62.5% to 72% with a mean of 66%. The overall sum of oil and protein of the ZEM1 parental line was 65% with the Vniimk351 and Vniimk405 populations having sum value of 68 and 69% respectively.

Table 4.1: Statistical analysis of the percentage oil content, protein content and sum of oil and protein for the three parental lines used in the development of both the ZEM1xVniimk351 and ZEM1xVniimk405 populations.

Variable	Line	Mean	SE Mean	StDev	Minimum	Maximum
PROTEIN	**Vniimk351**	22.527	0.258	1.484	20.453	25.304
	Vniimk405	23.601	0.216	1.479	21.316	27.971
	ZEM1	26.602	0.103	1.429	22.532	31.347
OIL	**Vniimk351**	46.013	0.399	2.292	42.16	49.905
	Vniimk405	46.063	0.416	2.851	38.468	52.088
	ZEM1	39.152	0.133	1.842	34.276	43.659
SUM	**Vniimk351**	68.479	0.282	1.623	65.361	73.228
	Vniimk405	69.664	0.246	1.684	65.689	73.404
	ZEM1	65.754	0.0849	1.173	63.211	68.502

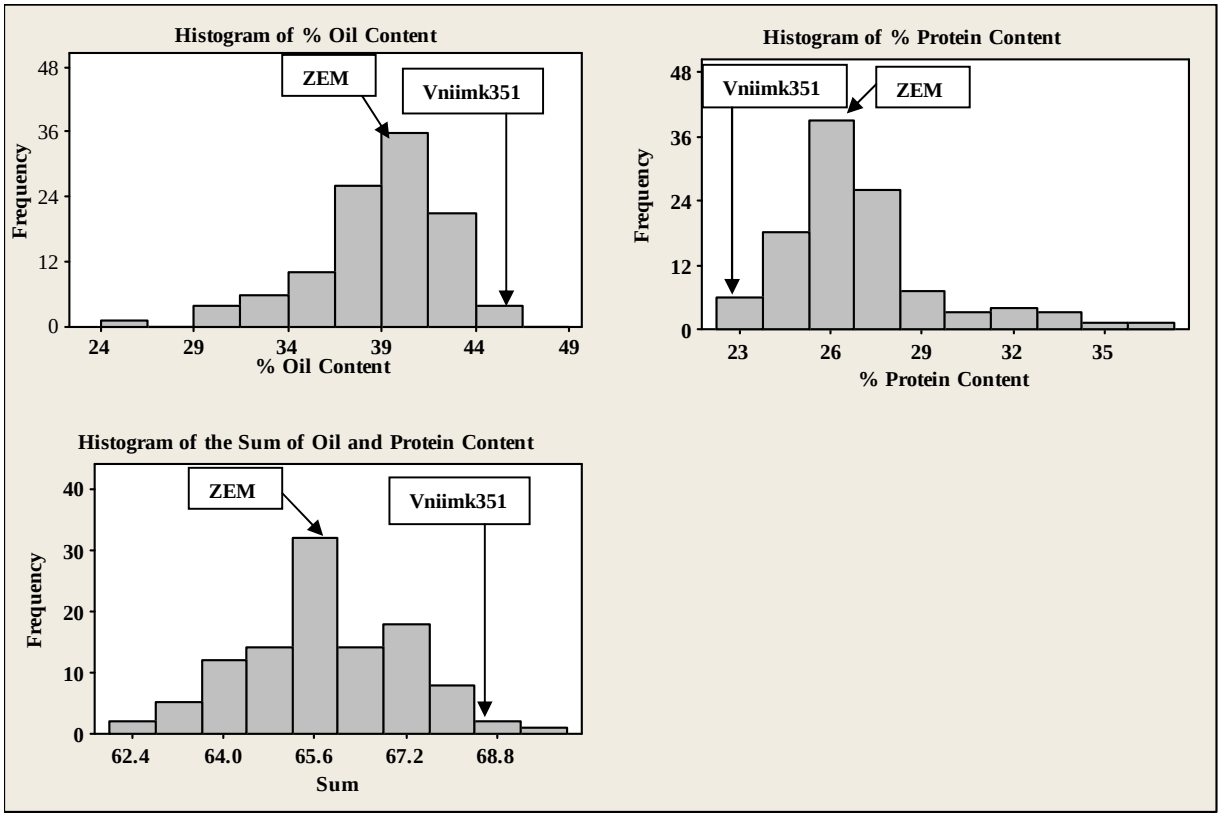


Figure 4.1: Distribution of the percentage of oil, protein and the sum of oil and protein content in the ZEM1xVniimk351 population.

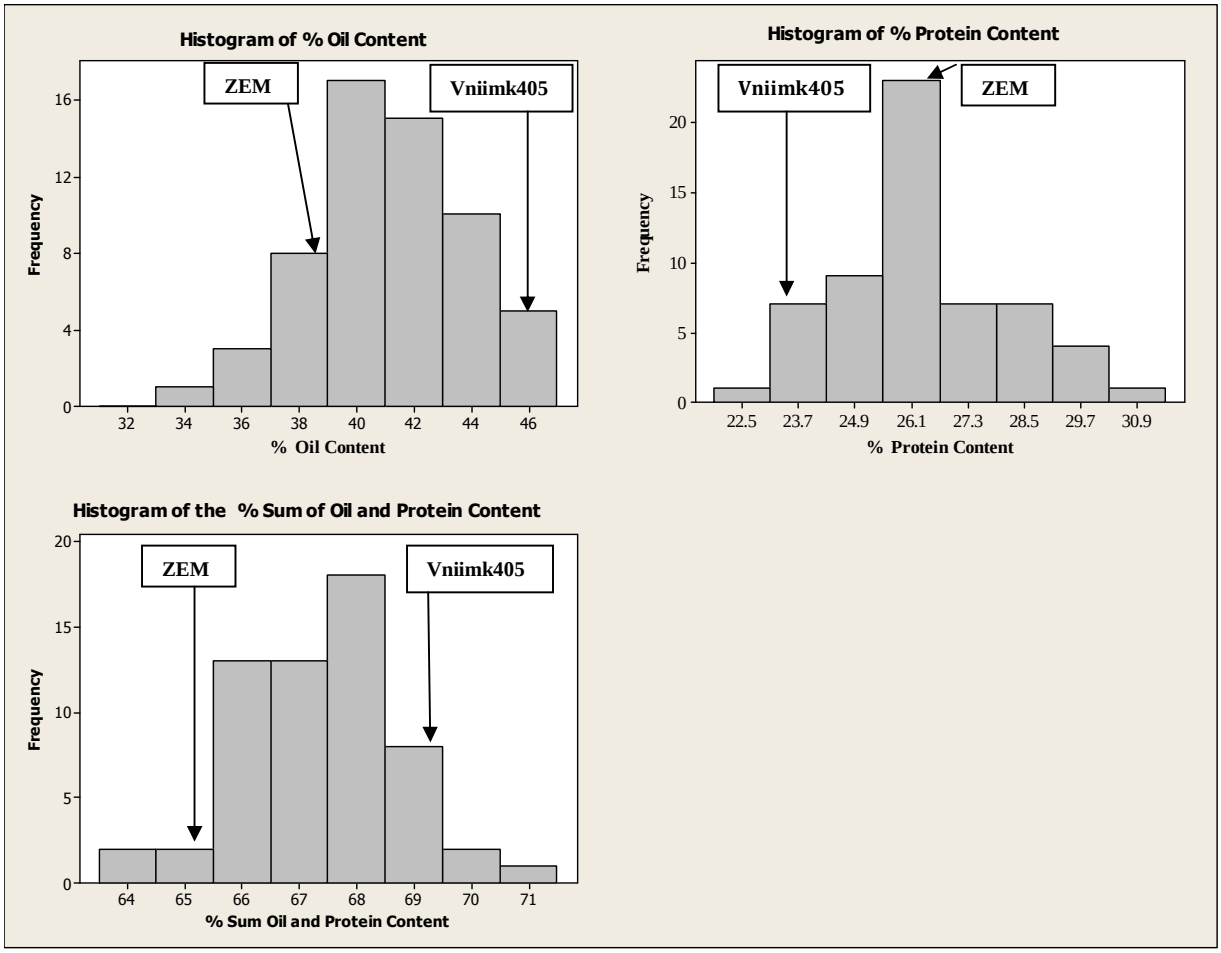


Figure 4.1: Distribution of the percentage of oil, protein and Sum of oil and protein content of the ZEM1xVniimk405 population.

4.4.2 Fatty acid variation

The fatty acid variation showed similar profiles but significant variation within the two DH populations. Three fatty acids (oleic, eicosenoic and erucic) had shown wide variation between lines in both populations. The observed erucic acid levels show variation within the population of a minimum of 0.17% to a maximum value of 31%. Similar levels of oleic acid were observed within the two populations which range from 16% to 49% in the ZEM1xVniimk351 population and 15% to 49% in the ZEM1xVniimk405 population. The data collected from the three nursery sites showed a distinct bimodal distribution in oleic, linoleic, eicosenoic, erucic and arachidic fatty acids. This distribution is observed in monogenic inheritance within the populations

following a Mendelian 1:1 segregation ratio observed with the inheritance of a single gene in a DH population.

Using a chi-squared goodness of fitness test the segregation within the populations are evaluated to determine if they fall within the expected 1:1 segregation that would be expected in a DH population. The erucic acid inheritance within both populations follows a distinct bimodal distribution. Chi square tests on the segregation of the erucic acid within both populations show significant evidence that the population is distributed by a 1:1 ratio. The low erucic acid was defined in both populations as having an erucic acid content >5% and high erucic acid was >5%.

A bimodal distribution was observed in oleic acid also indicating a monogenic inherence. The low oleic acid content was defined as being <25% oleic acid content for both populations and conversely the high oleic acid content was defined as having >25%. A chi squared analysis of the distribution of high and low oleic content in both populations supported the expected 1:1 segregation ratio in the DH populations.

Like the distribution of erucic and oleic fatty acid distribution, bimodal distribution patterns were observed in the Linoleic, eicosenoic and arachidic fatty acids content. Low level of linoleic fatty acid content was determined to be less than 30% of the overall oil content. Low levels of eicosenoic and nervonic acid content in both populations were determined to be <5% and <1% respectively. All were shown to follow the expected 1:1 segregation ratio as determined by chi square analysis. Both were determined to be segregating at a 1:1 Mendelian segregation ratio as expected in a DH population

The distribution of linolenic fatty acid within both the ZEM1 x Vniimk351 and ZEM1xVniimk405 populations follows a more normally distributed pattern than observed with

the other fatty acids observed in the population. The distribution of linolenic acid ranged between 7% and 14% of the overall oil content in the ZEM1xVniimk351 population and 7% and 16% of the oil contented in the ZEM1xVniimk405 population. The parental variation within the population for linoleic acid was 11%, 9% and 10% for ZEM1, Vniimk351 and Vniimk405 respectively.

Table 4.2: Statistical analysis of the percentage of oil content, protein content and sum of oil and protein for the three parental lines used in the development of both the ZEM1xVniimk351 and ZEM1xVniimk405 populations.

Variable	Line	Mean	SE Mean	StDev	Minimum	Maximum
C16:0	**Vniimk351**	2.7828	0.016	0.0531	2.7121	2.8781
	Vniimk405	2.2697	0.0228	0.0912	2.0968	2.401
	ZEM1	3.2935	0.0183	0.1185	3.0272	3.4973
C18:0	**Vniimk351**	1.1681	0.018	0.0596	1.1028	1.3088
	Vniimk405	0.9473	0.01	0.0401	0.8614	1.0103
	ZEM1	1.9734	0.024	0.1559	1.6016	2.2676
C18:1	**Vniimk351**	23.215	0.245	0.812	21.728	24.514
	Vniimk405	17.98	0.249	0.995	16.078	19.325
	ZEM1	43.148	0.436	2.825	36.489	47.349
C18:2	**Vniimk351**	23.727	0.159	0.529	22.787	24.353
	Vniimk405	20.923	0.326	1.305	18.606	22.73
	ZEM1	36.059	0.25	1.618	32.154	38.198
C18:3	**Vniimk351**	11.048	0.136	0.452	10.464	11.757
	Vniimk405	10.763	0.0817	0.327	10.004	11.163
	ZEM1	9.078	0.172	1.116	8.152	13.354
C20:0	**Vniimk351**	0.6956	0.0142	0.0472	0.6113	0.7575
	Vniimk405	0.64637	0.00812	0.03247	0.59384	0.70813
	ZEM1	0.55732	0.0062	0.0402	0.43633	0.66191
C20:1	**Vniimk351**	12.301	0.107	0.356	11.898	13.022
	Vniimk405	10.148	0.213	0.852	8.397	10.947
	ZEM1	2.47	0.148	0.959	1.33	4.422
C22:0	**Vniimk351**	0.3246	0.00842	0.02794	0.28214	0.36152
	Vniimk405	0.5217	0.0193	0.0771	0.4351	0.697
	ZEM1	0.25477	0.00475	0.0304	0.18588	0.32455
C22:1	**Vniimk351**	21.334	0.229	0.761	20.134	22.167
	Vniimk405	31.546	0.65	2.6	28.59	37.331
	ZEM1	2.239	0.277	1.751	0.334	7.098
C24:1	**Vniimk351**	1.3608	0.0131	0.0435	1.2982	1.4322
	Vniimk405	1.8051	0.0146	0.0584	1.7059	1.9092
	ZEM1	0.4849	0.0182	0.1139	0.3065	0.7255

Table 4.3: Statistical analysis of the saturated, monounsaturated and polyunsaturated fatty acids of the parental lines of the ZEM1xVniimk351 and ZEM1xVniimk405 populations.

Variable	Line	Mean	SE Mean	StDev	Minimum	Maximum
SFA	**Vniimk351**	4.9711	0.0373	0.1237	4.7606	5.2276
	Vniimk405	4.3851	0.0225	0.0901	4.2363	4.523
	ZEM1	6.073	0.0347	0.2246	5.6242	6.516
MUFA	**Vniimk351**	58.723	0.254	0.842	57.173	59.843
	Vniimk405	61.68	0.383	1.532	59.695	64.296
	ZEM1	44.912	0.635	4.114	38.649	53.124
PUFA	**Vniimk351**	34.775	0.168	0.557	33.677	35.534
	Vniimk405	31.686	0.4	1.6	28.609	33.781
	ZEM1	45.137	0.221	1.429	40.778	48.112

Table 4.4: Analysis of the monogenic inheritance patterns of the fatty acid distribution within the two *B. juncea* DH populations. Chi Square analysis of the 1:1 phenotypic distribution of oleic, linoleic, eicosenoic, erucic and arachidic fatty acids.

Population		**Oleic Acid**		**Linoleic Acid**		**Eicosenoic**		**Erucic**		**Arachidic**	
ZEM1xVniimk351		χ^2	P	χ^2	P	χ^2	P	χ^2	P	χ^2	P
	Winnipeg	1.40291	0.0939	1.41177	0.2348	1.40291	0.0939	1.6408	0.2002	1.24742	0.26404
	Carman	0.94231	0.1698	1.385	0.2393	0.94231	0.1698	1.8846	0.1698	0.0989	0.75315
	Chile	0.96078	0.1657	0.03922	0.8430	0.96078	0.1657	3.1765	0.0747	2.80583	0.09392
ZEM1xVniimk405											
	Winnipeg	1.47273	0.2249	2.2	0.1380	1.47273	0.2249	0.8909	0.3452	0.69231	0.40537
	Carman	1.47273	0.2249	2.2	0.1380	1.47273	0.2249	0.8909	0.3452	1.2549	0.26261
	Chile	2.2	0.1380	3.07273	0.0796	2.2	0.1380	0.667	0.4142	2.61818	0.10564

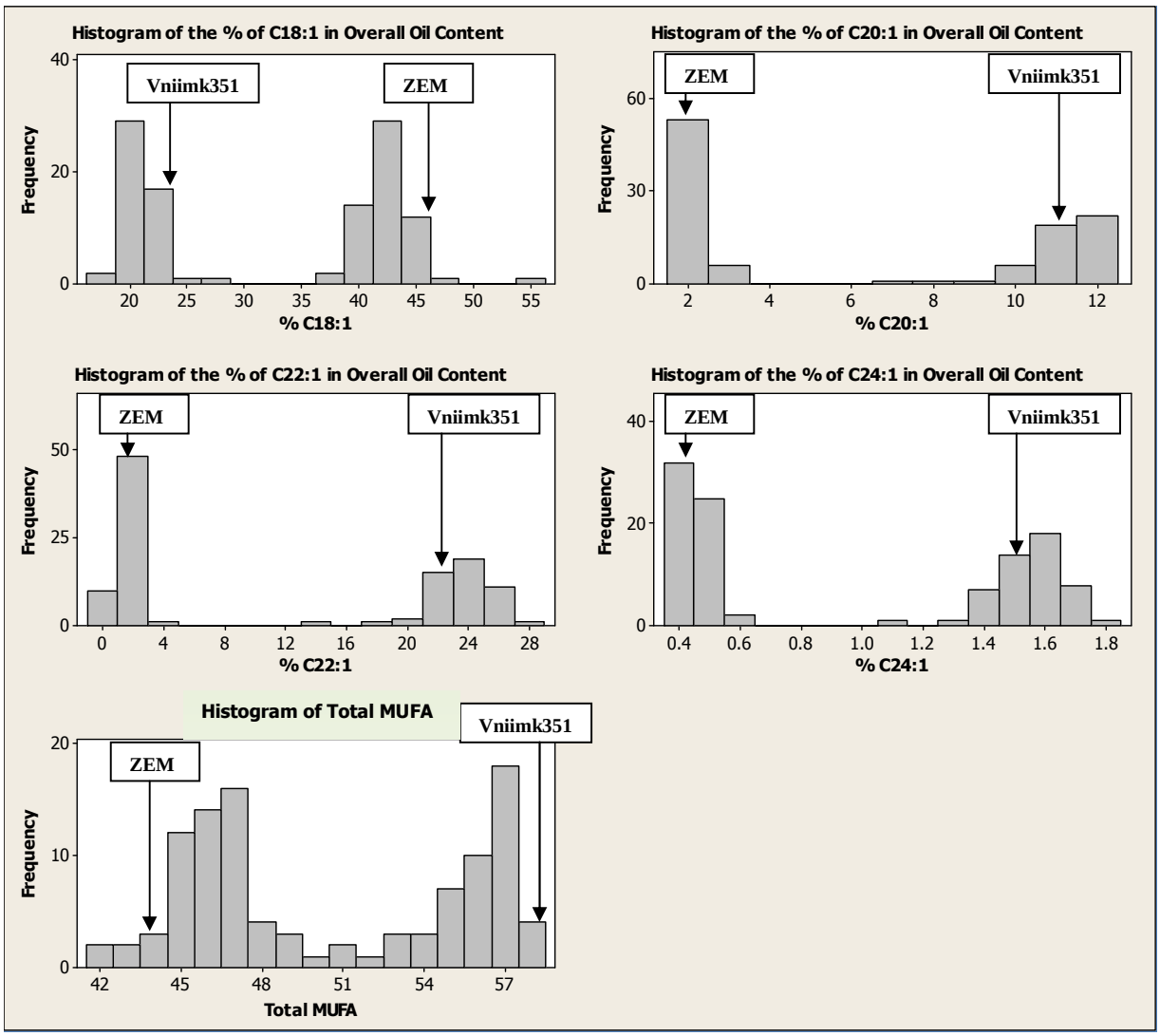


Figure 4.2 Distribution of the monounsaturated fatty acids in the ZEM1xVniimk351 population

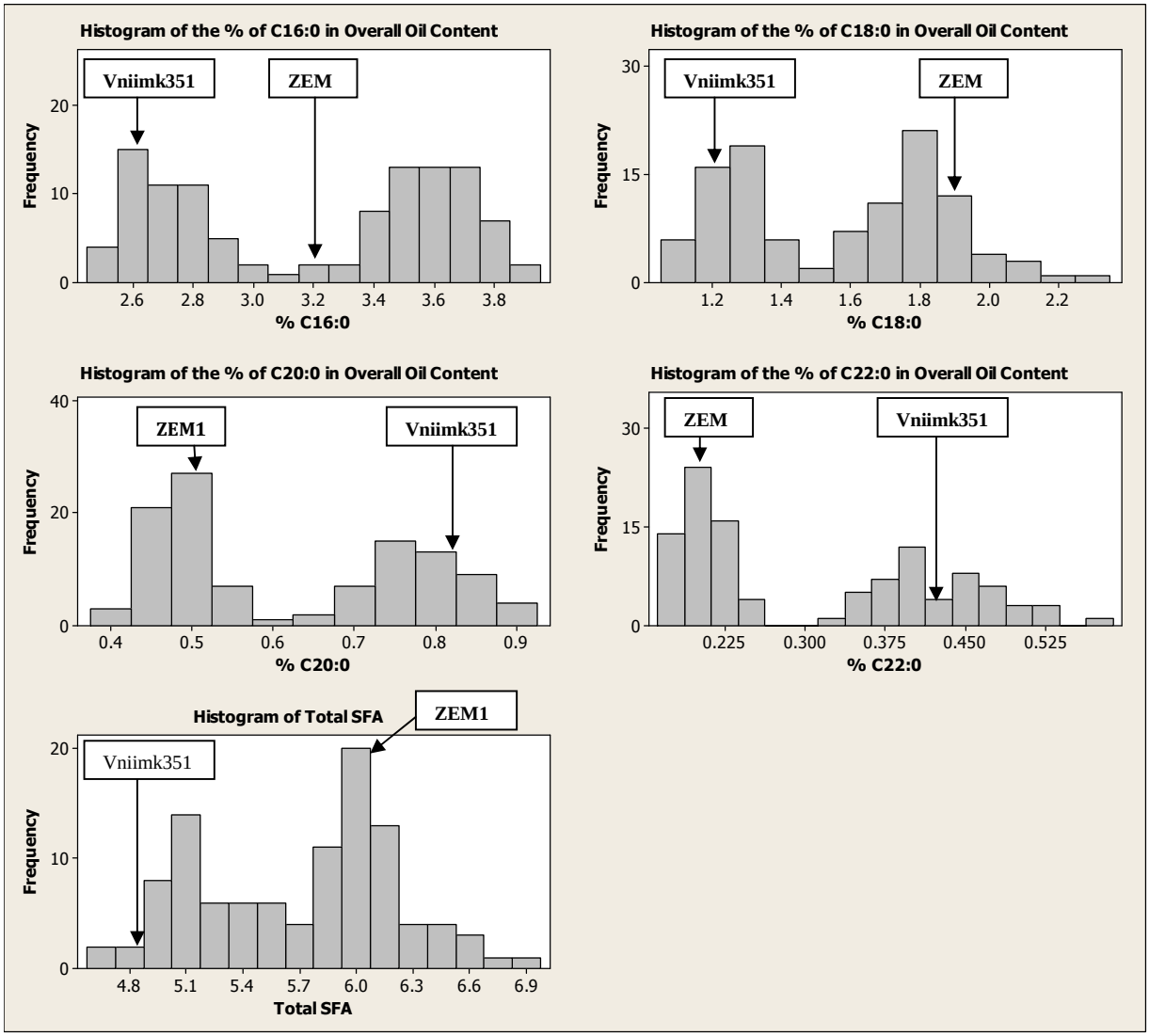


Figure 4.3: Distribution of saturated fatty acid in the ZEM1xVniimk351 population

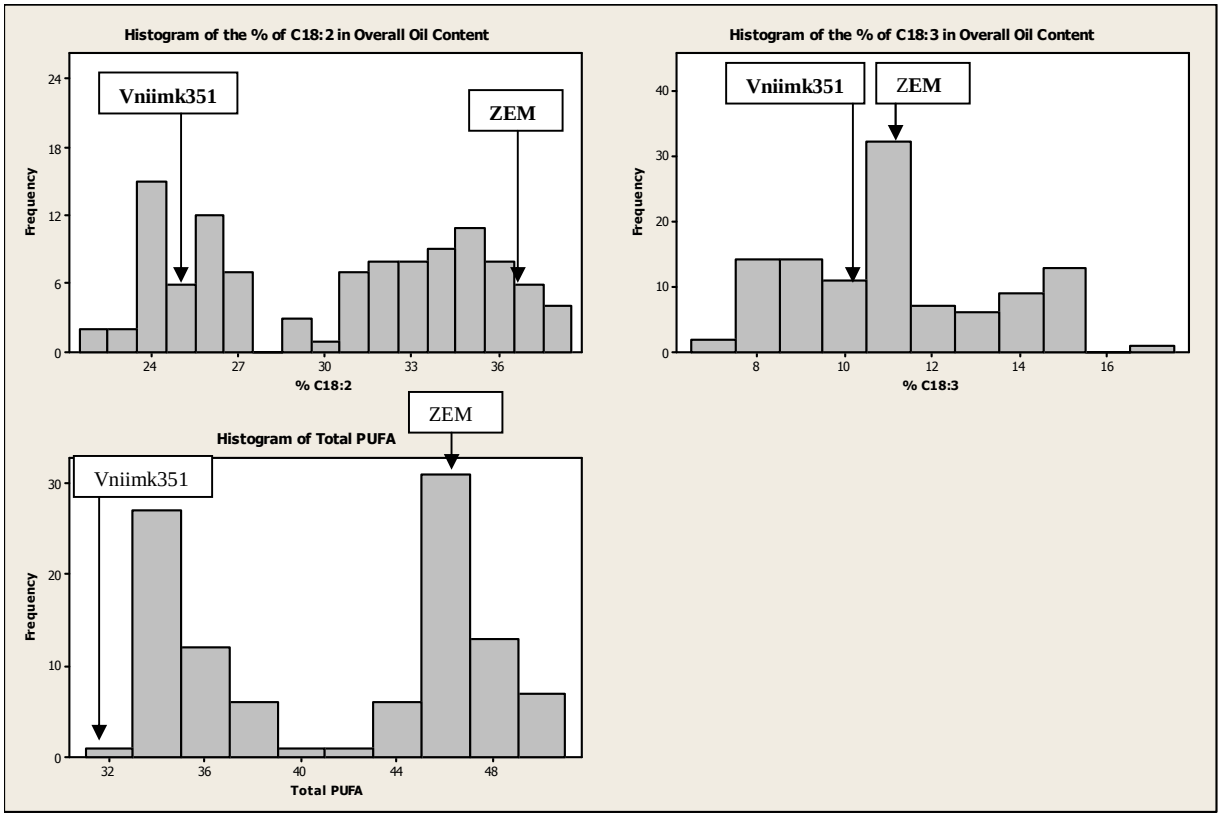


Figure 4.4: Distribution of Polyunsaturated fatty acid in the ZEM1xVniimk351 population

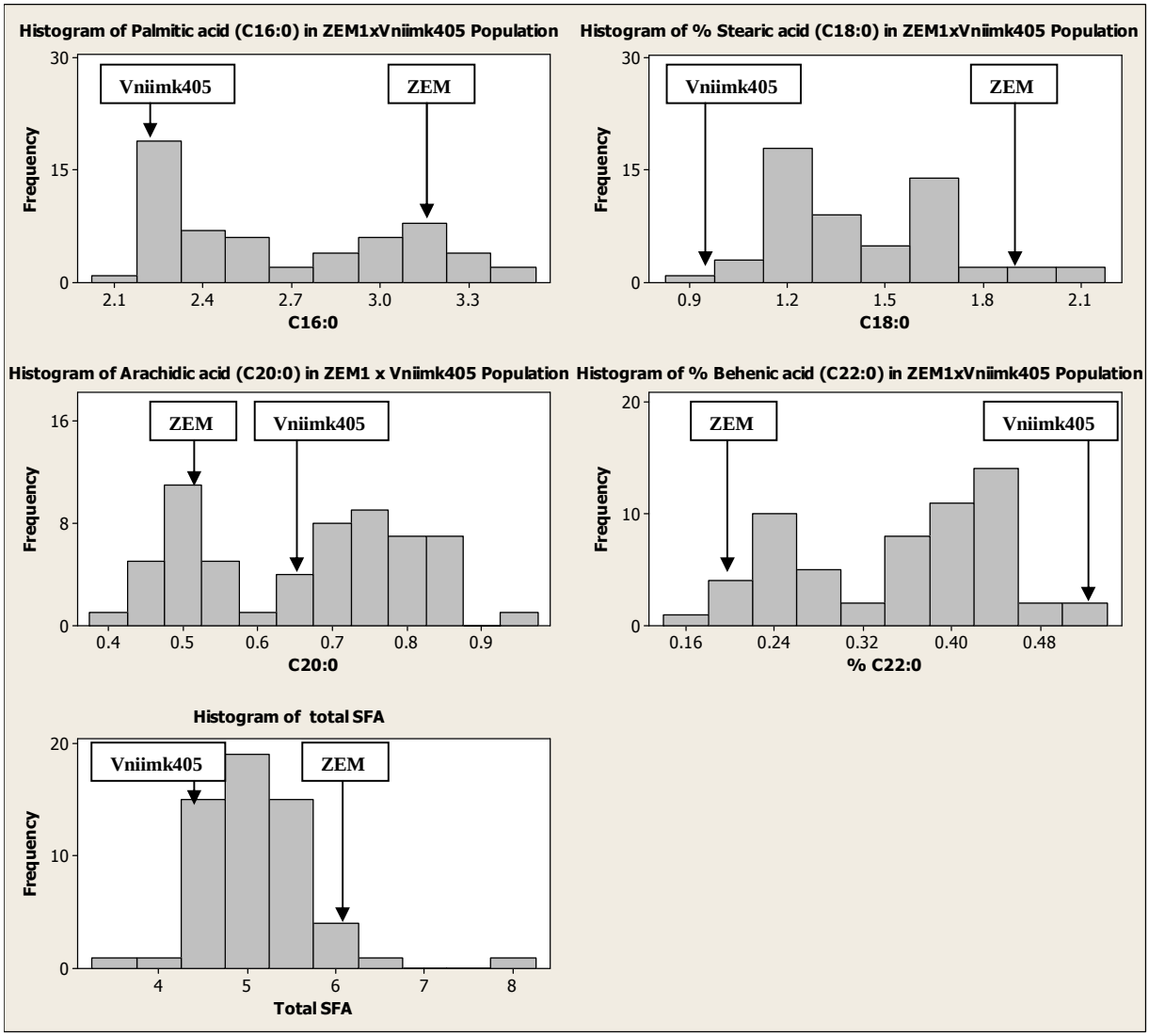


Figure 4.5: Histograms illustrating the distribution of saturated fatty acids within the ZEM1xVniimk405 population.

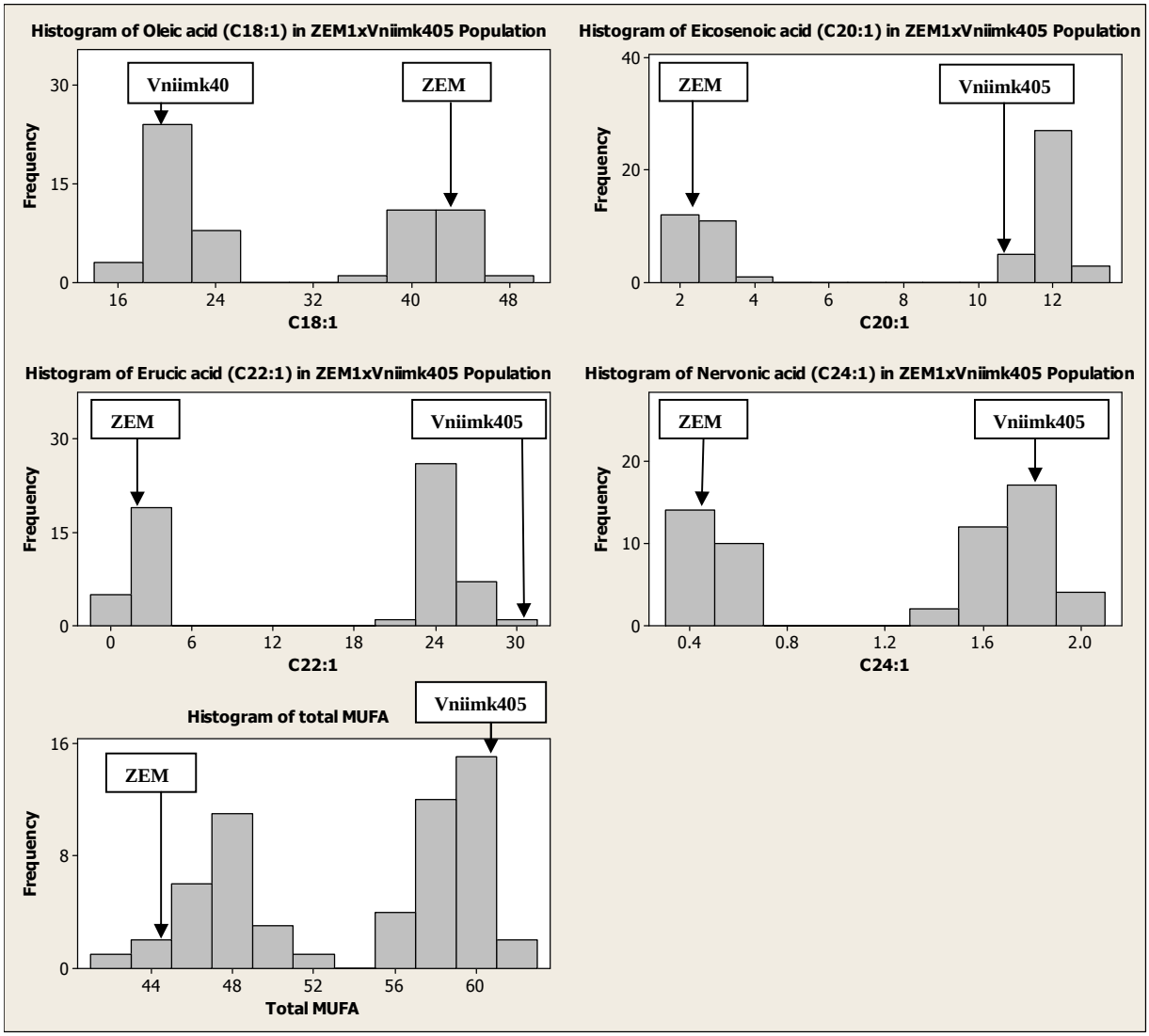


Figure 4.6: Histograms illustrating the distribution of the monounsaturated fatty acids within the ZEM1xVniimk405 population.

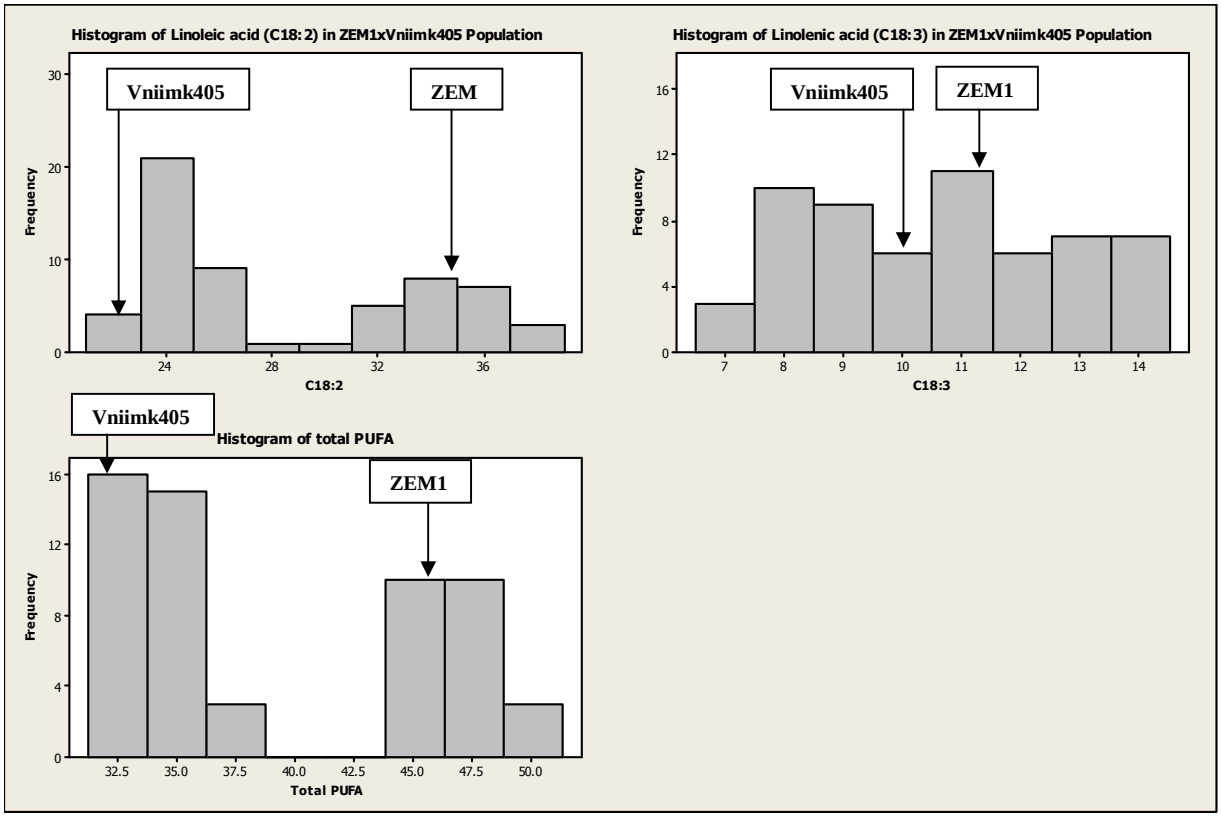


Figure 4.7: Histogram illustrating the distribution of polyunsaturated fatty acids of the ZEM1xVniimk405.

4.4.3 Correlation of traits

The analysis of the correlation between fatty acid composition and oil content showed a number of significant associations between the traits. Correlations for both populations were completed using the data collected from the three nursery based phenotypic evaluation trials and combined to determine the correlation between each of the traits collected. There was a significant negative correlation observed between oil and protein content in both populations. Overall oil content had a negative correlation with stearic, oleic, linoleic and linolenic oil content. This was contrasted with the high positive correlation between eicosenoic and erucic acid. A negative correlation between long chain fatty acids eicosenoic acid and erucic acid and shorter chain fatty

acids stearic, oleic, linolenic and linoleic fatty acids confirms that the accumulation of erucic acid and eicosenoic acid are derived from other fatty acids accumulated during biosynthesis.

Table 4.5: Pearson Correlation of the oil, protein and fatty acid content of the ZEM1xVniimk351 DH population

**** p value <0.001, * p value <0.05, >0.001**

	Protein	Oil	Sum	C16:0	C18:0	C18:1	C18:2	C18:3	C20:0	C20:1	C22:0	C22:1	C24:1	SFA	MUFA
Oil	-0.887**														
Sum	-0.404**	0.745**													
C16:0	0.261**	-0.449**	-0.573**												
C18:0	0.141**	-0.259**	-0.366**	0.839**											
C18:1	0.097*	-0.177**	-0.251**	0.842**	0.853**										
C18:2	0.294**	-0.401**	-0.428**	0.876**	0.81**	0.852**									
C18:3	0.109*	-0.149**	-0.154**	0.501**	0.395**	0.476**	0.216**								
C20:0	-0.089*	0.141**	0.182**	-0.776**	-0.674**	-0.913**	-0.794**	-0.519**							
C20:1	-0.207**	0.304**	0.353**	-0.909**	-0.863**	-0.974**	-0.899**	-0.528**	0.908**						
C22:0	0.079	-0.025	0.079	-0.736**	-0.722**	-0.89**	-0.754**	-0.484**	0.911**	0.854**					
C22:1	-0.163**	0.258**	0.321**	-0.898**	-0.878**	-0.984**	-0.897**	-0.527**	0.907**	0.986**	0.986**				
C24:1	-0.102*	0.219**	0.335**	-0.885**	-0.861**	-0.979**	-0.878**	-0.507**	0.9**	0.979**	0.979**	0.981**			
SFA	0.239**	-0.422**	-0.557**	0.956**	0.939**	0.814**	0.847**	0.436**	-0.659**	-0.869**	-0.788**	-0.868**	-0.852**		
MUFA	-0.302**	0.32**	0.234**	-0.589**	-0.482**	-0.519**	-0.568**	-0.441**	0.608**	0.581**	0.581**	0.57**	0.523**	-0.521**	
PUFA	0.289**	-0.394**	-0.418**	0.931**	0.834**	0.901**	0.92**	0.582**	-0.87**	-0.961**	-0.961**	-0.959**	-0.934**	0.881**	-0.65**

Table 4.6: Pearson Correlation table for oil, protein and fatty acid content of the ZEM1xVniimk351 DH population.

**** p value <0.001, * pvalue <0.05, >0.001**

	Protein	Oil	Sum	C16:0	C18:0	C18:1	C18:2	C18:3	C20:0	C20:1	C22:0	C22:1	C24:1	SFA	MUFA
Oil	-0.87														
Sum	-0.364**	0.775**													
C16:0	0.272**	-0.529**	-0.216**												
C18:0	0.015	-0.022	-0.001	0.134**											
C18:1	0.092	-0.263**	-0.125*	0.815**	0.147*										
C18:2	0.248**	-0.429**	-0.149*	0.868**	0.142*	0.894**									
C18:3	0.2**	-0.229**	-0.086	0.344**	-0.017	0.432**	0.292**								
C20:0	-0.065	0.154*	0.049	-0.639**	-0.081	-0.813**	-0.747**	-0.564**							
C20:1	-0.2**	0.372**	0.137*	-0.858**	-0.126*	-0.972**	-0.937**	-0.499**	0.832**						
C22:0	-0.039	0.154*	0.064	-0.679**	-0.105	-0.824**	-0.738**	-0.444**	0.859**	0.8**					
C22:1	-0.165**	0.342**	0.146**	-0.849**	-0.14*	-0.982**	-0.933**	-0.505**	0.825**	0.985**	0.819**				
C24:1	-0.102	0.303**	0.168*	-0.841**	-0.142*	-0.971**	-0.905**	-0.45**	0.777**	0.959**	0.808**	0.976**			
SFA	0.073	-0.148*	-0.061	0.344**	0.942**	0.263**	0.287**	-0.009	-0.11*	-0.257**	-0.142*	-0.261**	-0.275**		
MUFA	-0.301**	0.475**	0.166*	-0.852**	-0.107	-0.872**	-0.935**	-0.586**	0.788**	0.945**	0.736**	0.945**	0.903**	-0.239**	
PUFA	0.279**	-0.441**	-0.155*	0.848**	0.112*	0.902**	0.941**	0.598**	-0.826**	-0.962**	-0.775**	-0.96**	-0.918**	0.238**	-0.99**

4.4.5 QTL analysis

Using the data collected from field trials at all three locations a QTL analysis was completed on the two DH populations. The analysis was completed independently using data from each location for the oil, protein and fatty acid traits. This was due to significant location effects for the traits under study showing quantitative phenotypic distribution.

4.4.6 ZEM1xVniimk351 Population QTL Analysis

Analysis of data collected from the trial in Carman identified a total of 10 QTLs for the ZEM1xVniimk351 population for oil, protein and oleic, linoleic, linolenic and arachidic fatty acid content. As seen in Table 4.7, three QTL for protein were identified from the data collected in Carman found on LG 1, 7 and 12 which accounts for 20% of the variation in the population. The QTL region on LG 7 was associated with C16:0 and C18:2 levels in the populations. An additional QTL for C18:2 levels were identified on LG6. Analysis of the sum of seed oil and protein content identified 2 QTL on LG 4 and LG 11 with an R^2 value of 16%. QTL for C18:3 and C20:0 were identified on LG 10 and LG12 respectively.

Table 4.7: QTL analysis for ZEM1xVniimk351 using data collected from Carman nursery trial. QTL region identified by flanking markers and r^2 value noted for overall QTL effect.

Flanking Markers	LG	Protein	Sum	C16:0	C18:2	C18:3	C20:0
SA7PM32-167-NA14C12A	1	0.13[a]					
BG23BG45-242-FC1FE8-334	4		0.09[a]				
Na12D08-SR094a	6				0.11[a]		
BG23PM29-99-NA10A11	7	0.03[a]		0.12[a]	0.13[a]		
BRMS40-Na10G10A	10						0 .08[a]
SR027-CB10524C	11		0.07[a]				
CB10288-CB10524a	12	0.04[a]					
MD053-CB10288	12					0 .09[a]	

[a] R^2 value explaining the phenotypic variation explained.

QTL analysis for ZEM1xVniimk351 for the data collected from the trial conducted in Chile identified 14 QTL overall. In Table 4.8 the QTL analysis of the protein content identified 5 QTL on 4 linkage groups (LG 3, LG 7, LG 11 and LG 12). One of the two QTL on LG 3 correlated with a QTL identified for oil content. An additional QTL for oil was identified on LG 4 with both QTL accounting for 18% of the variation observed in the population. A single QTL for C16:0 FA content was identified on LG 11 with an R^2 value of 0.11 accounting for 11% of the variation. 3 QTL for C18:3 were identified on LG 4, 7 and 11 which accounted for 47% of the variation within the population. Finally 1 QTL for C20:0 and two QTL C22:0 was identified on LG 1 and LG 7.

Table 4.8: QTL analysis for ZEM1xVniimk351 using data collected from Chile nursery trial. QTL region identified by flanking markers and r^2 value noted for overall QTL effect.

Flanking markers	LG	Protein	Oil	C16:0	C18:3	C20:0	C22:0
SA7PM32-167-NA14C12A	1						0.14[a]
FADPM32-176-PM88ODD13-385	1					0.15[a]	
Na12H04A-BG23PM32-213	3	0.09[a]	0.04[a]				
Na10B07-CB10403a	3	0.11[a]					
BG23BG45-226-FADPM32-60	4				0.10[a]		
BG23BG45-242-FC1FE8-334	4		0.14[a]				
Na10A08A-BG23PM29	7						0.12[a]
BG23PM29-99-NA10A11	7	0 .11[a]			0.22[a]		
CB10524C-PM88FE8-141	11			0.11[a]			
CB10099A-ODD3PM33-175	11	0.17[a]					
PM88FE8-141-Na10H02	11				0.15[a]		
PM88PM45-364-CB10524A	12	0.07[a]					

[a] R^2 value explaining the phenotypic variation explained.

Table 4.9 shows the results of the QTL analysis of the ZEM1xVniimk351 population identified a total of 10 QTL for the data collected in Winnipeg at the Point research station. 3 QTL were identified for protein on LG 1, 3 and 10 which accounts for 24% of the variation observed in the trial. 2 QTL for oil content were identified on LG 1 and LG 4. The QTL on LG 1 correlated with one of 2 QTL for C20:0 with the second QTL being identified on LG 13. The QTL on LG 13 correlated with one of the 2 QTL for C18:3 with the other being located on LG 12. The final QTL identified was found on LG7 for C18:2.

Table 4.9: QTL analysis for ZEM1xVniimk351 using data collected from Winnipeg nursery trial. QTL region identified by flanking markers and r^2 value noted for overall QTL effect.

Flanking Markers	LG	Protein	Oil	C18:2	C18:3	C20:0
SR120-BG23PM32-102	1	0.10[a]				
SR044-CB10277	1		0.08[a]			0.15[a]
Na10B07-CB10403a	3	0.06[a]				
BG23BG45-242-FC1FE8-334	4		0.09[a]			
CB10092B-NA10A11	7			0.19[a]		
BRAS029-NA10G10a	10	0.08[a]				
PM88PM45-364-CB10524A	12				0.13[a]	
PROPM113-182-Na10B8	13				0.16[a]	0.18[a]

[a] R^2 value explaining the phenotypic variation explained.

A comparison of QTLs between the experiment locations identifies some common QTLs between locations. Similar QTL for protein content and oil content were identified on LG 3 and LG 4 respectively in the QTL analysis completed in Chile and Winnipeg. The QTL identified at both locations was correlated with a QTL for the sum of oil and protein content in the analysis completed using the data obtained from the Carman location. QTLs for protein were identified on LG 7 in the QTL analysis completed on the data collected from Chile and Carman locations. This region also contains a QTL for linoleic acid content identified using the data collected from the Point research farm in Winnipeg. Similar QTL identified for C22:0 and protein were identified on LG 1 from the analysis completed for Chile and Carman respectively. Similar correlations between protein and C18:3 were identified on LG 12 between the data collected in Chile and Winnipeg respectively. Additional similarities were identified on LG 7 with QTL identified for C18:2 and C22:0 between Carman and Winnipeg locations.

4.4.7 ZEM1xVniimk405 population QTL analysis

Analysis of the ZEM1xVniimk405 population identified a total of 13 QTL as shown in Table 4.10. One QTL for protein was identified on LG 1 which accounts for 13% of the variation. One QTL for C16:0 levels where identified on LG 12 which accounted for 20% of the variation in the population. This QTL region was shared with C18:0 and C18:1 levels which accounted for 28% and 24% of the variation respectively. An additional QTL for C18:0 was identified on LG3 and LG 1 respectively. 2 QTLs for C18:2 were identified on LG 4 and LG 11 which accounts for 24% of the variation observed. For C18:3 QTLs for on LG 1 and LG 3 which accounted for 28% of the variation observed in the population. Two additional QTLs were identified on LG 3 and LG 9 for C22:0 which accounted for 18% of the variation. Finally 1 QTL for C22:0 was identified on LG 1 sharing the region where the QTL for protein was identified.

Table 4.10: QTL analysis for ZEM1xVniimk405 using data collected from Carman nursery trial. QTL region identified by flanking markers and r2 value noted for overall QTL effect.

Carman	LG	Protein	C16:0	C18:0	C18:2	C18:3	C20:0	C22:0
Na10C06B-- BG23PM29-457	1							0.21[a]
CB10373bA-- BG23PM29-457	1	0.13[a]						
Na10C06A-- CB10504	1					0.16[a]		
EM1BG45-56-- CB578	3					0.12[a]		
SR072-- ODD3PM33-177	3			0.12[a]				
BG23BG45-93-- MD053	4				0.18[a]			
Ra2G08-- SR087	4						0.10[a]	
SR071-- CB10534	9						0.08[a]	
SA7PM32-56-- BRAS158	11				0.06[a]			
Ni3G05-- BG23PM32-275	12		0.27[a]	0.28[a]				

[a] R^2 value explaining the phenotypic variation explained.

In Table 4.11 the QTL analysis of the ZEM1xVniimk405 population for the data collected for the trials conducted at the University of Manitoba research farm in Winnipeg identified shows total of 9 QTL. One QTL was identified for oil and protein content on LG 1 and LG 3 respectively. One QTL for C16:0 was identified on LG1 which accounted for 26% of the overall variation observed. This QTL region was correlated with a QTL identified for C18:2. An additional QTL on LG 4 was identified for C18:2 with the two QTL accounting for 45% of the variation observed. One QTL for the sum of oil and protein content was observed on LG 1 which correlated with a QTL for C18:3. An additional QTL for C18:3 were observed on LG 3 which correlated with the QTL identified for protein.

Table 4.11: QTL analysis for ZEM1xVniimk405 using data collected from Winnipeg nursery trial. QTL region identified by flanking markers and r2 value noted for overall QTL effect.

Point	LG	Protein	Oil	Sum	C16:0	C18:2	C18:3
SR068-- CB10501B	1				0.26[a]	0.21[a]	
CB10373bA-- BG23PM29-457	1		0.14[a]				
BG23PM29-457-- CB10504	1			0.06[a]			0.23[a]
EM1ODD13-303-- NA12H09	3	0.10[a]					0.13[a]
Ni4D12-- CB10234b	4						
Ra2G08-- SR087	4					0.24[a]	
CB10079-- BRAS029	7						

[a] R^2 value explaining the phenotypic variation explained.

For the ZEM1xVniimk405 population a total of 8 QTL were identified from the trial conducted in Chile as shown in Table 4.12. One QTL for protein content was identified on LG 1 which accounted for 26% of the variation observed. 2 QTL for oil content were identified on LG 1 and LG 4 which covers 29% of the overall variation. The QTL on LG 1 correlated with one of the 2

QTL identified for C16:0 with the other QTL identified on LG 9. One QTL for sum of oil and protein was observed on LG6. Finally one QTL for C18:3 were identified on LG2 which accounts for 17% of the variation observed.

Table 4.12: QTL analysis for ZEM1xVniimk405 using data collected from Chile nursery trial. QTL region identified by flanking markers and r2 value noted for overall QTL effect.

Chile	LG	Protein	Oil	Sum	C16:0	C18:3
MR013-- PM88ODD13-271	1	0.26[a]				
ODD3PM33-45-- EM1PM113-356	1		0.18[a]		0 .08[a]	
FADPM32-296-- CB10092B	2					0.17[a]
MD053-- Ra2G08	4					
MD053-- SR087	4		0.11[a]			
SR021-- NA12D03	6			0.19[a]		
SR071-- CB10534	9				0.05[a]	

[a] R^2 value explaining the phenotypic variation explained.

A comparison of the QTL analysis completed on the data collected at the three experiment locations identified similar regions where QTL were identified on LG 1, LG 4 and LG 9 for the ZEM1xVniimk405 population. QTL were identified on LG 4 in both the Winnipeg and Carman location data for the C20:0 FA content. This region also contained a QTL identified for oil at the Chile location. Similar regions were identified on LG1 between these two sites linked to protein and C22:0 in Carman and oil in Winnipeg. Additional similar regions were identified between QTL for sum of oil at the Winnipeg location and C18:3 at the Carman location. A QTL located on LG 9 in the Winnipeg and Chile analysis identified QTL for C20:0 and C16:0 respectively.

4.5 Discussion

Upon initial inspection of the phenotypic data collected of the population the most significant variations were observed in the fatty acid profiles in each of the populations. Evaluation of the parental lines of each population showed a significant variation of the fatty acid profiles between each of the parental lines. The ZEM1 parental lines were identified as a zero erucic acid line maintaining erucic acid levels less than 2% of the overall fatty acid profile (Kirk and Oram, 1981). The two Russian mustard lines Vniimk351 and Vniimk405 lines used to develop the separate DH populations observed mean erucic acid levels of 21.7% and 30.4% of the overall fatty acid profile respectively. Although there were significant variations observed in the fatty acids between the parental lines in both populations, it appeared that the variation of the erucic acid content provided the most significant variation in fatty acid content. This was observed in the variation between the erucic acid and oleic acid, which is the precursor in erucic acid synthesis.

Previous studies completed on erucic acid inheritance within *Brassica* species has identified two alleles acting additively resulting in low, moderate and high erucic acid levels (Anand and Downey, 1981; Chen and Heneen, 1989; Chen et al., 1988; Dorrell and Downey, 1964; Harvey and Downey, 1964; Kondra and Stephansson, 1965). The inheritance of erucic acid levels has shown that the zero erucic acid phenotype is due to homozygous recessive alleles of both genes in *Brassica juncea* (Bhat et al., 2002; Kirk and Oram, 1981; Kirk and Hurlstone, 1983; Labana et al., 1987). The two *FAE* genes that have been determined to result in the variation of erucic acid within *B. juncea* have been defined as *FAE1.1* and *FAE1.2* (Gupta et al., 2004; Xu et al., 2010). These genes produce a fatty acid elongase enzyme that has been shown to result in the elongation of oleic acid to produce eicosenoic (C20:1) and erucic acid (C22:1) in *Brassica juncea* (Gupta et

al., 2004; Kanrar et al., 2005). The distribution of erucic acid phenotypes in both populations indicate a single dominate alleles segregating within each of the populations resulting in the expression of moderate and low erucic acid at a 1:1 segregation ratio. This was also observed in studies of erucic acid inheritance in *B. juncea* where a single dominant gene was found responsible for moderate erucic acid content within populations segregating for erucic acid (Bhat et al., 2002; Kirk and Hurlstone, 1983).

This single gene inheritance was also observed in the expression of eicosenoic acid levels in the populations. Moderate erucic acid levels correlated with higher concentrations of eicosenoic acid which is the precursor to erucic acid levels. Conversely low erucic acid levels results in a decrease in the concentration of eicosenoic acid levels which was observed in both populations. This positive correlation supports early findings that showed similar response (Cheung et al., 1998; Lionneton et al., 2002; Mahmood et al., 2006b). The segregation within the eicosenoic acid also follows a 1:1 segregation ratio indicating the effects of a single gene. This suggests that the *FAE1* gene in these populations was responsible for the elongation of oleic acid to eicosenoic then eicosenoic to erucic acid. This is supported by the negative correlation observed in the oleic acid with both erucic and eicosenoic FA levels observed in both populations as it is the substrate for the synthesis of both fatty acids.

Variations in both populations showed a significant negative correlation between oleic and erucic acid levels. The negative correlation has been observed in a number of *Brassica* species looking at the association between these fatty acids (Alemayehu and Becker, 2001; Bhat et al., 2002; Burns et al., 2003; Ecke et al., 1995; Zhao et al., 2008). The analysis of the distribution of the oleic acid in both populations showed that there was a 1:1 segregation ratio which correlates to the segregation of erucic acid and eicosenoic acid content.

The distribution of the fatty acid in the two DH populations shows that the linoleic fatty acids follow a similar monogenic inheritance patterns as observed in the erucic acid segregation whereas the linolenic acid followed a more normal distribution. Studies in *Brassica* have shown that the *FAD2* and *FAD3* genes are involved in the desaturation of oleic acid with respect to these two fatty acids (Auld et al., 1992; Barret et al., 1999; Hu et al., 2006; Jourdren et al., 1996a; Okuley et al., 1994). The FAD2 gene has been associated with the desaturation of oleic acid to product linoleic acid (Okuley et al., 1994; Peng et al., 2010; Sivaraman et al., 2004). Similarly the *FAD3* gene has been linked to the further desaturation of linoleic acid to produce linolenic acid (Barret et al., 1999; Jourdren et al., 1996a); an undesirable component that leads to reduced shelf life with the extracted oil (Scarth and McVetty, 1999; Uppstrom, 1995). The two genes required to produce linolenic acid result in the more normal distribution of the oil profile due to the interaction between the two genes required for the synthesis of linolenic acid.

QTL analysis for the ZEM1xVniimk351 population identified 1 QTL for linoleic acid on LG 7 through data analyzed from both the Carman and Winnipeg. A single QTL on LG 7 for Linolenic acid was identified from the data collected from Chile. These regions could be associated to the *FAD2* or *FAD3* which have been to increase both fatty acids. An analysis of the data collected for linoleic acid levels in both populations follows a monogenic inheritance pattern indicating that a single gene accounts for the variation which falls within the range of the parental lines. The data shows that the population is also segregating at a 1:1 ration supporting the single *FAD2* gene responsible for the variation of linoleic acid in the population. For linolenic acid, 2 QTL were detected for both populations at all locations except for Carman for ZEM1xVniimk351 population and Chile for ZEM1xVniimk405 population. These two QTL

may be linked to the *FAD2* and *FAD3* genes that produce the fatty acid desaturase enzymes that result in the synthesis of linolenic acid.

Though observed in small quantities, the inheritance between the overall saturated fatty acids (SFA) in Vniimk405 population follows a more normal distribution than observed in the ZEM1xVniimk351 population which reflects the multiple genes involved in SFA biosynthesis (Baud et al., 2008; Ohlrogge and Jaworski, 1997). The distribution of the palmitic (C16:0) and stearic acid (C18:0) content in the ZEM1xVniimk405 population showed no distinct distribution pattern however there appears be a monogenic pattern emerging from the C20:0 and C22:0 fatty acids. This difference could be due to the small population size in the ZEM1xVniimk405 population as compared to the ZEM1xVniimk351 population which has almost twice as many individuals in the population. A distinct bimodal distribution was observed in the SFA in the ZEM1xVniimk351 population. QTLs for SFA were identified in both populations however little correlation between the Chile, Winnipeg and Carman field nursery sites was observed.

Very little research has been conducted on SFA in *B. juncea* as it is not a significant component to the fatty acid profile only making up 5% to 6% of the overall fatty acids. However there have been studies identifying plant stearic acid as a healthy saturated fat showing little to no increase in plasma cholesterol levels (Grundy, 1994; Yu et al., 1995). There has been development of high stearic acid lines in other oilseed crops (Graef et al.; Osorio et al.; Pérez-Vich et al., 2006) however there has been little success in developing conventional *Brassica* high stearic lines and only transgenic approaches had significant effects on the fatty acid content (Hawkins and Kridl, 1998; Karim Zarhloul et al., 2006; Knutzon et al., 1992). There is room for research in this area and may offer opportunities for future research on the area of fatty acid manipulation in the *Brassica* species. QTL analysis completed on the two populations identified some correlations

between genes associated with fatty acid synthesis. A single QTL for palmitic acid (C16:0) was identified at each location with the exception of ZEM1xVniimk405 population data from Chile. This single QTL suggests a small number of genes responsible for the variation within the populations. As palmitic acid is the precursor to stearic acid which is elongated by *KASII* gene within the cytoplasm, this QTL may be associated with this gene (Ohlrogge and Browse, 1995; Pidkowich et al., 2007). A single QTL on LG7 was identified in the ZEM1xVniimk351 population was observed at the Carman and Winnipeg for the locations which may be correlated with the *FAD2* gene which is responsible for the desaturation of oleic acid to produce linoleic acid (Okuley et al., 1994).

The populations used in this study were characterized for variation of oil and protein levels in the seed in addition to the fatty acids profile of each of the populations. Analysis of other *Brassica napus* populations segregating for erucic acid have shown a strong correlation between fatty acid profile and oil and protein content (Burns et al., 2003; Ecke et al., 1995; Qiu et al., 2006) and similar results have been shown in *Brassica juncea* (Cheung et al., 1998; Mahmood et al., 2006b; Ramchiary et al., 2007b). Strong correlations ware observed between the monounsaturated and polyunsaturated fatty acid and the oil and protein content within both populations. This correlation between fatty acid and oil content observed within the two populations could explain much of the variation of the oil content based on the observation of previous studies.

Previous studies of oil content in *B. juncea* have identified 2 to 7 QTL within the populations studied (Lionneton et al., 2002; Mahmood et al., 2006b; Ramchiary et al., 2007b; Sharma et al., 1999). The QTL analysis completed on the two DH populations in this study identified a small number of QTL for oil content. The data collected from the Winnipeg and Chile locations for the ZEM1xVniimk351 population identified 2 QTL for oil content. For the ZEM1xVniimk405

population, one QTL was identified for oil content from the data collected in Winnipeg and 2 QTL identified from the data collected in Chile. The high correlation of fatty acid profile and oil content identified from previous studies in *B. napus* (Ecke et al., 1995), *B. rapa* (Tanhuanpaa et al., 1996) and *B. juncea* ((Cheung et al., 1998; Lionneton et al., 2002; Ramchiary et al., 2007b; Sharma et al., 2002b) suggests that these QTL may be linked to regions that control fatty acid content. Similar observations were observed in the QTL analysis completed by Lionneton which identified only 2 QTL for oil in these populations. The DH population in Lionneton et al. (2002) study was developed from a cross between a moderate erucic acid and high erucic acid *B. juncea* lines which showed a monogenic segregation similar to those observed in the two DH populations in this study (Lionneton et al., 2002). The other studies that identified a higher number of QTL for oil content have used populations that are segregating for high erucic acid and low erucic acid (Mahmood et al., 2006b; Ramchiary et al., 2007b). This added variation and the correlation with erucic acid in these populations could result in the identification of additional oil QTL.

The analysis of the protein content data collected identified QTL for each of the locations it was tested in. The analysis of the ZEM1xVniimk351 population identified more QTL than the ZEM1xVniimk405 population with 3 QTL identified in the Carman and Winnipeg locations and 5 QTL from the data collected in Chile. A number of these QTL were related to QTL identified for fatty acids. Additionally there is a significant negative correlation between protein and fatty acid content which would suggest that these QTL could be related to genes linked to fatty acid synthesis. This negative correlation was observed in a number of studies analyzing oil and protein content in *B. juncea* (Cheung et al., 1998; Mahmood et al., 2006b; Ramchiary et al., 2007b). In addition there is a significant environmental effect to the expression of oil content

observed in the *Brassica sp.* which makes it difficult to identify the regions which account for the accumulation of oil content (Chauhan et al., 2010; Chauhan et al., 2011; Gunasekera et al., 2006; Pritchard et al., 2000; Si et al., 2003; Zhao et al., 2005). Of the QTL identified in other studies of *B. juncea* their overall QTL affect observed was less than 12.5% which makes it difficult to identify minor effects with the significant effects of environmental factors.

The molecular maps used in this study covered only 13 of the 18 linkage groups identified in *B. juncea*. This limited coverage of these maps can limit the detection of significant QTL that are found on other parts of the genome (Jansen, 1993; Zeng, 1994). The maps developed cover a distance of 791 cM in the ZEM1xVniimk351 population and 624 in the ZEM1xVniimk405 population. Current published maps covering the 18 linkage groups of the *B. juncea* genome have map distance ranging from 1564 cM to 1922 cM (Lionneton et al., 2002; Mahmood et al., 2003; Panjabi et al., 2008a; Pradhan et al., 2003; Ramchiary et al., 2007b). With these maps covering approximately 50% of the distance of other publish maps the ability to identify all QTL for these traits was limited.

4.6 Conclusion

From this study we were able to characterize the inheritance of oil, protein and fatty acids within two *B. juncea* DH populations using linkage maps and QTL analysis. The QTL analysis for seed oil content conducted on the two DH populations yielded little correlation between QTL identified between locations and populations. This supports the significant effects of environment factors on the expression of seed oil. A similar effect was observed with seed protein content in both populations as there was no consistency in the region of which the QTL were identified between locations and populations. Further analysis of the two populations identified multiple QTL for oleic, linoleic and linolenic fatty acid in both populations however, little correlation was observed between the locations or between the two populations. For both DH populations under study, pattern of monogenic inheritance in the erucic acid levels was evident. The segregation of erucic acid content within both of these populations lends further support to the inheritance of *FAE1.1* and *FAE1.2* genes that have been identified as being a key factor in the accumulation of erucic acid. These genes have been identified as showing the additive epistatic inheritance patterns which accounts for the variation in erucic acid as identified first in *B. napus* and later in *B. juncea*. The segregation of a single homozygous dominant gene inherited from the Vniimk (moderate erucic acid) parental lines and a homozygous recessive gene inherited from the ZEM1 (zero erucic acid) parent corresponds to the variation observed in both DH populations which correlates with previous studies. In addition, significant correlations observed between fatty acids and the accumulation of oil and protein was identified which suggests that erucic acid is responsible for the variation observed in seed oil and protein content within the two DH *B. juncea* populations. Further investigation to isolate and identify the *FAE* gene involved in the accumulation of erucic acid can be used to develop gene specific markers to

be incorporated into the linkage map developed for these DH populations. This will help identify seed oil content QTL which are not linked to the *FAE* gene which can be used for the specific improvement of seed oil content within *B. juncea.*

5.0 General Discussion

The increase of oilseed production in Canada requires additional options for producers to increase overall production and the improvement of oil content within *Brassica* species. *B. juncea* offers an opportunity to improve production using both strategies. *Brassica juncea* production in western Canada is primarily dedicated to condiment mustard and has been shown to grow successfully in the drier regions of Saskatchwan and Alberta where Canola production is limited (Skrypetz, 2003). The introduction of oilseed *B. juncea* lines these regions provide opportunities to increase *Brassica* oilseed production in Canada to help meet the growing demand.

Oil accumulation in *Brassica* has shown to be a complex inherited traits which involve multiple pathways (Baud et al., 2008). Studies to identify specific genes that can increase the accumulation of fatty acids within the seed have identified a number of transcription factors that are thought to control fatty acid accumulation (Gutierrez et al., 2007; Mendoza et al., 2008; Mu et al., 2008; To et al., 2006). Transgenic approaches have been used in *Brassica napus* which have shown an increase of oil accumulation through the over expression of genes involved in fatty acid synthesis (Thelen and Ohlrogge, 2002), genes involved in the acyltransferase genes in the Kennedy pathway (Jain et al., 2000; Jako et al., 2001; Taylor et al., 2002; Zou et al., 1997) and cytosolic gly3PDH (Vigeolas et al., 2007). The use of transgenic approaches has allowed for the increase of oil content however the use of GMO technology has limited application for current cultivars due to the high cost of registration.

Through the use of QTL analysis and MAS it is possible to select favorable genotypes that have an increase in fatty acid accumulation. Screening for favorable genotypes can improve the efficiency of breeding. By selecting the desired genotypes the number of lines can be reduced to

identify the desired phenotypes. Molecular markers have been linked to a number of important traits such as disease resistance (Kole et al., 2002a; Pilet et al., 2001; Zhao and Meng, 2003), improvement of fatty acid content (Hu et al., 2006; Rahman et al., 2008) and abiotic stress tolerance (Kole et al., 2002b). These can be developed to allow for the profiling of lines by selecting desired genotypes prior to phenotypic screening.

In this study, the development of the molecular maps for the two DH populations demonstrated the effectiveness the use of SSR and SRAP markers for the genomic study of *Brassica juncea*. The majority of the published linkage maps of *Brassica juncea* has been developed using AFLP and RFLP markers with have limited application for high throughput screening for the use of MAS. The reproducibility and high rate of polymorphism makes SSR markers a good candidate for MAS and genetic studies. Development of these maps offers opportunity to increase the opportunities to use SSR and SRAP markers for screening *B. juncea* population as their use has been limited to date.

The linkage maps developed in this study provided a limited coverage of the entire *B. juncea* genome. In the case of both of the DH populations studied there linkage maps developed covered 13 of the 18 linkage groups. The use of *Brassica napus* derived SSR and SRAP markers may have contributed to the limited coverage of the genome. The *B. nigra* derived B genome which contributes 8 chromosomes to *B. juncea* has been shown to have wide genetic distance that the A and B genomes of *B. napus* derived from *B. rapa* and *B. oleracea* respectively. *B. napus* and *B. juncea* share the *B. rapa* derived A genome which suggests the markers used in this study may have favored this genome and resulted in the limited genome coverage. Further marker screening would need to be completed to allow for the complete coverage of the *B. juncea* genome using these markers.

Through the phenotypic screening of the two populations the fatty acid profiles and oil and protein content of each of the lines was evaluated. From these studies it was shown that both populations were segregating for erucic acid. Early studies into the heritability of erucic acid have shown that the trait is controlled by two genes in *B. juncea* with additive effects resulting in high, moderate and low erucic phenotypic expression. The analysis of the two populations developed from the cross between moderate erucic lines (Vniimk351 and Vniimk405) and the low (zero) erucic acid line ZEM1. There resultant DH populations showed a distinct 1:1 segregation of moderate and low erucic lines. This suggests the segregation of a single dominant gene responsible for the moderate erucic acid phenotypes within both populations.

Using phenotypic data collected from nursery trials at two site locations in Manitoba and one in Chile the oil and protein content were studied using QTL analysis. Using the linkage maps developed the analysis of the fatty acid profile and oil and protein content of each of the lines identified a number of QTL within both populations. Little correlation was observed between the two populations despite sharing a common parental line. Significant correlations between specific fatty acids and oil and protein accumulation were identified in the analysis of the lines within the two populations. This suggests that much of the variation within oil content is dependent of the fatty acids within the seed. There has been evidence that supports the increase of erucic acid within the seed results in an increase in overall oil contents. This correlation has been identified in other studies and supported with the significant positive correlation between erucic acid content and oil content within the seed. Additionally there is a significant negative correlation between oil and protein content which supports previous studies related to *Brassica* oil and protein accumulation.

This studies shows that *B. juncea* can contribute to the increase in oilseed production within Canada. Further research will be required to improve the availability of molecular markers available for the B genome to aid in the genetic studies of *B. juncea*. This will help in the development of improved linkage maps coverage and marker density within the maps. This will further the ability to screen *B. juncea* genotypes for desired alleles. Further conclusions derived from this study suggest that populations derived from parental lines with varying fatty acid profiles reduces the identification of oil content QTL due to the significant correlations between fatty acids and overall oil accumulation. This correlation causes a large amount of genetic "noise" in the QTL analysis inhibiting the ability to identify discrete regions responsible for the increase in overall fatty acid accumulation.

6.0 Appendix

6.1 Appendix 1: Analysis of Variance (ANOVA) analysis of the combined data measured in field trial at two sites in Manitoba and one in Chile for ZEM1xVniimk351 DH population.

Source of Variation	DF	Sum of Squares	Mean Square	Expected Mean Square	F Value	Pr > F
Sum Oil and Protein (Sum)						
Line	107	3093.77173	28.913754	Var(Residual) + Q(Line,Location*Line)	50.59	<.0001
Location	2	1498.976887	749.488443	Var(Residual) + Q(Location,Rep(Location),Location*Line)	1311.3	<.0001
Rep(Location)	9	136.928501	15.214278	Var(Residual) + Q(Rep(Location))	26.62	<.0001
Location*Line	206	441.515208	2.143278	Var(Residual) + Q(Location*Line)	3.75	<.0001
Residual	1069	610.982343	0.571546	Var(Residual)	.	.
Seed oil content (Oil)						
Line	107	14368	134.283563	Var(Residual) + Q(Line,Location*Line)	63.89	<.0001
Location	2	2353.193417	1176.596709	Var(Residual) + Q(Location,Rep(Location),Location*Line)	559.77	<.0001
Rep(Location)	9	144.794662	16.088296	Var(Residual) + Q(Rep(Location))	7.65	<.0001
Location*Line	206	1153.319958	5.598641	Var(Residual) + Q(Location*Line)	2.66	<.0001
Residual	1069	2246.969668	2.101936	Var(Residual)	.	.
Seed protein content (Protein)						
Line	107	6733.489888	62.929812	Var(Residual) + Q(Line,Location*Line)	38.43	<.0001
Location	2	164.686859	82.343429	Var(Residual) + Q(Location,Rep(Location),Location*Line)	50.28	<.0001
Rep(Location)	9	103.999057	11.555451	Var(Residual) + Q(Rep(Location))	7.06	<.0001
Location*Line	206	685.35369	3.32696	Var(Residual) + Q(Location*Line)	2.03	<.0001
Residual	1069	1750.674471	1.637675	Var(Residual)	.	.

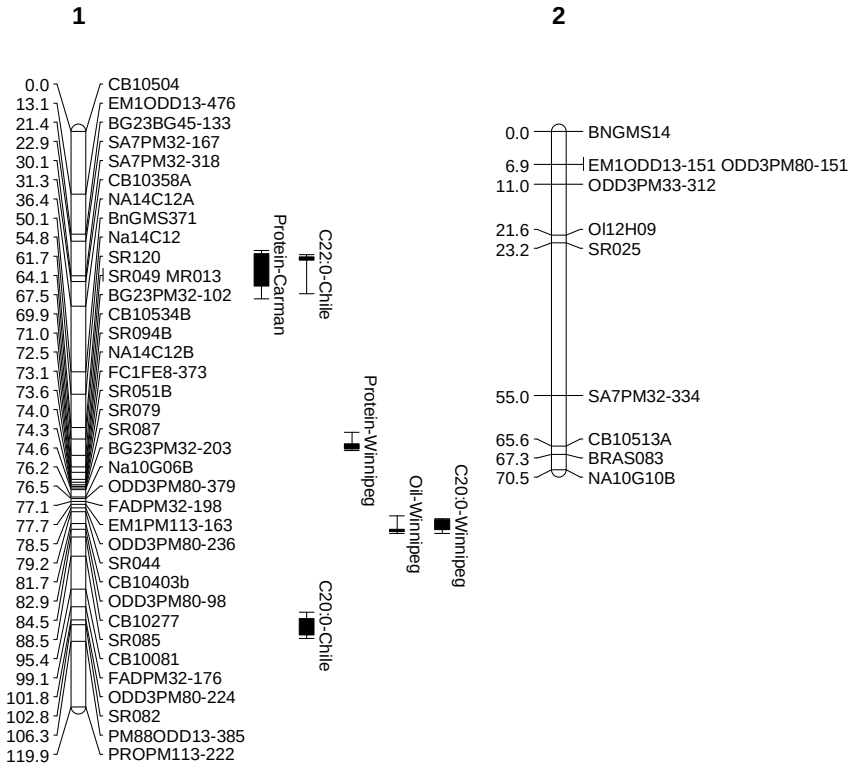
Linolenic acid content (C18:3)						
Line	107	2581.937432	24.130256	Var(Residual) + Q(Line,Location*Line)	26.1	<.0001
Location	2	108.109922	54.054961	Var(Residual) + Q(Location,Rep(Location),Location*Line)	58.46	<.0001
Rep(Location)	3	0.706002	0.235334	Var(Residual) + Q(Rep(Location))	0.25	0.858
Location*Line	190	90.192259	0.474696	Var(Residual) + Q(Location*Line)	0.51	1
Residual	273	252.412315	0.924587	Var(Residual)	.	.
Palmitic acid content (C16:0)						
Line	107	108.238122	1.011571	Var(Residual) + Q(Line,Location*Line)	81.62	<.0001
Location	2	6.951068	3.475534	Var(Residual) + Q(Location,Rep(Location),Location*Line)	280.42	<.0001
Rep(Location)	3	0.125809	0.041936	Var(Residual) + Q(Rep(Location))	3.38	0.019
Location*Line	190	4.035761	0.021241	Var(Residual) + Q(Location*Line)	1.71	<.0001
Residual	273	3.383551	0.012394	Var(Residual)	.	.

6.2 Appendix 1.1: Analysis of Variance (ANOVA) analysis of the combined data measured in field trial at two sites in Manitoba and one in Chile for ZEM1xVniimk405 DH population.

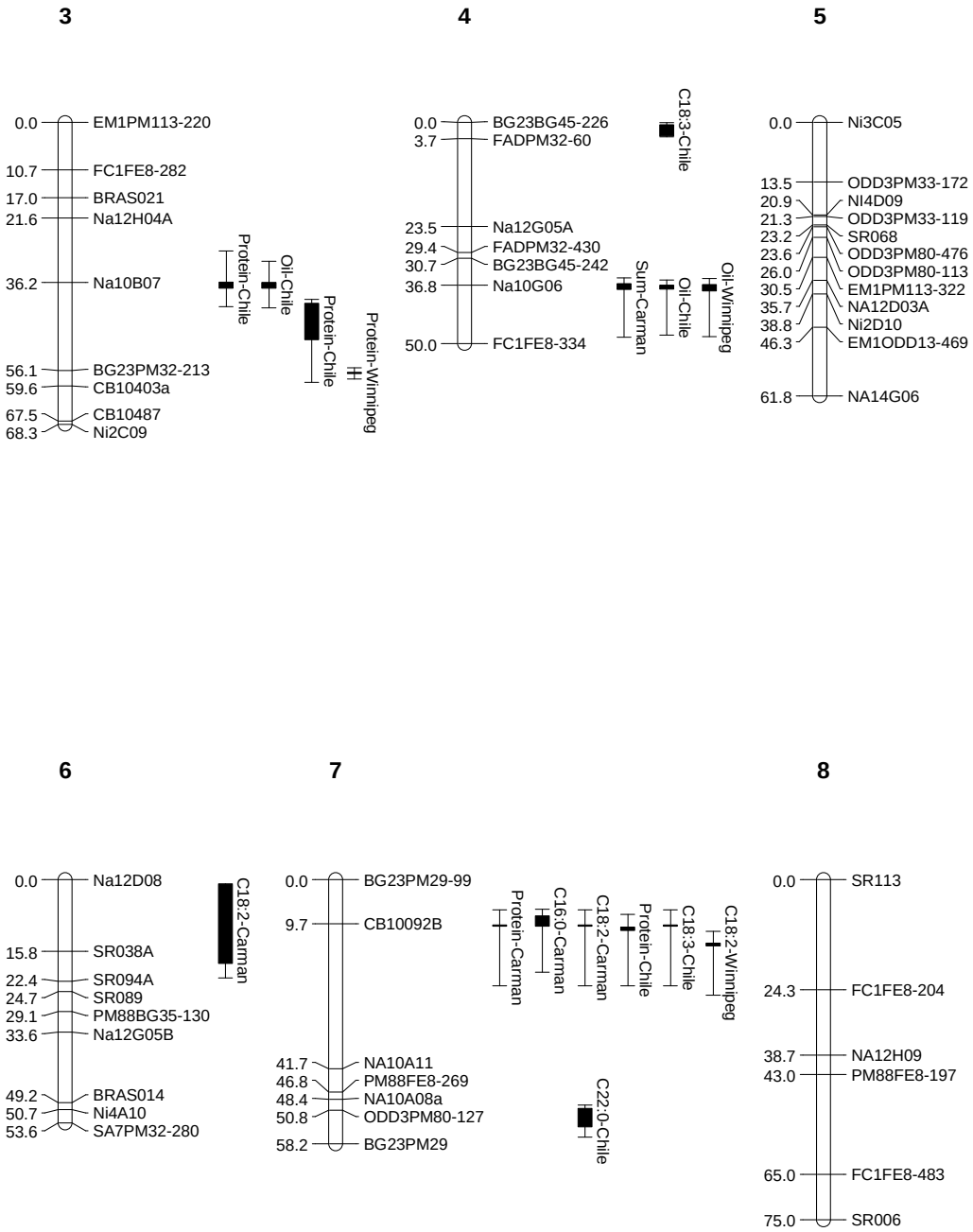
Source of Variance	DF	Sum of Squares	Mean Square	Expected Mean Square	F Value	Pr > F
Linolenic acid content (C18:3)						
Line	60	1489.564	24.82606	Var(Residual) + Q(Line,Location*Line)	22.73	<.0001
Location	2	148.1014	74.05072	Var(Residual) + Q(Location,Rep(Location),Location*Line)	67.8	<.0001
Rep(Location)	3	2.18236	0.727453	Var(Residual) + Q(Rep(Location))	0.67	0.5738
Location*Line	110	128.8157	1.171052	Var(Residual) + Q(Location*Line)	1.07	0.3304
Residual	214	233.7315	1.092203	Var(Residual)	.	.
Palmitic acid content (C16:0)						
Line	60	59.75777	0.995963	Var(Residual) + Q(Line,Location*Line)	36.49	<.0001
Location	2	4.952907	2.476453	Var(Residual) + Q(Location,Rep(Location),Location*Line)	90.73	<.0001
Rep(Location)	3	0.114693	0.038231	Var(Residual) + Q(Rep(Location))	1.4	0.2436
Location*Line	110	3.898279	0.035439	Var(Residual) + Q(Location*Line)	1.3	0.0538
Residual	214	5.841025	0.027295	Var(Residual)	.	.
Sum Oil and Protein (Sum)						
Line	59	1381.552	23.41613	Var(Residual) + Q(Line,Location*Line)	24.25	<.0001
Location	2	228.4646	114.2323	Var(Residual) + Q(Location,Rep(Location),Location*Line)	118.32	<.0001
Rep(Location)	9	16.67073	1.852303	Var(Residual) + Q(Rep(Location))	1.92	0.0468
Location*Line	112	264.3224	2.360022	Var(Residual) + Q(Location*Line)	2.44	<.0001
Residual	616	594.7387	0.965485	Var(Residual)	.	.
Seed oil content (Oil)						
Line	59	5059.108	85.74759	Var(Residual) + Q(Line,Location*Line)	31.13	<.0001
Location	2	727.9726	363.9863	Var(Residual) + Q(Location,Rep(Location),Location*Line)	132.15	<.0001
Rep(Location)	9	29.47037	3.274485	Var(Residual) + Q(Rep(Location))	1.19	0.2991
Location*Line	112	875.8586	7.820166	Var(Residual) + Q(Location*Line)	2.84	<.0001
Residual	616	1696.62	2.754253	Var(Residual)	.	.

Seed protein content (Protein)						
Line	59	1971.635	33.41755	Var(Residual) + Q(Line,Location*Line)	20.13	<.0001
Location	2	142.094	71.04698	Var(Residual) + Q(Location,Rep(Location),Location*Line)	42.81	<.0001
Rep(Location)	9	45.86383	5.095981	Var(Residual) + Q(Rep(Location))	3.07	0.0013
Location*Line	112	448.7902	4.007055	Var(Residual) + Q(Location*Line)	2.41	<.0001
Residual	616	1022.374	1.659699	Var(Residual)	.	.

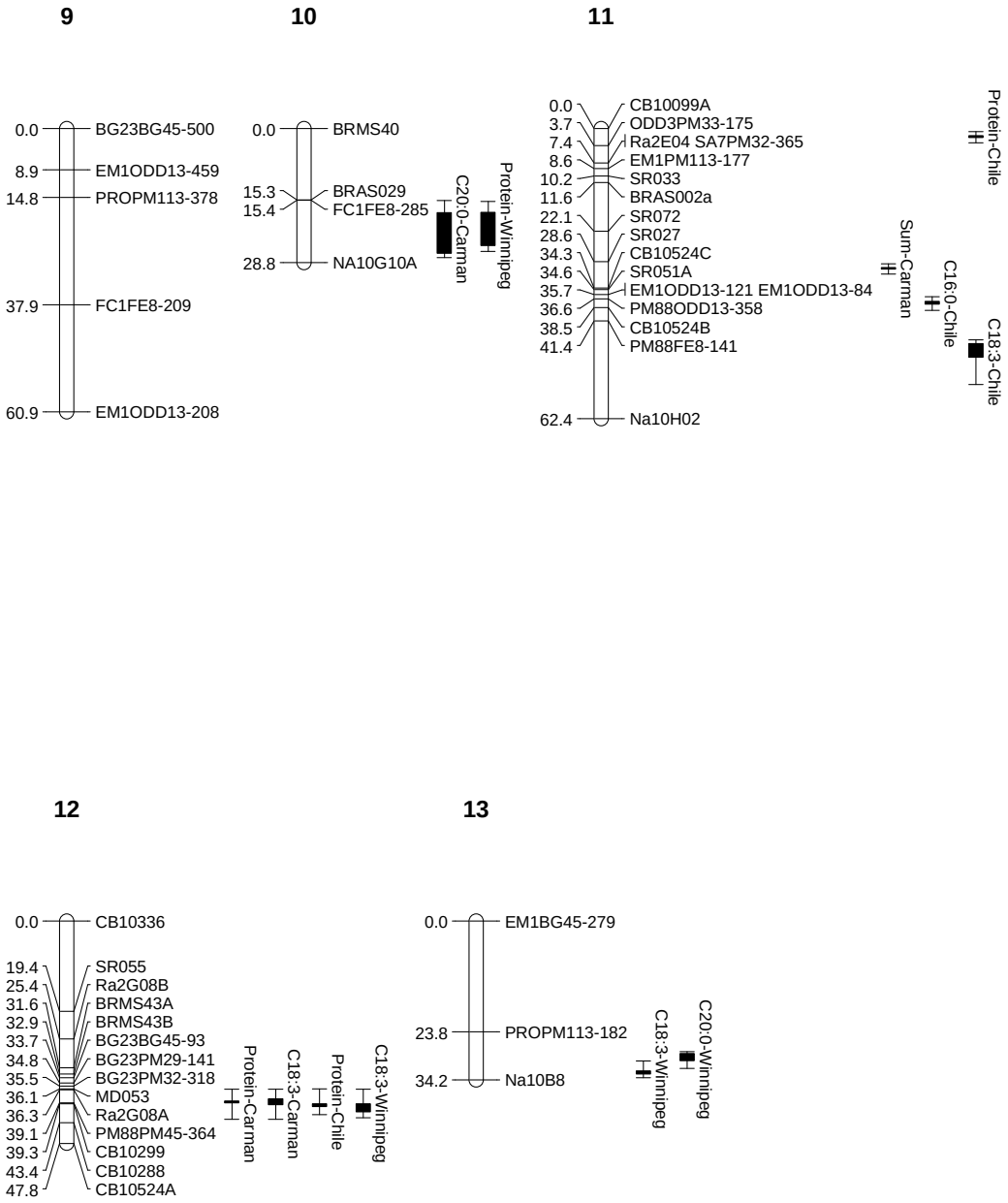
6.3 Appendix 2.1: QTL locations on the 13 linkage groups of the linkage map developed using SSR and SRAP markers based on the ZEM1xVniimk351 DH population.



Appendix 2.1 con't

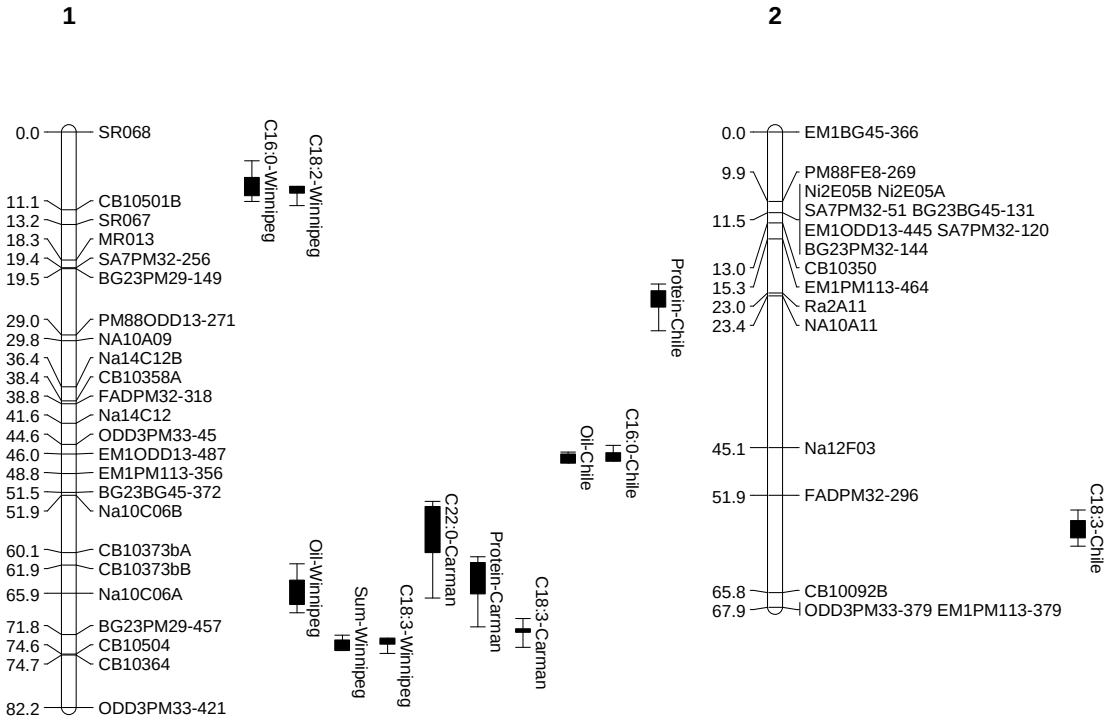


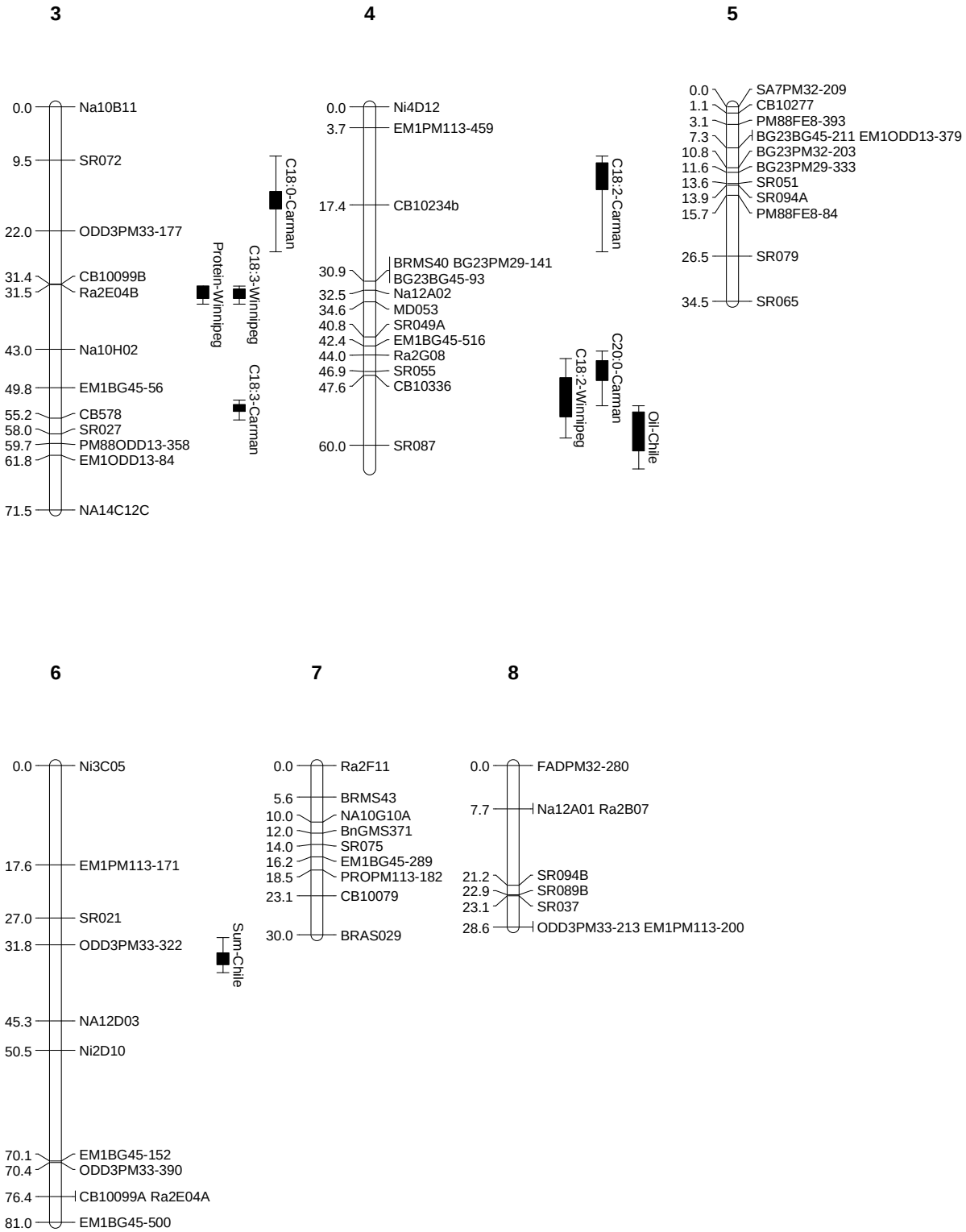
Appendix 2.1 con't



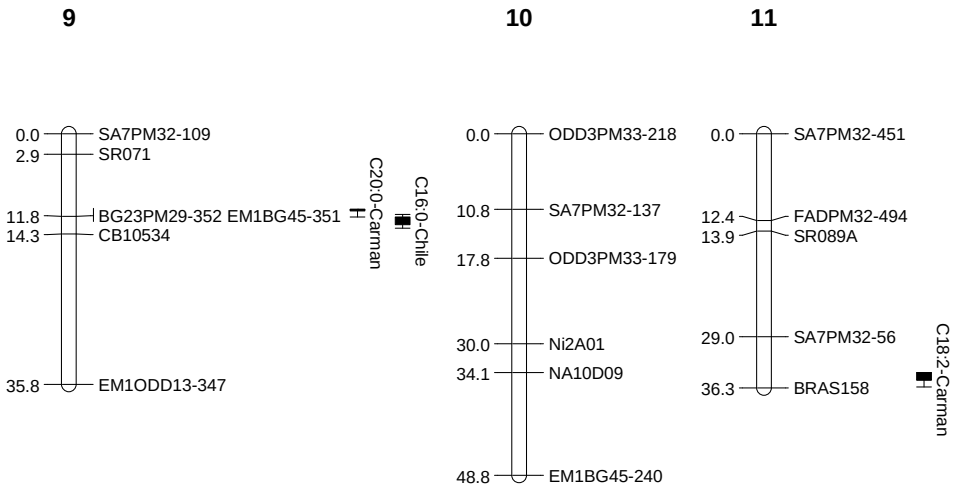
Appendix 2.1 con't

6.4 Appendix 2.2: QTL locations on the 13 linkage groups of the linkage map developed using SSR and SRAP markers based on the ZEM1xVniimk405 DH population.

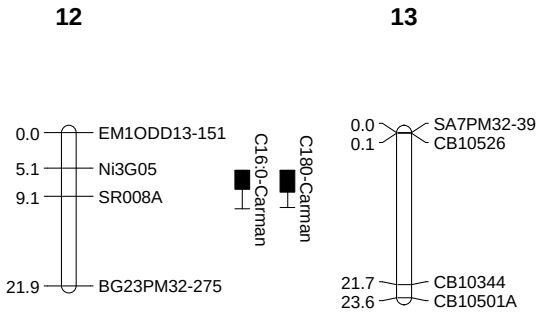




Appendix 2.2 con't



Appendix 2.2 con't



Appendix 2.2 con't

7.0 References

AAFC. (2009) Overview of the Canadian Special Crops Industry, Agri-Food Trade Service - Special Crops Section. pp. 23.

AAFC. (2010) Canada Canola Sector Profile - Part One: Situation and Outlook, Market Outlook Report, Agriculture and Agri-Food Canada.

Agarwal M., Shrivastava N., Padh H. (2008) Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Reports 27:617-631. DOI: 10.1007/s00299-008-0507-z.

Aggarwal R.A.K., Sharma R., Kumar R., Mohapatra T., Sharma R.P. (2003) Molecular mapping of loci affecting the contents of three major fatty acids in Indian mustard (*Brassica juncea* L). Journal of Plant Biochemistry and Biotechnology 12:131-137.

Alemayehu N., Becker H.C. (2001) Variation and inheritance of erucic acid content in *Brassica carinata* germplasm collections from Ethiopia. Plant Breeding 120:331-335. DOI: 10.1046/j.1439-0523.2001.00623.x.

Anand I.J., Downey R.K. (1981) A study of erucic acid alleles in digenomic rapeseed (*Brassica napus* L.). Can J Plant Sci 61:199-203.

Attia T., RÖbbelen G. (1986) Meiotic pairing in haploids and amphihaploids of spontaneous versus synthetic origin in rape, *Brassica napus* L. Canadian Journal of Genetic Cytology 28:330-334.

Auld D.L., Heikkinen M.K., Erickson D.A., Sernyk J.L., Romero J.E. (1992) Rapeseed Mutants with Reduced Levels of Polyunsaturated Fatty Acids and Increased Levels of Oleic Acid. Crop Sci. 32:657-662. DOI: 10.2135/cropsci1992.0011183X003200030016x.

Axelsson T., Shavorskaya O., Lagercrantz U. (2001) Multiple flowering time QTLs within several Brassica species could be the result of duplicated copies of one ancestral gene. Genome 44:856-864. DOI: 10.1139/gen-44-5-856.

Axelsson T., Bowman C.M., Sharpe A.G., Lydiate D.J., Lagercrantz U. (2000) Amphidiploid *Brassica juncea* contains conserved progenitor genomes. Genome 43:679-688.

Barker G.C., Larson T.R., Graham I.A., Lynn J.R., King G.J. (2007) Novel insights into seed fatty acid synthesis and modification pathways from genetic diversity and quantitative trait loci analysis of the Brassica C genome. Plant Physiology (Rockville) 144:1827.

Barret P., Delourme R., Brunet D., Jourdren C., Horvais R., Renard M. (1999) Low linolenic acid level in rapeseed can be easily assessed through the detection of two single base substitution in *fad3* genes., 10th international rapeseed congress, Canberra, Australia. pp. 26-29.

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

Basten C., Weir B., Zeng B. (2001) QTL Cartographer, Department of Statistics, North Carolina State University, Raleigh, NC.

Baud S., Dubreucq B., Miquel M., Rochat C., Lepiniec L. (2008) Storage Reserve Accumulation in *Arabidopsis*: Metabolic and Developmental Control of Seed Filling. The Arabidopsis Book 63:1. DOI: 10.1199/tab.0113.

Beare J., Campell, JA, Youngs, CG, Craig, BN,. (1963) Effects of saturated fat in rats fed rapeseed oil. Canadian journal of physiology and pharmacology 41:605-612.

Becker H.C., Engqvist G.M., Karlsson B. (1995) Comparison of rapeseed cultivars and resynthesized lines based on allozyme and RFLP markers. TAG Theoretical and Applied Genetics 91:62-67. DOI: 10.1007/bf00220859.
Bell J.M. (1982) From rapeseed to canola: A brief history of research for superior meal and edible oil. Poultry Science:613-622.
Bhat M.A., Gupta M.L., Banga S.K., Raheja R.K., Banga S.S. (2002) Erucic acid heredity in *Brassica juncea* - some additional information, Plant Breeding. pp. 456-458.
Bhat S., Sarla N. (2004) Identification and overcoming barriers between *Brassica rapa* L. em. Metzg. and *B. nigra* (L.) Koch crosses for the resynthesis of *B. juncea* (L.) Czern. Genetic Resources and Crop Evolution 51:455-469. DOI: 10.1023/B:GRES.0000024154.19867.cd.
Bille N., Eggum B.O., Jacobsen I., Olsen O., Sørensen H. (1983) Antinutritional and toxic effects in rats of individual glucosinolates (± myrosinases) added to a standard diet. Zeitschrift für Tierphysiologie Tierernährung und Futtermittelkunde 49:195-210. DOI: 10.1111/j.1439-0396.1983.tb00801.x.
Bonfils A.C., Sproule A., Webb J., Keller W. (1992) Plant regeneration from stem cortex expiant and protoplast cultures of *Brassica juncea* (mustard). Plant Cell Reports 11:614-617. DOI: 10.1007/bf00236384.
Burns M.J., Barnes S.R., Bowman J.G., Clarke M.H.E., Werner C.P., Kearsey M.J. (2003) QTL analysis of an intervarietal set of substitution lines in *Brassica napus*: (i) Seed oil content and fatty acid composition. Heredity 90:39-48.
Burton W., Salisbury P., Potts D. (2003) The potential of canola quality *Brassica juncea* as an oilseed crop for Australia, in: J. Edwards (Ed.), Thirteenth Biennial Australian Research Assembly on Brassicas. Proceedings of a conference, Tamworth, New South Wales, Australia. pp. 62-64.
Burton W.A., Ripley V.L., Potts D.A., Salisbury P.A. (2004) Assessment of genetic diversity in selected breeding lines and cultivars of canola quality *Brassica juncea* and their implications for canola breeding. Euphytica; Netherlands journal of plant breeding 136:181-192.
Burton W.A., Flood R.F., Norton R.M., Field B., Potts D.A., Robertson M.J., Salisbury P.A. (2008) Identification of variability in phenological responses in canola-quality *Brassica juncea* for utilisation in Australian breeding programs. Australian Journal of Agricultural Research 59:874-881. DOI: 10.1071/ar07415.
Canola Council of Canada. (2007) From Farm to Fuel - The impact of a Canola-Based Biodiesel Industry in Canada, Canola Council of Canada.
Canola Council of Canada. (2011) The Economic Impact of Canadian Grown Canola and its End Products on the Canadian Economy, LMC International. pp. 63.
Chauhan J.S., Meena S.S., Singh M., Singh K.H. (2010) Estimating genotype x environment interaction and stability parameters for oil and seed meal quality, seed yield and its contributing characters in Indian mustard (*Brassica juncea*). Indian Journal of Agricultural Sciences 80:110-115.
Chauhan J.S., Singh K.H., Meena S.S., Singh M. (2011) Environmental effects on oil, fatty acid profile, protein and glucosinolate content in Indian mustard (*Brassica juncea*). The Indian Journal of Agricultural Sciences 81.

Chen B.Y., Heneen W.K. (1989) Fatty acid composition of resynthesized *Brassica napus* L., *B.campestris* L. and *B. alboglabra* Bailey with special reference to the inheritance of erucic acid content. Heredity 63:309-314.
Chen B.Y., Heneen W.K., Jonsson R. (1988) Independant inheritance of erucic acid content and flower colour in the C-genome of *Brassica napus* L. . Plant Breeding 100:147-149.
Chen W., Zhang Y., Liu X., Chen B., Tu J., Fu T. (2007) Detection of QTL for six yield-related traits in oilseed rape (*Brassica napus*) using DH and immortalized F-2 populations. Theoretical and Applied Genetics 115:849-858. DOI: 10.1007/s00122-007-0613-2.
Cheng B.F., Seguin-Swartz G., Somers D.J., Rakow G. (2001) Low glucosinolate *Brassica juncea* breeding line revealed to be nullisomic. Genome 44:738-741.
Cheng X., Xu J., Xia S., Gu J., Yang Y., Fu J., Qian X., Zhang S., Wu J., Liu K. (2009) Development and genetic mapping of microsatellite markers from genome survey sequences in *Brassica napus*. Theoretical and Applied Genetics 118:1121-1131. DOI: 10.1007/s00122-009-0967-8.
Cheung W.Y., Champagne G., Hubert N., Landry B.S. (1997) Comparison of the genetic maps of *Brassica napus* and *Brassica oleracea*. Theoretical and Applied Genetics 94:569-582.
Cheung W.Y., Landry B.S., Raney P., Rakow G.F.W. (1998) Molecular mapping of seed quality traits in *Brassica juncea* L. Czern. and Coss, in: G. Thomas and A. A. Monteiro (Eds.), Brassica 97 - International Symposium on Brassicas. pp. 139-147.
Cheung W.Y., Friesen L., Rakow G.F.W., Seguin-Swartz G., Landry B.S. (1997b) A RFLP-based linkage map of mustard (*Brassica juncea* (L.) Czern. and Coss.). Theoretical and Applied Genetics 94:841-851.
Christie W.W. (1989) Gas Chromatography and Lipids A Practical Guide The Oily Press Ltd.
Delourme R., Falentin C., Huteau V., Clouet V., Horvais R., Gandon B., Specel S., Hanneton L., Dheu J.E., Deschamps M., Margale E., Vincourt P., Renard M. (2006) Genetic control of oil content in oilseed rape (*Brassica napus* L.). Theoretical and Applied Genetics 113:1331-1345. DOI: 10.1007/s00122-006-0386-z.
Dorrell D.G., Downey R.K. (1964) The inheritance of erucic acid content in rapeseed (*Brassica campestris*). Can J Plant Sci 44:499-504.
Downey R.K., Harvey B.L. (1963) Methods of breeding for oil quality in rape. Canadian Jour Plant Sci 43:271-275.
Downey R.K., Röbbelen, G. (1989) Oilcrops of the World: Their Breeding and Utilization McGraw-Hill Publishing Company, New York.
Ecke W., Uzunova M., Weissleder K. (1995) Mapping the genome of rapeseed (*Brassica napus* L.): II. Localization of genes controlling erucic acid synthesis and seed oil content. Theoretical and Applied Genetics 91:972-977.
Edwards D., Salisbury P., Burton W., Hopkins C., Batley J. (2007) Indian Mustard, in: C. K. (Ed.) (Ed.), Genome Mapping and Molecular Breeding in Plants Volume 2 Oilseeds, Springer-Verlag, Berlin Heidelberg. pp. 179-210.
F.W. Hougen, Bodo V. (1973) Extraction and Methanolysis of Oil from Whole or Crushed Rapeseed for Fatty Acid Analysis." Journal of American Oil Chemists. Journal of American Oil Chemists' Society 50:230-234.
FAO. (2010) Food Outlook Global Market Analysis 2010 - Oilseeds, Oils and Meals, FAO Corporate Document Repository, FAO. pp. 104.
Ferreira M.E., Williams P.H., Osborn T.C. (1994) RFLP mapping of *Brassica napus* using doubled haploid lines. Theoretical and Applied Genetics 89:615-621.

Foisset N., Delourme R., Barret P., Hubert N., Landry B.S., Renard M. (1996) Molecular-mapping analysis in *Brassica napus* using isozyme, RAPD, RFLP markers on a double-haploid progeny. Theoretical and Applied Genetics 93:1017-1025.
Fondon J.W., Garner H.R. (2004) Molecular origins of rapid and continuous morphological evolution. Proceedings of the National Academy of Sciences 101:18058-18063. DOI: 10.1073/pnas.0408118101.
Fourmann M., Barret P., Renard M., Pelletier G., Delourme R., Brunei D. (1998). Theor Appl Genet 96:852.
Gomez-Campo C., Prakash, Shyam. (1999) Origin and Domestication, in: C. Gomez-Campo (Ed.), Biology of Brassica Coenospecies, Elseveier, Amsterdam. pp. 33-58.
Gopalan C.K., D, Shenolikar, IS, Krishnamachari, K. (1974) Myocardial changes in monkeys fed mustard oil. Nutritional Metabolism:352-365.
Graef G.L., Fehr W.R., Hammond E.G. Inheritance of Three Stearic Acid Mutants of Soybean. Crop Sci. 25:1076-1079. DOI: 10.2135/cropsci1985.0011183X002500060043x.
Grundy S. (1994) Influence of stearic acid on cholesterol metabolism relative to other long-chain fatty acids. The American Journal of Clinical Nutrition 60:986S-990S.
Gul M.K. (2003) QTL mapping and analysis of QTLXnitrogen interactions for some yield components in *Brassica napus* L. Turkish Journal of Agriculture and Forestry 27:71-76.
Gunasekera C., Martin L., Siddique K., Walton G. (2006) Genotype by environment interactions of Indian mustard (*Brassica juncea* L.) and canola (*B. napus* L.) in Mediterranean-type environments1. Crop growth and seed yield. European Journal of Agronomy 25:1-12. DOI: 10.1016/j.eja.2005.08.002.
Gupta V., Mukhopadhyay A., Arumugam N., Sodhi Y.S., Pental D., Pradhan A.K. (2004) Molecular tagging of erucic acid trait in oilseed mustard (*Brassica juncea*) by QTL mapping and single nucleotide polymorphisms in FAE1 gene. TAG Theoretical and Applied Genetics 108:743-749. DOI: 10.1007/s00122-003-1481-z.
Gutierrez L., Wuytswinkel O.V., Castelain M., Bellini C. (2007) Combined networks regulating seed maturation. Trends in Plant Science 12:294-300.
Hamada H., Petrino M.G., Kakunaga T. (1982) A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proceedings of the National Academy of Sciences 79:6465-6469.
Harvey B.L., Downey R.K. (1964) The inheritance of erucic acid content in rapeseed (*Brassica napus* L.). Can J Plant Sci 44:104-111.
Hawkins D.J., Kridl J.C. (1998) Characterization of acyl-ACP thioesterases of mangosteen (*Garcinia mangostana)* seed and high levels of stearate production in transgenic canola. The Plant Journal 13:743-752. DOI: 10.1046/j.1365-313X.1998.00073.x.
He C., Poysa V., Yu K. (2003) Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars. Theor Appl Genet 106:363 - 373.
Hemingway J.S. (1995) The Mustard Species: Condiment and Food Ingredient Use and Potential as Oilseed Crops., in: D. a. M. Kimber, D.I (Ed.), Brassica Oilseeds Production and Utilization, CAB International, Wallingford.
Hu X., Sullivan-Gilbert M., Gupta M., Thompson S. (2006) Mapping of the loci controlling oleic and linolenic acid contents and development of *fad2* and *fad3* allele-specific markers in canola (*Brassica napus* L.). TAG Theoretical and Applied Genetics 113:497-507. DOI: 10.1007/s00122-006-0315-1.

Hulan H., Thompson, B, Kramer, JKG, Sauer, FD, Corner, AH. (1977) An evaluation of the incidence of heart lesions in rats fed rapeseed oil. Canadian Institute of Food Science Technology Journal 10:23-26.

Innan H., Terauchi R., Miyashita N.T. (1997) Microsatellite Polymorphism in Natural Populations of the Wild Plant *Arabidopsis thaliana*. Genetics 146:1441-1452.

Jagannath A., Sodhi Y.S., Gupta V., Mukhopadhyay A., Arumugam N., Singh I., Rohatgi S., Burma P.K., Pradhan A.K., Pental D. (2010) Eliminating expression of erucic acid-encoding loci allows the identification of "hidden" QTL contributing to oil quality fractions and oil content in *Brassica juncea* (Indian mustard). Theoretical and Applied Genetics 122:1091-1103. DOI: 10.1007/s00122-010-1515-2.

Jain R.K., Coffey M., Lai K., Kumar A., Mackenzie S.L. (2000) Enhancement of seed oil content by expression of glycerol-3-phosphate acyltransferase genes. Biochem Soc Trans. 28:958-961.

Jako C., Kumar A., Wei Y., Zou J., Barton D.L., E.M.Giblin, Covello P.S., Taylor D.C. (2001) Seed-specific over-expression of an *Arabidopsis* cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. Plant Physiology 126:861-874.

James D.W., Lim E., Keller J., Plooy I., Ralston E., 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.

Jansen R.C. (1993) Interval Mapping of Multiple Quantitative Trait Loci. Genetics 135:205-211.

Jansen R.C. (1994) Controlling the Type-I and Type II Errors in Mapping Quantitive Trait Loci. Genetics 138:871-881.

Johnston J.S., Pepper A.E., Hall A.E., Chen Z.J., Hodenett G., Drabek J., Lopez R., Price H.J. (2005) Evolution of Genome Size in Brassicaceae. Annals of Botany 95:229-235. DOI: 10.1093/aob/mci016.

Jourdren C., Barret P., Brunel D., Delourme R., Renard M. (1996a) Specific molecular marker of the genes controlling linolenic acid content in rapeseed. Theoretical and Applied Genetics 93:512-518.

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

Kanrar S., Venkateswari J., Dureja P., Kirti P.B., Chopra V.L. (2005) Modification of erucic acid content in Indian mustard (*Brassica juncea*) by up-regulation and down-regulation of the Brassica juncea FAT TY ACID ELONGATION1 (*BjFAE1*) gene. Plant Cell Reports 25:148-155. DOI: 10.1007/s00299-005-0068-3.

Karim Zarhloul M., Stoll C., Lühs W., Syring-Ehemann A., Hausmann L., Töpfer R., Friedt W. (2006) Breeding high-stearic oilseed rape (*Brassica napus*) with high- and low-erucic background using optimised promoter-gene constructs. Molecular Breeding 18:241-251. DOI: 10.1007/s11032-006-9032-3.

Kim K.S., Park S.H., Choung M.G., Jang Y.S. (2007) Use of Near-Infrared Spectroscopy for Estimating Fatty Acid Composition in Intact Seeds of Rapeseed. Journal of Crop Science and Biotechnology 10:15-20.

Kirk J.T.O., Oram R.N. (1981) Isolation of Erucic Acid-Free Lines of *Brassica juncea* - Indian Mustard Now a Potential Oilseed Crop in Australia. Journal of the Australian Institute of Agricultural Science 47:51-52.

Kirk J.T.O., Hurlstone C.J. (1983) Variation and inheritance of erucic acid content in *Brassica juncea*. Zeitschrift fur Pflanzenzuchtung 90:331-338.
Knutzon D.S., Thompson G.A., Radke S.E., Johnson W.B., Knauf V.C., Kridl J.C. (1992) Modification of Brassica seed oil by antisense expression of a stearoyl-acyl carrier protein desaturase gene. Proceedings of the National Academy of Sciences 89:2624-2628.
Kole C., Williams P.H., Rimmer S.R., Osborn T.C. (2002a) Linkage mapping of genes controlling resistance to white rust (*Albugo candida*) in *Brassica rapa* (syn. campestris) and comparative mapping to *Brassica napus* and *Arabidopsis thaliana*. Genome 45:22-27.
Kole C., Thormann C.E., Karlsson B.H., Palta J.P., Gaffney P., Yandell B., Osborn T.C. (2002b) Comparative mapping of loci controlling winter survival and related traits in oilseed *Brassica rapa* and *B. napus*. Molecular Breeding 9:201-210.
Kondra Z.P., Stephansson B.R. (1965) Inheritance of erucic acid and eicosenoic acid content of seed oil of rapeseed (Brassica napus). Canadian journal of Cytology 7:505-510.
Kondra Z.P., Stefansson B.R. (1970) Inheritance of Mayor Glucosinolates of Rapeseed (*Brassica napus*) Meal. Canadian Journal of Plant Science 50:643-&.
Konishi T., Shinohara K., Yamada K., Sasaki Y. (1996) Acetyl-CoA carboxylase in higher plants: Most plants other than gramineae have both the prokaryotic and the eukaryotic forms of this enzyme. Plant and Cell Physiology 37:117-122.
Koundal V., Parida S.K., Yadava D.K., Ali A., Koundal K.R., Mohapatra T. (2008) Evaluation of Microsatellite Markers for Genome Mapping in Indian Mustard (*Brassica juncea* L.). Journal of Plant Biochemistry and Biotechnology 17:69-72.
Kumar A., Singh D.P., Singh P. (1987) Genotypic variation in the responses of Brassica species to water deficit. Journal of Agricultural Science 109.
Kumar A., Singh D.P., Singh H., Sharma H.C. (1984) Differences in osmoregulation in Brassica species. Annals of Botany 54:537-541.
Kumar A., Singh D.P., Singh. P., Res L.F.C. (1994) Influence of water stress on photosynthesis, transpiration, water use efficiency and yield of *Brassica juncea*. L. Field Crops Res 37:95-101.
La Mura M., Norris C., Sporle S., Jayaweera D., Greenland A., Lee D. (2010) Development of genome-specific 5S rDNA markers in Brassica and related species for hybrid testing. Genome 53:643-649. DOI: 10.1139/g10-033.
Labana K.S., Chaurasia B.D., Sangha A.S. (1987) INHERITANCE OF OIL CONTENT IN INDIAN MUSTARD. Indian Journal of Agricultural Sciences 57:539-543.
Lagercrantz U. (1998) Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that Brassica genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics 150:1217-1228.
Lagercrantz U., Lydiate D.J. (1996) Comparative genome mapping in Brassica. Genetics 144:1903-1910.
Lander E.S., Botstein D. (1989) Mapping Mendelian Factors Underlying Quantitative Traits Using RFLP Linkage Maps. Genetics 121:185-199.
Landry B.S., Hubert N., Etho T., Harada J.J., Lincoln S.E. (1991) A Genetic Map for *Brassica napus* Based on Restriction Fragment Length Polymorphisms Detected with Expressed DNA Sequences. Genome 34:543-552.

Lecerf J.M. (2009) Fatty acids and cardiovascular disease. Nutrition Reviews 67:273-283. DOI: 10.1111/j.1753-4887.2009.00194.x.
Li G., 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. Theoretical and Applied Genetics 103:455-461.
Lin X., Kaul S., Rounsley S., Shea T.P., Benito M.-I., Town C.D., Fujii C.Y., Mason T., Bowman C.L., Barnstead M., Feldblyum T.V., Buell C.R., Ketchum K.A., Lee J., Ronning C.M., Koo H.L., Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L., Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M., Venter J.C. (1999) Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. Nature 402:761-768.
Lincoln S.M.D.E.L. (1993) Mapping genes controlling quantitative traits using MAPMAKER/QTL. Version 1.1. Whitehead Institute for Biomedical Research Technical Report, 2nd.
Lincoln S.M.D.E.L. (1993) Constructing genetic linkage maps with MAPMAKER/EXP. Version 3.0. Whitehead Institute for Biomedical Research Technical Report, 3rd.
Lionneton E., Aubert G., Ochatt S., Merah O. (2004) Genetic analysis of agronomic and quality traits in mustard (*Brassica juncea*). Theoretical and Applied Genetics 109:792-799. DOI: 10.1007/s00122-004-1682-0.
Lionneton E., Ravera S., Sanchez L., Aubert G., Delourme R., Ochatt S. (2002) Development of an AFLP-based linkage map and localization of QTLs for seed fatty acid content in condiment mustard (*Brassica juncea*). Genome 45:1203-1215. DOI: 10.1139/g02-095.
Liu A.-h., Wang J.-b. (2006) Genomic Evolution of Brassica Allopolyploids Revealed by ISSR Marker. Genetic Resources and Crop Evolution 53:603-611. DOI: 10.1007/s10722-004-2951-0.
Liu B.-H. (1998) Statistical genomics: linkage, mapping, and QTL analysis. 1 ed. CRC Press., Boca Raton.
Liu K.-S. (1984) Review-Preparation of Fatty Acid Methyl Esters for Gas-Chromatographic Analysis of Lipids in Biological Materials. Journal of American Oil Chemists' Society 71:1179-1187.
Lombard V., Delourme R. (2001) A consensus linkage map for rapeseed (*Brassica napus* L.): Construction and integration of three individual maps from DH populations. Theoretical and Applied Genetics 103:491-507.
Love H., Rakow G., Raney J., Downey R. (1991) Development of low glucosinolate mustard. Canadian Journal of Plant Science 71:419-424.
Lowe A.J., Moule C., Trick M., Edwards K.J. (2004) Efficient large-scale development of microsatellites for marker and mapping applications in Brassica crop species. Theoretical and Applied Genetics 108:1103-1112.
Lysak M.A. (2005) Chromosome triplication found across the tribe Brassiceae. Genome Research 15:516-525. DOI: 10.1101/gr.3531105.
Lysak M.A., Lexer C. (2006) Towards the era of comparative evolutionary genomics in Brassicaceae. Plant Systematics and Evolution 259:175-198. DOI: 10.1007/s00606-006-0418-9.
Mahmood T., Ekuere U., Yeh F., Good A.G., Stringam G.R. (2003) RFLP linkage analysis and mapping genes controlling the fatty acid profile of *Brassica juncea* using reciprocal DH

populations. TAG Theoretical and Applied Genetics 107:283-290. DOI: 10.1007/s00122-003-1244-x.
Mahmood T., Rahman M.H., Stringam G.R., Yeh F., Good A. (2005a) Molecular markers for yield components in *Brassica juncea* - do these assist in breeding for high seed yield? Euphytica 144:157-167. DOI: 10.1007/s10681-005-5339-0.
Mahmood T., Rahman M.H., Stringam G.R., Yeh F., Good A.G. (2006a) Quantitative trait loci for early maturity and their potential in breeding for earliness in *Brassica juncea*. Euphytica 154:101-111. DOI: 10.1007/s10681-006-9276-3.
Mahmood T., Rahman M.H., Stringam G.R., Yeh F., Good A.G. (2006b) Identification of quantitative trait loci (QTL) for oil and protein contents and their relationships with other seed quality traits in *Brassica juncea*. Theoretical and Applied Genetics 113:1211-1220. DOI: 10.1007/s00122-006-0376-1.
Mendoza M.S., Dubreucq B., Baud S., Parcy F., Caboche M., Lepiniec L. (2008) Deciphering gene regulatory networks that control seed development and maturity in *Arabidopsis*. Plant Journal 50:825-838.
Morgante M., Olivieri A.M. (1993) PCR-amplified microsatellites as markers in plant genetics. The Plant Journal 3:175-182. DOI: 10.1046/j.1365-313X.1993.t01-9-00999.x.
Morin P., Smith D. (1995) Nonradioactive detection of hypervariable simple sequence repeats in short polyacrylamide gels. Biotechniques 19:223-228.
Mu J., Tan H., Zheng Q., Fu F., Liang Y., Zhang J., Yang X., Wang T., Chong K., Wang X.J., Zuo J. (2008) LEAFY COTYLEDON1 Is a Key Regulator of Fatty Acid Biosynthesis in Arabidopsis. Plant Physiology 148:1042-1054. DOI: 10.1104/pp.108.126342.
Murphy D.J., Vance J. (1999) Mechanisms of lipid body formation. Trends in Biochemical Sciences 24:109-115.
Nelson J.C. (1997) QGENE: software for marker-based genomic analysis and breeding. Molecular Breeding 3:239-245. DOI: 10.1023/a:1009604312050.
Oetting W.S., Lee H.K., Flanders D.J., Wiesner G.L., Sellers T.A., King R.A. (1995) Linkage Analysis with Multiplexed Short Tandem Repeat Polymorphisms Using Infrared Fluorescence and M13 Tailed Primers. Genomics 30:450-458. DOI: 10.1006/geno.1995.1264.
Ohlrogge J., Browse J. (1995) Lipid Biosynthesis. The Plant Cell 7:957-970.
Ohlrogge J.B., Jaworski J.G. (1997) REGULATION OF FATTY ACID SYNTHESIS. Annual Review of Plant Physiology and Plant Molecular Biology 48:109-136. DOI: doi:10.1146/annurev.arplant.48.1.109.
Okuley J., Lightner J., Feldmann K., Yadav N., Lark E., Browse J. (1994) Arabidopsis *FAD2* Gene Encodes the Enzyme That Is Essential for Polyunsaturated Lipid Synthesis. The Plant Cell 6:147-158. DOI: http://dx.doi.org/10.1105/tpc.6.1.147
Osorio J., Fernández-Martínez J., Mancha M., Garcés R. Mutant Sunflowers with High Concentration of Saturated Fatty Acids in the Oil. Crop Sci. 35:739-742. DOI: 10.2135/cropsci1995.0011183X003500030016x.
Palmer J.D., Shields C.R., Cohen D.B., Orton T.J. (1983) Chloroplast DNA evolution and the origin of amphidiploid Brassica species. Theoretical and Applied Genetics 65:181-189. DOI: 10.1007/bf00308062.
Panjabi P., Jagannath A., Bisht N.C., Padmaja K.L., Sharma S., Gupta V., Pradhan A.K., Pental D. (2008) Comparative mapping of *Brassica juncea* and *Arabidopsis thaliana* using Intron Polymorphism (IP) markers: homoeologous relationships, diversification and

evolution of the A, B and C Brassica genomes. Bmc Genomics 9. DOI: 11310.1186/1471-2164-9-113.
Parkin I.A.P., Sharpe A.G., Lydiate D.J. (2003) Patterns of genome duplication within the Brassica napus genome. Genome 46:291.
Parkin I.A.P., Gulden S.M., Sharpe A.G., Lukens L., Trick M., Osborn T.C., Lydiate D.J. (2005) Segmental structure of the Brassica napus genome based on comparative analysis with Arabidopsis thaliana. Genetics 171:765-781. DOI: 10.1534/genetics.105.042093.
Pearson C.E., Edamura K.N., Cleary J.D. (2005) Repeat instability: mechanisms of dynamic mutations. Nat Rev Genet 6:729-742. DOI: http://www.nature.com/nrg/journal/v6/n10/suppinfo/nrg1689_S1.html.
Peng Q., Hu Y., Wei R., Zhang Y., Guan C., Ruan Y., Liu C. (2010) Simultaneous silencing of FAD2; and FAE1; genes affects both oleic acid and erucic acid contents in Brassica napus; seeds. Plant Cell Reports 29:317-325. DOI: 10.1007/s00299-010-0823-y.
Pérez-Vich B., Leon A., Grondona M., Velasco L., Fernández-Martínez J. (2006) Molecular analysis of the high stearic acid content in sunflower mutant CAS-14. TAG Theoretical and Applied Genetics 112:867-875. DOI: 10.1007/s00122-005-0188-8.
Pidkowich M.S., Nguyen H.T., Heilmann I., Ischebeck T., Shanklin J. (2007) Modulating seed Beta-ketoacyl-acyl carrier protein synthase II level converts the composition of a temperate seed oil to that of a palm-like tropical oil., Proceedings of the National Academy of Science, USA. pp. 4742-4747.
Pilet M.L., Duplan G., Archipiano H., Barret P., Baron C., Horvais R., Tanguy X., Lucas O., Renard M., Delourme R. (2001) Stability of QTL for field resistance to blackleg across two genetic backgrounds in oilseed rape. Crop Science 41:197-205.
Piquemal J., Cinquin E., Couton F., Rondeau C., Seignoret E., doucet I., Perret D., Villeger M.J., Vincourt P., Blanchard P. (2005) Construction of an oilseed rape (Brassica napus L.) genetic map with SSR markers. Theoretical and Applied Genetics 111:1514-1523.
Plieske J., Struss D. (2001) STS markers linked to Phoma resistance genes of the Brassica B-genome revealed sequence homology between Brassica nigra and Brassica napus. Theoretical and Applied Genetics 102:483-488.
Prabhu K.V., Somers D.J., Rakow G., Gugel R.K. (1998) Molecular markers linked to white rust resistance in mustard Brassica juncea. TAG Theoretical and Applied Genetics 97:865-870. DOI: 10.1007/s001220050966.
Pradash S., Hinata, K. (1980) Taxonomy, Cytogenetics and Origin of Crop Brassicas - A review. Opera Botanica:1-57.
Pradhan, Gupta, Mukhopadhyay, Arumugam, Sodhi, Pental. (2003) A high-density linkage map in Brassica juncea (Indian mustard) using AFLP and RFLP markers. TAG Theoretical and Applied Genetics 106:607-614. DOI: 10.1007/s00122-002-1083-1.
Pritchard F.M., Eagles H.A., Norton R.M., Salisbury P.A., Nicolas M. (2000) Environmental effects on seed composition of Victorian canola. Australian Journal of Agricultural Research 40:679.
Qiu D., Morgan C., Shi J., Long Y., Liu J., Li R., Zhuang X., Wang Y., Tan X., Dietrich E., Weihmann T., Everett C., Vanstraelen S., Beckett P., Fraser F., Trick M., Barnes S., Wilmer J., Schmidt R., Li J., Li D., Meng J., Bancroft I. (2006) A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. Theoretical and Applied Genetics 114:67-80. DOI: 10.1007/s00122-006-0411-2.

Quijada P.A., Udall J.A., Lambert B., Osborn T.C. (2006) Quantitative trait analysis of seed yield and other complex traits in hybrid spring rapeseed (*Brassica napus* L.): 1. Identification of genomic regions from winter germplasm. Theoretical and Applied Genetics 113:549.

Quiros C.F., Hu J., This P., Chevre A.M., Delseny M. (1991) Development and chromosomal localization of genome-specific markers by polymerase chain reaction in Brassica. Theoretical and Applied Genetics 82:627-632.

Rahman M., McVetty P.B.E., Li G. (2007) Development of SRAP, SNP and Multiplexed SCAR molecular markers for the major seed coat color gene in *Brassica rapa* L. Theoretical and Applied Genetics 115:1101-1107. DOI: 10.1007/s00122-007-0636-8.

Rahman M., Sun Z., McVetty P.B.E., Li G. (2008) High throughput genome-specific and gene-specific molecular markers for erucic acid genes in *Brassica napus* (L.) for marker-assisted selection in plant breeding. Theoretical and Applied Genetics 117:895-904. DOI: 10.1007/s00122-008-0829-9.

Rajcan I., Kasha K.J., Kott L.S., Beversdorf W.D. (1999) Detection of molecular markers associated with linolenic and erucic acid levels in spring rapeseed (*Brassica napus* L.). Euphytica 105:173-181.

Rakow G., Woods. D. (1987) Outcrossing in rape and mustard under Saskatchewan prairie conditions. Canadian Journal of Plant Science 67:147-151.

Ramchiary N., Bisht N.C., Gupta V., Mukhopadhyay A., Arumugam N., Sodhi Y.S., Pental D., Pradhan A.K. (2007a) QTL analysis reveals context-dependent loci for seed glucosinolate trait in the oilseed *Brassica juncea*: importance of recurrent selection backcross scheme for the identification of 'true' QTL. Theoretical and Applied Genetics 116:77-85. DOI: 10.1007/s00122-007-0648-4.

Ramchiary N., Padmaja K.L., Sharma S., Gupta V., Sodhi Y.S., Mukhopadhyay A., Arumugam N., Pental D., Pradhan A.K. (2007b) Mapping of yield influencing QTL in *Brassica juncea*: implications for breeding of a major oilseed crop of dryland areas. Theoretical and Applied Genetics 115:807-817. DOI: 10.1007/s00122-007-0610-5.

Ren Z.L., Lelley T., Robbelen G. (1990) The Use of Monosomic Rye Addition Lines for Transferring Rye Chromatin into Bread Wheat. I. The Occurrence of Translocations. Plant Breeding 105:257-264. DOI: 10.1111/j.1439-0523.1990.tb01283.x.

Robenek H., Hofnagel O., Buers I., Robenek M.J., Troyer D., Severs N.J. (2006) Adiphphilin-enriched domains in the ER membrane are sites of lipid droplets biogenesis. Journal of Cell Science 119:4215-4224.

Sakova L., Paclik R., Curn V. (1995) The drought tolerance of four Brassica species. Sbornik Jihoceska Univerzita Zemedelska Fakulta Ceske Budejovice Fytotechnicka Rada 12:77-86.

Saskatchwan Agriculture. (2011) 2010 Specialty Crop Report, Saskatchwan Minister of Agriculture. pp. 23.

Scarth R., McVetty P. (1999) Designer oil canola — A reveiw of new food-grade *Brassica* oils with focus on high Oleic, low Linolenic types., 10th International Rapeseed Congress, The Regional Institute Inc., Canberra, Australia. pp. 100-117.

Scarth R., McVetty P.B.E., Rimmer S.R., Stefansson B.R. (1988) Stellar Low Linolenic-High Linoleic Acid Summer Rape. Canadian Journal of Plant Science 68:509-511.

Sharma R., Mohapatra T., Mukherjee A.K., Krishanpal, Sharma R.P. (1999) Molecular markers for seed oil content in Indian mustard. Journal of Plant Biochemistry and Biotechnology 8:99-102.
Sharma R., Aggarwal R.A.K., Kumar R., Mohapatra T., Sharma R.P. (2002) Construction of an RAPD linkage map and localization of QTLs for oleic acid level using recombinant inbreds in mustard (*Brassica juncea*). Genome 45:467-472.
Sharpe A.G., Parkin I.A.P., Keith D.J., Lydiate D.J. (1995) Frequent nonreciprocal translocations in the amphidiploid genome of oilseed rape (*Brassica napus*). Genome 38:1112-1121.
Shenk J.S., Westerhaus M.O. (1991) Population definition, sample selection and calibration procedures for near infrared reflectance spectroscopy. Crop Sci. 31:469-474.
Si P., Mailer R.J., Galwey N., 54:397–407 D.W.T.A.J.A.R. (2003) Influence of genotype and environment on oil and protein concentrations of canola (*Brassica napus* L.) grown across southern Australia. Australian Journal of Agricultural Research 54:397-407.
Sinha S., Jha J., Maiti M., Basu A., Mukhopadhyay U., Sen S. (2007) Metabolic engineering of fatty acid biosynthesis in Indian mustard (*Brassica juncea*) improves nutritional quality of seed oil. Plant Biotechnology Reports 1:185-197. DOI: 10.1007/s11816-007-0032-5.
Sivaraman I., Arumugam N., Sodhi Y.S., Gupta V., Mukhopadhyay A., Pradhan A.K., Burma P.K., Pental D. (2004) Development of high oleic and low linoleic acid transgenics in a zero erucic acid *Brassica juncea* L. (Indian mustard) line by antisense suppression of the fad2 gene. Molecular Breeding 13:365-375.
Skrypetz S. (2003) Mustard seed: Situation and Outlook In Agriculture and Agri-foods Canada Bi-weekly Bulletin. pp. 9.
Snowdon R., Lühs W., Friedt W. (2007) Oilseed Rape, in: C. Kole (Ed.), Oilseeds, Springer Berlin Heidelberg. pp. 55-114.
Somers D.J., Demmon G. (2002) Identification of repetitive, genome-specific probes in crucifer oilseed species. Genome 45:485-492. DOI: 10.1139/g02-006.
Song K., Lu P., Tang K., Osborn T.C. (1995) Rapid genome change in synthetic polyploids of Brassica and its implications for polyploid evolution. Proceedings of the National Academy of Sciences 92:7719-7723.
Song K.M., Osborn T.C., Williams P.H. (1988) Brassica taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs). II: Preliminary analysis of subspecies within *B. rapa* (syn. campestris) and *B. oleracea* Springer, Heidelberg, Allemagne.
Srivastava A., Gupta V., Pental D., Pradhan A.K. (2001) AFLP-based genetic diversity assessment amongst agronomically important natural and some newly synthesized lines of *Brassica juncea*. TAG Theoretical and Applied Genetics 102:193-199. DOI: 10.1007/s001220051635.
StatCan. (2011) Preliminary Estimates of Principle Field Crop Areas, in: S. Canada (Ed.), Field Crop Reporting Series.
StatCan. (2012) Stocks of Principle Field Crops at December 31, 2011, in: S. Canada (Ed.), Field Crop Reporting Series, Minister of Industry, Ottawa.
Stefansson B.R., Kondra Z.P. (1975) Tower Summer Rape. Canadian Journal of Plant Science 55:343-344.
Struss D., Bellin U., Roebbelen G. (1991) Development of B-Genome Chromosome Addition Lines of *Brassica napus* using Different Interspecific Brassica Hybrids. Plant Breeding 106:209.

Struss D., Plieske J., Quiros C.F. (1995) Evidence of intergenomic B-genome translocation in B. napus using molecular markers., Proceeding of the 9th International Rapeseed Congress, GCIRC, Cambridge, England. pp. 1125-1127.

Struss D., Quiros C.F., Plieske J., Roebbelen G. (1996) Construction of Brassica B genome synteny groups based on chromosomes extracted from three different sources by phenotypic, isozyme and molecular markers. Theoretical and Applied Genetics 93:1026-1032.

Stymne S., Stobart A.K. (1993) Triacylglycerol Biosynthesis, in: P. R. Shewry and K. Stobart (Eds.), Seed Storage Compounds: Biosynthesis, Interactions and Manipulations. pp. 96-114.

Sun Z., Wang Z., Tu J., Zhang J., Yu F., McVetty P.B.E., Li G. (2007) An ultradense genetic recombination map for *Brassica napus*, consisting of 13551 SRAP markers. Theoretical and Applied Genetics 114:1305-1317. DOI: 10.1007/s00122-006-0483-z.

Suwabe K., Morgan C., Bancroft I. (2008) Integration of Brassica A genome genetic linkage map between Brassica napus and B. rapa. Genome 51:169-176. DOI: 10.1139/g07-113.

Tang J., Baldwin S., Jacobs J., Linden C.G., Voorrips R., Leunissen J., van Eck H., Vosman B. (2008) Large-scale identification of polymorphic microsatellites using an in silico approach. BMC Bioinformatics 9:374.

Tanhuanpaa P.K., Vilkki J.P., Vilkki H.J. (1996) Mapping of a QTL for oleic acid concentration in spring turnip rape (*Brassica rapa* ssp. oleifera). Theoretical and Applied Genetics 92:952-956.

Tanksley S.D. (1993) Mapping Polygenes. Annual Review of Genetics 27:205-233.

Tautz S. (1994) Simple sequences. Current Opinion in Genetics and Development 4:832-837.

Taylor D.C., Katavic V., Zou J.T., MacKenzie S.L., Keller W.A., An J., Friesen W., Barton D.L., Pedersen K.K., E.M.Giblin, Y.Ge, M.Dauk, C.Sonntag, T.Luciw, D.Males. (2002) Field testing of transgenic cv. Hero transformed with a yeast sn-2 acyltransferase results in increased oil content, erucic acid content and seed yield. Molecular Breeding 8:317-322.

Thelen J.J., Ohlrogge J.B. (2002) Metabolic engineering of fatty acid biosynthesis in plants. Metabolic Engineering 4:12-21.

To A., Valon C., Savino G., Guilleminot J., Devic M., Giraudat J., Parcy F. (2006) A network of local and redundant gene regulators governs Arabidopsis seed maturation. The Plant Cell 18:1642-1651.

Truco M.J., Hu J., Sadowski J., Quiros C.F. (1996) Inter- and intra-genomic homology of the Brassica genomes: Implications for their origin and evolution. Theoretical and Applied Genetics 93:1225-1233.

U N. (1935) Genomic analysis in Brassica with special reference to the experimental formation of *B. napus* and its peculiar mode of fertilization. Japanese Journal of Genetics 7:784-794.

Udall J.A., Quijada P.A., Lambert B., Osborn T.C. (2006) Quantitative trait analysis of seed yield and other complex traits in hybrid spring rapeseed (*Brassica napus* L.): 2. Identification of alleles from unadapted germplasm. Theoretical and Applied Genetics 113:597-609.

Uppstrom B. (1995) Seed Chemistry, in: D. a. M. Kimber, D.I (Ed.), Brassica Oilseeds: Production and Utilization, CAB International, Wallingford. pp. 217-243.

Uzunova M., Ecke W., Weissleder K., Robbelen G. (1995) Mapping the genome of rapeseed (*Brassica napus* L.). I. Construction of an RFLP linkage map and localization of QTLs for seed glucosinolate content. Theoretical and Applied Genetics 90:194-204.
Uzunova M.I., Ecke W. (1999) Abundance, polymorphism and genetic mapping of microsatellites in oilseed rape (*Brassica napus* L.). Plant Breeding 118:323-326. DOI: 10.1046/j.1439-0523.1999.00371.x.
Van Ooien J.W., Boer M.P., Jansen R.C., Maliepaard C. (2002) MapQTL 4.0: Software For The Calculation Of QTL Positions On Genetic Maps. Plant Research Int.
Van Ooijen J.W.a.R.E.V. (2001) Joinmap3.0, Software for the calculation of genetic linkage maps. . Plant Research Int.
Varshney R., Graner A., Sorrells M. (2005) Genic microsatellite markers in plants: features and applications. Trends in Biotechnology 23:1.
Vergnaud G., Mariat D., Zoroastro M., Lauthier V. (1991) Detection of single and multiple polymorphic loci by synthetic tandem repeats of short oligonucleotides. Electorphorisis 12:134-140.
Vigeolas H., Waldeck P., Zank T., Geigenberger P. (2007) Increasing seed oil content in oil-seed rape (*Brassica napus* L.) by over-expression of a yeast glycerol-3-phosphate dehydrogenase under the control of a seed-specific promoter. Plant Biotechnology Journal 5:431-441. DOI: 10.1111/j.1467-7652.2007.00252.x.
Warner K., Orr P., Parrott L., Glynn M. (1994) Effects of Frying Oil Composition on Potato Chip Stability. Journal of the American Oil Chemists Society 71:1117-1121.
Warwick S.I., Black L.D. (1991) Molecular systematics of Brassica and allied genera (Subtribe Brassicinae, Brassiceae) - Chloroplast genome and cytodeme congruence. Theoretical and Applied Genetics 82:81-92.
Weber J.L., Wong C. (1993) Mutation of human short tandem repeats. Human Molecular Genetics 2:1123-1128. DOI: 10.1093/hmg/2.8.1123.
Windham W.R., Mertens D.R., FE. F.E.B. (1989) Protocol for NIRS calibration: sample selection and equation development and validation. In Near infrared reflectance spectroscopy (NIRS): Analysis of forage quality, in: G. C. Marten (Ed.), USDA Agriculture Handbook.
Woods D., Capcara J., Downey R. (1991) The potential of mustard (*Brassica juncea* (L.) Coss) as an edible oil crop on the Canadian Prairies. Canadian Journal of Plant Science 71:195-198.
Wright P.R., Morgan J.M., Jessop R.S., Cass A. (1995) Comparative adaptation of canola (*Brassica napus*) and Indian mustard (*B. juncea*) to soil water deficits: yield and yield components. Field Crops Reports 42:1-13.
Wu X.-m., Chen B.-y., Lu G., Wang H.-z., Xu K., Guizhan G., Song Y. (2009) Genetic diversity in oil and vegetable mustard (*Brassica juncea*) landraces revealed by SRAP markers. Genetic Resources and Crop Evolution 56:1011-1022. DOI: 10.1007/s10722-009-9420-8.
Xu A., Huang Z., Ma C., Xiao E., Zhang X., Tu J., Fu T. (2010) FAE1 sequence characteristics and its relationship with erucic acid content in *Brassica juncea*. Acta Agronomica Sinica 36:794-800. DOI: 10.3724/sp.j.1006.2010.00794.
Yu S., Derr J., Etherton T., Kris-Etherton P. (1995) Plasma cholesterol-predictive equations demonstrate that stearic acid is neutral and monounsaturated fatty acids are hypocholesterolemic. The American Journal of Clinical Nutrition 61:1129-1139.

Zeng Z.-B. (1993) Theoretical Basis for Separation of Multiple Linked Gene Effects in Mapping Quantitative Trait Loci. Proceedings of the National Academy of Sciences of the United States of America 90:10972-10976.
Zeng Z.B. (1994) Precision Mapping of Quantitative Trait Loci. Genetics 136:1457-1468.
Zhao J., Meng J. (2003) Genetic analysis of loci associated with partial resistance to Sclerotinia sclerotiorum in rapeseed (*Brassica napus* L.). Theoretical Applied Genetics 106:759-764.
Zhao J., Becker H.C., Zhang D., Zhang Y., Ecke W. (2005) Oil content in a European x Chinese rapeseed population: QTL with additive and epistatic effects and their genotype-environment interactions. Crop Science 45:51-59.
Zhao J., Becker H.C., Zhang D., Zhang Y., Ecke W. (2006) Conditional QTL mapping of oil content in rapeseed with respect to protein content and traits related to plant development and grain yield. Theoretical and Applied Genetics 113:33-38. DOI: 10.1007/s00122-006-0267-5.
Zhao J., Dimov Z., Becker H.C., Ecke W., Moellers C. (2008) Mapping QTL controlling fatty acid composition in a doubled haploid rapeseed population segregating for oil content. Molecular Breeding 21:115.
Zou J., Katavic V., Giblin E.M., Barton D.L., MacKenzie S.L., Keller W.A., Hu X., Taylor D.C. (1997) Modification of seed oil content and acyl compostion in the Brassicaceae by expression of an yeast sn-2 acyltransferase gene. . Plant Cell 9:909-923.