# Development of New Molecular Markers and Quantitative Trait Loci for Oil Content of *Brassica napus* L.

by

WIN JIAN CHAN

A Thesis

Submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

Department of Plant Science

University of Manitoba

Winnipeg, Manitoba

## ABSTRACT

**Chan, Win Jian. M.Sc. The University of Manitoba, June 2020.**

**Development of New Molecular Markers and Quantitative Trait Loci for Oil Content of *Brassica napus* L.**

*Brassica napus* L. from Brassicaceae consisting AACC genome is one of the most important oil crops that contributes to the Canadian economy. The oil produced by *B. napus* can be made into high profit oil-related products such as cooking oil and machinery lubricant, and it can be classified as either rapeseed or canola depending on the amount of erucic acid and glucosinolates in the seed. Numerous studies on mapping quantitative trait loci (QTL) associated with the seed oil content (SOC) in *B. napus* have been conducted. Studies have identifed the location of oil-related genes such as *BnFAE* genes located on chromosomes A8 and C3 which are responsible for the production of erucic acid in the seed. One doubled haploid (DH) population ZT (Topas x Zhongyou821) consisting of 94 DH genotypes was used to collect SOC data in field experiments and under controlled environmental conditions in the greenhouse. A new molecular marker detection method that includes PCR-based sequencing library preparation and Illumina sequencing detection (PISD) was used to identify single nucleotide polymorphisms (SNPs). Two genetic maps, named PISD and combined (standard genotyping-by-sequencing, GBS), were created and 1,092 and 1,821 SNPs molecular markers were used to identify QTL for SOC in the *B. napus* DH ZT genotypes. In total, six and 8 major QTL were identified using the two genetic maps respectively. In addition, QTL for SOC on chromosomes A8 and C3 were identified as being closely linked to *BnFAE1.1* and *BnFAE1.2*, which are responsible for the production of erucic acid. These results show a better understanding of gene control and can be useful for *B. napus* breeding industries for erucic acid content in the future.

## ACKNOWLEDGEMENTS

I would like to thank Dr. Genyi Li for being patient and supportive, and for his excellent advice and guidance for the completion of my project and thesis writing.

I would love to thank my advisory committee members Dr. Robert Duncan from the Department of Plant Science, University of Manitoba for providing me the opportunity to learn about canola from a canola breeder's perspective, Dr. Dana Schroeder from the Department of Biological Sciences, University of Manitoba for providing me a lot of good advice during the study, and Dr. Frank You from Agriculture and Agri-Food Canada (AAFC). I would like to give my gratitude to Dr. Yanglong Mou who provided me bioinformatics knowledge on how to deal with genetic data and Dr. Kede Liu who helped me to create a PCR-based Illumina SNP detection (PISD) genotypic map.

I acknowledge the help and guidance from my lab technician Dr. Yaping Wang, research associate Dr. Yongxia Wang, Ph.D. candidates Keval Shah, Kenny So, Yang Lin, Jia Sun, and Master student An Bo; without them, I would not be able to complete this project.

I am also truly grateful to the greenhouse technicians Mr. Rob Visser and Ms. Cathy Bay who took care of the plant material for me and Ms. Judith Nugent-Rigby and Mr. Lynn Coleman for helping me conduct my field experiments.

I would like to acknowledge the financial support from The GETS fund of the University of Manitoba as well as an NSERC Discovery grant and CRD grants awarded to Dr. Genyi Li.

Finally, I thank my family and my girlfriend Hannah Froese for giving me unconditional support during my studies.

**TABLE OF CONTENTS**

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| AFLP | Amplified fragment length polymorphism |
| DH | Doubled haploid |
| GBS | Genotyping-by-sequencing |
| GWAS | Genome-wide association study |
| LOD | Logarithm of the odd |
| MAS | Marker-assisted selection |
| NIR | Near-infrared spectroscopy |
| PISD | PCR-based Illumina sequencing detection |
| QTL | Quantitative trait locus/loci |
| RE | Restriction enzyme |
| RFLP | Restriction fragment length polymorphism |
| SNP | Single nucleotide polymorphism |
| SOC | Seed oil content |
| SRAP | Sequence-related amplified polymorphism |
| SSR | Simple-sequence repeat |

## Chapter 1.0 GENERAL INTRODUCTION

One of the most important oil crops in the world is *Brassica napus* L. It is cultivated in almost every part of the world (Leff et al., 2004; Kramer, 2012; Friedt et al., 2018). In Canada, although there was a slight decline in *B. napus* production from 21.3 million tonnes in 2017 to 20.3 million tonnes in 2018 due to a decrease in average bushels per acre, the demand for *B. napus* and oil-related products remains high (Statistics Canada, 2019). Compared to other countries, Canada has the highest production of *B. napus*, with China in the second place producing 13 million tonnes and India producing 7.9 million tonnes in 2018 (FAO, 2019; Sawe, 2019).

On a global scale, *B. napus* oil provides 13% of the vegetable oil supply in 2018, making it one of the most produced oil crops, behind oil palm and soybean (Murphy, 2014; FAO, 2019). *Brassica napus* oil is useful because of its heat stability, oxidative stability, and several positive nutritional aspects (Eskin et al., 2005). With the good heat stability of *B. napus* oil, the seed can be crushed to produce oil-related products, such as cooking oil and machinery lubricant (Saeidnia and Gohari, 2012). Its use is dependent on the amount of erucic acid the seed contains. Oil with high erucic acid content is used for machinery lubricant, while *B. napus* with little or no erucic acid (<1%) is used for products such as cooking oil (Kramer, 2012). High erucic acid oil can also be made into bio-diesel, and with the depletion of non-renewable resources such as fossil fuels, bio-diesel made from *B. napus* may be useful in the future (Issariyakul et al., 2008). Öztürk (2015) compared the performance, emission, and combustion of fossil fuels with biodiesel made from canola oil-hazelnut soapstock. The carbon monoxide emissions were similar while the nitrous oxide emissions were observed to be lower in canola oil-hazelnut soapstock. Thus, they concluded that the biodiesel from canola oil is slightly better than the fossil fuel.

The commercialization of *Brassica napus* oil took place in the early 19th century in Canada (Raymer, 2002). In the past 40 years, private sector investment in *B. napus* research has increased, and the oil crop has been extensively studied (Heisey et al., 2002; Cowling, 2007). Extensive research work has contributed to the development of new *B. napus* cultivars that are tolerant to drought, high temperatures, herbicides, diseases, and pests, and all of the improved traits have a positive correlation with the seed oil production of *B. napus* (Wan et al., 2009; Salisbury and Barbetti, 2011; Zhou et al., 2014). In addition, genomic, transcriptomic, proteomic, and metabolomic studies have been conducted to improve the traits in *B. napus* (Snowdon and Iniguez, 2012). One particular phenotypic trait that interests researchers is the seed oil content (SOC) in *B. napus* which links to quantitative trait locus (QTL) studies (Heisey et al., 2002).

Mapping quantitative trait loci assists in the identification of the location of genes controlling phenotypic traits in *B. napus* based on the distance in centimorgan (cM) from specific molecular markers (Chen et al., 2010). A single nucleotide polymorphism (SNP) is a change of one or more nucleotides in DNA to create genetic variation in the genotype (Qu et al., 2017).

A large number of genetic map studies have been performed on SOC in *B. napus*, and many have identified the QTL for SOC (Ecke et al., 1995; Thormann et al., 1996; Chen et al., 2010). For example, Zhao et al., (2006) identified QTL with a strong relationship between the SOC and protein content. They have identified six major QTL, and two out of six QTL were linked to oil-related genes using 282 $F_1$ genotypes from the German cultivar Sollux and Chinese cultivar Gaoyou. In the study, they concluded that the SOC of *B. napus* had a genetic relationship with the number of seeds per silique but no evidence was found for SOC with flowering time and seed weight. A study by Chen et al., (2010) found that the SOC of *B. napus* is affected by the environment, seed colour, and erucic acid content by using SRAP and SSR markers. Six SOC experiments were evaluated based on location and year and 19 genetic

groups were identified, covering 1,868 cM of the *B. napus* genome. Additionally, a study by Javed et al., (2016) using a doubled haploid population derived from Polo and Topas had identified 14 QTL that were associated to SOC, and 131 QTL were linked to six fatty acids.

Quantitative trait locus mapping for SOC in this study aims to build upon previous research by using new SNP markers for the ZT doubled haploid (DH) population. The objectives of this experiment are: to create two genetic maps from PCR-based Ilumina SNP detection (PISD) and standard GBS method using SNP markers and 94 ZT DH genotypes; and to identifiy QTL for SOC in *B. napus*.

# Chapter 2.0 LITERATURE REVIEW

## 2.1 *Brassica napus* L.

Ancient Sanskrit records suggest *Brassica nigra* L. (black mustard), which belongs to Brassicaceae like *Brassica napus* L. (oilseed rape), was first cultivated in India as early as 2000 B.C. (Kramer, 2012; Canola Council of Canada, 2019). In contrast, *B. napus* was only introduced into China, Canada, Europe, America, Japan, and Korea for oil production in the early 20th century (Nagaharu, 1935). *Brassica napus* is classified into two different groups called rapeseed and canola, depending on the erucic acid and glucosinolate content (Ghazani and Marangoni, 2013).

Rapeseed was produced in northern latitudes such as Europe as early as the 13th century because rapeseed generated profitable oil yields (Friedt et al., 2018). In the 19th and 20th centuries, rapeseed oil was mainly used for cooking and fuel for lighting lamps (Appelqvist and Loof, 1972; Canola Council of Canada, 2019). Rapeseed oil was superior to mineral oil as a lubricant due to its poor oxidation property and the fact that it came from a renewable resource (Sharma et al., 2015). The ability of rapeseed oil to combine with water or steam to cling to metal surfaces made rapeseed oil a useful marine engine lubricant. Rapeseed oil reached a critical shortage for the production of marine engine and merchant ship lubricants during World War II in the early 1940s because a significant proportion of rapeseed oil was diverted to Germany for cooking (Ramadhas et al., 2004; Kramer, 2012). The discovery of rapeseed oil use as a lubricant led to the first commercial Canadian rapeseed processing plant in Hamilton, Ontario in 1943. It was operated by J. Gordon Ross under the name Prairie Vegetable Oils Ltd (Kramer, 2012; Canola Council of Canada, 2019).

The selection of a low erucic acid ‘Oro’ cultivar derived from a cross between a rapeseed variety ‘Nugget’ and a forage cultivar ‘Liho’ in 1968 and low glucosinolates content ‘Bronowski’ had led to the creation of the first ‘double low’ canola variety, ‘Tower’, in 1974 (Juska et al., 1997; Bell, 1982). Subsequently, in 1977, the first ‘double low’ Polish variety

was developed and registered as 'Candle' by AAFC Saskatoon, after which an Argentine variety called Regent was developed at University of Manitoba (Stefansson et al., 1961; Seyis 2013; Canola Council of Canada, 2019). Canola was registered as a trademark in Canada back in 1978 (Raymer, 2002). It was derived from the words "Canada" and "ola", which means Canada's oil (Burton et al., 2004). Canola was developed by Canadian scientists B. R. Stefansson, F. W. Hougen, and R. K. Downey via conventional plant breeding (Eskin et al., 2005). The term "canola" is used when the oil has a low erucic acid content, which is considerably healthier for human consumption (Laryea et al., 1992), and the canola has a low glucosinolate content, which is good for animal consumption (Mithen et al., 2000). Erucic acid content in other economically important *Brassica* species, such as *B. juncea*, varies from 21 to 47%, while *B. rapa* ranges from 22 to 55% (Stefansson et al., 1961; Canola Council of Canada, 2019). Current canola oil has a standard reqirement that the erucic acid and glucosinolate content is less than 0.1% and 8.5 $\mu molg^{-1}$ respectively, according to Ghazani and Marangoni (2013). The erucic acid content of unimproved *B. napus* oil is between 28 and 48% of the total seed oil content (SOC), while the equivalent content in canola is less than 1% (Canola Council of Canada, 2019).

*Brassica napus* is an allotetraploid species, and its relationship with diploid *Brassica* species was first summarized using U's Triangle in the early 20th century, based on cytogenetic studies (Figure 2.1; Nagaharu, 1935). The triangle is made up of the three diploid species, *B. rapa*, *B. oleracea*, and *B. nigra*, the genomes of which are represented as AA, BB, and CC respectively, while the three allotetraploid species *B. juncea*, *B. carinata*, and *B. napus* are positioned along the sides of the triangle between their respective progenitor diploid parents (Nagaharu, 1935). *Brassica napus* (AACC; n = 19) descended from an interspecific hybrid of *B. rapa* (AA, n = 10) and *B. oleracea* (CC; n = 9), which then went through whole-genome duplication (Nagaharu, 1935). The relationship between the diploids *B. rapa* and *B. oleracea*

and the allotetraploid *B. napus* was recently affirmed by whole-genome sequencing (Chalhoub et al., 2014).

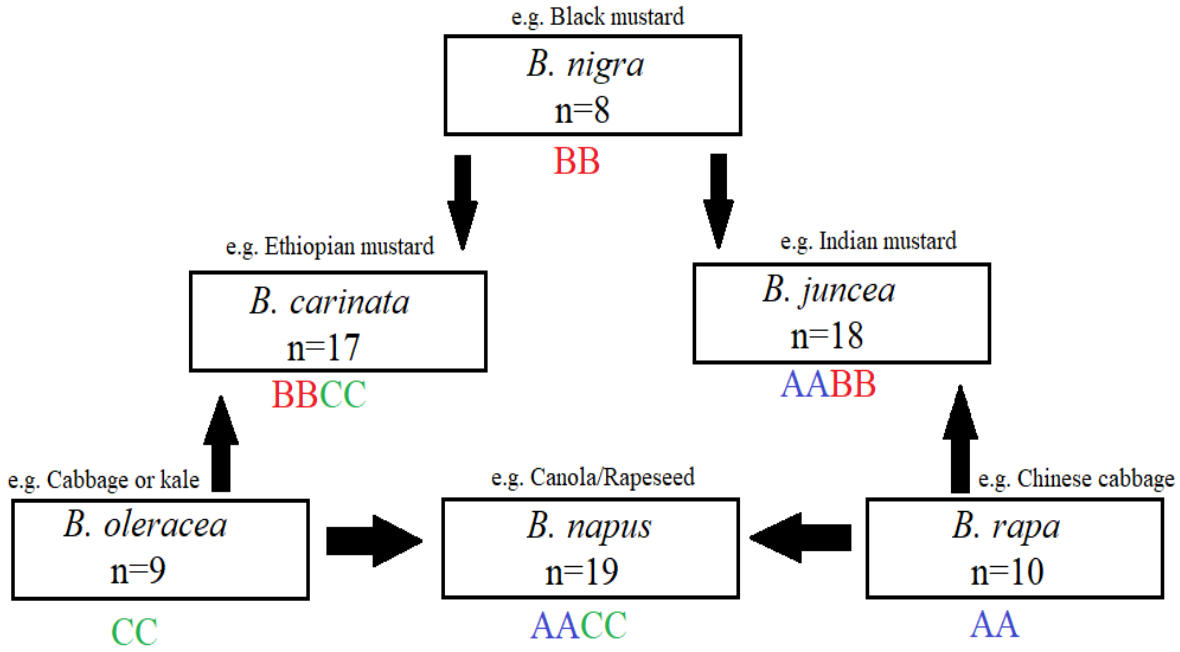


Figure 2.1**:** Relationship between the six *Brassica* species in a U's Triangle diagram and their common names (adapated from Nagaharu, 1935).

## 2.2 Characteristics of *Brassica napus*

*Brassica napus* is a self-pollinating plant (Williams et al., 1986). However, cross-pollination can occur with the help of insects such as flies, bees, butterflies, and other pollinators (Hayter and Cresswell, 2006; Hudewenz et al., 2014). This is why *B. napus* flowers are categorized as entomophilous, capable of self-pollination and cross-pollination (Kevan and Eisikowitch, 1990; Hayter and Cresswell, 2006). The cross-pollination process is important because the pollinators can help increase the seed yield in *B. napus*, and it is the major mechanism through which the genetic variations in *B. napus* and other important crops occur (Fenster, 1991). Even without the pollinator, however, shaking adjacent plant stems results in an 80% seed set, while there is only a 30% seed set without shaking (Hayter and Cresswell, 2006; Australian Government Office of the Gene Technology Regulator, 2008).

*Brassica napus* is categorized into spring-type (annual) genotypes, which can be grown on the Canadian prairies from spring to fall, producing flowers and seeds in the same year. The winter-type (biennial) varieties require vernalization at 4 to 10 degrees Celsius ($^{o}$C) for least two to three months before flowering and seed set (Gulden et al., 2008; Paridaen and Kirkegaard, 2015; Gacek et al., 2017). The mean time for plant maturity ranges from 76 to 140 days, depending on weather and location. In Winnipeg, the mean time for maturity is 96 days (Canola Council of Canada, 2019b). Approximately 45-55% of flowers develop into pods, and 75% of the pods develop within 14 days after flowering (Tayo and Morgan, 1975; Agriculture and Agri-Food Canada, 2019). In Canada, farmers on the prairies plant canola and rapeseed six weeks after the mean frost, which can be as early as late April or as late as the end of May, depending on the mean temperature and sowing date (Luo et al., 2018). Canola and rapeseed can be grown on any soil, but for better yield and oil content, a dark brown or black prairie soil called chernozemic, which contains approximately 4 to 16% humus in the topsoil (at least 10 cm thick), is preferred because it contains sources of macronutrients such as nitrogen, phosphorus, and other organic matter (CFIA, 2019; Agriculture and Agri-Food Canada, 2019). The optimum pH for rapeseed ranges from pH 5 to 9, preferably slightly alkaline well-drained soil (Krupinsky et al., 2002; Jabran et al., 2008; Farooq et al., 2011). Pods are susceptible to shattering at low moisture levels, so *B. napus* is swathed when the seed moisture is approximately 35% (Vera et al., 2007). A swath is an old process of harvesting rapeseed if yield is low to prevent the pods from breaking and the seeds from falling on the field if the *B. napus* does not tolerate shattering. After swathing, the cut plants are left to dry horizontally over the stubble and harvested after drying to minimize yield loss (Berglund et al., 2002).

Another distinct feature of *B. napus* is the presence of glucosinolates. The glucosinolates in *B. napus* have a pungent and distinctive flavour and are especially useful in plant defence against herbivores and fungi (Mithen et al., 2000; Talalay et al., 2001; Tierens et al., 2001; Shroff et al., 2008; Bednarek et al., 2009). China and India have a long history of

consuming “double high” rapeseed, which is produced from rapeseed with the high content of glucosinolates and erucic acid. The consumption of high erucic acid oil was found to induce haematological changes in ducklings (Abdellatif, 1972; Abdellatif et al., 1972), while two other studies by Deuel et al., (1949) and Vaisey et al., (1973) found that rapeseed oil can be easily digested by humans and other animals. A recent study by Cartea and Velasco (2008) found that glucosinolates in *B. napus* can be converted in the human body into isothiocyanate and indole products, which are chemically active compounds linked to the inhibition of cancer cell development. The therapeutic mechanism of these compounds is speculated to be the result of regulating enzymes, apoptosis, and blocking or stopping the cancer cell cycle. Glucosinolates contain a β-thioglucoside with a side chain and a β-D-glucopyranose moiety which can be hydrolyzed depending on requirements such as pH value, the amount of ferrous ion, substrate concentration, and presence of the protein epithiospecifier (Grubb and Abel, 2006; Cartea and Velasco, 2008). The gene responsible for biosynthesis pathway of glucosinolates were first identified in *Arabidopsis thaliana* L., *AtMYB28* (Lu et al., 2014; Liu et al., 2020). The oxidation of the aliphatic and aromatic amino acids by cytochrome P450 monooxygenase is the first step, followed by the secondary modification of the side chains (Mikkelsen et al., 2002; Wittstock and Halkier, 2002). However, a high level of glucosinolate consumption is toxic to insects, animals and humans (Mithen et al., 2000; Wittstock et al., 2004; Cornell University, 2019).

### 2.3 *Brassica napus* seed oil content

*Brassica napus* on average can yield an average of 42% SOC, which is a cumulative component due to the environment, pod formation, and seed production (Walton et al., 1999; Canola Council of Canada, 2019). *Brassica napus* stores the carbon and hydrogen energy in a form of triacyglycerols (TAGs) in the seed as reserve energy for supporting germinating at the early stage of growing (Baud and Lepiniec, 2010). The oil synthesis occurs in the seed, governed by enzymes such as acyl-coA synthetase (Ding et al., 2020) and diacylglycerol

cholinephosphotransferase (PDCT) (Bai et al., 2020). Oil synthesis gene *BnROD1* with two members on chromosomes A3 and C3 is strongly expressed during the oil synthesis of the seed, acting as the activator for PDCT (Bai et al., 2020). Many enzymes encoded by the oil systhesis genes play an important role in mediating the production of polyunsaturated fatty acids.

Triacyglycerols in *Brassica napus* oil consist of many types of fatty acids, including palmitate (C16:0), stearate (C18:0), oleic (C18:1), linolenic (18:2), α-linolenic (C18:3), arachidic (C20:0), eicosenoic (C20:1), and erucic acid (C22:1) (Rahman et al., 2008). Approximately 50-66% of fatty acids in *B. napus* oil is monounsaturated and 25-29%, polyunsaturated. The common names for polyunsaturated fatty acids C18:2 and C18:3 are omega-6 and omega-3 fatty acids (Ghazani and Marangoni, 2013). *Brassica napus* oil is considered one of the healthiest oils that can be consumed by humans and animals due to the high amount of polyunsaturated fat (Lin et al., 2013).

## 2.4 Quantifying *Brassica napus* seed oil content

### 2.4.1 Hexane extraction

One common SOC quantifying method that can be used in the laboratory is the hexane extraction or solvent extraction method. Hexane is a great solvent for *B. napus* oil and other plant oils because the boiling point of hexane is relatively low. The boiling point of hexane is 69 $^{o}$C wheareas unrefined *B. napus* oil is approximately 107 $^{o}$C (Lohani et al., 2015). The seed is dried in the oven to remove water content and crushed before dissolving with hexane. The *B. napus* oil is washed, dissolved in hexane, and finally the oil is recovered from boiling the mixture. The end product after the hexane distillates is the *B. napus* oil from the seed (Kumar et al., 2017). However, the solvent extraction method is less precise. A study by Lohani et al., (2015) showed that a sample for SOC in *B. napus* measurement was 21-36% under three levels of extraction temperature and time. Therefore, the hexane extraction method is sensitive toward

ranges of temperature and time and so this method is less likely than NIR to be used in the laboratory scale to measure SOC for precise and accurate measurement.

### 2.4.2 Near-infrared spectroscopy

Near-infrared spectroscopy (NIR or NIRS) is a non-destructive method to quantify SOC. (Veras et al., 2010; Mba et al., 2014). The *B. napus* seed sample is exposed to the NIR laser, the laser signal is detected, and the SOC from the sample is quantified (Daun et al., 1994). According to the American Oil Chemists' Society (AOCS), the whole seed NIR analysis for the SOC is highly precise for *B. napus* with a standard error of prediction less than 0.5% if the sample is dry (Daun et al., 1994). One benefit of using the NIR is that the *B. napus* seed is not destroyed in the process of analysis so the seed can be reused for further research. Additionally, a relatively small amount of oilseed is required (approximately 5 g) and NIR is cheaper and less time consuming in the laboratory scale (Mba et al., 2014).

## 2.5 Biotic and abiotic effects on *Brassica napus* oil

### 2.5.1 Biotic factors

*Brassica napus* SOC can be easily affected by biotic stresses. These biotic stresses include fungal diseases such as blackleg by *Leptosphaeria maculans* (Desmazieres) Cesati et de Notaris, 1863, Sclerotinia stem rot by *Sclerotinia sclerotiorum* (Libert) de Bary, 1837, insects and pests, and weeds, all of which compromise seed yield, SOC, and limit *B. napus* production.

#### 2.5.1.1 Fungi

One of the most important fungal diseases for *B. napus* is the blackleg disease which can be easily spread by infected neigbouring plants on the field (Gugel and Petrie, 1992). The pathogen infects the cotyledons, leaves, stems, and roots of *B. napus* (Fernando and Chen, 2003). The presence of pycnidia and symptoms such as necrosis of the the leaf tissue and stem

lesions can be observed, usually resulting in plant death due to the disease preventing the flow of nutrients and minerals (Zhou et al., 1999). Maximum yield losses can be up to 30-50% and this disease has been documented worldwide (Hwang et al., 2016).

Sclerotinia stem rot is another disease that causes approximately a 7-15% reduction of yield in *B. napus* (Del Rio et al., 2007). The pathogen targets plant flowers which can trasmit easily from plant to plant through wind and pollinators (Morrall and Dueck, 1982). Symptoms such as growth of white mycelium that affect the xylem of *B. napus* and cause stem necrosis can be observed (Gugel, 1986).

#### 2.5.1.2 Insects

In addition to fungal pathogens, *B. napus* is also susceptible to damage by insects such as *Schizaphis graminum* Rondani (aphids) (Buntin and Raymer, 1994) and *Psylliodes chrysocephala* L. Linnaeus, 1758 (flea beetles) (Dosdall and Stevenson, 2005). *Schizaphis graminum* feeds directly by injecting a needle-like elongated appendage called proboscis into the *B. napus* phloem vessel (Douglas, 2003). The prescence of *S. graminum* in different growth stages of *B. napus* results in different types of damage to the plant; the damage at the early stages to *B. napus* results in no flowering, and the plants are dead at the later stages of damage, causing a 20-30% yield reduction (Sarwar, 2013).

*Psylliodes chrysocephala* is one of the most common and important pests for *B. napus*. It can be observed throughout all growth stages but the pest does the most damage during the seedling stage of *B. napus* (Ferguson et al., 2006). *Psylliodes chrysocephala* feeds on the cotyledons, stems, and young leaf tissue of the plant which causes tiny (<3 mm) circular pits. Severe damages to the plant tissue by *P. chrysocephala* during the early growth stage of *B. napus* result in plant death while later stage damage reduces the seed size, weight and SOC of *B. napus* (Elliott et al., 2008).

#### 2.5.1.3 Weeds

Weeds are another major biotic stress that limits *B. napus* production in the field (Beckie et al., 2008). Interspecific (different plant species) and intraspecific (same plant species) competition can be easily identified in the field when the weed species emerge from the soil (Majnoun et al., 2006). A few examples of important weeds for *B. napus* that are found in Northern America are *Cirsium arvense* L. (Canada thistle) (Gulden et al., 2008), *Taeaxacum officinale* L. (common dandelion) (Froese and Van Acker, 2003), and *Galium aparine* L. (cleavers) (Malik and Born, 1988). These weeds lead to a weaker plant or plant death following resource competition, which decreases the overall SOC and production.

### 2.5.2 Abiotic factors

Abiotic factors in the environment play a major role in *B. napus* seed oil production and SOC. Similar to biotic factors, *B. napus* is sensitive to abiotic stresses such as water availability, temperature, light, macro- and micronutrients (Gates, 1968; Morrison and Steward, 2002).

#### 2.5.2.1 Water

Water in the plant system serves as an important role in transporting nutrients, mineral salts, and plant growth regulators (Park et al., 2017). When soil is lacking in water content, drought occurs. Drought has the most significant impact during the early flowering stage on *B. napus* (Elferjani and Soolanayakanahally, 2018). Without the transportation of water to the apical metistem of the plant, the plant was not be able to produce flowers, form pods, produce seeds or adequate SOC (Din et al., 2011; Park et al., 2017). An increase of electrons in the plant system during photosynthesis and drought leads to an increase of reactive oxygen species such as superoxide radical, hydrogen peroxide, nitric oxide and hydroxyl radical (Maliba et al., 2018). An increase of reactive oxygen species in the plant system causes oxidation of the protein cross-linkages and damages the nucleic acid which is important for synthesizing SOC

(Moghadam et al., 2011; Elferjani and Soolanayakanahally, 2018). A study by Moghadam et al., (2011) observed a decrease in SOC and an increase in glucosinolates of *B. napus* in response to water stress.

The role of water is important throughout the life cycle of *B. napus*. Without water, *B. napus* is not be able to germinate and grow; with too much water during senescence stage, the *B. napus* pod is not be able to dry up quickly for harvesting.

#### 2.5.2.2 Temperature

It is important to understand how the temperature affects the growing stages and SOC of *B.* napus. Similar to all plants, *B. napus* is sensitive toward exctreme change of temperature (Elferjani and Soolanayakanahally, 2018). High temperatures induce flower sterility, seed development, and fatty acid composition in *B. napus* (Polowick and Sawhney, 1988; Angadi et al., 2000). A constant 38 $^{o}$C for 5 days during seed development results in a 52% reduction of the total yield and SOC due to loss in seed weight (Aksouh-Harradj et al., 2006). A study by Morrison and Steward (2002) observed that temperatures above 32 $^{o}$C over a seven-day period during flowering negatively affects the development of flowers, seed set, and pod formation. The lower the seed set, the lower the yield and SOC. Another study by Baux et al., (2008) observed how *B. napus* performs under a minimal temperature threshold of 16 $^{o}$C for 60 days. The study found that fatty acids such as oleic acid (C18:1) were negatively influenced by low temperatures but no correlation was observed between low temperatures and linoleic acid (18:2).

#### 2.5.2.3 Light

Other than water and temperature, light is one of the most important environmental factors that supports plant growth and life (Bohr, 1933). Many factors, such as the distance of the light source, type of light, and duration of exposure, affect plant growth and yield. *Brassica napus* grows best in the wavelengths between the red and the far-red spectrum (600 to 800 nm), and between the red and the blue spectrum (400-500 nm) (Hart, 2012; Rondanini et al., 2014).

Photosynthesis takes place in *B. napus* during the cotyledon stage until the end of the plant life cycle. Lengthy exposure of intense light results in wilting of the plant or even death, due to the plant excessive reaction to oxygen (Zhu, 2016). Conversely, poor or low light intensity results in yellowing of the leaf and lower yields. One study by Shengxin et al., (2016) found that different plant species require different amounts of light spectra. The chlorophyll content of *B. napus* under 100% red light was the lowest, while balancing the ratio between red and blue as well as between the red and far-red showed higher chlorophyll, sucrose, and starch content, signifying healthy plant growth. Without light, the process of photosysnthesis in *B. napus* does not occur which leads to premature death and a reduction in the overall yield.

#### 2.5.2.4 Micro- and macronutients

In addition to temperature, water and light stresses, nutrients in the soil play major roles in supporting the growth of *B. napus* and subsequent seed oil production (Hua et al., 2012). Essential macronutrients and micronutrients for *B. napus* are carbon, nitrogen, phosphorus, potassium, sulphur, iron, manganese, boron, and many others. In *B. napus,* deficiency symptoms such as chlorosis of the leaf and stem can be easily identified (Uchida, 2000). If macronutrients such as nitrogen (N), phosphorus (P), potassium (K), and sulphur are limited in the soil, the germination to senescence process takes more than 96 days (Jackson, 2000; Cheema et al., 2001; Brennan and Bolland, 2007; Canola Council of Canada, 2019). Nitrogen, P, and K are the most important macronutrients during the *B. napus* flowering stage because these three macronutrients increase the rate of cell division and cell differentiation in the meristematic tissue of the plant (MacAdam, 1989; Hua et al., 2012). Heavy metals such as aluminium in the soil inhibit *B. napus* root formation. Clune and Copeland (1999) found that in soil with more than 60 µmol of aluminium compound, root formation in *B. napus* is inhibited and cellular damage is observed at the root apical meristem and result in plant death. Therefore, fertlization with macro and micro-nutrients during the germination and flowering stages of *B. napus* is essential for better plant health and development.

#### 2.5.2.5 Microclimate

A microclimate is an area or a climate zone in which the temperature, water, light, micro- and macronutrients, pollinators, diseases, and pests differ from the surrounding environment (Ma et al., 2019). This uneven distribution of biotic and abiotic factors affects plant development which can result in a change in SOC of *B. napus*. A study by Veromann et al., (2013) had observed the effect of a difference in macronutrients on *B. napus* growth development. An uneven distribution of macronutrients on the field or greenhouse may greatly affect plant development which results in either a lower or higher SOC of *B. napus* and affects the result during oil quantification. In contrast, optimum microclimate conditions create a better environment to allow the *B. napus* to grow. A study by Omidi et al., (2010) observed an increase in absorbtion of crop residue from the previous cropping year such as stem, leaves, and roots from the soil as nutrients by *B. napus*, resulting in healthier plants, improved yield and SOC. *Brassica napus* yield and SOC are both complex phenotypic traits that are affected by the environment and governed by many genes. In conclusion, abiotic stress critically impacts seed oil production of *B. napus*.

## 2.6 Genetic control of seed oil content in *Brassica napus*

Since the oil produced from *B. napus* is economically and agriculturally useful, understanding the genetic control of oil production is beneficial and can improve SOC in the seed through traditional and modern breeding techniques. Seed oil content in *B. napus* is a quantitative trait governed by many genes, each with a small effect on the trait (Downey and Craig, 1964; Gacek et al., 2017). Oil synthesis is controlled by multiple genes in the A and C genomes of *B. napus*, such as the β-ketoacyl-ACP synthase gene (*KASIII-2*), located on chromosome C9 for long-chain fatty acids, and fatty acid elongase gene (*FAE1*) located on chromosome A8 and C3 for the production of erucic acid (Gacek et al., 2017). The diacylglycerol O-acyltransferase 1 gene (*DGAT1-2*), also located on chromosome C9, plays a

role in the esterification of fatty acids into glycerol, which affects oil accumulation in the seed (Vigeolas et al., 2007). The glycerol-3-phosphate acyltransferase gene (*GPAT*) and glycerol-3-phosphate dehydrogenase gene (*GPD1*), both located on chromosome C3, function during oil synthesis (Fourmann et al., 2002; Mietkiewska et al., 2004; Li et al., 2015). *Brassica napus* fatty acid dehydrogenase 2 and *BnFAD3* are involved in fatty acid biosynthesis (Xue et al., 2018). *Brassica napus* fatty acid dehydrogenase 2 genes are located on chromosomes A1, A5, C1, and C5, while *BnFAD3* genes are on chromosomes A3, A4, A5, C3, and C4 (Gacek et al., 2017). Besides these few examples, many other genes are still unknown and need to be investigated (Zhao et al., 2008).

## 2.7 Molecular markers

Historically, crop breeding has relied only on repeated selection for desired phenotypic traits over multiple generations (Sakhno, 2010). Technological advances have since allowed breeders to improve their efficiency by using molecular markers to help them make selections early in the growth cycle and enable the screening of larger populations (Ribaut and Hoisington, 1988; Asghari et al., 2011; Singh, 2017). A molecular marker is a DNA molecule and a tool used by geneticists and breeders to identify or locate certain genes of interest that correlate to the phenotype of the organisms (Winter and Khal, 1995).

Molecular markers are classified into different groups based on their detection methods. These include restriction fragment length polymorphisms (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphisms (AFLP), variable number tandem repeat (VNTR), simple sequence repeat (SSR), sequence-related amplified polymorphisms (SRAP), inter-retrotransposon amplified polymorphisms (IRAP), oligonucleotide polymorphisms (OP), allele specific associated primers (ASAP), and single nucleotide polymorphism (SNP) markers (Liu et al., 2013b; Liu et al., 2016). Currently, the most popular molecular marker systems look at SNP at positions all across the plant genome

(Hu et al., 2006). With the help of computer programs and bioinformatics knowledge, molecular markers used for a specific target gene can be easily identified (Shendure and Ji, 2008; Deschamps et al., 2012). In most cases, high throughput SNP detection is necessary. The most commonly used methods in plants are Illumina Bead Array and genotyping-by-sequencing (GBS) methods (Deschamps et al., 2012; Elshire et al., 2011). The Illumina Bead Array technology is popular and has proven to be effective in plant species such as *B. napus* (Clarke et al., 2016) and other crop species (Song et al., 2013; Winfield et al., 2016) because multiple SNPs across the whole genome can be screened simultaneously.

## 2.8 Next generation sequencing

Next generation sequencing (NGS) technologies including Roche 454, Illumina Solexa, ABI SOLiD, HeliScope, Ion Torrent, Pacific Biosciences SMRT, Oxford Nanopore MinION, and Qiagen Gene Reader are technologies used only for genetic sequencing (Mardis, 2017). In contrast to the first generation sequencing, Sanger sequencing, the preparation of the NGS library is faster, simpler, and more efficient (Sanger, 1988; Ari and Arikan, 2016).

One of the most utilized NGS technologies is the Illumina Solexa or Illumina sequencing. The Illumina sequencing is used by many researchers due to its high-throughput because many sequencing reactions can occur simultaneously (Metzker, 2010; Mardis, 2017). This enables Illumina sequencing to be applied to human, animal, plant, microorganism, and other genomic studies since the technology enables the production of large genetic data sets which are useful for various genomic analyses (Shendure and Ji, 2008).

The Illumina sequencing produces pair-end sequencing reads. The paired-end sequencing allows scientists to sequence both ends of DNA fragments to generate high quality and useful genetic data from *B. napus* (Chen et al., 2008). The genetic data from Illumina sequencing is particularly useful for the development of molecular markers for mapping quantitative trait loci (QTL) (Chen et al., 2017).

## 2.9 Genotyping-by-sequencing

Genotyping-by-sequencing (GBS) is a method used for detecting the SNP from the genetic data developed from the NGS (Elshire et al., 2011). In GBS, SNP detection is performed when sequencing data are compared to the reference genome (Rafalski, 2002). The standard GBS library construction protocol (Figure 2.2) requires the use of restriction enzymes (REs) to digest plant genomic DNA into smaller fragments during the GBS library construction to reduce the complexity of plant genomic DNA (Elshire et al., 2011). Next generation sequencing (NGS) generates data that have smaller read lengths than Sanger sequencing, so the risk of alignment of these short fragments to incorrect positions in the reference genome increases, which leads to errors in SNP calling (Poland and Rife, 2012). Although the use of the REs in GBS is highly specific, quick, and easy, the selection of the REs used in the GBS protocol may be challenging and may not cover the entire region in the plant genomic DNA (Elshire et al., 2011).

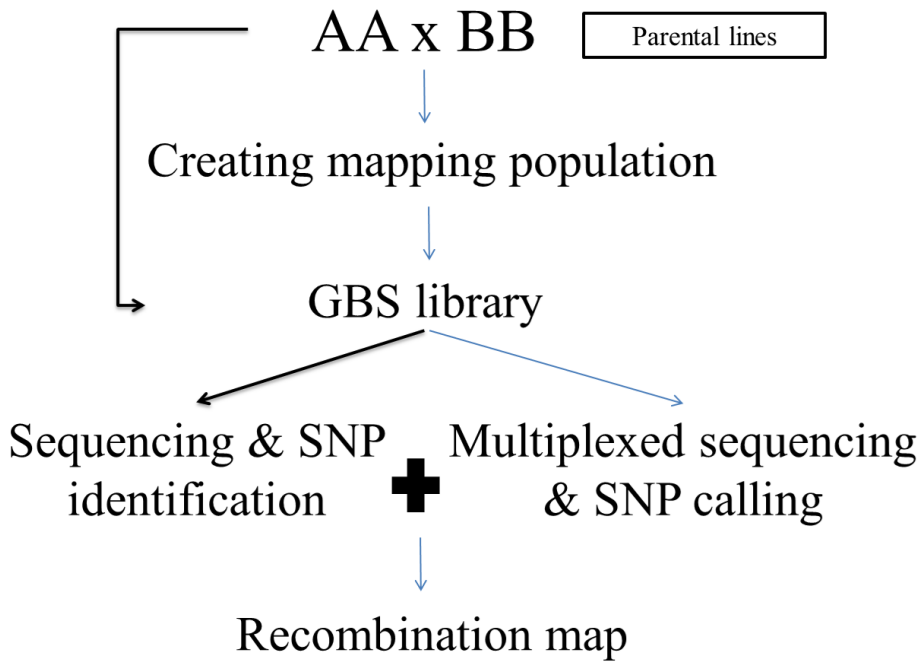


Figure 2.2: Schematic diagram of genotyping-by-sequencing (GBS) protocol with AA and BB representing the hypothetical parental genotpyes for mapping population and SNP calling to produce a recombination map (adapted from Poland and Rife, 2012).

## 2.10 Mapping quantitative trait loci

*Brassica napus* seed yield and SOC are typically inherited quantitatively (Tanksley, 1993). Quantitative trait loci (QTL) represent a segment of a chromosome or specific genetic locus that contributes to the variation in the phenotype of an organism (Brummer et al., 1997). Single nucleotide polymorphisms can be linked or pinpointed to a specific phenotypic trait. Multiple QTL can be correlated with individual quantitative traits, and the accuracy of QTL mapping can be affected by environmental conditions, which can vary across replicated field plots and experimental sites (Chen et al., 2010; Yu et al., 2016). In general, QTL mapping requires a segregating population from distinct parental phenotypes such as differences in SOC, and the progenies of those parental genotypes are used for producing a genotype map with a specific molecular marker for detection. Quantitative trait locus studies of crop plants are conducted using phenotypic data collected in field experiments across multiple locations and multiple years to account for the environment effect on phenotypes. The combined susceptibility of *B. napus* to both biotic and abiotic stresses (section 2.5) makes QTL mapping studies in this crop challenging.

Two major components of genetic complexity are linkage and recombination during meiosis, which is related to the correlation and linkage between phenotypic traits and the genes underlying the traits in chromosomes (John, 2005; Benfey, 2014). The linkage is measured by the genetic distance of the genes representing the frequency of crossing-overs during homologous recombination. The units used in the measurement are called centiMorgan (cM) (Benfey, 2014). To study linkage in *B. napus*, there must be genetic differences between maternal and paternal parents. For instance, the phenotypic difference of the SOC of two parents is related to the genes functioning in the seed oil biosynthesis, as well as the differences between pod length, protein content, and other traits. For mapping QTL for complex traits, a doubled haploid (DH) population with a phenotypic trait following a normal

distribution is preferable (Benfey, 2014). Studies have been performed to identify and map the QTL that correlate with desirable traits such as flowering time (Osborn et al., 1997; Axelsson et al., 2001), SOC (reviewed further in section 2.9.1 and 2.9.2) (Burns et al., 2003; Delourme et al., 2006; An et al., 2019), seed coat colour (Yan et al., 2009), seed glucosinolate content (Uzunova et al., 1995; Howell et al., 2003), erucic acid content (Ecke et al., 1995), clubroot disease resistance (Manzanares-Dauleux et al., 2000), sclerotinia disease resistance (Zhao and Meng, 2003), pod shatter resistance (Raman et al., 2014), protein content (Zhao et al., 2006), as well as seed weight and silique length (Zhang et al., 2011; Li et al., 2014).

Composite interval mapping (CIM) and simple interval mapping (SIM) had been developed for QTL analysis to identify major QTL for QTL studies. Both methods have their advantages and disadvantages and both methods were discussed by Wu et al., (2007) and Alam et al., (2016).

### 2.10.1 Studies for seed oil content

The goal of QTL studies is to identify genetic loci on the chromosomes that correlate with quantitative phenotypes (Kearsey, 1998). Mapping QTL associated with SOC in *B. napus* needs phenotypic data collected over multiple years of field experiments to improve the accuracy of the mapped QTL (Parkin et al., 1995; Delourme et al., 2006; Chen et al., 2010). Generally speaking, it is hard to identify oil synthesis genes that correspond to the mapped QTL since these genes often have minor effects and are easily affected by various environmental conditions. Other challenges to the identification of QTL for SOC in *B. napus* include the production of large mapping populations, selection of adequate genetic stocks, construction of genetic maps with a high density of molecular markers, and collection of accurate phenotypic data (Kearsey and Farquhar, 1998; Fu et al., 2017).

Burns et al., (2003) identified two major QTL that may correspond to *BnFAE1.1* and *BnFAE1.2,* and one minor QTL for erucic acid on chromosome C8 by using backcross population derived from Victor and Tapidor. Both *BnFAE1.1* and *BnFAE1.2* are responsible

for the synthesis of erucic acid in *B. napus*. Burns et al., (2003) located the QTL for *BnFAE1.1* on chromosome A8 at the proximal end with a confident interval of 12 cM while the QTL for *BnFAE1.2* was identified on chromosome C3 at the distal end with a confident interval of 26 cM.

In another report by Chen et al., (2010), SOC data were collected from three different locations within Manitoba (The Point Research Station at University of Manitoba, Carman Research Station, and Portage La Prairie) and 353 SRAP and 34 SSR markers were used to identify QTL for SOC using 150 *B. napus* DH genotypes derived from a high-oil content and a low-oil content line. Twenty-seven loci were identified, and the highest logarithm of the odd (LOD) of 5.9 was located at linkage group 2 (LG2, chromosome A2), followed by LOD values of 5.1 and 5.0 at LG13 (C3) and LG2 (A2) respectively at the Point Research Station. In Carman, different QTL on LG3 (A3), LG4 (A4) and LG14 (C4) were identified. Similar results were observed at Portage La Prairie in 2008, and a LOD value of 4.5 on LG2 (A2) at 77 cM was observed after composite interval mapping (CIM) analysis with 1,000 time permutation test to determine the major QTL threshold (Chen et al., 2010). Twenty-seven QTL were identified for SOC in *B. napus*, explaining the highest phenotypic variance of 30.2% on chromosome A1 of the *B. napus*. Among the 27 QTL, four of the QTL were reproducible in two different environments, whereas 23 QTL were only identified only in one environment.

Quantitative trait loci were studied by Wang et al. (2013) using 348 DH *B. napus* genotypes that were derived from KenC-8 and N53-2 varieties. The QTL experiment was conducted in eight different environments and a total of 63 QTL were identified, explaining phenotypic variance ranging from 2.64 to 17.88%. Eight individual genetic maps were combined into a consensus map with a length of 2395.2 cM and average marker interval of 1.8 cM. Composite interval mapping was used for QTL analysis with a 1,000 permutation test. The result of the study shows the QTL that was identified on chromosome C9 of the *B. napus* is only expressed in genotypes with high oil content.

Javed et al., (2016) conducted a study using DH population derived from Polo and Topas in four different environments. A total of 14 QTL were identified for SOC in *B. napus*. The study had identified QTL on chromosome A10 of *B. napus*, explaining 26.99% of phenotypic variance that is positively correlated to Topas. Further investigation on two *B. napus* genome C3 positions, 147.83 and 154.55 cM, showed that these positions were associated with fatty acid production and contributed to the overall SOC of *B. napus*. Quantitative trait locus analysis was performed using composite interval method based on SSR markers with an average genetic distance of 3.7 cM.

Another QTL study for SOC by An et al., (2019) took place at the University of Manitoba in 2015 and 2016. An et al., (2019) had identified two QTL associated to SOC on *B. napus* chromosomes C3 and C7 using SRAP and SNP molecular markers. Logarithm of the odd scores of 2.8 and 3.3 for chromosomes C3 and C7, explaining 9% and 10.7% of genotypic variances were observed. This demonstrates how identified QTL can differ depending upon the phenotyping environment.

Quantitative trait locus studies for SOC in *B. napus* are not typically conducted in controlled environmental conditions such as greenhouses because oil-related genes may or may not be expressed (Raboanatahiry et al., 2018). Additionally, the greenhouse is not an ideal place to grow large numbers of *B. napus* plants with significant seed production. However, a study by Bahrani and McVetty, (2008) shows that an average temperature cycle of 20 $^{o}$C in the greenhouse is close to the optimum temperature for production of *B. napus* which results in an increase in the SOC, protein, and glucosinolates content, and therefore it is valuable to conduct *B. napus* experiments in controlled environmental conditions such as greenhouses.

Further, QTL and SOC experiments for *A. thaliana* can be easily established in the greenhouse due to the rapid life cycle of the plant, the plant is a great model for *B. napus*, and the plant is a fully sequenced genome (Marra et al., 1999). Hobbs et al., (2004) had identified two QTL associated to SOC with a peak LOD score of 7.06 and 7.29 at LG1 and LG2. Major

QTL on LG2 were further investigated and they managed to observe that on chromosomal position 85-96 cM, the AFLP molecular markers were replaced with one of the parental genotypes by recombination which resulted in a lower SOC. They suspected it was due to the gene control of acyl-CoA synthetase of *A. thaliana* which is responsible for the synthesis of oil in the seed.

Given the lack of mapped QTL for SOC in *B. napus* in a controlled environmental condition such as a greenhouse, the goal of this research is to identify QTL for SOC in *B. napus* population in both greenhouse and field experiments. Previous SOC studied by An et al., (2019) focused on using standard GBS method to identify SNP.

# Chapter 3.0 MOLECULAR MARKERS FOR OIL CONTENT IN *BRASSICA NAPUS* WITH PCR-BASED ILLUMINA SNP DETECTION (PISD)

## 3.1 Abstract

Many molecular markers, such as restriction fragment length polymorphisms (RFLP), single sequence repeats (SSR), amplified fragment length polymorphisms (AFLP), and sequence-related amplified polymorphisms (SRAP), have been used for studying traits such as seed oil content (SOC) in *Brassica napus*. High throughput detection methods of single nucleotide polymorphism (SNP) markers, such as Illumina Bead Array and genotyping-by-sequencing (GBS), have been developed and are currently used for various genomic applications. In previous research, a doubled haploid (DH) *B. napus* population ZT was used, derived from a cross of Topas and ZY821, utilizing SRAP and GBS molecular marker technologies. In this study, a PCR-based Illumina SNP detection (PISD) method was used to construct genetic maps. Using the PISD, 1,092 SNPs were used to construct a genetic map (PISD map). The PISD SNPs were also combined with our previously identified SNPs from GBS to construct a combined genetic map with 1,821 SNPs (combined map). The first genetic map covered 1,870 cM while the combined genetic map had a size of 2,057 cM. In the combined genetic map, 985 SNPs were located in the A subgenome and 836 SNPs in the C subgenome of *B. napus.*

## 3.2 Introduction

*Brassica napus* L., also known as oilseed rape or canola, consists of the AACC genome (2n=38) and is one of the most important oilseed crops in the world (Piquemal et al., 2005). Many molecular markers such as restriction fragment length polymorphisms (RFLP) amplified fragment length polymorphisms (AFLP), simple sequence repeats (SSR), sequence-related amplified polymorphisms (SRAP), and single nucleotide polymorphisms (SNP), have been developed for the construction of genetic maps for quantitative trait locus (QTL) studies on

important traits for *B. napus* such as seed oil content (SOC), flowering day, day to maturity, and many other traits (Semagn et al., 2006; Dhingani et al., 2015).

In order to obtain accurate results in QTL studies, sufficiently large population sizes (Brummer et al., 1997), multiple experiments, and high density molecular marker genetic maps need to be used (Yu et al., 2011). The Illumina Bead Array technology uses microbeads for genotyping the genotype in the population which targets a specific locus of the genome in *B. napus* (Clarke et al., 2016). The genotyping-by-sequencing (GBS) is a method for detecting SNP through restriction enzyme (RE) to reduce the complexity of the *B. napus* genome. After RE digestion, PCR is performed to increase DNA fragments which are sequenced by Illumina sequencing (Rafalski, 2002; Elshire et al., 2011). High throughput SNP detection methods such as Illumina Bead Array and GBS increase the precision and accuracy of the studies, as SNPs are dispersed across the genome and can therefore be used to construct high density genetic maps (Deschamps et al., 2012).

The genome size of *B. napus* with approximately 1,200 to 1,280 Mb is relatively large, and the resequencing of whole genome for SNP detection is expensive and time consuming (Song et al., 2020). To reduce the complexity of *B. napus* genome, RE or double digestion GBS (ddGBS) of genomic DNA coupled with DNA barcoded adapters in GBS provides reasonable coverage of the genome in a cost-effective manner (Chen et al., 2013). The reduced representation by RE cutting and size selection is achieved in the GBS method which only sequences a small portion of the *B. napus* genome. Commonly used REs in the construction of GBS library are *Eco*RI, *Hin*dIII, *Pst*I, and *Ape*KI (Elshire et al., 2011). The REs digest genomic DNA into fragments with their specific sticky ends.

Genetic diversity in *B. napus* is relatively large and thousands or even millions of SNPs can be identified following resequencing. For example, over 24 million SNPs were identified to develop Illumina Bead Array in *B. napus* (Clarke et al., 2016; Song et al., 2020). In a recent report, 1.2 million SNPs were identified between the first whole genome of Darmor-*bzh* and

newly assembled whole genome of ZS11 (Sun et al., 2017). Since a few thousand SNPs are adequate for most genomic applications in *B. napus*, partial sequencing of individual DNA samples in GBS can be used to detect adequate SNPs in *B. napus* (Elshire et al., 2011; Chen et al., 2013).

In addition to QTL studies, high throughput SNP marker detection methods have been successfully used for genome-wide association studies (GWAS), marker-assisted selection (MAS), and phylogenetic analyses (Wang et al., 2011; Snowdon and Iniguez-Luy, 2012; Liu et al., 2013a). Delourme et al., (2013) reported that 5,764 SNPs and 1,603 PCR markers were used to generate a genetic map covering 2,250 cM for *B. napus*. Chen et al., (2013) also used a modified double digestion restriction-associated DNA sequencing approach to detect SNPs. Despite a significant decline in the cost of next-generation sequencing, genotyping costs could be confined to smaller breeding projects such as those in the public domain. A simplified SNP detection method reduced the cost and increase the efficiency of genetic analysis.

The objectives of this work are: firstly, to develop a simple SNP detection method using PCR through PCR-based Illumina SNP detection (PISD); and secondly, to construct genetic maps from PISD and copare these genetic maps with previous studies.

## 3.3 Materials and Methods

### 3.3.1 Mapping population

A doubled haploid (DH) population (ZT) was generated from a cross between spring type canola Topas and winter type rapeseed ZY821. A total of 94 DH genotypes were used for SNP genotyping at Huazhong Agricultural University, Wuhan and Cornell University, New York. A total of 2,093 SNP markers were identified in Wuhan and 869 in GBS data from Cornell University which were used for map construction.

### 3.3.2 DNA extraction and preparation of modified GBS library

A modified cetyltrimethylammonium bromide (CTAB) DNA extraction method (Li and Quiros, 2001) was used to extract genomic DNA from leaves to prepare libraries for Illumina sequencing in the PISD protocol. Approximately 0.5 g of fresh, true-leaf tissue was collected from 94 genotypes in the ZT DH population, 2 weeks after seed germination. The fresh leaf tissue was harvested into a 1.5 mL micro-centrifuge tube and then crushed into a fine powder in the presence of liquid nitrogen. After the liquid nitrogen had dissipated, the tissue was resuspended in 500 µL of 2x CTAB buffer (2% CTAB, 20 mM ethylenediaminetetraacetic acid, 100 mM tris-hydroxymethyl-aminomethane, and 1.4 M sodium chloride, pH 7.5). The samples were then incubated in a water bath at 65$^{o}$C for 60 to 90 min. After incubation, 400 µL of chloroform was added to each tube after which the tube was capped and rocked back and forth for 10 min. The tubes were then centrifuged at 13,000 rpm for 10 min at room temperature until 400 µL of supernatant was pipetted into a new 1.5 mL tube. The DNA was then precipitated by adding 200 µL isopropanol and the supernatant was again shaken gently back and forth. To recover the DNA, the samples were centrifuged at 8,000 rpm for one min at room temperature to form a pellet which was washed once with 800 µL 70% ethanol. The ethanol was discarded and the pellet was air dried before adding 200 µL of water to dissolve the DNA. The quality of the DNA was checked using an ABI 3130 genetic analyzer (Life technologies, USA) and Genographer ® v1.6.0 (Genographer) software. The best quality DNA samples were selected from the ZT DH population and individually added in a 96-well plate along with the parents.

After the quality check, the DNA was used to prepare the Illumina sequencing library. Originally, the standard protocol for GBS library preparation required digesting the DNA with a combination of restriction enzymes (Elshire et al., 2011), but for the PCR-based library preparation protocol, this digestion step was omitted. First, DNA quality was tested using a set

of primers to optimize the PCR conditions. Four pairs of primer combinations (two P5 end primers and two P7 end primers) were then combined with 96 DNA samples to fill a 384-well plate and a total of 32 different unique primer combinations were included in the first round of PCR. Polymerase chain reaction products were then precipitated using 3 µL of 3M sodium acetate (NaOAc), pH 5.6, and 16.5 µL of isopropanol. The plates were then covered with adhesive aluminum foil and centrifuged at 6,500 rpm for at least five minutes at room temperature to pellet the precipitated DNA. Sodium acetate and isopropanol were gently discarded without losing the DNA pellet in the 384-well plate by inverting the plate and tapping the liquid out. The DNA pellet was washed with 20 µL 70% ethanol to remove salts and the plates were centrifuged at 6,000 rpm for one minute at room temperature. After the ethanol was discarded, the plates were placed inside the vacuum centrifuge without covering and dried for ten minutes. The DNA was then ready for the second round of PCR.

In the second round of PCR, 96 index primers, P7M1-96, were used for 96 DNA samples. A 96-needle replicator was used to add the index primers – four times for four primer pair combinations in a 384-well plate. The primer, P5M, was added in the second round of PCR. After the second round of PCR, PCR products were checked with 1% agarose gels and all the PCR products pooled into a 1.5 mL micro-centrifuge tube. The pooled DNA was purified using Omega magnetic beads (Omega Bio-tek Institute, Norcross, GA, United States) following the manufacturer's protocol. The magnetic bead solution was taken from the fridge and kept on the bench until it reached room temperature; for every 50 µL of PCR products in a 1.5 mL tube, 0.8 volume of magnetic bead solution was gently mixed by tapping the tube at room temperature after incubation. The resultant samples were incubated for 15 minutes at room temperature before the tubes were placed on a magnetic rack to collect the magnetic beads with bound DNA. The tubes were kept on the rack for more than 30 seconds until the solution was clear. The solution was discarded. The beads were washed once using 200 µL of 80% ethanol and then air dried for three to five minutes. After removing the tube from the rack,

50 µL of water was added to the tube to dissociate the DNA fragments from the beads. The 300 to 400 bp DNA fragments were checked using ABI 3150 and identified on 1% agarose gels after three PCR rounds. Equal quantities of DNA sample were pooled for Illumina sequencing (Figure 3.1).

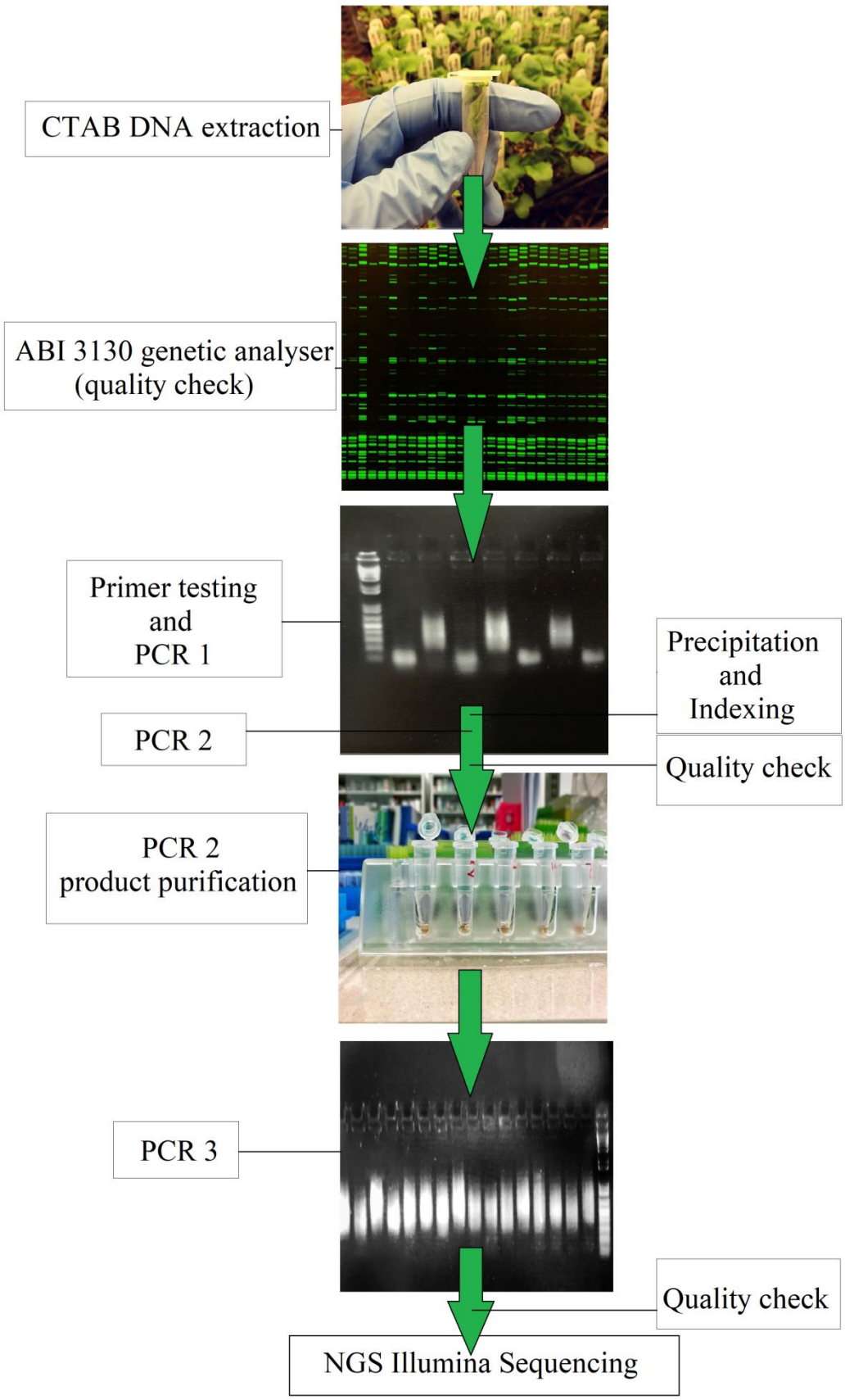


Figure 3.1: Summary of library preparation in the PCR-based Illumina SNP detection (PISD) and quality check using 1% agarose gel.

### 3.3.3 PCR-based Illumina SNP detection

A DNA library consisting of 94 ZT DH genotypes along with the two parental genotypes Topas and ZY821 were sent to Huazhong Agricultural University, Wuhan, China for Illumina sequencing using HiSeq2000 system following the manufacture's protocol (GenoSeq, Wuhan) to produce 150 bp paired-end sequence reads (Hu et al., 2018). All reads from 96 individual DNA samples were mapped with the reference *B. napus* genome 'Darmor-*bzh*' v4.1 (Chalhoub et al., 2014). Single nucleotide polymorphisms were detected using custom Perl scripts and the ANNOVAR program (Wang et al., 2018). The four identities 0/0 (no SNP was found in this locus), 1/1 (SNP was found in this locus), 0/1 (heterozygous), and ./. (missing data), were assigned to each locus. A total of 2,093 SNP markers were identified and sorted based on their chromosome position and location of the SNP in an excel spreadsheet which was completed at the Huazhong Agricultural University as previously described (Hu et al., 2018).

The precision and accuracy of the SNPs obtained from the Huazhong Agricultural University were confirmed by extracting the raw paired-end FASTQ files for alignment in HISAT2 (Kim and et al., 2015). The DNA sequences from the FASTQ files were compared with the *Brassica_napus*_v4.1. reference genome (http://www.genoscope.cns.fr/*Brassica*napus/data/) and BLASTED against the NCBI database to identify the SNP locations. Non-polymorphic SNP markers for both parents and genotypes with more than 10% missing data were removed. The identities (0/0, 1/1, and ./.) were changed to a readable format in JoinMap ® (version 4.0). A total of 1,239 SNP markers were selected from PISD data for genetic map construction.

### 3.3.4 Construction of genetic maps

For genetic map construction, 1,239 SNPs from 94 ZT DH genotypes were uploaded to JoinMapv4 to produce the first genetic map (this map is referred to as the PISD genetic map).

The 1,239 PISD SNP markers were then combined with our previously identified 869 SNPs from the data collected using conventional restriction enzyme digestion GBS (An et al., 2019), and the combined genetic map was constructed (referred to as the combined genetic map).

The JoinMapv4 user manual was the same as described previously (Van Ooijeb, 2006). The population parameter was set to DH population for both the PISD and combined genetic maps to create a population node tree. Within the population group tree, the SNP markers associated with each other were grouped together based on the specificity in each ZT DH genotype, and the missing data was calculated and aligned to the adjacent marker by JoinMapv4. Nineteen major PISD and combined genetic maps were generated, and minor genetic groups showing SNP markers with no alignment were generated. After the minor genetic groups were eliminated, a total of 1,092 SNP markers for the PISD genetic map and 1,821 SNP markers for the combined genetic map were aligned to the 19 genetic groups in both genetic maps. JoinMapv4 programme with Kosambi mapping function estimated the genetic intervals of each genetic group in cM. Nineteen genetic groups were extracted from JoinMapv4 into an excel spreedsheet. The number of PISD and combined genetic maps have been summarized. MapChart (version 2.32) (Voorrips, 2002) was used to produce genetic group diagrams with the results provided from JoinMapv4. Both the PISD and the combined genetic maps were used for QTL analysis in Chapter 4.

## 3.4 Results

A total of 729 SNPs were integrated into the PISD genetic map to generate a combined genetic map using the GBS data from a previous study (An et al., 2019). Four hundred and thirteen and 316 SNPs were assigned to the A and C genomes, respectively. The highest and lowest numbers of SNPs in the A genome were located on chromosomes A5 and A2 respectively (Figure 3.2). In the C genome, the highest number of SNPs was 71 on chromosome C1 while the lowest number of SNPs was 18 on chromosome C9 (Table 3.1).

A total of 1,092 SNPs were assigned to all 19 genetic groups of *B. napus* using the PISD data. Five hundred and seventy-two and 520 SNPs were assigned to the A and C genomes, respectively. On the A genome of the PISD map, the highest number of SNPs was 86 on chromosome A5 while the lowest number of SNPs was 25 on chromosome A8. In the C genome of the PISD genetic map, the highest number of SNPs was 95 on chromosome C2, while the lowest number of SNPs (32) was on chromosome C6 (Table 3.1). The PISD genetic map had a total length of 1,870.45 cM as calculated by JoinMapv4. Nine hundred and forty-two and 928 cM were assigned to the A and C genomes, respectively. In the A genome, chromosome A3, which is the longest in the PISD genetic map, was 138.64 cM, while A2 was the shortest. In the C genome, C9 had the largest size of 127.78 cM, while C6, the smallest, had a size of 66.08 cM.

In the combined genetic map, a total of 1,821 SNPs were assigned to all 19 genetic groups of *B. napus*. Nine hundred and eighty-five and 833 SNPs were assigned to the A and C genomes respectively. The highest number of SNPs was assigned to chromosome A5 which contained 175 SNPs while the lowest number of SNPs was 64 on chromosome A8 in the A genome. In the C genome, the highest number of SNPs was 154 on chromosome C2 while the lowest number of SNPs was 50 on chromosome C6. A total size of 2,057.97 cM was obtained for the combined genetic map. The assembled sizes of the A and C genomes were 1,094 and 963 cM, respectively. In the A genome, the largest genetic group was 181.64 cM on chromosome A3 and the smallest was 76.55 cM on A10. In the C genome, the largest and smallest sizes of 136.58 cM and 92.24 cM were chromosomes C1 and C4, respectively.

Overall, similar results of the A and C genomes were observed in both the PISD and combined genetic maps. The size of the A genome for both the PISD and the combined genetic map was estimated to be larger than the C genome (Table 3.1 and Figure 3.2 to Figure 3.5).

Table 3.1: The genotyping-by-sequencing (GBS) from previous studies (An et al., 2019), PCR-based Illumina SNP detection (PISD) and combined (GBS and PISD) genetic maps with

number of SNP markers aligned on the chromosomes of the A and C genomes. The total length for each chromosome was calculated using JoinMap ®.

| Chromosome | SNPs from previous study | SNPs from PISD | Combined SNPs | PISD Genetic Map (cM) | Combined Genetic Map (cM) |
|---|---|---|---|---|---|
| A1 | 42 | 54 | 96 | 96.87 | 114.90 |
| A2 | 22 | 49 | 71 | 55.73 | 104.27 |
| A3 | 51 | 68 | 119 | 138.64 | 181.64 |
| A4 | 29 | 57 | 86 | 130.21 | 84.33 |
| A5 | 89 | 86 | 175 | 104.47 | 127.47 |
| A6 | 32 | 60 | 92 | 75.73 | 77.81 |
| A7 | 40 | 44 | 84 | 69.95 | 88.91 |
| A8 | 39 | 25 | 64 | 58.30 | 76.55 |
| A9 | 30 | 65 | 95 | 100.68 | 92.83 |
| A10 | 39 | 64 | 103 | 111.79 | 144.84 |
| Total | 413 | 572 | 985 | 942.38 | 1094.55 |
| | | | | | |
| C1 | 71 | 40 | 111 | 96.87 | 136.58 |
| C2 | 59 | 95 | 154 | 126.23 | 99.19 |
| C3 | 41 | 56 | 97 | 109.16 | 126.67 |
| C4 | 20 | 63 | 83 | 93.38 | 92.24 |
| C5 | 21 | 76 | 97 | 106.23 | 98.12 |
| C6 | 18 | 32 | 50 | 66.08 | 79.56 |
| C7 | 46 | 41 | 87 | 86.41 | 124.86 |
| C8 | 22 | 58 | 80 | 114.93 | 106.40 |
| C9 | 18 | 59 | 77 | 127.78 | 99.80 |
| Total | 316 | 520 | 836 | 928.08 | 963.42 |
| | | | | | |
| TOTAL | 729 | 1,092 | 1821 | 1,870.45 | 2,057.97 |

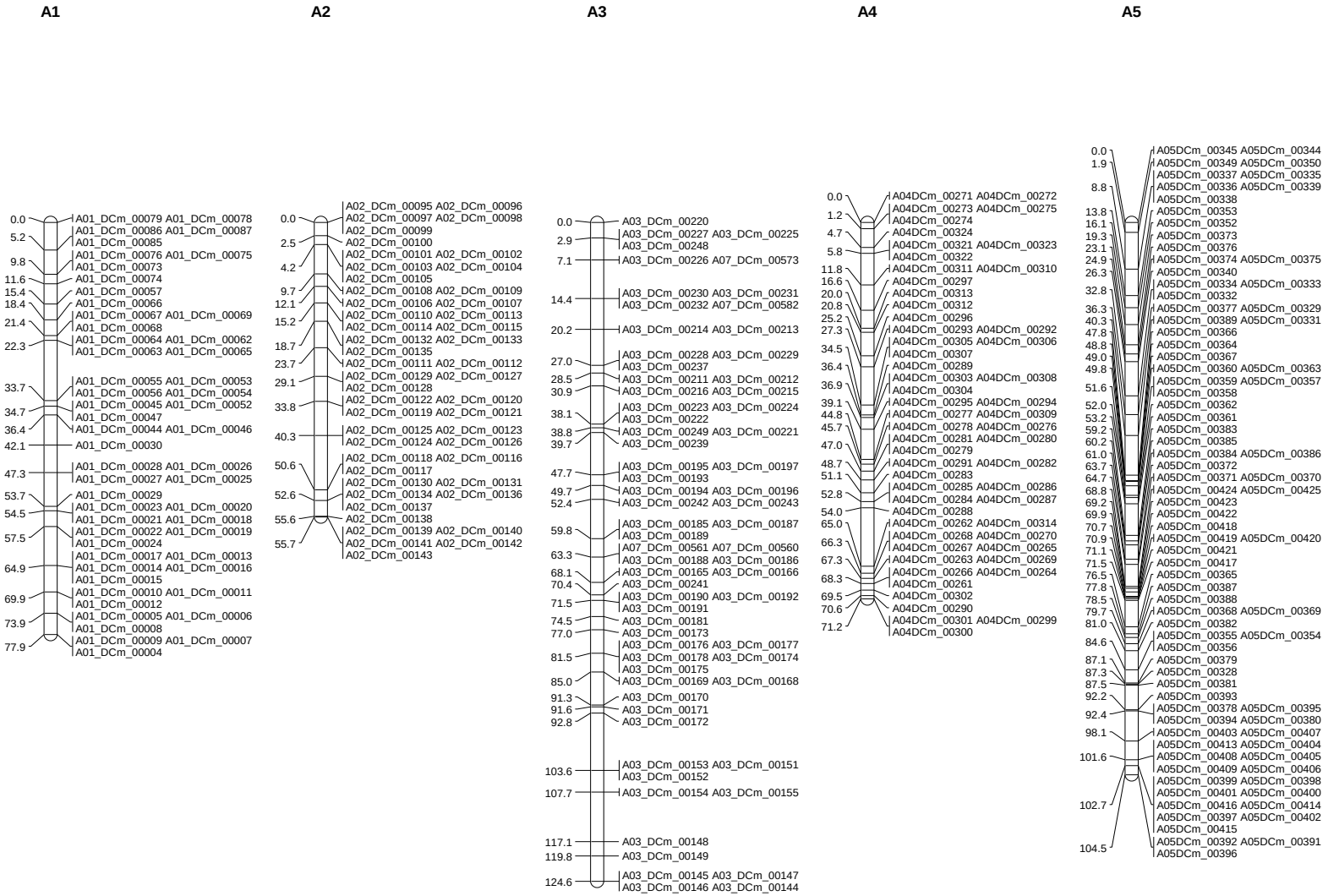


Figure 3.2: The PCR-based Illumina SNP detection (PISD) genetic map with SNP markers on chromosomes A1 to A10 in ZT doubled haploid population.

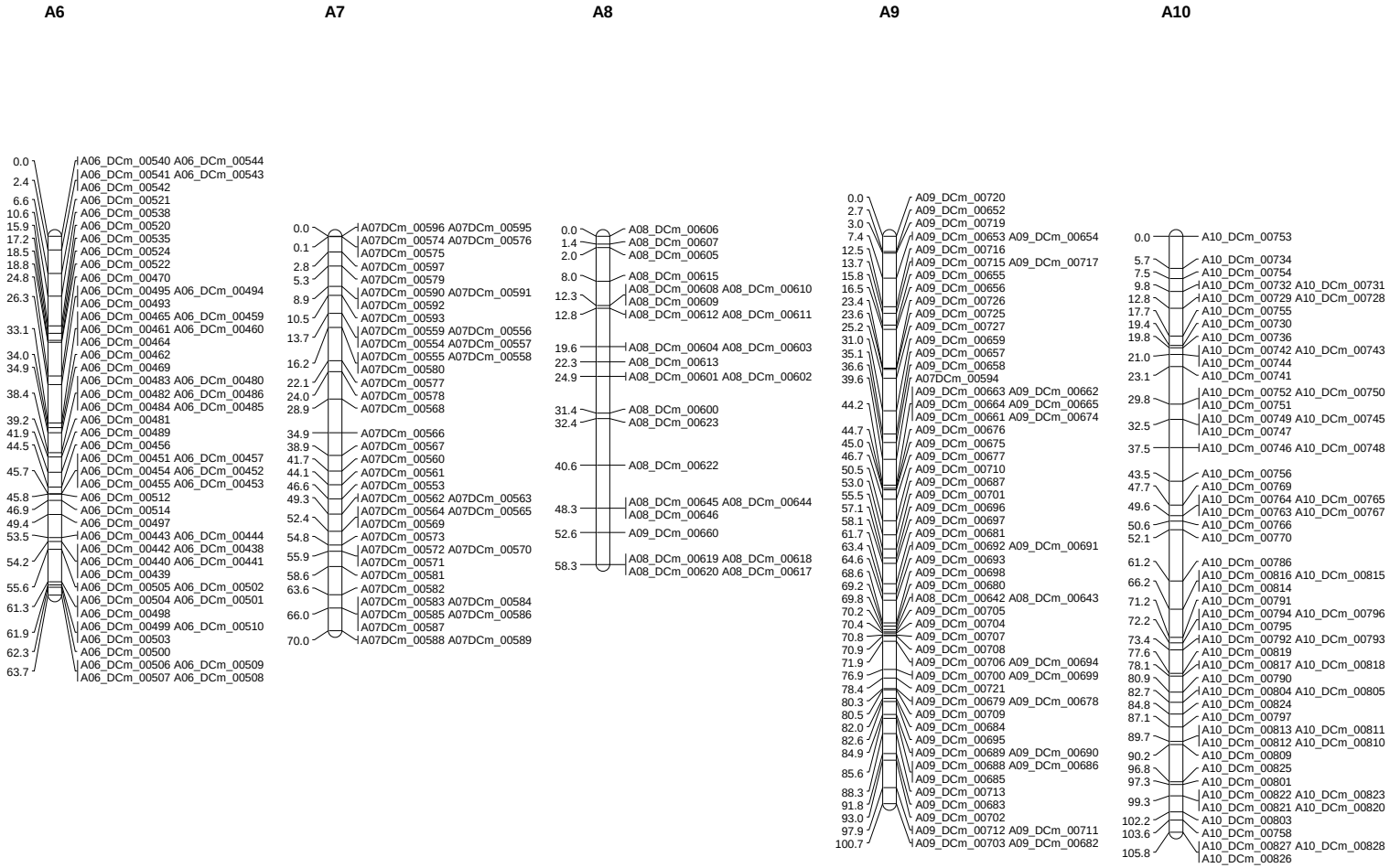


Figure 3.3: Continued.

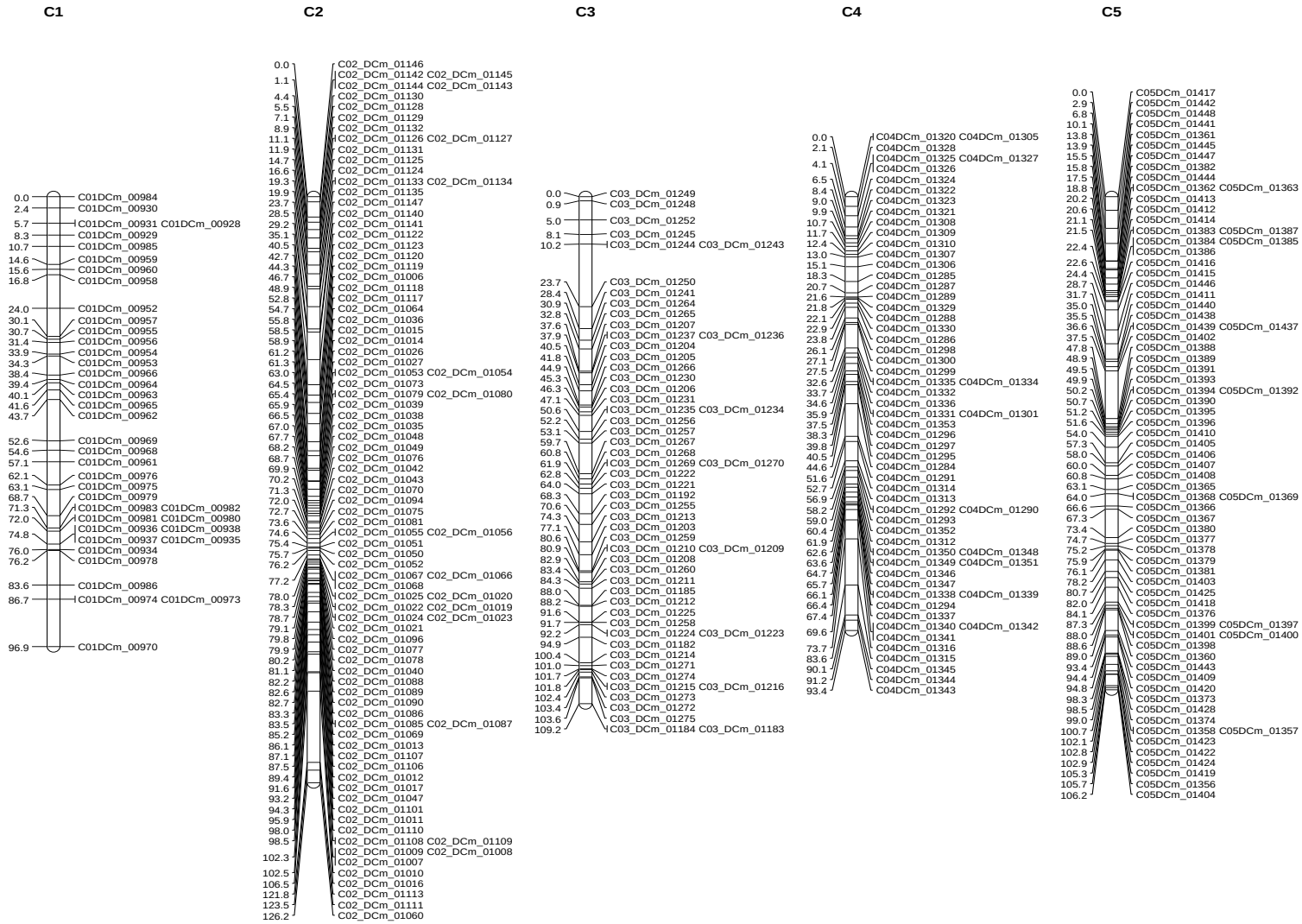


Figure 3.4: The PCR-based Illumina SNP detection (PISD) genetic map with SNP markers on chromosomes C1 to C9 in ZT doubled haploid population.

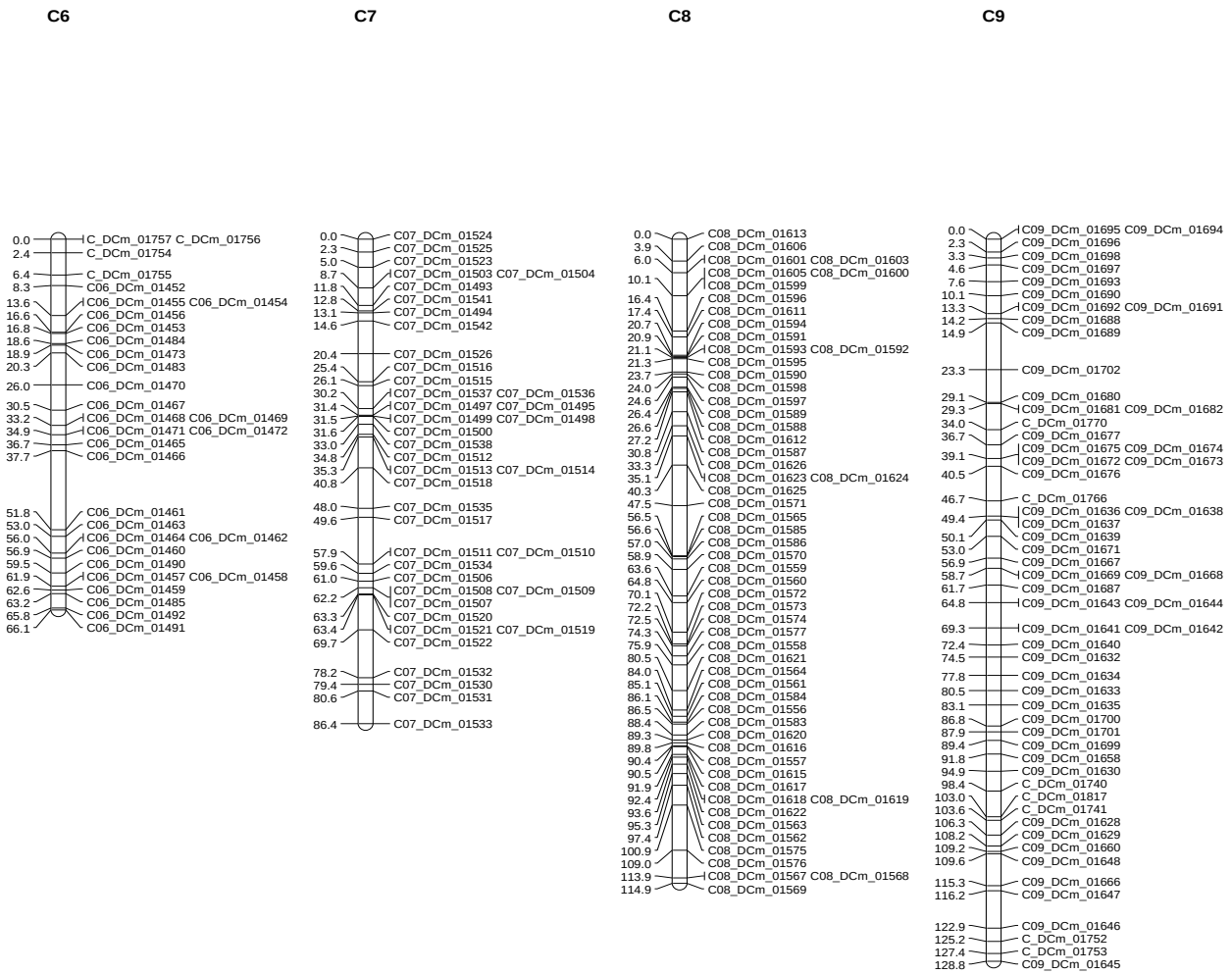


Figure 3.5: Continued.

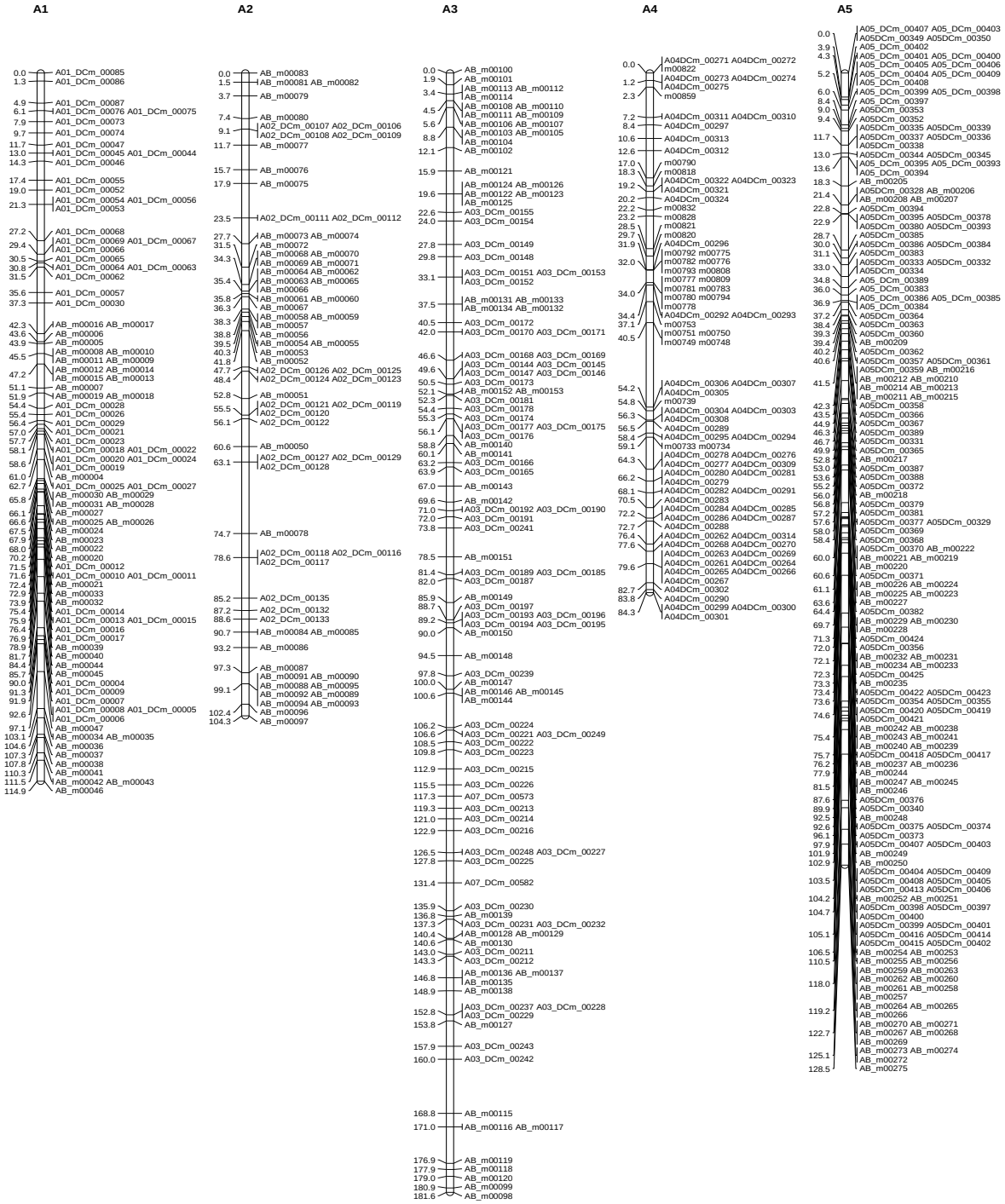


Figure 3.6: The combined genetic map (PISD and GBS) of ZT doubled haploid population with SNP markers on chromosomes A1 to A10.

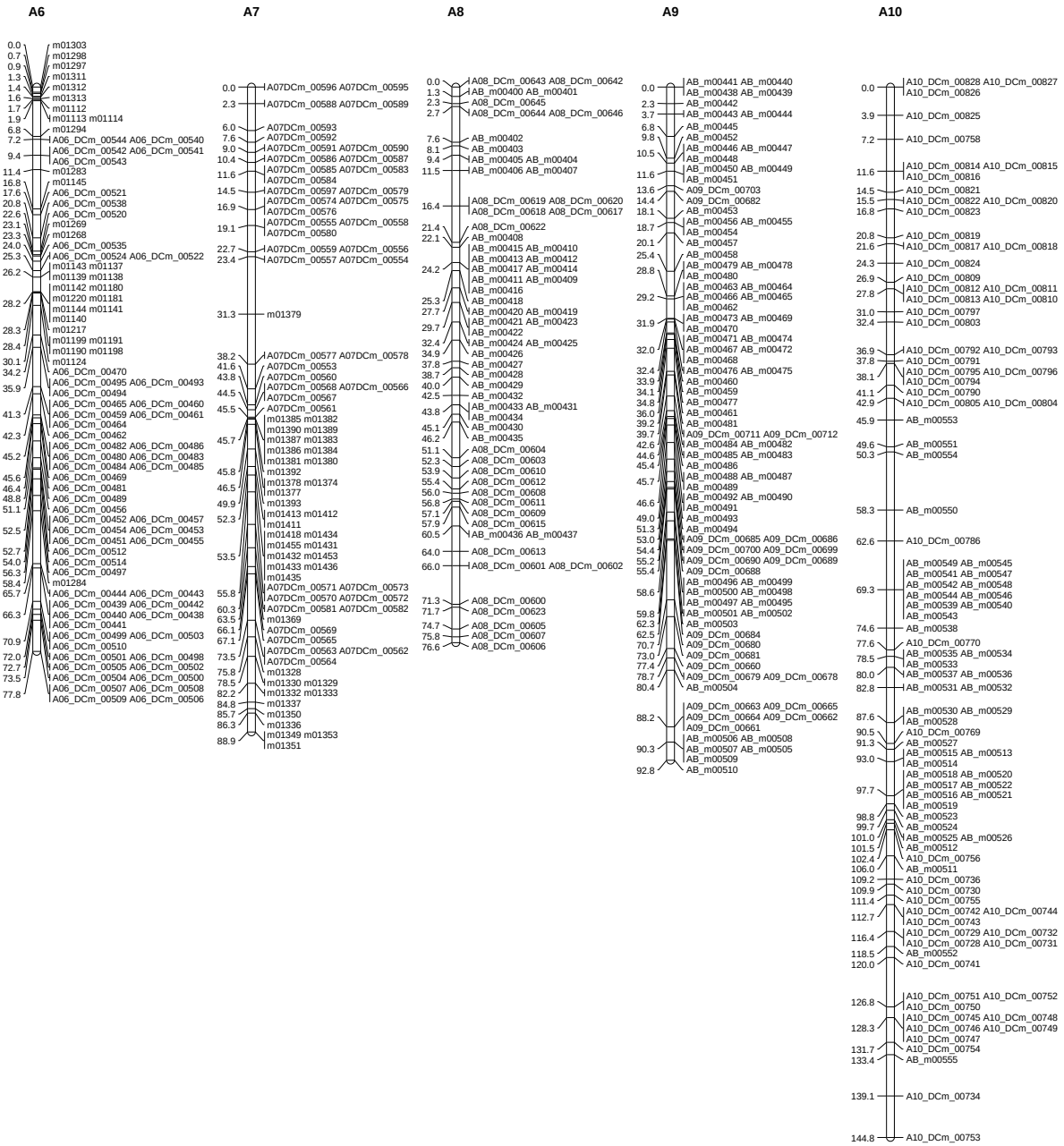


Figure 3.7: Continued.

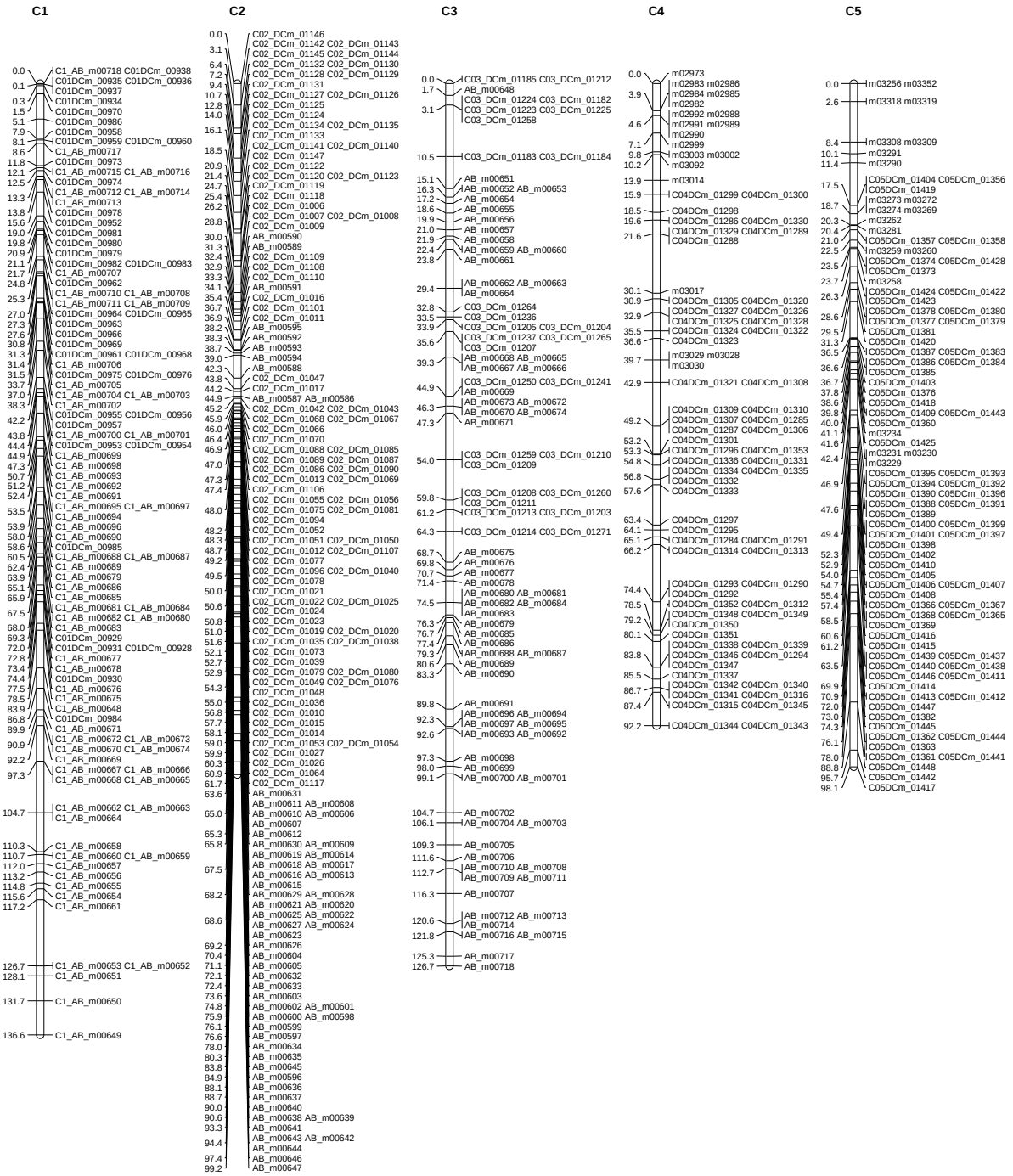


Figure 3.8: The combined genetic map (PISD and GBS) of ZT doubled haploid population with SNP markers on chromosomes C1 to C9.

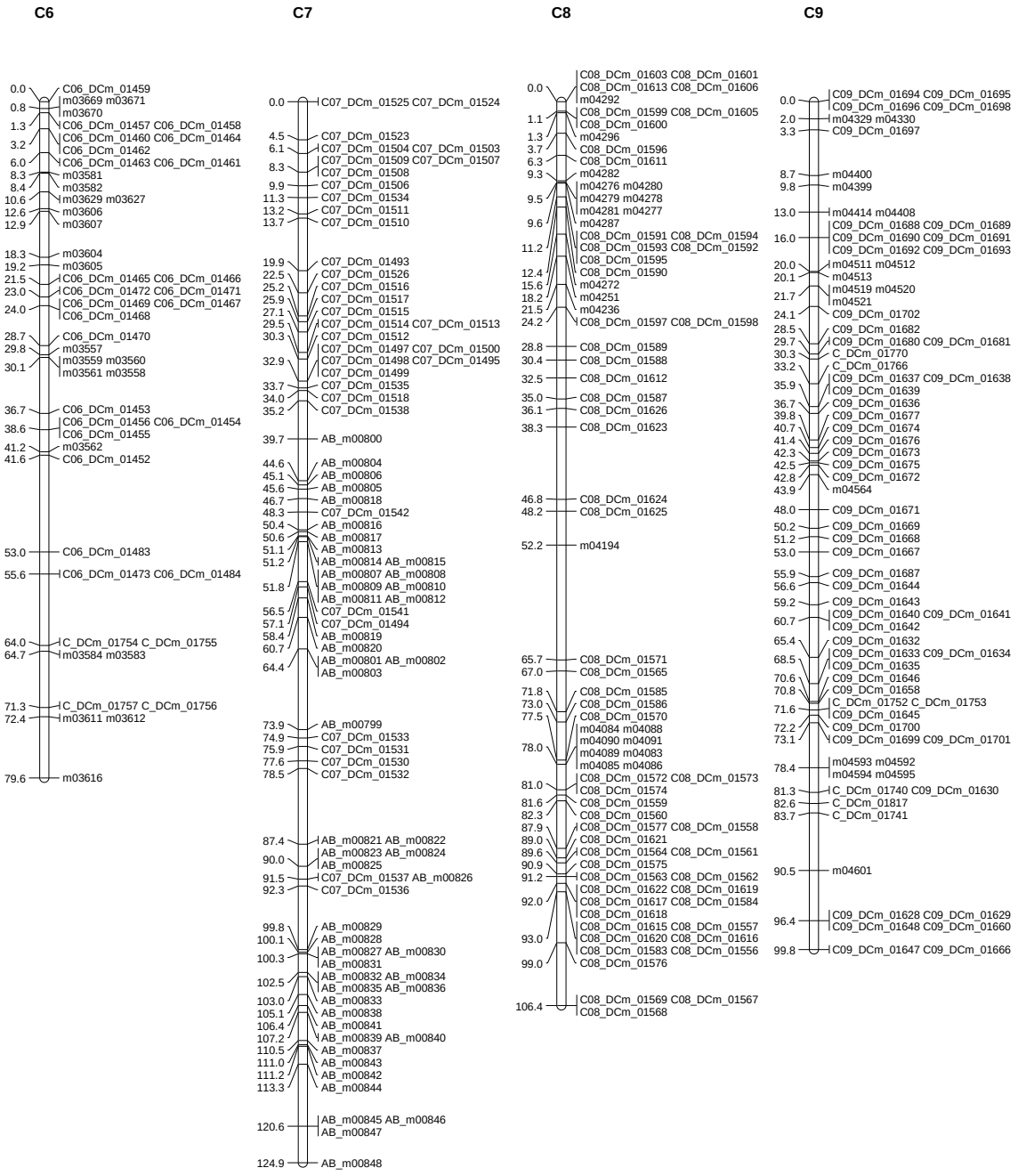


Figure 3.9: Continued.

## 3.5 Discussion

There are several SNP detection technologies such as Illumina Bead Array and GBS that are available for the analysis of *B. napus* (Chen et al., 2010). Since SNPs are not evenly distributed in the *B. napus* genome, the fixed SNPs in Illumina Bead Array do not evenly cover the whole genome, so genetic maps constructed with the Illumina Bead Array technology often contain many gaps. For example, genetic maps constructed with Illumina Bead Array had only nine SNP bins on C9 (Liu et al., 2013b). Single nucleotide polymorphism physical positions in the *B. napus* A and C genomes in Illumina Bead Array technology were selected based on their original sequences among the sequenced materials. Although Clarke et al., (2013) had shown that the bead array technology assay covers the *B. napus* genome with SNPs distributed evenly and the SNPs had a density of 1 SNP per 15 Kb of the *B. napus* genome, a high density of SNPs had been observed only in the A genome and not in the C genome of *B. napus*, and large intervals in the C genome were identified (Liu et al., 2013b). In this study, the results showed that SNPs were distributed evenly across all the A and C chromosomes with no large gap larger than 15 cM. For chromosome C9, 47 bins were identified on the PISD genetic map while there were only 9 SNP bins on C9 using the Illumina Bead Array technology (Liu et al., 2013b). Another challenge in the Illumina Bead Array technology is that the SNPs may be observed at different orthologous positions in different *B. napus* genotypes due to prevalent homologous recombination events. The array does not focus on functioning SNPs in the *B. napus* genome, which results in bias of the SNP marker distribution (Clarke et al., 2013).

Since resequencing *B. napus* whole genome for SNP identification is challenging and time consuming, GBS is performed using partial genome sequencing through adequate selection of restriction enzymes (RE) such as *Eco*RI, *Bam*HI, *Pst*I, and *Ape*KI (Deschamps et al, 2012; Arif Uz Zaman, 2014). However, multiple GBS libraries are necessary, which increases labour, experimental time, and cost (Kim et al., 2015). In the preparation of GBS

libraries, DNA digestion with REs and size selection are necessary (Elshire et al., 2011). In addition, complete RE digestion is affected by DNA methylation, which results in incomplete digestion (He et al., 2014). Incomplete RE digestion reduces repeatability among DNA samples; thus, reducing the efficiency in GBS. Many SNPs are eliminated due to missing data in some samples since cross sample comparison is not possible if allelic sequence reads are missing in some samples (Elshire et al., 2011; He et al., 2014).

For a more simplified and convenient set up for PISD library construction, DNA samples from 94 ZT DH genotypes and two parental genotypes (ZY821 and Topas) were selected based on the previous study (An et al., 2019). In the previous study, digestion of ZT DNA by REs at the beginning of the GBS library construction protocol was performed to reduce the genome complexity of the plant. In comparison to GBS, the construction of PISD libraries was achieved through two rounds of PCR. In addition, the size selection of DNA was already fixed during the PCR steps. In the PISD method, specific DNA size fragments are amplified during PCR.

The data from the PISD were compared with the data collected through the standard GBS protocol. With similar data, 1,092 SNPs from the PISD were assigned on the genetic map while only 729 SNPs from the standard GBS were identified. We observed that with the combination of SNPs from the standard GBS and the PISD, the overall size of the combined genetic map was larger than the genetic map constructed with the SNPs using only the PISD protocol, suggesting that there were score errors in the data of standard GBS as the number of SNPs increased.

Compared to the previous data (Sun et al., 2007), 13,551 SRAP markers were assigned and the sizes of the A genome and C genomes were 829.1 cM and 775.7 cM, respectively. Similar results for both the PISD and combined genetic maps showed that the A genome was larger than the C genome. In another study, 5,306 SNPs were assembled into a genetic map for *B. napus* with the sizes of 1,333.80 cM and 1,180.99 cM for the A and C genomes, respectively

(Raman et al., 2014). The PISD and combined maps in this study were similar to others (Sun et al., 2007; Raman et al., 2014). In contrast, Liu et al., (2013) constructed a genetic map with 1,220 bins to perform QTL mapping for glucosinolates and Sclerotinia (*Sclerotinia sclerotiorum* (Libert) de Bary, 1837) using Illumina Bead Array. Although 8,839 SNPs were identified, 39 SRAP markers were necessary to fill the gaps to produce the genetic map with 1,220 bins. This means, although there were many SNPs used in the experiment, only 1,220 markers were allocated to specific loci for QTL mapping.

Even with the low number of genotypes in the ZT DH population (<100 genotypes), the PISD protocol was able to generate a genetic map without using the REs to reduce the complexity of the plant genome. The size selection was based on PCR. The PISD protocol is less complex compared to the preparation of libraries in GBS, however, it is cost effective and able to produce similar results compared to GBS genotyping. Moreover, this is the first study of GBS without using the REs, a simple modification from the GBS protocol.

# Chapter 4.0 IDENTIFICATION OF QUANTITATIVE TRAIT LOCI FOR OIL CONTENT IN *BRASSICA NAPUS* L.

## 4.1 Abstract

Using single nucleotide polymorphism (SNP) data of a doubled haploid (DH) population generated from PCR-based Illumina SNP detection (PISD) and genotyping-by-sequencing (GBS), two genetic maps with 1,092 and 1,821 SNPs, respectively, were constructed and used for quantitative trait locus (QTL) mapping. The seeds of the DH population and the parent genotypes were harvested from four field experiments at the Point Research Station (University of Manitoba), and four replicates of greenhouse plantings in the Department of Plant Science, University of Manitoba. Seed oil content (SOC) was determined using near-infrared spectroscopy (NIR). Six and 8 major QTL for SOC were identified using the PISD and combined genetic maps with 10,000 permutation tests, respectively. Major QTL were identified at chromosomes A8, C2, C3, C7, and two at C8 for PISD, and A8, A10, three at C1, C2, and two at C3 for the combined genetic map, respectively. In the A genome, QTL with logarithms of the odd (LOD) scores of 3.30 and 3.33 were identified on chromosome A8 and there was a 4 cM difference between the PISD genetic map and the combined genetic map. On the C genome, the highest LOD score of 4.67 was observed on chromosome C1 with the negative additive effect of -1.37 from the Field2017 combined genetic map. The QTL on chromosomes A8 and C3 identified in the PISD and combined genetic maps overlapped with the physical positions of *FAE1.1* and *FAE1.2* for erucic acid, indicating that these two genes may contribute to the SOC in the ZT DH population. Two major QTL identified on chromosome C2 had different additive effects which indicates the data might be affected solely by the environment.

## 4.2 Introduction

*Brassica napus* L. is one of the most important oilseed crops in the world. This plant species contains ten pairs of A and nine pairs of C chromosomes, named as the AACC genome (2n=38) (Stefansson et al., 1961; Parkin et al., 1995; Parkin et al., 2005). It was produced through natural interspecific hybridization between two closely related species – *B. rapa* L., which contributed the A genome, and *B. oleracea* L., which contributed the C genome (Nagaharu, 1935). *Brassica napus* oilseed is also known as canola or rapeseed depending on the erucic acid content (C22:1) and glucosinolate content in the seed. It is an economically important plant as the oil that can be extracted from its seed and made into products such as cooking oil, lubricant, paints and other products (Canola Council of Canada, 2019).

The biosynthesis of oil which is important for the study of SOC in *B. napus,* has been well characterized in *Arabidopsis thaliana* L., (Zhao et al., 2012). The fatty acid elongase 1 or 3-ketoacyl-CoA synthase gene (*FAE1*) was first discovered in *A. thaliana* where the *FAE1* genes are located in several different loci on chromosome four. Fatty acid elongase 1 catalyses the elongation of erucic acid during erucic acid synthesis in the seed (Kunst et al., 1992). Two homologs of *FAE1* were identified in *B. napus* and other species, such as *B. rapa*, *B. oleracea*, *Zea mays* L. (corn), and *Oryza sativa* L. (rice) (Fourmann et al., 1998; Rossak et al., 2001; Schreiber et al., 2005; Ying et al., 2012). The *FAE1* gene in *Arabidopsis* functions during the elongation of fatty acids from oleic acid (C18:1) to erucic acid (C22:1), from arachidic acid (C20:1) to erucic acid, and during the biosynthesis of very-long fatty acid (VLCFA) from arachidic acid to cerotic acid (C26) (Rossak et al., 2001). In *B. napus, BnFAE1.1* and *BnFAE1.2* were identified on chromosomes A8 and C3 respectively (Qiu et al., 2006; Zhu et al., 2019). Other genes discovered in *A. thaliana* that play important roles in the fatty acid biosynthesis pathway are the omega-6 fatty acid desaturase and omega-3 desaturase genes (*AtFAD2* and *AtFAD3*) (Okuley et al., 1994; Shah and Xin, 1997). The *AtFAD2* gene is responsible for 90% of the polyunsaturated fatty acid (PUFA) synthesis during seed development in oilseed crops

like *B. napus*, *Glycine max* L. (soybean), *Gossypium hirsutum* L. (cottonseed), and *Helianthus annuus* L. (sunflower) (Canvin, 1965). The *AtFAD3* gene encodes omega-3 desaturase, which functions in the biosynthesis of linolenic acid (C18:3) by introducing a double bond into oleic acid in *A. thaliana* (Shah and Xin, 1997).

Quantitative trait locus (QTL) analysis is a statistical analysis method that is used to perform a correlation between each marker and phenotype (Kearsey and Faequhar, 1998). The logarithm of the odd (LOD) is used to indicate which correlations are significant (Chen et al., 2010). The differences in the phenotypic traits must be present across the two parents to generate a segregating population for QTL studies. Many QTL studies have been conducted to identify the locations on genetic maps that are related to SOC by using different molecular markers such as restricted fragment length polymorphism (RFLP), simple-sequence repeat (SSR), sequence-related amplified polymorphism (SRAP), and single nucleotide polymorphism (SNP) (Burns et al., 2003; Qiu et al., 2006; Liu et al., 2013b; Liu et al., 2016). A combination of QTL analysis and genome-wide association studies (GWAS) by Menard et al., (2017) showed that *AtFAD2* and *AtFAD3* were inhibited by high temperatures and the QTL analysis for heat stress responsiveness showed that they were located in an approximately 450 kb region of chromosome three in *A. thaliana*. Four orthologs of *FAD2* and six of *FAD3* were identified in *B. napus* using restriction fragment length polymorphisms (RFLP) analysis (Scheffler et al., 1997; Yang et al., 2012). A recent genome distribution study of the *Brassica* species by Yang et al., (2012) found *BnFAD2* genes on chromosomes A1, A5, C1, and C5 while *BnFAD3* genes were identified on chromosomes A3, A4, A5, C3, and C4. A previous study on *B. napus* by Chen et al., (2010) was conducted with the DH population derived from the high and low SOC genotypes for genotyping, genetic map construction, and QTL analysis for *B. napus* SOC by using SRAP and SSR markers. Chen et al., (2010) detected a large number of QTL with minor effects for SOC in *B. napus*. However, detections of those minor effects in the QTL analysis occurred because *B. napus* is highly sensitive to environmental conditions.

Twenty-seven QTL were identified for SOC in different environments in the field; only 4 QTL related to the SOC were identified as reproducible in at least two environments (Chen et al., 2010).

Even with significant SOC QTL studies previously completed, the populations and maps available affect the QTL results. Therefore, objective of this research is to utilize the newly developed genetic maps produced from PCR-based Illumina sequencing detection (PISD) along with combined SNP markers from An et al., (2019) to conduct a QTL analysis for *B. napus* SOC in field and controlled environmental condition studies.

## 4.3 Materials and Methods

### 4.3.1 Plant material and seed production

Ninety-four ZT DH genotypes from a cross between Zhongyou821 and Topas were phenotyped for SOC in the field and greenhouse. The ZT DH population was produced from microspore culture as described by Weber et al., (2005). The female parent Topas is a spring-type canola cultivar developed in Sweden with 43% SOC (Oplinger et al., 1989). Zhongyou821 (ZY821) is a semi-winter-type double high rapeseed Chinese cultivar with 37% SOC. This semi-winter-type cultivar requires two months of vernalization to induce flowering. Preliminary testing of 94 ZT DH genotpyes in a controlled environmental greenhouse conditions was conducted to identify the semi-winter genotype in the DH population.

The semi-winter type ZT DH genotypes that were identified previously were seeded in 96-cell flats in soilless mix (Sunshine® Mix#4, Sun Gro Horticulture Canada Ltd, Canada) for two weeks under a 22/16$^{o}$C, and 16/8hr day/night controlled environmental growth room conditions. The semi-winter plants were fertilized with 20-20-20 water-soluble fertilizer (Plant-Prod, Leamington, Canada) at a rate of 200 ppm once every two weeks. The rate of watering was twice every week to prevent the soil from drying. After fertilization, the semi-winter plant flats were transferred into an 8±2$^{o}$C cold room with fluorescent light for two

months. The 96-well flats in the cold room were checked frequently for decolouration, aphids, and powdery mildew. After two months of cold treatment, the flats were transferred into the greenhouse for light, humidity, and temperature acclimatization for a week.

The ZT DH genotypes along with the parental genotypes, Topas and ZY821, were seeded in 96-cell flats for seed multiplication. The plants were transplanted into six-inch standard pots in a 2:2:1 mix of topsoil, sand, and spahagnum peat moss. The plants were grown in the Department of Plant Science greenhouse at the University of Manitoba, Winnipeg, Manitoba. The maintenance of the plants was the same as previously described. Once the plants began to flower, they were covered with pollination bags to ensure self-pollination and seeds were harvested at maturity.

### 4.3.2 Field experiments

Field experiments were conducted at the Point Research Station, University of Manitoba (49.82$^{o}$N/ 97.12$^{o}$W). The soil at the Point Research Station is classified as sandy clay loam. The experiments were set up in a completely randomized block design with two replicates in 2015, 2016, 2017, and 2018, for a total of four site years (Field2015, Field2016, Field2017, and Field2018). The two parental genotypes, Topas and ZY821, were planted in each replicate as controls. Semi-winter plant maintenance was the same as previously mentioned in section 4.3.1. As for the spring-type ZT DH genotypes, the seeds were seeded directly to the field. The semi-winter ZT DH genotypes and ZY821 were hand-transplanted to the field at the Point Research Station after two months of vernalization. Eight plants from each semi-winter ZT DH genotype and ZY821 were hand-transplanted to the field where spring-type DH genotypes and Topas were seeded earlier. All hand-transplanted plants were hand-watered with approximately 2 L of water per experimental row every morning for three weeks during late May to early June in 2018. Compared to previous studies, in this study vernalization and hand-transplanted methods were used only for semi-winter genotypes in 2018 because this

method allowed for a better flowering rate when compared with plants that did not go through vernalization in previous years.

Approximately 0.35 g of seeds from spring-type DH genotypes and Topas were seeded in a three-meter row with 0.4 m row spacing for both replicates. When the semi-winter DH plants were ready to be hand-transplanted to the field, they were placed in empty rows of both two replicates based on the experimental design. Electrical fences were installed to prevent herbivory by deers, Helix-treated corn grit (Syngenta Basel, Switzerland) was used to control flea beetles, and Muster ® (DuPont, Wilmington, DE, USA) and Lontrel ® (Dow Agroscience, Midland, MI, USA) were used for weed control. Lance ® (BASF Canada Inc, Mississauga, ON, CA) was applied during the 20 to 50% flowering stage to control for Alternaria black spot by *Alternaria Brassicaceae* Berk. and Sclerotinia stem rot by *Sclerotinia sclerotiorum* (Libert) de Bary, 1837. All pesticides were applied according to the manufacturer's specified rates. At the beginning of senescence stage when the moisture content of the plant was approximately 10-15%, 10-15 g sample of seed was hand-harvested from each experimental row at the end of the summer by crushing the pods in paper bags.

### 4.3.3 Greenhouse experiments

Ninety-four ZT DH genotypes and both parental genotypes, Topas and ZY821, were planted in a completely randomized block design, and the one experiment was replicated across two greenhouses at the Department of Plant Science greenhouse, University of Manitoba. All semi-winter DH genotypes and ZY821 were vernalized in the same fashion as the semi-winter plants in field experiments after two months in the cold room and were hand-transplanted in six-inch pots filled with the same ratio of two parts sand, two parts topsoil, and one part sphagnum peat moss mixed with approximately 30 g of slow-release fertilizer (PlantProd ®, ON, Canada) per one half-filled pot of soil in a 150 L concrete mixer.

The spring type DH genotypes and Topas were seeded in 96-cell flats 14 days before the winter plants were ready to be transplanted, so all plants were transplanted into pots at the

same time. There was a total of four replicates for the greenhouse experiment and each replicate consisted of four potted plants of each DH genotype in order to collect approximately 10 g of seed per replicate for near-infrared spectroscopy (NIR). The fertilization of the plant was described previously in section 4.3.1. The plants were watered with an automatic dripping irrigation system twice per day for one hour at 10:00 a.m and 6:00 p.m. Approximately 5 mL of Verimark $^{TM}$ insecticide (Dupont, Canada) per 250 plants was applied after the transplant. *Aphelinus abdominalis* (Dalman) and sulphur fumigation were applied at all growing stages to control for aphids, thrips, and powdery mildew. Approximately 6 g of the seed of all four plants from the same line were harvested, combined and placed in put in paper bags, and the seeds were dried in an oven at 40$^{o}$C for one day to reduce seed moisture content to 1-5% before SOC analysis using NIR.

### 4.3.4 Seed oil content measurement

Seeds harvested from field and greenhouse experiments were cleaned using a seed blower. A gentle air was blown through the seed to remove the debris and approximately 10 g of seed was collected for near-infrared spectroscopy (NIR) analysis. Near-infrared spectroscopy is the Seed oil content was measured using a FOSS 6500 NIR system (Foss NIRSystems Inc., Maryland, USA) following the approved guidelines from American Oil Chemist's Society (AOCS) for SOC in the quality control laboratory at the Department of Plant Science, University Of Manitoba. The electromagnetic spectrum was set at 780 to 2500 nm, and SOC was read in percentage using the protocol for *B. napus* described previously (Daun, 1994; Daun et al., 1994; Kim et al., 2007). Five grams of seed from each replicate ZT DH and the parental genotypes were used to determine SOC.

### 4.3.5 Statistical evaluation of field and greenhouse seed oil content

The SOC obtained from the NIR analyses from the four field experiments and controlled environmental greenhouse experiments were analyzed separately using ANOVA

SAS ® University Edition 2019 (SAS Institute, NC, and USA). Data were combined based on the replicates effect and significant different from ANOVA.

However, not all ZT DH genotypes were able to produce seed at the end of the growing season due to pest damage or plant death which led to missing data. Regression analysis is used during data analysis by predicting the missing data from other replicates and if both data from the replicates were missing, the data was eliminated and not included in the ANOVA.

### 4.3.6 Quantitative trait locus analysis

The genotype data (from Chapter 3) and the SOC data were pooled and uploaded into JoinMap (version 4) as a quick data file (QDF) format excel spreadsheet. After the ANOVA, five sets of SOC data (Field2015, Field2016, Field2017, Field2018, and greenhouse (GH)) were analysed with both the PISD and combined (GBS and PISD) genetic maps (Tinker and Mather, 1995; Wang et al., 2013). Quantitative trait locus analysis was conducted using Qgene (version 4.4.0), following the protocols described in Joehanes and Nelson, (2008). Simple interval mapping (SIM) was applied, and the additive effects of both the parents were added into the QDF file. The positive additive effect was set as Topas while the negative additive effect was set as ZY821. The parameter of the recombination rate was set to two cM. The threshold LOD value for determining the significance of QTL peaks for the PISD and combined genetic maps were calculated using 10,000 permutation tests ($\alpha$=0.05). Based on the alpha value of 0.05, QTL with a mean LOD score greater than or equal 3.0 to 3.2 were considered major QTL.

## 4.4 Results

### 4.4.1 Seed oil content

The parental genotypes, Topas and ZY821, had SOC means of 43.21% and 37.80% across all field experiments and 41.39% and 35.70% in the greenhouse experiments, respectively (Figure 4.1). Mean SOC value 38.89% for the GH experiment and 38.87 to 39.58%

from all the field experiments were observed (Table 4.1). The minimum SOC data for the GH experiment was observed at 32.27%, and SOC data ranges from 32.71 to 33.60% were observed from the four field experiments. The maximum SOC GH experiment and four field experiments were 43.94 and 45.11%, respectively.

For the standard error of the mean (SE), the lowest SE in GH was 0.25 and the highest SE was observed in Field2017 experiment at 0.29. The highest coefficient of variation (CV) in Field2017 was 6.56%, and the lowest CV of 6.21% was in the GH experiment (Table 4.1).

In the ANOVA analysis, if the P-value is less than the P critical value or the F-value is larger than the F-critical value, the null hyphothesis is rejected. Based on Table 4.2 and the calculation from ANOVA, the null hypothesis was rejected. In consequence, the data were significantly different. All the data from Table 4.2 did not happen by chance.

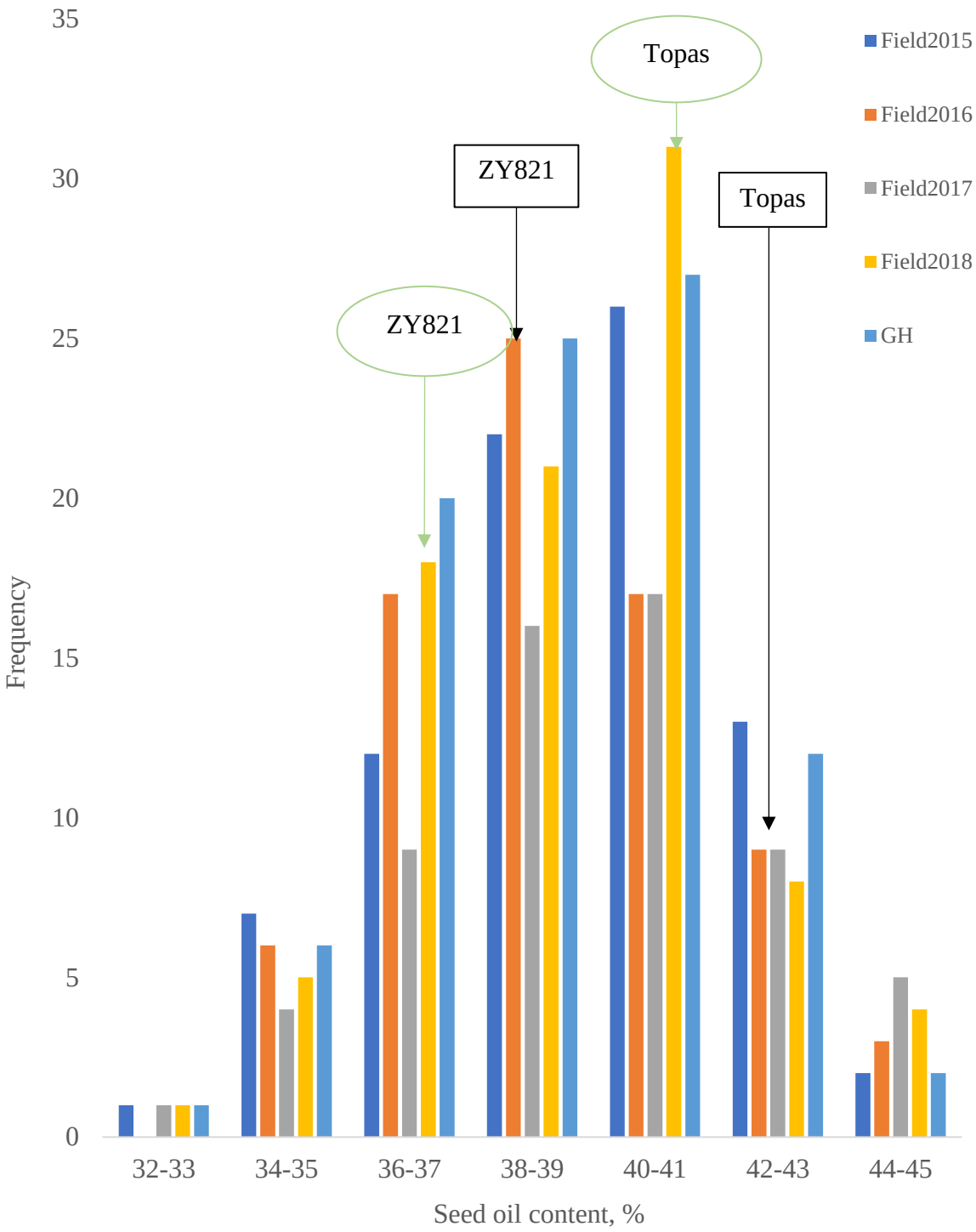


Figure 4. 1: Frequency distribution of the mean seed oil content for all the experiments in the greenhouse and field experiments for ZT doubled haploid population. The arrow indicates the mean seed oil content of the parents, ZY821 and Topas from the field and GH experiments. Green circles indicate the parental genotypes from the greenhouse and the black boxes indicate the parental genotypes from the field.

Table 4.1: The seed oil content data for both two replicates in four-year field experiments and four replicates greenhouse (GH) and for ZT DH population.

| Experiments | Field2015 | Field2016 | Field2017 | Field2018 | GH |
|---|---|---|---|---|---|
| n | 90 | 91 | 73 | 91 | 93 |
| Mean | 39.23 | 38.87 | 39.58 | 39.13 | 38.89 |
| Std dev | 2.50 | 2.52 | 2.60 | 2.44 | 2.41 |
| Variance | 6.26 | 6.33 | 6.75 | 5.97 | 5.83 |
| CV | 6.38 | 6.47 | 6.56 | 6.24 | 6.21 |
| SE | 0.27 | 0.28 | 0.29 | 0.27 | 0.25 |
| | | | | | |
| Min | 32.71 | 33.60 | 33.44 | 33.11 | 32.27 |
| Q1 | 37.49 | 37.18 | 38.30 | 37.27 | 36.86 |
| Median | 39.33 | 38.69 | 39.56 | 39.32 | 39.21 |
| Q3 | 41.10 | 40.81 | 41.41 | 40.92 | 40.75 |
| Max | 44.28 | 43.94 | 44.88 | 45.11 | 43.95 |

Table 4.2: Analysis of variance for seed oil content (%) for four field experiments and a controlled environmental experiment for ZT doubled haploid genotypes of *Brassica napus*.

| Source | DF | Sum of Squares | Mean Square | F-value | P-value |
|---|---|---|---|---|---|
| Field 2015 | | | | | |
| Genotype | 89 | 1096.01 | 16.81 | 95.74 | <.0001** |
| Rep | 1 | 2.59 | 2.59 | 19.62 | <.0001** |
| Error | 89 | 13.93 | 0.75 | . | . |
| Field 2016 | | | | | |
| Genotype | 90 | 1236.32 | 13.74 | 10.78 | <.0001** |
| Rep | 1 | 7.05 | 7.05 | 5.53 | 0.0023* |
| Error | 65 | 46.57 | 1.27 | . | . |
| Field 2017 | | | | | |
| Genotype | 72 | 802.00 | 13.87 | 68.68 | <.0001** |
| Rep | 1 | 1.60 | 1.60 | 9.24 | 0.0040* |
| Error | 64 | 15.07 | 0.72 | . | . |
| Field 2018 | | | | | |
| Genotype | 90 | 1097.28 | 15.53 | 58.60 | <.0001** |
| Rep | 1 | 2.23 | 2.23 | 11.99 | <.0001** |
| Error | 80 | 17.27 | 0.22 | . | . |
| GH | | | | | |
| Genotype | 92 | 2073.34 | 20.11 | 79.08 | <.0001** |
| Rep | 3 | 0.80 | 0.66 | 2.55 | 0.0567* |
| Error | 232 | 72.16 | 0.24 | . | . |

****** Statistically significant at the 1% level; * at the 5% level.

### 4.4.2 Quantitative trait locus analysis

Based on QTL analysis (Table 4.3), six and eight major QTL that correlated with *B. napus* SOC were identified using simple interval mapping for the PISD and combined (GBS and PISD) genetic maps, respectively. For the PISD genetic map, six major QTL were identified on chromosomes A8, C2, C3, C7, and C8 from Field2015, Field2016, and Field2017 – two out of six of major QTL were identified on chromosome C8. Eight major QTL were detected for the combined genetic map on chromosomes A8, A10, C1, C2, and C3 from Field2015, Field2016, Field2017, Field2018, and GH – three out of eight of the major QTL were identified on chromosome C1 and two out of the eight were identified on chromosome C3, one QTL each was identified on the remaining chromosomes. Overlapping QTL for both the PISD and combined genetic maps on chromosomes A8 and C3 were also observed. Peak LOD scores lower than 3.0 or 3.2 in some cases were not considered QTL based on 10,000 fold permutation tests. The experiments at different years with peaks lower than the threshold were also identified based on Figures 4.2 and 4.3. The phenotypic variance (R2) ranging from 16-27% for PISD and 11-24% for the combined genetic map was explained. Major QTL from the GH were only indentified in the combined genetic map and 12-13% of phenotypic variance was explained.

For the PISD genetic map, a peak LOD score of 3.76 on chromosome C3 and the lowest LOD score of 3.03 on chromosome C2 were observed (Table 4.3). The largest confidence interval was observed at 4 cM on chromosomes A8, C3, C7, and C8 while the smallest difference was observed to be 2 cM on chromosome C2 and C8 for the PISD genetic map. No positive additive effect was identified and the negative additive effects ranged from -1.38 to -1.04 for the PISD genetic map QTL analysis. All six major QTL were identified from the field experiments – none were identified on Field2018 – while no QTL were identified from the greenhouse experiments. The highest phenotypic variance, 27%, was explained by the QTL on

chromosome C3 in Field2017 while the lowest phenotypic variance (16%) was explained by the QTL on chromosome C2 in Field2016. Two different major QTL were identified on chromosome C8 at positions 26 cM and 32 cM.

For the combined genetic map, a peak LOD score of 4.67 on chromosome C1 and the lowest LOD score of 3.0 on chromosome C2 were observed (Table 4.3). The largest confidence interval was identified on chromosome C1, ranging from 46-56 cM, while the lowest difference was observed at 2 cM. The negative additive effects for the combined genetic map ranged from -1.37 to -0.93. One positive additive effect (0.83) was identified at chromosome C2 from Field2018. Major QTL were identified from all experiments with the phenotypic variance ranging from 11-24% explained. The highest phenotypic variance, 24% was explained by the QTL on chromosome C1 from Field2017 while the lowest 11% was explained by the QTL on chromomes C2 from Field2018. On chromosome C1, two different major QTL peaks were observed at position 36 cM and 46 cM for Field2017 and GH experiments.

The QTL for the PISD and combined genetic map on chromosome A8 were compared (Tables 4.3). Major QTL identified from chromosomes A8 and C3 from the PISD and combined genetic maps were found in similar locations. The difference of 4 cM was found for the PISD and combined genetic map on chromosome A8. As for chromosome C3, 4 cM and 2 cM differences for the PISD and combined genetic map were observed. Similar location and peak LOD scores were observed on chromosome A8. On chromosome C3, QTL located at similar positions were also observed. The SNP marker locus of the PISD genetic map was covered up to 4 cM, in contrast to a mean locus of 2 cM for the combined genetic map.

Table 4.3: Simple interval QTL analysis for seed oil content of ZT doubled haploid population using single nucleotide marker for PISD and combined genetic map. Field indicated the location of The Point Research Station at University of Manitoba and GH indicated as Department of Agricultural and Food Sciences greenhouse experimental loacation.

| Chromosome | Experiments | LOD | Additive effect | R2 (%) | Confidence interval (cM) | QTL peak position (cM) | Total length (cM) |
|---|---|---|---|---|---|---|---|
| PISD genetic map | | | | | | | |
| A8 | Field2015 | 3.30 | -1.04 | 17 | 22-26 | 24 | 58 |
| C2 | Field2016 | 3.03 | -1.12 | 16 | 14-16 | 16 | 126 |
| C3 | Field2017 | 3.76 | -1.38 | 27 | 92-96 | 94 | 108 |
| C7 | Field2015 | 3.35 | -1.05 | 17 | 18-22 | 20 | 86 |
| C8 | Field2016 | 3.27 | -1.22 | 17 | 25-27 | 26 | 114 |
| | Field2016 | 3.25 | -1.12 | 17 | 30-34 | 32 | 114 |
| Combined genetic map | | | | | | | |
| A8 | Field2015 | 3.33 | -1.03 | 14 | 40-44 | 42 | 76 |
| A10 | Field2016 | 3.54 | -1.12 | 14 | 82-84 | 84 | 144 |
| C1 | Field2017 | 4.67 | -1.37 | 24 | 46-56 | 46 | 136 |
| | Field2017 | 3.24 | -1.10 | 18 | 36-38 | 36 | 136 |
| | GH | 3.20 | -0.93 | 12 | 46-48 | 46 | 136 |
| C2 | Field2018 | 3.00 | 0.83 | 11 | 74-76 | 76 | 98 |
| C3 | Field2018 | 3.68 | -0.97 | 15 | 100-102 | 100 | 126 |
| | GH | 3.50 | -0.96 | 13 | 100-102 | 100 | 126 |

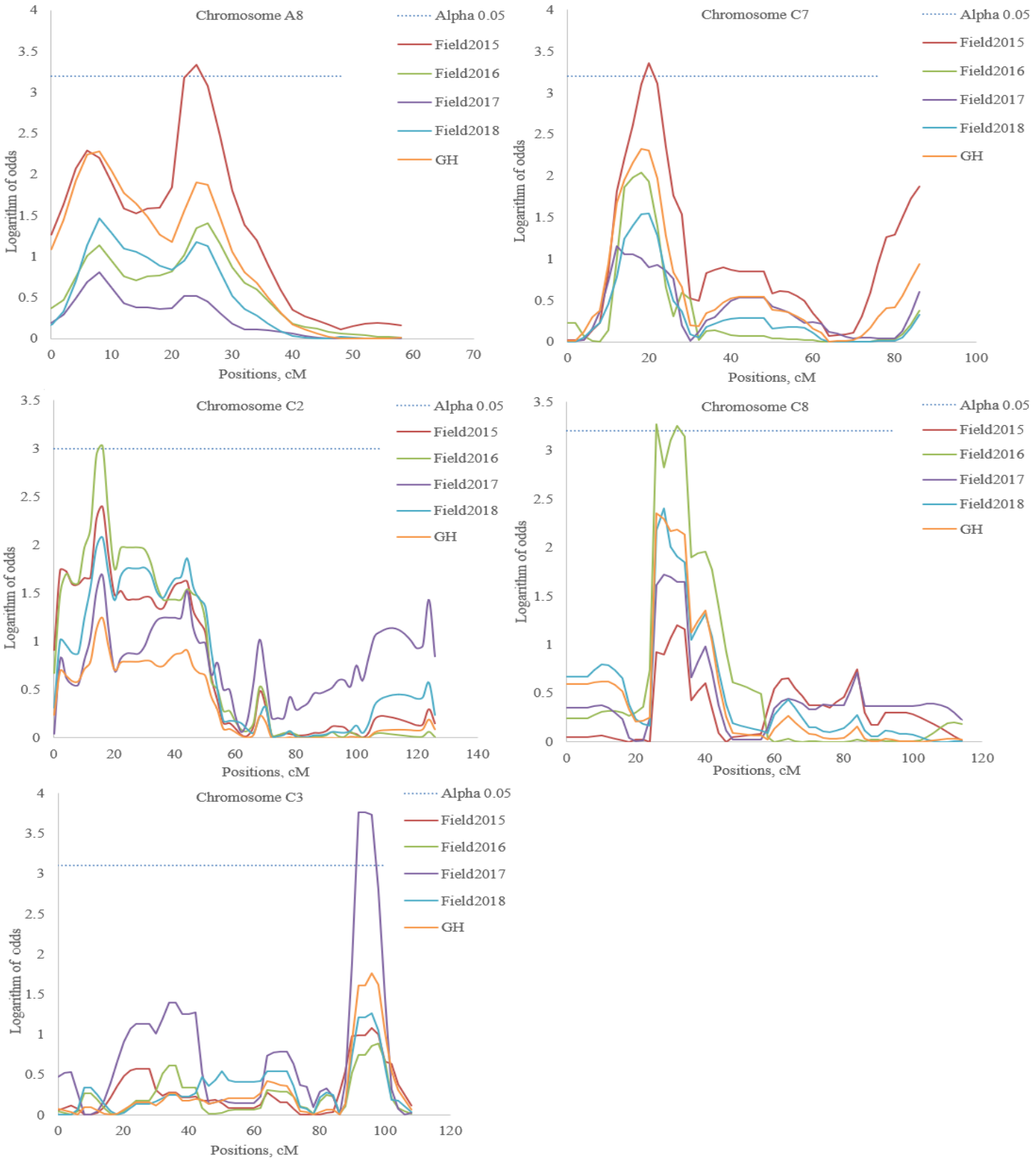


Figure 4. 2: The simple interval QTL analysis of the PISD genetic map for seed oil content for four field experiments and one greenhouse (GH) experiment using Qgene v4.4 on chromosomes A8, C2, C3, C7 and C8 for ZT doubled haploid population population. Dotted lines indicate the threshold at alpha 0.05 for each QTL on the graph.

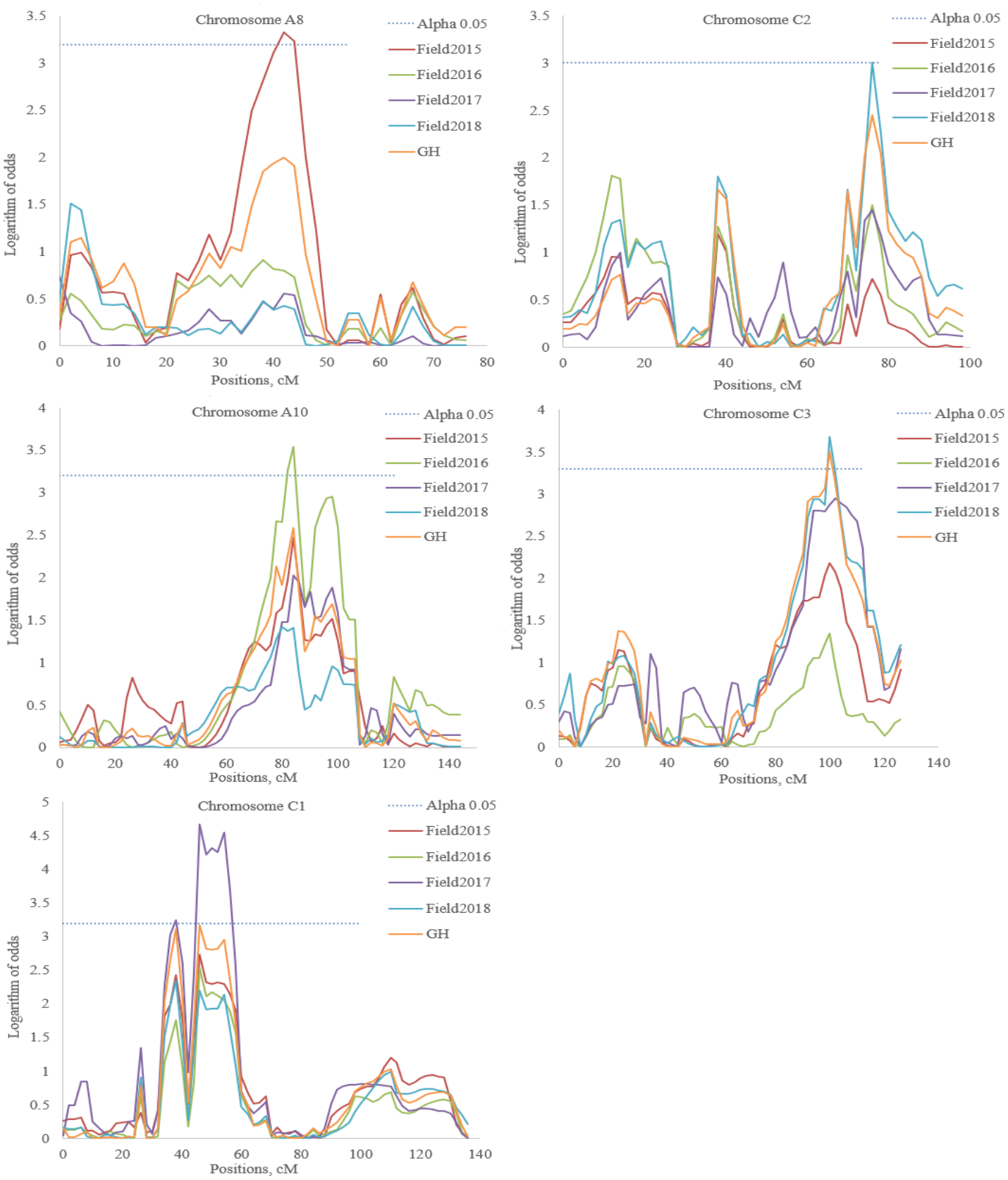


Figure 4. 3: The simple interval QTL analysis of the combined genetic map for seed oil content for four and one greenhouse (GH) experiments from Qgene v4.4 on chromosomes A8, A10, C1, C2, and C3 for ZT doubled haploid population. Dotted line indicated the threshold at alpha 0.05 for each QTL on the graph.

## 4.5 Discussion

### 4.5.1 Seed oil content

In this SOC study of *Brassica napus*, the ZT DH population was grown in field experiments and in controlled environmental conditions in the greenhouse (GH). Poor flowering rates of the semi-winter genotypes were observed during growing season in 2015, 2016, and 2017. Therefore, the process of vernalization was implemented in 2018 to guarantee higher seed production from the ZT DH genotypes. This is the first time that the vernalized semi-winter type ZT DH genotypes were tested in the field in Winnipeg, Manitoba, Canada. A lack of rainfall in early summer was the main environmental stress encountered by the newly transplanted semi-winter genotypes. Therefore, the plants were hand-watered to prevent excess wilting during the severe drought. There were many challenges, including high temperature in the summer inhibiting the flower formation of the ZT DH and parental genotypes, high humidity in the GH causing powdery mildew disease caused by *Erysiphe cruciferarum* Opiz ex L., Junell, 1967, and infestation of pests such as *Schizaphis graminum* Rondani (aphids) and Thysanoptera Haliday, 1836 (thrips), all of which reduced SOC in *B. napus*.

Pest damage by *Psylliodes chrysocephala* L. (flea beetle) was observed on the field experiments throughout the growing season from 2016 to 2018. Elliott et al., (2008) had showed the effect of *P. chrysocephala* during the early growth stage of *B. napus* resulting in plant death while later stage damage by the pest reduced the seed size, weight and SOC of *B. napus*. Plants that were severely damaged by *P. chrysocephala* during the cotyledon stage for the four-year field experiments were not harvested from the field to prevent the bias of reducing the overall SOC from a specific DH genotype. For the GH experiment, *P. chrysocephala* was not observed throughout the experiment. However, powdery mildew disease and infestation of *S. graminum* and

Thysanoptera pests were observed. Similar *S. graminum* damage of the ZT DH genotypes were observed compared to Douglas, (2003) and Sarwar, (2013). The infested ZT DH genotype in the GH experiment was removed and discarded from the experiment because damaged plants results in no production of SOC.

Another observation between the GH and the four-year field experiment was the light. Shengxin et al., (2016) and Zhu, (2016) had proven *B. napus* requires specific a spectrum of blue and red balance light to promote an optimum condition for *B. napus* growth, and poor or low light intensity results in yellowing of the leaf and lower yields and SOC. An optimum amount of sunlight was obtained for the four-year growth season but not in the greenhouse. The use of fluorescent light and non-uniform light exposure in the greenhouse did not promote an optimum growth for the ZT DH genotypes which resulted in lower SOC. A slight decrease of the parental genotypes Topas and ZY821 SOC were observed when compared to the GH and four-year field experiments.

The *B. napus* seed yield and SOC are easily affected by environmental conditions such as temperature (Young et al., 2004; Koscielny et al., 2018). Fayyaz-ul-Hassan et al., (2005) conducted an experiment using different genotypes of *B. napus* at three different locations and observed that the plants produced a higher SOC in cooler temperatures during seed development. The current GH experiment was conducted in controlled environmental conditions, the temperature being set to 22/16$^{o}$C, which may have decreased the normal distribution of SOC. A fluctuation of temperature from 22$^{o}$C to 16$^{o}$C might affect the production of SOC. In contrast, in the Bahrani and McVetty, (2008) study, the GH average temperature experiment was 20$^{o}$C to achieve optimum growth for *B. napus.*

The constant temperature in the greenhouse in this study provided more consistent SOC between replicates. In this study, differences in SOC using genotypic effects ANOVA analysis

were significant from year to year in three field experiments, due to the difference in the environment during day/night cycles in different years.

### 4.5.2 Quantitative trait locus analysis for seed oil content

Fatty acid elongase genes (*BnFAE1.1* and *BnFAE1.2*) are responsible for the production of erucic acid in oilseed rape (Fourmann et al., 1998). The *BnFAE1* sequence from cultivar ZS11 was obtained from NCBI (accession number NC_027764.2) and BLASTn using the *B. napus* genome sequence (http://www.genoscope.cns.fr/*Brassica*napus/) to locate the physical locations on the chromosomes. The physical location of *BnFAE1.1* on chromosome A8 is at 10,187,439 to 10,189,121 bp, and the total physical length of A8 is 18,961,841 bp. The *BnFAE1.2* sequence is located on chromosome C3 from 55,684,024 to 55,685,679 bp and the total physical length of the chromosome is 60,573,294 bp.

By comparing the *B. napus* chromosomal locations of A8 and C3 to the PISD genetic map, a total of two major QTL were identified, one on chromosome A8 and one on C3. The total estimated length for chromosomes A8 and C3 were 58 cM and 108 cM, respectively. The peak LOD score positions were identified at 24 cM and 94 cM, respectively. For the combined (GBS and PISD) genetic map, the total estimated length for chromosomes A8 and C3 were 76 cM and 126 cM; the peak LOD score positions were located at 42 cM and 100 cM, respectively. By comparing our estimated position data with the physical location of the *BnFAE1* gene, we were able to observe that the PISD genetic map showed higher precision on chromosome A8 than the combined genetic map and higher precision for a combined genetic map on chromosome C3 when compared to the previous study by An et al., (2019). Chromosomes A8 and C3 were identified as linked to the *BnFAE1* gene that is responsible for the biosynthesis of erucic acid in *B. napus*.

In contrast to the sudy by Burns et al., (2003), both QTL from PISD and combined genetic maps on chromosomes A8 and C3 were identified. Both locations on the chromosomes are associated to the production of erucic acid content in *B. napus*. Based on the observation, we were able to observe that QTL identified on chromosome A8 was located at the proximal end and chromosome C3 at the distal end as discussed by Burns et al., (2003). In comparison to Burns et al., (2003) the confident interval identified with our QTL analysis on chromosome A8 was 4 cM, versus a 12 cM confident interval from Burns et al., (2003). On chromosome C3, a confident interval range of 2 to 4 cM were identified and 26 cM was observed from Burns et al., (2003).

A study by Qiu et al (2006) identified QTL for SOC on chromosomes N1, N3, N4, N8, N12, N13, and N17 and QTL that were linked to erucic acid on chromosomes N1, N2, N6, N8, N10, and N13. Chromosomes N8 and N13 were identified as linked to *BnFAE1*, which was responsible for the biosynthesis of erucic acid in their study. Qiu et al., (2006) found that the QTL that were associated with SOC were located at the position 18 cM out of 89 cM and 108 cM out of 120 cM on chromosomes N8 and N13 respectively. After comparing with the PISD and combined genetic maps, similar positions of the major QTL were identified at chromosomes A8 and C3, which are linked to *BnFAE1*. Additionally, Qiu et al., (2006) also reported that oil-related genes on chromosome C2 can be very sensitive toward the environment. By comparing with this study, we observed two major QTL on chromosome C2 at different positions on different genetic maps (PISD and combined).

When comparing our data with Chen et al., (2010), six major QTL for SOC in *B. napus* were identified from the PISD genetic map and eight major QTL were identified from the combined genetic map while Chen et al., (2010) had identified 27 QTL related to SOC. The highest LOD of 3.76 from PISD genetic map QTL analysis was identified at chromosome C3 and LOD score of 4.67 in chromosome C1 from the combined genetic map QTL analysis. The highest

LOD score of 5.9 was identified from chromosome A2 from the Chen et al., (2010) study. No QTL were identified on chromosome A2 from PISD and combined QTL analysis and as Chen et al., (2010) had mentioned that QTL on chromosome A2 were correlated to both high SOC and early-flowering QTL. A negative additive effect was observed by Chen et al., (2010) for experiments 07WIN and 08WIN with PISD and combined genetic maps at chromosome C3 suggesting that it is related to one of the parent genotypes – which means that one of our parental genotypes might have been similar to Chen et al., (2010). Although Chen et al., (2010) did not further discuss the QTL identified on chromososme C3, the location of the QTL was similar to our study. Among the QTL that were identified from Chen et al., (2010) that were related to SOC, one interesting observation was about the QTL identified at chromosmes C2 compared to QTL from the combined genetic map. The QTL on chromosome C2 was further discussed by Chen et al., (2010) who observed that the QTL was identified in a nonvernalized condition. In contrast to the combined genetic map, the opposite result was observed where vernalized condition was applied, particularly to Field2018. However, the QTL from chromosome C2 was identified from the PISD genetic map under a nonvernalized condition to Field2016 which is similar to Chen et al., (2010) where the oil-related gene was expressed. With the opposite result of QTL from combined genetic map at chromosome C2, we can conclude that the PISD genetic map was more accurate than the combined genetic map.

A previous QTL study by An et al., (2019) identified two major SOC QTL based on SRAP and SNP markers on chromosomes C7 and C3 for the ZT DH population. In contrast, six major QTL were identified on chromosomes A8, C3, C7, C8, and C9 from the PISD genetic map using a modified GBS protocol. More QTL were identified with a close relationship between the gene position of *BnFAE1* and the estimated chromosomal position of A8 and C3. This indicates that the PISD method was more accurate than the GBS from the previous study. With the increase in SNP

markers in the combined genetic map, more QTL that are linked to SOC were identified. This resulted in a larger confidence interval range.

A genome-wide association study (GWAS) by Zhu et al., (2019) had identified 30 SNP loci associated to SOC on chromosomes A8 and C3 of *B. napus* using 300 inbred lines with four important fatty acid traits including erucic acid, oleic acid, linoleic acid, and linolenic acid. In the study, 201,187 SNP markers were developed for the identification of the specific oil gene responsible for the synthesis of SOC in *B. napus.* More than 80% of the SNPs were identified on chromosome A8 and C3 that were associated to the production of SOC. Followed by a discussion of these genomic regions, both the A8 and C3 genomic regions in *B. napus* were related to the biosynthesis of the fatty acid which linked to the *BnFAE1.1* on chromosome A8 and *BnFAE1.2* on chromosome C3. Similarly to our study, QTL for A8 and C3 were identified in both PISD and combined genetic maps. Therefore, it is highly possible that this specific region of the *B. napus* genome for ZTDH population was associated to the *BnFAE1.1* and *BnFAE1.2*. Furthermore, Zhu et al., (2019) GWAS also identified a few other important regions for SOC of *B. napus* – this includes genomic regions A6, A9, A10, C2, C6, C7, and C8. In comparison to our study, QTL were identifed on chromosomes C2, C7, and C8 for the PISD genetic map and A10, C1, and C2 for the combined genetic map which could be importance for the sysnthesis of the SOC in ZTDH population. Further investigation of these genomic region is required to accurately identify the specific gene responsible for biosynthesis of the SOC in *B. napus*.

Potential SOC QTL were identified on chromosomes A10, C1, C2, C7, and C8 from the PISD and combined genetic maps which requires further genetic study to identify the traits in *B. napus*. Although QTL were not identified from all experiments, LOD values were observed at the same location on chromosomes A8 and C3 for all experiments. Most likely due to environmental differences, the QTL were not identifiable and did not reach a level of statistical significance.

## Chapter 5.0 GENERAL DISCUSSION

*Brassica napus* L. contains both the A and C genomes (2n=38), and is an important oil crop for making edible oil, biofuel, machinery lubricant, and animal feed (Nagaharu, 1935; Issariyakul et al., 2008). Developing a cultivar to produce a high amount of seed oil content (SOC) in *B. napus* is a general objective in canola or rapeseed breeding programs. With the many benefits of canola oil consumption, the demand for the oil is increasing (Ecke et al., 1995; Thormann et al., 1996; Burns et al., 2003; Delourme et al., 2006; Chen et al., 2010).

In this study, single nucleotide polymorphism (SNP) markers are utilized because they represent high-density, low cost and high-throughput markers. Single nucleotide polymorphism markers are the main sources used for quantitative trait loci (QTL) studies when dissecting complex quantitative traits such as SOC (Liu et al., 2013a). Two sequencing data sets were constructed from PCR-based Illumina SNP detection (PISD) and genotyping-by-sequencing (GBS) methods. The QTL analysis using the combined genetic map showed a higher number of major QTL peaks when compared to the PISD genetic map QTL analysis. This was mainly due to the increase in high throughput SNP markers from the previous study. There were two main observations in this study. First, the PISD protocol gave better results than the genotyping-by-sequencing (GBS) protocol used in previous studies based on the confidence interval of the QTL analysis (An et al., 2019). In contrast to Chen et al., (2010), we observed that PISD QTL analysis on chromosome C2 was similar. The PISD genetic map for genomes A and C was similar to other studies (Sun et al., 2007; Liu et al., 2013b; Javed et al., 2016). Second, the increased SNP markers did not show better or worse results based on the PISD and combined genetic map. It did not significantly affect the overall size of the genetic map. The main difference between the PISD and

combined genetic maps were the number of SNP markers. The combined genetic map QTL analysis had an increase in SNP markers, and therefore more QTL were identified.

Temperature is a major factor that affects the SOC in *B. napus*, which led to differences in the field and the greenhouse experiments. (Fayyaz-ul-Hassan et al., 2005). Following the method by Bahrani and McVetty, (2008) to sustain an average greenhouse temperature of 20$^{o}$C, this improved the SOC data. Other than that, the light intensity, precipitation, amount of macro and micronutrients in the soil, and biotic factors from the field and greenhouse were totally different, which resulted in minor differences in SOC when an ANOVA was conducted (Maheswaran, 1998). Further experimenting and improvement of light bulbs in the greenhouse with wavelengths closer to the blue and red spectrum improves the quality of seed, and therefore, improves the SOC of *B. napus* (Hart, 2012; Rondanini et al., 2014).

Two important genes that contribute to the biosynthesis of erucic acid, *BnFAE1.1* and *BnFAE1.2,* were identified in both PISD and combined genetic maps on chromosomes A8 and C3. Unfortunately, no QTL were identified on the chromosomes associated with *FAD2* and *FAD3* (normally linked with the production of linolenic acid) in this study. A further investigation and breakdown of the complexity of the genes is required to identify the genes associated with SOC.

The discovery of the QTL location for *BnFAE* gene on chromosomes A8 and C3 through the PISD method allows a better understanding of the region on the chromosome. There are still many oil-related genes in *B. napus* that are unknown (Li et al., 2006; Li et al., 2015). With the help of advanced technologies, new methods are being developed, and more research work will be conducted to dissect the genetic basis and foundation for the synthesis of SOC in *B. napus* or other *Brassica* species.

## Chapter 6.0 FUTURE RESEARCH RECOMMENDATIONS

Since the quantitative trait loci (QTL) for the seed oil content (SOC) in *Brassica napus* is easily affected by environmental factors, increasing the DH genotypes in the ZT population for future QTL study will be required. Increasing the number of genotypes in the ZT DH population may increase the accuracy and precision of the QTL analysis. Repeating the greenhouse study with the same ZT DH population to compare with older studies can prove the usefulness of greenhouse data for QTL study. Furthermore, conducting the experiment in different fields will increase the accuracy and precision of the SOC data for QTL analysis. The experiment needs to be conducted at numerous field environments across western Canada.

Another recommendation would be to repeat the PCR-based Illumina SNP detection (PISD) method on the same DH population and compare it to previous studies. This would prove whether or not the PISD protocol will produce consistent results. The PISD method can be used on other populations, such as recombinant inbred lines (RILs) and other DH populations that have differences in SOC. Furthermore, repeating the PISD on *B. napus* for different phenotypic traits, such as flowering date, seed coat colour, disease resistance, protein content, and pod length, will prove that the protocol works on other phenotypic traits for *B. napus*.

In conclusion, these recommendations will ensure that the QTL study for SOC or other phenotypic traits in *B. napus* coupled with the PISD method will be cost-effective and less time consuming. The knowledge gained and these recommendations will provide the framework for future study on SOC, which will be able to assist and benefit the *B. napus* breeding programms.

## REFERENCES


Abdellatif, A. M. (1972). Cardiopathogenic effects of dietary rapeseed oil. *Nutrition Reviews*, 30(1), 2-6.

Abdellatif, A. M. M., Starrenburg, A., and Vles, R. O. (1972). Effects of hardened palm oil and protein on the pathological and hematological characteristics of ducklings fed rapeseed oil. *Annals of Nutrition and Metabolism*, 14(1), 17-27.

Agriculture and Agri-Food Canada. (2019, June 11). Retrieved from http://sis.agr.gc.ca/cansis/taxa/cssc3/chpt05.html

Aksouh-Harradj, N. M., Campbell, L. C., and Mailer, R. J. (2006). Canola response to high and moderately high temperature stresses during seed maturation. *Canadian Journal of Plant Science*, 86(4), 967-980.

Alam, M. J., Alamin, M., Hossain, M. R., Islam, S. M. S., and Mollah, M. N. H. (2016). Robust linear regression based simple interval mapping for QTL analysis with backcross population. *Journal of Biological Science*, 24, 75-81.

An, B., Li, G., Ayele, B., Duncan. R. W., and Schroeder. D. (2019). Identification of QTL for seed weight, flowering time, and oil content in oilseed rape (*Brassica napus* L.). (Master's thesis, University of Manitoba). Retrieved from https://mspace.lib.umanitoba.ca/handle/1993/33865

Angadi, S. V., Cutforth, H. W., Miller, P. R., McConkey, B. G., Entz, M. H., Brandt, S. A., and Volkmar, K. M. (2000). Response of three *Brassica* species to high temperature stress during reproductive growth. *Canadian Journal of Plant Science*, 80(4), 693-701.

Appelqvist, L. Å., and Loof, B. (1972). Post-harvest handling and storage of rapeseed. *Rapeseed (Eds L. Appelqvist, R. Ohlson). Elsevier, Amsterdam*, 60-70.

Ari, Ş., and Arikan, M. (2016). Next-generation sequencing: Advantages, disadvantages, and future. *In Plant Omics: Trends and Applications*, 109-135.

Arif Uz Zaman, M. (2014). Genetic Dissection and Molecular Approaches of Root Morphology of Canola (Master's thesis, North Dakota State University). Retrieved from https://library.ndsu.edu/ir/bitstream/handle/10365/27276/Genetic%20Dissection%20and%20Molecular%20Approaches%20of%20Root%20Morphology%20of%20Canola.pdf?sequence=1andisAllowed=y

Asghari, A., Shokrpour, M., Chamanabad, H. M., and Sofalian, O. (2011). Evaluating genetic diversity of canola cultivars using morphological traits and molecular markers. *Romanian Biotechnological Letters*, 16(4), 6305-6312.

Australian Government Office of the Gene Technology Regulator. (2008, February). Retrieved from http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/content/canola-3/$FILE/biologycanola08_2.pdf

Axelsson, T., Shavorskaya, O., and Lagercrantz, U. (2001). Multiple flowering time QTL within several *Brassica* species could be the result of duplicated copies of one ancestral gene. *Genome,* 44(5), 856-864.

Bahrani, J., and McVetty, P. B. E. (2008). Relationship of seed quality traits for greenhouse-grown versus field-grown high erucic acid rapesseed: Is seed quality trait selection for greenhouse-grown seed worthwhile?. *Canadian Journal of Plant Science,* 88(3), 419-423.

Bai, S., Wallis, J. G., Denolf, P., Engelen, S., Bengtsson, J. D., Van Thournout, M., Dierickx, J., Haesendonckx, B., and Browse, J. (2020). The biochemistry of headgroup exchange during triacylglycerol synthesis in canola. *The Plant Journal: For Cell and Molecular Biology*, 103(1), 83.

Basunanda, P., Radoev, M., Ecke, W., Friedt, W., Becker, H. C., and Snowdon, R. J. (2010). Comparative mapping of quantitative trait loci involved in heterosis for seedling and yield traits in oilseed rape (*Brassica napus* L.). *Theoretical and Applied Genetics*, 120(2), 271-281.

Baud, S., and Lepiniec, L. (2010). Physiological and developmental regulation of seed oil production. *Progress in Lipid Research*, 49(3), 235-249.

Baux, A., Hebeisen, T., and Pellet, D. (2008). Effects of minimal temperatures on low-linolenic rapeseed oil fatty-acid composition. *European Journal of Agronomy*, 29(2-3), 102-107.

Beckie, H. J., Johnson, E. N., Blackshaw, R. E., and Gan, Y. (2008). Productivity and quality of canola and mustard cultivars under weed competition. *Canadian Journal of Plant Science*, 88(2), 367-372.

Bedbrook, J. R., and Bogorad, L. (1976). Endonuclease recognition sites mapped on *Zea mays* chloroplast DNA. *Proceedings of the National Academy of Sciences*, 73(12), 4309-4313.

Bednarek, P., Piślewska-Bednarek, M., Svatoš, A., Schneider, B., Doubský, J., Mansurova, M., Humphry, M., Consonni, C., Panstruga, R., Sanchez-Vallet, A. and Molina, A. (2009). A glucosinolate metabolism pathway in living plant cells mediates broad-spectrum antifungal defense. *Science,* 323(5910), 101-106.

Bell, J. M. (1982). From rapeseed to canola: A brief history of research for superior meal and edible oil. *Poultry Science,* 61(4), 613-622.

Benfey, P. N. (2014). *Quickstart Molecular Biology: An Introduction for Mathematics, Physicsts, and Computational Scientists.* Cold Spring Harbor Laboratory Press.

Berglund, D. R., McKay, K., and Knodel, J. J. (2002). Canola production. *NDSU Extension Service- North Dakota State University (USA)*, 2-11.

Bohr, N. (1933). Light and life. *Nature*, 131, 457-459

Brennan, R. F., and M. D. A. Bolland. (2007). Influence of potassium and nitrogen fertiliser on yield, oil and protein concentration of canola (*Brassica napus* L.) grain harvested in south-western Australia. *Australian Journal of Experimental Agriculture*, 47(8), 976-983.

Brummer, E. C., Graef, G. L., Orf, J., Wilcox, J. R., and Shoemaker, R. C. (1997). Mapping QTL for seed protein and oil content in eight soybean populations. *Crop Science*, 37(2), 370-378.

Buntin, D. G., and Raymer, P. L. (1994). Pest status of aphids and other insects in winter canola in Georgia. *Journal of Economic Entomology*, 87(4), 1097-1104.

Burns, M. J., Barnes, S. R., Bowman, J. G., Clarke, M. H. E., Werner, C. P., and 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(1), 39.

Burton, W. A., Ripley, V. L., Potts, D. A., and 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*, 136(2), 181-192.

Campbell, P.J., Stephens, P.J., Pleasance, E.D., O'Meara, S., Li, H., Santarius, T., Stebbings, L.A., Leroy, C., Edkins, S., Hardy, C. and Teague, J.W. (2008). Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nature Genetics*, 40(6), 722.

Canadian Food Inspection Agency. (2019, May 1). Retrieved from http://www.inspection.gc.ca/plants/plants-with-novel-traits/applicants/directive-94-08/biology-documents/*Brassica*-napus-l-/eng/1330729090093/1330729278970#b25

Canola Council of Canada. (2019a, March 6). Retrieved from https://www.canolacouncil.org/canola-encyclopedia/crop-development/history-of-varietal-development/#*Brassica*-species-and-relatives

Canola Council of Canada. (2019b, May 11). Retrieved from https://www.canolacouncil.org/canola-encyclopedia/crop-development/growth-stages/

Canvin, D. T. (1965). The effect of temperature on the oil content and fatty acid composition of the oils from several oil seed crops. *Canadian Journal of Botany*, 43(1), 63-69.

Cartea, M. E., and Velasco, P. (2008). Glucosinolates in *Brassica* foods: Bioavailability in food and significance for human health. *Phytochemistry Reviews*, 7(2), 213-229.

Chahal, G. S., and Gosal, S. S. (2002). *Principles and Procedures of Plant Breeding: Biotechnological and Conventional Approaches.* Alpha Science Int'l Ltd, 1-104.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I.A., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B. and Corréa, M. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science*, 345(6199), 950-953.

Cheema, M. A., Malik, M. A., Hussain, A., Shah, S. H., and Basra, S. M. A. (2001). Effects of time and rate of nitrogen and phosphorus application on the growth and the seed and oil yields of canola (*Brassica napus* L.). *Journal of Agronomy and Crop Science*, 186(2), 103-110.

Chen, G., Geng, J., Rahman, M., Liu, X., Tu, J., Fu, T., Li, G., McVetty, P.B. and Tahir, M. (2010). Identification of QTL for oil content, seed yield, and flowering time in oilseed rape (*Brassica napus*). *Euphytica*, 175(2), 161-174.

Chen, J., Wang, B., Zhang, Y., Yue, X., Li, Z., and Liu, K. (2017). High-density ddRAD linkage and yield-related QTL mapping delimits a chromosomal region responsible for oil content in rapeseed (*Brassica napus* L.). *Breeding Science*, 16116.

Chen, W., Kalscheuer, V., Tzschach, A., Menzel, C., Ullmann, R., Schulz, M.H., Erdogan, F., Li, N., Kijas, Z., Arkesteijn, G. and Pajares, I.L. (2008). Mapping translocation breakpoints by next-generation sequencing. *Genome Research*, 18(7), 1143-1149.

Chen, X., Li, X., Zhang, B., Xu, J., Wu, Z., Wang, B., and Wu, J. (2013). Detection and genotyping of restriction fragment associated polymorphisms in polyploid crops with a pseudo-reference sequence: A case study in allotetraploid *Brassica napus*. *BMC Genomics*, 14(1), 346.

Clarke, W.E., Higgins, E.E., Plieske, J., Wieseke, R., Sidebottom, C., Khedikar, Y., Batley, J., Edwards, D., Meng, J., Li, R. and Lawley, C.T. (2016). A high-density SNP genotyping array for *Brassica napus* and its ancestral diploid species based on optimised selection of single-locus markers in the allotetraploid genome. *Theoretical and Applied Genetics*, 129(10), 1887-1899.

Clune, T. S., and Copeland, L. (1999). Effects of aluminium on canola roots. *Plant and Soil*, 216(1-2), 27-33.

Corley, R. H. V., and Tinker, P.B. (2008). *The Oil Palm.* John Wiley and Sons, 3-11.

Cornell University. (2019, March 1). Retrieved from http://poisonousplants.ansci.cornell.edu/toxicagents/glucosin.html

Cowling, W. A. (2007). Genetic diversity in Australian canola and implications for crop breeding for changing future environments. *Field Crops Research*, 104(1-3), 103-111.

Daun, J. K. (1994). Instrumental Methods for Determination of Oil, Protein, and Other Constituents of Oilseeds and Meals. *Journal of the American Oil Chemists Society.* 71(10).

Daun, J. K., Clear, K. M., and Williams, P. (1994). Comparison of three whole seed near-infrared analyzers for measuring quality components of canola seed. *Journal of the American Oil Chemists' Society*, 71(10), 1063-1068.

Del Rio, L.E., Bradley, C.A., Henson, R.A., Endres, G.J., Hanson, B.K., McKay, K., Halvorson, M., Porter, P.M., Le Gare, D.G. and Lamey, H.A. (2007). Impact of Sclerotinia stem rot on yield of canola. *Plant Disease*, 91(2), 191-194.

Delourme, R., Falentin, C., Fomeju, B.F., Boillot, M., Lassalle, G., André, I., Duarte, J., Gauthier, V., Lucante, N., Marty, A. and Pauchon, M. (2013). High-density SNP-based genetic map development and linkage disequilibrium assessment in *Brassica napus* L. *BMC Genomics*, 14(1), 120.

Delourme, R., Falentin, C., Huteau, V., Clouet, V., Horvais, R., Gandon, B., Specel, S., Hanneton, L., Dheu, J.E., Deschamps, M. and Margale, E. (2006). Genetic control of oil content in oilseed rape (*Brassica napus* L.). *Theoretical and Applied Genetics*, 113(7), 1331-1345.

Deschamps, S., Llaca, V., and May, G. D. (2012). Genotyping-by-sequencing in plants. *Biology*, 1(3), 460-483.

Deuel Jr, H. J., Johnson, R. M., Calbert, C. E., Gardner, J., and Thomas, B. (1949). Studies on the Comparative Nutritive Value of Fats: XII. The Digestibility of Rapeseed and Cottonseed Oils in Human Subjects. *The Journal of Nutrition*, 38(3), 369-379.

Din, J., Khan, S. U., Ali, I., and Gurmani, A. R. (2011). Physiological and agronomic response of canola varieties to drought stress. *Journal of Animal and Plant Science*, 21(1), 78-82.

Ding, L. N., Gu, S. L., Zhu, F. G., Ma, Z. Y., Li, J., Li, M., Wang, Z., and Tan, X. L. (2020). Long-chain acyl-CoA synthetase 2 is involved in seed oil production in *Brassica napus*. *BMC Plant Biology*, 20(1), 1-14.

Dhingani, R. M., Umrania, V. V., Tomar, R. S., Parakhia, M. V., and Golakiya, B. (2015). Introduction to QTL mapping in plants. *Annual Plant Science*, 4(4), 1072-1079.

Dosdall, L. M., and Stevenson, F. C. (2005). Managing flea beetles (*Phyllotreta* spp.) (Coleoptera: Chrysomelidae) in canola with seeding date, plant density, and seed treatment. *Agronomy Journal*, 97(6), 1570-1578.

Douglas, A. E. (2003). The nutritional physiology of aphids. *Advances in Insect Physiology*, 31(31), 73-140.

Downey, R. K., and Craig, B. M. (1964). Genetic control of fatty acid biosynthesis in rapeseed (*Brassica napus* L.). *Journal of the American Oil Chemists' Society*, 41(7), 475-478.

Dow-Walsh, D. S., Mahadevan, S., Kramer, J. K. G., and Sauer, F. D. (1975). Failure of dietary erucic acid to impair oxidative capacity or ATP production of rat heart mitochondria isolated under controlled conditions. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 396(1), 125-132.

Dressman, D., Yan, H., Traverso, G., Kinzler, K. W., and Vogelstein, B. (2003). Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proceedings of the National Academy of Sciences*, 100(15), 8817-8822.

Ecke, W., Uzunova, M., and 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(6-7), 972-977.

Elferjani, R., and Soolanayakanahally, R. (2018). Canola responses to drought, heat, and combined stress: shared and specific effects on carbon assimilation, seed yield, and oil composition. *Frontiers in Plant Science*, 9, 1224.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., and Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLOS One*, 6(5), e19379.

Elliott, R. H., Franke, C., and Rakow, G. F. W. (2008). Effects of seed size and seed weight on seedling establishment, vigour and tolerance of Argentine canola (*Brassica napus*) to flea beetles, *Phyllotreta* spp. *Canadian Journal of Plant Science*, 88(1), 207-217.

Eskin, M. N., Aladedunye, F., Unger, E. H., Shah, S., Chen, G., and Jones, P. J. (2005). Canola oil. *Bailey's Indusexperiment Oil and Fat Products*, 1-63.

Farooq, M., Jabran, K., Cheema, Z. A., Wahid, A., and Siddique, K. H. (2011). The role of allelopathy in agricultural pest management. *Pest Management Science*, 67(5), 493-506.

Fenster, C. B. (1991). Gene flow in *Chamaecrista fasciculata* (Leguminosae) I. Gene dispersal. *Evolution*, 45(2), 398-409.

Ferguson, A. W., Barari, H., Warner, D. J., Campbell, J. M., Smith, E. T., Watts, N. P., and Williams, I. H. (2006). Distributions and interactions of the stem miners *Psylliodes chrysocephala* and *Ceutorhynchus pallidactylu*s and their parasitoids in a crop of winter oilseed rape (*Brassica napus*). *Entomologia Experimentalis et Applicata*, 119(2), 81-92.

Fernando, W. G. D., and Chen, Y. (2003). First Report on the presence of *Leptosphaeria maculans* pathogenicity Group-3, the causal agent of Blackleg of canola in Manitoba. *Plant Disease*, 87(10), 1268-1268.

Food and Agriculture Organization of the United Nations. (2019, November). Retrieved from FAOSTAT: http://www.fao.org/faostat/en/#data/QC

Fourmann, M., Barret, P., Froger, N., Baron, C., Charlot, F., Delourme, R., and Brunel, D. (2002). From *Arabidopsis thaliana* to *Brassica napus*: development of amplified consensus genetic markers (ACGM) for construction of a gene map. *Theoretical and Applied Genetics*, 105(8), 1196-1206.

Fourmann, M., Barret, P., Renard, M., Pelletier, G., Delourme, R., and Brunel, D. (1998). The two genes homologous to *Arabidopsis* FAE1 co-segregate with the two loci governing erucic acid content in *Brassica napus*. *Theoretical and Applied Genetics*, 96(6-7), 852-858.

Friedt, W., Tu, J., and Fu, T. (2018). Academic and economic importance of *Brassica napus* rapeseed. *Springer*. 1-20.

Froese, N. T., and Van Acker, R. C. (2003). Distribution and interference of dandelion (*Taraxacum officinale*) in spring canola. *Weed Science*, 51(3), 435-442.

Fu, Y., Zhang, D., Gleeson, M., Zhang, Y., Lin, B., Hua, S., Ding, H., Frauen, M., Li, J., Qian, W. and Yu, H. (2017). Analysis of QTL for seed oil content in *Brassica napus* by association mapping and QTL mapping. *Euphytica*, 213(1), 17.

Gacek, K., Bayer, P. E., Bartkowiak-Broda, I., Szala, L., Bocianowski, J., Edwards, D., and Batley, J. (2017). Genome-wide association study of genetic control of seed fatty acid biosynthesis in *Brassica napus*. *Frontiers in Plant Science*, 7, 2062.

Gates, D. M. (1968). Transpiration and leaf temperature. *Annual Review of Plant Physiology*, 19(1), 211-238.

Gaut, B. S., and Clegg, M. T. (1993). Nucleotide polymorphism in the Adh1 locus of pearl millet (*Pennisetum glaucum*) (Poaceae). *Genetics*, 135(4), 1091-1097.

Ghazani, S. M., and Marangoni, A. G. (2013). Minor components in canola oil and effects of refining on these constituents: A review. *Journal of the American Oil Chemists' Society*, 90(7), 923-932.

Gilchrist, E. J., Sidebottom, C. H., Koh, C. S., MacInnes, T., Sharpe, A. G., and Haughn, G. W. (2013). A mutant *Brassica napus* (Canola) population for the identification of new genetic diversity via TILLING and next generation sequencing. *PLOS One*, 8(12).

Goodwin, S., McPherson, J. D., and McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, 17(6), 333.

Grubb, C. D., and Abel, S. (2006). Glucosinolate metabolism and its control. *Trends in Plant Science*, 11(2), 89-100.

Gugel, R. K. (1986). Inoculum-disease relationships in Sclerotinia stem rot of rapeseed in Saskatchewan. *Canadian Journal of Plant Pathology*, 8(1), 89-96.

Gugel, R. K., and Petrie, G. A. (1992). History, occurrence, impact, and control of blackleg of rapeseed. *Canadian Journal of Plant Pathology*, 14(1), 36-45.

Gulden, R. H., Warwick, S. I., and Thomas, A. G. (2008). The biology of Canadian weeds. 137. *Brassica napus* L. and B. rapa L. *Canadian Journal of Plant Science*, 88(5), 951-996.

Gunstone, F. D. (2004). *Rapeseed and canola oil: Production, processing, properties and uses.* CRC Press, 1-182.

Harker, K. N., Clayton, G. W., Blackshaw, R. E., O'Donovan, J. T., and Stevenson, F. C. (2003). Seeding rate, herbicide timing and competitive hybrids contribute to integrated weed management in canola (*Brassica napus*). *Canadian Journal of Plant Science*, 83(2), 433-440.

Hart, J. W. (2012). *Light and plant growth (Vol. 1).* Springer Science and Business Media, 1-51.

Hayter, K. E., and Cresswell, J. E. (2006). The influence of pollinator abundance on the dynamics and efficiency of pollination in agricultural *Brassica napus*: implications for landscape-scale gene dispersal. *Journal of Applied Ecology*, 43(6), 1196-1202.

He, J., Zhao, X., Laroche, A., Lu, Z. X., Liu, H., and Li, Z. (2014). Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Frontiers in Plant Science*, 5, 484.

Heisey, P. W., Srinivasan, C. D., and Thritle, C. (2002). Privatization of Plant Breeding in Indusexperimentized Countries: Causes, Consequences and Public Sector Response. *Agricultural Research Policy in an Era of Privatization.* CAB International Publishing, London, 177-198.

Howell, P. M., Sharpe, A. G., and Lydiate, D. J. (2003). Homoeologous loci control the accumulation of seed glucosinolates in oilseed rape (*Brassica napus*). *Genome*, 46(3), 454-460.

Hu, J., Guo, C., Wang, B., Ye, J., Liu, M., Wu, Z., Xiao, Y., Zhang, Q., Li, H., King, G.J. and Liu, K. (2018). Genetic properties of a nested association mapping population

constructed with semi-winter and spring oilseed rapes. *Frontiers in Plant Science*, 9, 1740.

Hu, X., Sullivan-Gilbert, M., Gupta, M., and Thompson, S. A. (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.). *Theoretical and Applied Genetics*, 113(3), 497-507.

Hua, S., Yu, H., Zhang, Y., Lin, B., Ding, H., Dongqing, D., Ren, Y., and Chen, Z. H. (2012). Variation of carbohydrates and macronutrients during the flowering stage in canola ('*Brasscia napus*' L.) plants with contrasting seed oil content. *Australian Journal of Crop Science*, 6(8), 1275.

Hudewenz, A., Pufal, G., Bögeholz, A. L., and Klein, A. M. (2014). Cross-pollination benefits differ among oilseed rape varieties. *Journal of Agriultural Sci*ence, 152(5), 770-778.

Hwang, S. F., Strelkov, S. E., Peng, G., Ahmed, H., Zhou, Q., and Turnbull, G. (2016). Blackleg (*Leptosphaeria maculans*) severity and yield loss in canola in Alberta, Canada. *Plants*, 5(3), 31.

Illumina microarray technology. (2020). Retrieved from http://www.illumina.com/science/technology/microarray.html

Issariyakul, T., Kulkarni, M. G., Meher, L. C., Dalai, A. K., and Bakhshi, N. N. (2008). Biodiesel production from mixtures of canola oil and used cooking oil. *Chemical Engineering Journal*, 140(1-3), 77-85

Jabran, K., Cheema, Z. A., Farooq, M., Basra, S. M. A., Hussain, M., and Rehman, H. . (2008). Tank mixing of allelopathic crop water extracts with pendimethalin helps in the management of weeds in canola (*Brassica napus*) field. *International Journal of Agriculture and Biology*, 10(3), 293-296.

Jackson, G. D. (2000). Effects of nitrogen and sulfur on canola yield and nutrient uptake. *Agronomy Journal*, 92(4), 644-649.

Javed, N., Geng, J., Tahir, M., McVetty, P. B. E., Li, G., and Duncan, R. W. (2016). Identification of QTL influencing seed oil content, fatty acid profile and days to flowering in *Brassica napus* L. *Euphytica*, 207(1), 191-211.

Joehanes, R., and Nelson, J. C. (2008). Qgene 4.0, an extensible Java QTL-analysis. *Bioinformatics*, 24, 2788-2789.

John, B. (2005). *Meiosis (Vol. 22).* Cambridge University Press.

Juska, A., Busch, L., and Wu, F. H. (1997). Producing genetic diversity in crop plants: The case of Canadian rapeseed, 1954–1991. *Journal of Sustainable Agriculture*, 9(4), 5-23.

Kearsey, M. J. (1998). The principles of QTL analysis (a minimal mathematics approach). *Journal of Experimental Botany*, 49(327), 1619-1623.

Kearsey, M. J., and Farquhar, A. G. L. (1998). QTL analysis in plants; where are we now? *Heredity*, 80(2), 137.

Kevan, P. G., and Eisikowitch, D. (1990). The effects of insect pollination on canola (*Brassica napus* L. cv. OAC Triton) seed germination. *Euphytica*, *45*(1), 39-41

Kim, S. U., Park, S. H., Choung, M. G., and 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(1), 13-18.

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, 12(4), 357.

Koscielny, C. B., Gardner, S. W., and Duncan, R. W. (2018). Impact of high temperature on heterosis and general combining ability in spring canola (*Brassica napus* L.). *Field Crops Research*, 221, 61-70.

Kramer, J. K. (2012). *High and Low Erucic Acid in Rapeseed Oils.* Academic Press. 144-311.

Kramer, P. J. (1963). Water Stress and Plant Growth 1. *Agronomy Journal*, 55(1), 31-35.

Krupinsky, J. M., Bailey, K. L., McMullen, M. P., Gossen, B. D., and Turkington, T. K. (2002). Managing plant disease risk in diversified cropping systems. *Agronomy Journal*, 94(2), 198-209.

Kumar, S. J., Prasad, S. R., Banerjee, R., Agarwal, D. K., Kulkarni, K. S., and Ramesh, K. V. (2017). Green solvents and technologies for oil extraction from oilseeds. *Chemistry Central Journal*, 11(1), 1-7.

Kunst, L., Taylor, D. C., and Underhill, E. W. (1992). Fatty acid elongation in developing seeds of *Arabidopsis thaliana*. *Plant Physiology and Biochemistry*, 30(4), 425-434.

Kwok, P. Y. (2001). Methods for genotyping single nucleotide polymorphisms. *Annual Review of Genomics and Human Genetics*, 2(1), 235-258.

Laryea, M. D., Jiang, Y. F., Xu, G. L., and Lombeck, I. (1992). Fatty acid composition of blood lipids in Chinese children consuming high erucic acid rapeseed oil. *Annals of Nutrition and Metabolism*, 36(5-6), 273-278.

Leff, B., Ramankutty, N., and Foley, J. A. (2004). Geographic distribution of major crops across the world. *Global Biogeochemical Cycles*, 18(1).

Li, G., and Quiros, C. F. (2001). Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: Its application to mapping and gene tagging in *Brassica*. *Theoretical and Applied Genetics*, 103(2-3), 455-461.

Li, N., Shi, J., Wang, X., Liu, G., and Wang, H. (2014). A combined linkage and regional association mapping validation and fine mapping of two major pleiotropic QTL for seed weight and silique length in rapeseed (*Brassica napus* L.). *BMC Plant Biology*, 14(1), 114.

Li, R. J., Wang, H. Z., Mao, H., Lu, Y. T., and Hua, W. (2006). Identification of differentially expressed genes in seeds of two near-isogenic *Brassica napus* lines with different oil content. *Planta*, 224(4), 952-962.

Li, Q., Shao, J., Tang, S., Shen, Q., Wang, T., Chen, W., and Hong, Y. (2015). Wrinkled1 accelerates flowering and regulates lipid homeostasis between oil accumulation and membrane lipid anabolism in *Brassica napus*. *Frontiers in Plant Science*, 6, 1015.

Lin, L., Allemekinders, H., Dansby, A., Campbell, L., Durance-Tod, S., Berger, A., and Jones, P. J. (2013). Evidence of health benefits of canola oil. *Nutrition Reviews*, 71(6), 370-385.

Liu, B., Wang, Y., Zhai, W., Deng, J., Wang, H., Cui, Y., Cheng, F., Wang, X. and Wu, J. (2013a). Development of InDel markers for *Brassica rapa* based on whole-genome re-sequencing. *Theoretical and Applied Genetics*, 126(1), 231-239.

Liu, J., Hirani, A. H., Li, Z., Wu, C., McVetty, P. B., Daayf, F., and Li, G. (2016). QTL controlling glucosinolate content in seeds of'*Brassica napus*' L. *Australian Journal of Crop Science*, 10(2), 152.

Liu, L., Qu, C., Wittkop, B., Yi, B., Xiao, Y., He, Y., Snowdon, R.J. and Li, J. (2013b). A high-density SNP map for accurate mapping of seed fibre QTL in *Brassica napus* L. *PLOS One*, 8(12), e83052.

Liu, S., Huang, H., Yi, X., Zhang, Y., Yang, Q., Zhang, C., Fan, C. and Zhou, Y. (2020). Dissection of genetic architecture for glucosinolate accumulations in leaves and seeds of *Brassica napus* by genome-wide association study. *Plant Biotechnology Journal*, *18*(6), 1472-1484.

Lohani, U. C., Fallahi, P., and Muthukumarappan, K. (2015). Comparison of ethyl acetate with hexane for oil extraction from various oilseeds. *Journal of the American Oil Chemists' Society*, 92(5), 743-754.

Luo, T., Zhang, J., Khan, M. N., Liu, J., Xu, Z., and Hu, L. (2018). Temperature variation caused by sowing dates significantly affects floral initiation and floral bud differentiation processes in rapeseed (*Brassica napus* L.). *Plant Science*, 271, 40-51.

Ma, D., Carpenter, N., Amatya, S., Maki, H., Wang, L., Zhang, L., Neeno, S., Tuinstra, M.R. and Jin, J. (2019). Removal of greenhouse microclimate heterogeneity with conveyor system for indoor phenotyping. *Computers and Electronics in Agriculture*, 166, 104979.

MacAdam, J. W., Volenec, J. J., and Nelson, C. J. (1989). Effects of nitrogen on mesophyll cell division and epidermal cell elongation in tall fescue leaf blades. *Plant Physiology*, 89(2), 549-556.

MacLean, D., Jones, J. D., and Studholme, D. J. (2009). Application of'next-generation'sequencing technologies to microbial genetics. *Nature Reviews Microbiology*, 7(4), 96-97.

Maheswaran, M. (1998). Establishing Marker-QTL Linkage: Principles, Requirements and Methodologies. Retrieved from http://bioinformatics.iasri.res.in/BAMAST/Book.html/EbookNew/web/book/module4/EstablishingMarker-QTLLinkage.pdf

MAJNOUN, H. N., Alizadeh, H. M., and MALEK, A. H. (2006). Effects of plant density and nitrogen rates on the competitive ability of canola (*Brassica napus* L.). *Journal of Agricultural Science and Technology*, 8, 281-291.

Malik, N., and Born, W. V. (1988). The biology of Canadian weeds.: 86. *Galium aparine* L. and *Galium spurium* L. *Canadian Journal of Plant Science*, 68(2), 481-499.

Maliba, B. G., Inbaraj, P. M., and Berner, J. M. (2018). The effect of ozone and drought on the photosynthetic performance of canola. *Journal of Integrative Agriculture*, 17(5), 1137-1144.

Manzanares-Dauleux, M. J., Delourme, R., Baron, F., and Thomas, G. (2000). Mapping of one major gene and of QTL involved in resistance to clubroot in *Brassica napus*. . *Theoretical and Applied Genetics*, 101(5-6), 885-891.

Mardis, E. R. (2017). DNA sequencing technologies: 2006–2016. *Nature Protocols*, 12(2), 213.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z. and Dewell, S.B. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057), 376.

Marra, M., Kucaba, T., Sekhon, M., Hillier, L., Martienssen, R., Chinwalla, A., Fedele, J., Grover, H., Gund, C., McCombie, W.R., and McDonald, K. (1999). A map for sequence analysis of the *Arabidopsis thaliana* genome. *Nature Genetics*, 22(3), 265-270.

Mba, O., Adewale, P., Dumont, M. J., and Ngadi, M. (2014). Application of near-infrared spectroscopy to characterize binary blends of palm and canola oils. *Indusexperiment Crops and Products*, 61, 472-478.

Men, A. E., Wilson, P., Siemering, K., and Forrest, S. (2008). Sanger DNA sequencing. *Next Generation Genome Sequencing: Towards Personalized Medicine*, 1-11.

Menard, G.N., Moreno, J.M., Bryant, F.M., Munoz-Azcarate, O., Kelly, A.A., Hassani-Pak, K., Kurup, S. and Eastmond, P.J. (2017). Genome wide analysis of fatty acid desaturation and its response to temperature. *Plant Physiology*, 173(3), 1594-1605.

Metzker, M. L. (2010). Sequencing technologies—the next generation. *Nature Reviews Genetics*, 11(1), 31-46.

Mietkiewska, E., Giblin, E.M., Wang, S., Barton, D.L., Dirpaul, J., Brost, J.M., Katavic, V. and Taylor, D.C. (2004). Seed-specific heterologous expression of a nasturtium FAE gene in *Arabidopsis* results in a dramatic increase in the proportion of erucic acid. *Plant Physiology*, 136(1), 2665-2675.

Mikkelsen, M. D., Petersen, B. L., Olsen, C. E., and Halkier, B. A. (2002). Biosynthesis and metabolic engineering of glucosinolates. *Amino Acids*, 22(3), 279-295.

Miles, C. M., and Wayne, M. (2008). Quantitative trait locus (QTL) analysis. *Nature Education*, 1(1), 208.

Mithen, R. F., Dekker, M., Verkerk, R., Rabot, S., and Johnson, I. T. (2000). The nutritional significance, biosynthesis and bioavailability of glucosinolates in human foods. *Journal of the Science of Food and Agriculture*, 80(7), 967-984.

Moghadam, H. R. T., Zahedi, H., and Ghooshchi, F. (2011). Oil quality of canola cultivars in response to water stress and super absorbent polymer application. *Pesquisa Agropecuária Tropical*, 41(4), 579-586.

Morrall, R. A. A., and Dueck, J. (1982). Epidemiology of Sclerotinia stem rot of rapeseed in Saskatchewan. *Canadian Journal of Plant Pathology*, 4(2), 161-168.

Morrison, M. J., and Stewart, D. W. (2002). Heat stress during flowering in summer *Brassica*. *Crop Science*, 42(3), 797-803.

Murphy, D. J. (2014). The future of oil palm as a major global crop: opportunities and challenges. *Journal of Oil Palm Research*, 26(1), 1-24.

Nagaharu, U. (1935). Genome analysis in *Brassica* with special reference to the experimental formation of B. napus and peculiar mode of fertilization. *Journal of Japanese Botany*, 7(7), 389-452.

NCBI. (2019). Retrieved from https://www.ncbi.nlm.nih.gov/genome/?term=*Brassica*+napus

NCBI. (2019). *Brassica napus cultivar ZS11 chromosome A8, Bra_napus_v2.0, whole genome shotgun sequence*. Retrieved from https://www.ncbi.nlm.nih.gov/nuccore/NC_027764.2?from=15858419andto=15860101andreport=fastaandstrand=true

Okuley, J., Lightner, J., Feldmann, K., Yadav, N., and Lark, E. (1994). *Arabidopsis* FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *The Plant Cell*, 6(1), 147-158.

Omidi, H., Tahmasebi, Z., Badi, H. A. N., Torabi, H., and Miransari, M. (2010). Fatty acid composition of canola (*Brassica napus* L.), as affected by agronomical, genotypic and environmental parameters. *Comptes Rendus Biologies*, 333(3), 248-254

Ooijen Van, J. W. (2006). JoinMap 4. Software for the calculation of genetic linkage maps in experimental populations. Wageningen, Netherlands, 33: Kyazma BV.

Oplinger, E. S., Hardman, L. L., Gritton, E. T., Doll, J. D., and Kelling, K. A. (1989). Canola (Rapeseed). *Alternative Field Crops Manual*, 7.

Osborn, T. C., Kole, C., Parkin, I. A. P., Sharpe, A. G., Kuiper, M., Lydiate, D. J., and Trick, M. (1997). Comparison of flowering time genes in *Brassica rapa*, B. napus and *Arabidopsis thaliana*. *Genetics*, 146(3), 1123-1129.

Öztürk, E. (2015). Performance, emissions, combustion and injection characteristics of a diesel engine fuelled with canola oil–hazelnut soapstock biodiesel mixture. *Fuel Processing Technology*, 129, 183-191.

Paridaen, A., and Kirkegaard, J. A. (2015). Forage canola (*Brassica napus*): spring-sown winter canola for biennial dual-purpose use in the high-rainfall zone of southern Australia. *Crop and Pasture Science*, 66(4), 275-286.

Park, J., Lee, Y., Martinoia, E., and Geisler, M. (2017). Plant hormone transporters: What we know and what we would like to know. *BMC Biology*, 15(1), 93.

Parkin, I. A. P., Sharpe, A. G., Keith, D. J., and Lydiate, D. J. . (1995). Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). *Genome*, 38(6), 1122-1131.

Parkin, I. A., Gulden, S. M., Sharpe, A. G., Lukens, L., Trick, M., Osborn, T. C., and Lydiate, D. J. (2005). Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics*, 171(2), 765-781.

Piquemal, J., Cinquin, E., Couton, F., Rondeau, C., Seignoret, E., Perret, D., Villeger, M.J., Vincourt, P. and Blanchard, P. (2005). Construction of an oilseed rape (*Brassica napus* L.) genetic map with SSR markers. *Theoretical and Applied Genetics*, 111(8), 1514-1523.

Poland, J. A., and Rife, T. W. (2012). Genotyping-by-sequencing for plant breeding and genetics. *The Plant Genome*, 5(3), 92-102.

Polowick, P. L., and Sawhney, V. K. (62(1), 83-86). High temperature induced male and female sterility in canola (*Brassica napus* L.). *Annals of Botany*, 1988.

Qiu, D., Morgan, C., Shi, J., Long, Y., Liu, J., Li, R., Zhuang, X., Wang, Y., Tan, X., Dietrich, E. and Weihmann, T. (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(1), 67-80.

Qu, C., Jia, L., Fu, F., Zhao, H., Lu, K., Wei, L., Xu, X., Liang, Y., Li, S., Wang, R. and Li, J. (2017). Genome-wide association mapping and identification of candidate genes for fatty acid composition in *Brassica napus* L. using SNP markers. *BMC Genomics*, 18(1), 1-17.

Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P. and Gu, Y. (2012). A tale of three next generation sequencing platforms: Comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*, 13(1), 341.

Raboanatahiry, N., Chao, H., Dalin, H., Pu, S., Yan, W., Yu, L., Wang, B. and Li, M. (2018). Quantitative trait loci alignment for seed yield and yield related traits in *Brassica napus*. *Frontiers in Plant Science*, 9, 1127.

Radford, A. D., Chapman, D., Dixon, L., Chantrey, J., Darby, A. C., and Hall, N. (2012). Application of next-generation sequencing technologies in virology. *The Journal of General Virology*, 93(Pt 9), 1853.

Rafalski, A. (2002). Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology*, 5(2), 94-100.

Rahman, M., Sun, Z., McVetty, P. B., and 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(6), 895-904.

Ramadhas, A. S., Jayaraj, S., and Muraleedharan, C. (2004). Use of vegetable oils as IC engine fuels—a review. *Renewable Energy*, 29(5), 727-742.

Raman, H., Dalton-Morgan, J., Diffey, S., Raman, R., Alamery, S., Edwards, D., and Batley, J. (2014). SNP markers-based map construction and genome-wide linkage analysis in *Brassica napus*. *Plant Biotechnology Journal*, 12(7), 851-860.

Raman, H., Raman, R., Kilian, A., Detering, F., Carling, J., Coombes, N., Diffey, S., Kadkol, G., Edwards, D., McCully, M. and Ruperao, P. (2014). Genome-wide delineation of natural variation for pod shatter resistance in *Brassica napus*. *PLOS One*, 9(7), e101673.

Raymer, P. L. (2002). Canola: An emerging oilseed crop. *Trends in New Crops and New Uses*, 1, 122-126.

Ribaut, J. M., and Hoisington, D. (1998). Marker-assisted selection: New tools and strategies. *Trends in Plant Science*, 3(6), 236-239.

Rocquelin, G., Leclerc, J., Bocon, D. (1969). Rapeseed oil rich in erucic acid and without erucic acid. 2. Comparative digestive utilisation by rats. *In Annales de Biologie Animale, Biochimie, Biophysique*, 9, 413-426.

Rondanini, D. P., del Pilar Vilariño, M., Roberts, M. E., Polosa, M. A., and Botto, J. F. (2014). Physiological responses of spring rapeseed (*Brassica napus*) to red/far-red ratios and irradiance during pre-and post-flowering stages. *Physiologia Plantarum*, 152(4), 784-794.

Rossak, M., Smith, M., and Kunst, L. (2001). Expression of the FAE1 gene and FAE1 promoter activity in developing seeds of *Arabidopsis thaliana*. *Plant Molecular Biology*, 46(6), 717-725.

Saeidnia, S., and Gohari, A. R. (2012). Importance of *Brassica napus* as a medicinal food plant. *Journal of Medicinal Plants Research*, 6(14), 2700-2703.

Sakhno, L. O. (2010). Variability in the fatty acid composition of rapeseed oil: Classical breeding and biotechnology. *Cytology and Genetics*, 44(6), 389-397

Salisbury, P. A., and Barbetti, M. J. (2011). Breeding oilseed *Brassica* for climate change. *Crop Adaptation to Climate Change*, 448-463.

Sanger, F. (1960). Chemistry of insulin. *British Medical Bulletin*, 16(3), 183-188.

Sanger, F. (1988). Sequences, sequences, and sequences. *Annual Review of Biochemistry*, 57(1), 1-29.

Sanger, F., and Tuppy, H. (1951). The amino-acid sequence in the phenylalanyl chain of insulin. 1. The identification of lower peptides from partial hydrolysates. *Biochemical Journal*, 49(4), 463.

Sarwar, M. (2013). The inhibitory properties of organic pest control agents against aphid (Aphididae: Homoptera) on Canola *Brassica napus* L.(Brassicaceae) Under field environment. *International Journal of Scientific Research in Environmental Sciences*, 1(8), 195.

Sawe, B. E. (2019, June 7). *World atlas*. Retrieved from The World's Top Rapeseed Producing Countries: https://www.worldatlas.com/articles/the-world-s-top-rapeseed-producing-countries.html

Scheffler, J.A., Sharpe, A.G., Schmidt, H., Sperling, P., Parkin, I.A.P., Lühs, W., Lydiate, D.J. and Heinz, E. (1997). Desaturase multigene families of *Brassica napus* arose through genome duplication. *Theoretical and Applied Genetics*, 94(5), 583-591.

Schreiber, L., Franke, R., and Lessire, R. (2005). Biochemical characterization of elongase activity in corn (*Zea mays* L.) roots. *Phytochemistry*, 66(2), 131-138.

Semagn, K., Bjørnstad, Å., and Ndjiondjop, M. N. (2006). An overview of molecular marker methods for plants. *African Journal of Biotechnology*, 5(25).

Seyis, F. (2013). Towards a canola quality resynthesized rapeseed (*Brassica napus* L.): B. oleracea genotypes as a basic resource. *The European Journal of Plant Science and Biotechnology (EJPSB)*, 33-37.

Shah, S., and Xin, Z. (1997). Overexpression of the FAD3 Desaturase Gene in a Mutant of *Arabidopsis*. *Plant Physiology*, 114(4), 1533-1539.

Sharma, R. V., Somidi, A. K., and Dalai, A. K. (2015). Preparation and properties evaluation of biolubricants derived from canola oil and canola biodiesel. *Journal of Agricultural and Food Chemistry*, 63(12), 3235-3242.

Shendure, J., and Ji, H. (2008). Next-generation DNA sequencing. *Nature Biotechnology*, 26(10), 1135.

Shendure, J., Porreca, G.J., Reppas, N.B., Lin, X., McCutcheon, J.P., Rosenbaum, A.M., Wang, M.D., Zhang, K., Mitra, R.D. and Church, G.M. (2005). Accurate multiplex polony sequencing of an evolved bacterial genome. *Science*, 309(5741), 1728-1732.

Shengxin, C., Chunxia, L., Xuyang, Y., Song, C., Xuelei, J., Xiaoying, L., Zhigang, X., and Rongzhan, G. (2016). Morphological, photosynthetic, and physiological responses of rapeseed leaf to different combinations of red and blue lights at the rosette stage. *Frontiers in Plant Science*, 7, 1144.

Shroff, R., Vergara, F., Muck, A., Svatoš, A., and Gershenzon, J. (2008). Nonuniform distribution of glucosinolates in *Arabidopsis thaliana* leaves has important consequences for plant defense. *Proceedings of the National Academy of Sciences*, 105(16), 6196-6201.

Singh, A. K. (2017). Discovery and role of molecular markers involved in gene mapping, molecular breeding, and genetic diversity. *Plant Bioinformatics*, 303-328.

Snowdon, R. J., and Iniguez Luy, F. L. (2012). Potential to improve oilseed rape and canola breeding in the genomics era. *Plant Breeding*, 131(3), 351-360.

Song, J, M., Guan, Z., Hu, J., Guo, C., Yang, Z., Wang, S., Liu, D., Wang, B., Lu, S., Zhou, R. and Xie, W. Z. (2020). Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. *Nature Plants*, 6(1), 34-45.

Song, Q., Hyten, D. L., Jia, G., Quigley, C. V., Fickus, E. W., Nelson, R. L., and Cregan, P. B. (2013). Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. *PLOS One*, 8(1), e54985.

Statistics Canada. (2019). Retrieved from https://www.google.com/search?q=statistic+canadaandrlz=1C1CHMO_en-GBCA583CA583andoq=statistic+canadaandaqs=chrome..69i57j0l5.3540j1j7andsourceid=chromeandie=UTF-8

Stefansson, B. R., Hougen, F. W., and Downey, R. K. (1961). Note on the isolation of rape plants with seed oil free from erucic acid. *Canadian Journal of Plant Science*, 41(1), 218-219.

Sun, F., Fan, G., Hu, Q., Zhou, Y., Guan, M., Tong, C., Li, J., Du, D., Qi, C., Jiang, L., and Liu, W. (2017). The high-quality genome of *Brassica napus* cultivar 'ZS 11' reveals the introgression history in the semi-winter morphotype. *The Plant Journal*, 92(3), 452-468.

Sun, Z., Wang, Z., Tu, J., Zhang, J., Yu, F., McVetty, P. B., and Li, G. (2007). An ultradense genetic recombination map for *Brassica napus*, consisting of 13,551 SRAP markers. *Theoretical and Applied Genetics*, 114(8), 1305-1317.

Taberlet, P., Gielly, L., Pautou, G., and Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, 17(5), 1105-1109.

Talalay, P., and Fahey, J. W. (2001). Phytochemicals from cruciferous plants protect against cancer by modulating carcinogen metabolism. *The Journal of Nutrition*, 131(11), 3027S-3033S.

Tanksley, S. D. (1993). Mapping polygenes. *Annual Review of Genetics*, 27(1), 205-233.

Tayo, T. M. (1975). Quantitative analysis of the growth, development and distribution of flowers and pods in oil seed rape (*Brassica napus* L.). *The Journal of Agricultural Science*, 85, 103-110.

Thormann, C. E., Romero, J., Mantet, J., and Osborn, T. C. (1996). Mapping loci controlling the concentrations of erucic and linolenic acids in seed oil of *Brassica napus* L. *Theoretical and Applied Genetics*, 93(1-2), 282-286.

Tierens, K.F.J., Thomma, B.P., Brouwer, M., Schmidt, J., Kistner, K., Porzel, A., Mauch-Mani, B., Cammue, B.P. and Broekaert, W.F. (2001). Study of the role of antimicrobial glucosinolate-derived isothiocyanates in resistance of *Arabidopsis* to microbial pathogens. *Plant Physiology*, 125(4), 1688-1699.

Tinker, N. A., and Mather, D. E. (1995). Methods for QTL analysis with progeny replicated in multiple environments. *Journal of Agricultural Genomics*, 1(1), 1-28.

Tollenaere, R., Hayward, A., Dalton-Morgan, J., Campbell, E., Lee, J.R., Lorenc, M.T., Manoli, S., Stiller, J., Raman, R., Raman, H. and Edwards, D. (2012). Identification and characterization of candidate Rlm4 blackleg resistance genes in *Brassica napus* using next-generation sequencing. *Plant Biotechnology Journal*, 10(6), 709-715.

Uchida, R. (2000). Essential nutrients for plant growth: Nutrient functions and deficiency symptoms. *Plant Nutrient Management in Hawaii's Soils*, 31-55.

Uzunova, M., Ecke, W., Weissleder, K., and Röbbelen, G. (1995). Mapping the genome of rapeseed (*Brassica napus* L.). I. Construction of an RFLP linkage map and localization of QTL for seed glucosinolate content. *Theoretical and Applied Genetics*, 90(2), 194-204.

Vaisey, M., Latta, M., Bruce, V. M., and McDonald, B. E. . (1973). Assessment of the intake and digestibility of high and low erucic acid rapeseed oils in a mixed Canadian diet. *Canadian Institute of Food Science and Technology Journal*, 6(3), 142-147.

Van Ooijen, J. W. (2006). JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations. *Kyazma BV, Wageningen*, 33(10.1371).

Vera, C. L., Downey, R. K., Woods, S. M., Raney, J. P., McGregor, D. I., Elliott, R. H., and Johnson, E. N. (2007). Yield and quality of canola seed as affected by stage of maturity at swathing. *Canadian Journal of Plant Science*, 87(1), 13-26.

Veras, G., de Araujo Gomes, A., da Silva, A. C., de Brito, A. L. B., de Almeida, P. B. A., and de Medeiros, E. P. (2010). Classification of biodiesel using NIR spectrometry and multivariate techniques. *Talanta*, 83(2), 565-568.

Veromann, E., Toome, M., Kännaste, A., Kaasik, R., Copolovici, L., Flink, J., Kovács, G., Narits, L., Luik, A. and Niinemets, Ü. (2013). Effects of nitrogen fertilization on insect pests, their parasitoids, plant diseases and volatile organic compounds in *Brassica napus*. *Crop Protection*, *43*, 79-88.

Vigeolas, H., Waldeck, P., Zank, T., and 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(3), 431-441.

Voorrips, R. (2002). MapChart: Software for the graphical presentation of linkage maps and QTL. *The Journal of Heredity*, 93 (1): 77-78.

Walton, G., Si, P., and Bowden, B. (1999). Environmental impact on canola yield and oil. *In Proceedings of the 10th International Rapeseed Congress*, 26-29.

Wan, J., Griffiths, R., Ying, J., McCourt, P., and Huang, Y. (2009). Development of drought-tolerant canola (*Brassica napus* L.) through the genetic modulation of ABA-mediated stomatal responses. *Crop Science*, 49(5), 1539-1554.

Wang, B., Wu, Z., Li, Z., Zhang, Q., Hu, J., Xiao, Y., Cai, D., Wu, J., King, G.J., Li, H. and Liu, K. (2018). Dissection of the genetic architecture of three seed-quality traits and consequences for breeding in *Brassica napus*. *Plant Biotechnology Journal*, 16(7), 1336-1348.

Wang, X., Long, Y., Yin, Y., Zhang, C., Gan, L., Liu, L., Yu, L., Meng, J. and Li, M. (2015). New insights into the genetic networks affecting seed fatty acid concentrations in *Brassica napus*. *BMC Plant Biology*, 15(1), 91.

Wang, X., Wang, H., Long, Y., Li, D., Yin, Y., Tian, J., Chen, L., Liu, L., Zhao, W., Zhao, Y. and Yu, L. (2013). Identification of QTLs associated with oil content in a high-oil *Brassica napus* cultivar and construction of a high-density consensus map for QTLs comparison in *B. napus*. *PloS one*, 8(12), e80569.

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., Bai, Y., Mun, J.H., Bancroft, I., Cheng, F. and Huang, S. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics*, 43(10), 1035.

Weber, S., Ünker, F., and Friedt, W. (2005). Improved doubled haploid production protocol for *Brassica napus* using microspore colchicine treatment in vitro and ploidy determination by flow cytometry. *Plant Breeding*, 124(5), 511-513.

Williams, I. H., Martin, A. P., and White, R. P. (1986). The pollination requirements of oil-seed rape (*Brassica napus* L.). *The Journal of Agricultural Science*, 106(1), 27-30.

Winfield, M.O., Allen, A.M., Burridge, A.J., Barker, G.L., Benbow, H.R., Wilkinson, P.A., Coghill, J., Waterfall, C., Davassi, A., Scopes, G., and Pirani, A. (2016). High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. *Plant Biotechnology Journal*, 14(5), 1195-1206.

Winter, P., and Kahl, G. (1995). Molecular marker technologies for plant improvement. *World Journal of Microbiology and Biotechnology*, 11(4), 438-448.

Wittstock, U., and Halkier, B. A. (2002). Glucosinolate research in the *Arabidopsis* era. *Trends in Plant Science*, 7(6), 263-270.

Wittstock, U., Agerbirk, N., Stauber, E.J., Olsen, C.E., Hippler, M., Mitchell-Olds, T., Gershenzon, J. and Vogel, H. (2004). Successful herbivore attack due to metabolic diversion of a plant chemical defense. *Proceedings of the National Academy of Sciences*, 101(14), 4859-4864.

Wu, R., Ma, C., and Casella, G. (2007). Statistical genetics of quantitative traits: linkage, maps and QTL. *Springer Science and Business Media*. 13, 287-293.

Xue, Y., Chen, B., Wang, R., Win, A. N., Li, J., and Chai, Y. (2018). Genome-wide survey and characterization of fatty acid desaturase gene family in *Brassica napu*s and its parental species. *Applied Biochemistry and Biotechnology*, 184(2), 582-598.

Yan, X. Y., Li, J. N., Fu, F. Y., Jin, M. Y., Chen, L., and Liu, L. Z. (2009). Co-location of seed oil content, seed hull content and seed coat color QTL in three different environments in *Brassica napus* L. *Euphytica*, 170(3), 355-364.

Yang, Q., Fan, C., Guo, Z., Qin, J., Wu, J., Li, Q., Fu, T. and Zhou, Y. (2012). Identification of FAD2 and FAD3 genes in *Brassica napus* genome and development of allele-specific markers for high oleic and low linolenic acid contents. *Theoretical and Applied Genetics*, 125(4), 715-729.

Ying, J. Z., Shan, J. X., Gao, J. P., Zhu, M. Z., Shi, M., and Lin, H. X. . (2012). Identification of quantitative trait loci for lipid metabolism in rice seeds. *Molecular Plant*, 5(4), 865-875.

Young, L. W., Wilen, R. W., and Bonham-Smith, P. C. (2004). High temperature stress of *Brassica napus* during flowering reduces micro-and megagametophyte fertility, induces fruit abortion, and disrupts seed production. *Journal of Experimental Botany*, 55(396), 485-495.

Yu, B., Boyle, K., Zhang, W., Robinson, S. J., Higgins, E., Ehman, L., Relf-Eckstein, J.A., Rakow, G., Parkin, I.A., Sharpe, A.G., and Fobert, P. R. (2016). Multi-trait and multi-environment QTL analysis reveals the impact of seed colour on seed composition traits in *Brassica napus*. *Molecular Breeding*, 36(8), 1-20.

Yu, X., and Sun, S. (2013). Comparing a few SNP calling algorithms using low-coverage sequencing data. *BMC Bioinformatics*, 14(1), 274.

Zhang, J., Mason, A. S., Wu, J., Liu, S., Zhang, X., Luo, T., Redden, R., and Yan, G. (2015). Identification of putative candidate genes for water stress tolerance in canola (*Brassica napus*). *Frontiers in Plant Science*, 6, 1058.

Zhang, L., Yang, G., Liu, P., Hong, D., Li, S., and He, Q. (2011). Genetic and correlation analysis of silique-traits in *Brassica napus* L. by quantitative trait locus mapping. *Theoretical and Applied Genetics*, 122(1), 21-31.

Zhao, F., Evans, E. J., Bilsborrow, P. E., and Syers, J. K. (1993). Influence of sulphur and nitrogen on seed yield and quality of low glucosinolate oilseed rape (*Brassica napus* L). *Journal of the Science of Food and Agriculture*, 63(1), 29-37.

Zhao, J., and Meng, J. (2003). Genetic analysis of loci associated with partial resistance to Sclerotinia sclerotiorum in rapeseed (*Brassica napus* L.). *Theoretical and Applied Genetics*, 106(4), 759-764.

Zhao, J., Becker, H. C., Zhang, D., Zhang, Y., and 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(1), 33-38.

Zhao, J., Dimov, Z., Becker, H. C., Ecke, W., and Möllers, C. (2008). Mapping QTL controlling fatty acid composition in a doubled haploid rapeseed population segregating for oil content. *Molecular Breeding*, 21(1), 115-125.

Zhao, J., Huang, J., Chen, F., Xu, F., Ni, X., Xu, H., Wang, Y., Jiang, C., Wang, H., Xu, A. and Huang, R. (2012). Molecular mapping of *Arabidopsis thaliana* lipid-related orthologous genes in *Brassica napus*. *Theoretical and Applied Genetics*, *124*(2), 407-421.

Zhou, Q. H., Fu, D. H., Mason, A. S., Zeng, Y. J., Zhao, C. X., and Huang, Y. J. (2014). In silico integration of quantitative trait loci for seed yield and yield-related traits in *Brassica napus*. *Molecular Breeding*, 33(4), 881-894.

Zhou, Y., Fitt, B. D., Welham, S. J., Gladders, P., Sansford, C. E., and West, J. S. (1999). Effects of severity and timing of stem canker (*Leptosphaeria maculans*) symptoms on yield of winter oilseed rape (*Brassica napus*) in the UK. *European Journal of Plant Pathology*, 105(7), 715-728.

Zhu, J. K. (2016). Abiotic stress signaling and responses in plants. *Cell*, 167(2), 313-324.

Zhu, Q., King, G.J., Liu, X., Shan, N., Borpatragohain, P., Baten, A., Wang, P., Luo, S. and Zhou, Q. (2019). Identification of SNP loci and candidate genes related to four important fatty acid composition in *Brassica napus* using genome wide association study. *PLOS One*, 14(8), e0221578.