Discovery of quantitative trait loci (QTL) associated with seed oil content, seed protein content, and amino acid content in *Brassica napus* L. meal

By

Danica Lynn Wilkinson Swaenepoel

A Thesis Submitted to the Faculty of Graduate Studies of

The University of Manitoba

In Partial Fulfillment of the Requirements of the Degree of

MASTER OF SCIENCE

Department of Plant Science

University of Manitoba

Winnipeg, Manitoba

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Rob Duncan, for the mentoring I have received over the course of this research. I have gained invaluable experience in plant breeding and your support and trust in my ability to succeed have allowed me to grow on a personal and professional level. I am incredibly grateful.

Many people greatly contributed to the success of this project. I would like to thank Dr. Curt McCartney, Dr. Isobel Parkin, and Erin Higgins for their expertise and guidance in the required data analysis, training, and execution of the QTL mapping portion of this study. I must also thank Anna Rogiewicz, Prakash Sharma, and Dr. Bogdan Slominski for their undertaking in accommodating, training, and assisting in the amino acid analysis portion of this project.

I would like to thank all members of the Brassica breeding group for their vast efforts and assistance in the greenhouse, laboratory, quality analysis, and field components of my project. In addition, I would like to thank my committee members Dr. Doug Cattani and Dr. Rotimi Aluko for their time and effort in the development of my project and thesis.

I would like to thank the University of Manitoba Graduate Fellowship (UMGF), NSERC, DL Seeds, Bunge Canada, and all other bursary and scholarship providers for the financial assistance and funding required to successfully complete this research.

This undertaking would not have been possible without the unwavering love and support I have received from my parents throughout my entire education thus far. I cannot thank my sisters, family, friends, colleagues, and coworkers enough for believing in me and supporting my passion for a higher education and a future in plant breeding and agriculture. It really means the world to me. Thank you.

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Canola (*Brassica napus* L.) is economically significant to the Canadian agricultural industry, largely due to the oil content of the seed. While seed oil content has been a large focus of past breeding efforts, the protein fraction of both the seed and the meal should also be considered of economic significance. Canola meal currently serves as a plant-based protein additive in feed sources for dairy, beef, swine, poultry, and aquaculture industries. Protein quantity and quality have suffered as a result of past breeding efforts to increase oil content, given the negative association between seed oil and seed protein content. Plant breeding efforts to improve the nutritional potential of canola's proteins require a better understanding of the genetic control of seed protein and amino acid contents and their relatedness to seed oil content. In this study, a quantitative trait loci (QTL) mapping procedure was conducted, successfully discovering genomic regions controlling seed oil content, seed protein content, and the contents of ten different amino acids in the meal of *B. napus*. Regions on chromosome A3 were discovered that control 4.6 – 7.2 % of seed oil content and 11.6 – 12.3 % of the phenotypic variation in seed protein content. Two major QTL were discovered on chromosomes C1 and C2 explaining 12.1 and 15.0 %, respectively, of the phenotypic variation in cysteine content. Regions on chromosome C5 was found to have major control of arginine, glutamine, isoleucine, leucine and lysine contents explaining 11.0 – 14.0 % of observed variation. The same region also had minor control over asparagine, phenylalanine, serine and valine content explaining 7.0 – 10.0 % of phenotypic variation. These newly discovered regions provide useful information for future efforts to enhance seed and meal quality from a protein perspective in *B. napus*. In addition, the negative correlation between seed oil and seed protein content was documented as well as the co-localization of genomic regions controlling the

two traits. All of these discoveries provide a deeper genetic understanding of seed quality in *B. napus* and serve to enhance end use opportunities for the canola industry in Canada.

TABLE OF CONTENTS

1.0 GENERAL INTRODUCTION

The canola (*Brassica napus* L.) industry currently contributes $26.7 billion to the Canadian economy (LMC International, 2016). Canola's seed quality is one of the largest factors driving the success and rapid growth of the industry over the last few decades (Morrison et al., 2016). The main focus of *B. napus* seed quality has been the oil, widely regarded as one of the healthiest edible oils for human consumption (Przybylski et al., 2005). Beyond oil, seed protein also plays an influential role when evaluating the end-use products of *B. napus* from an overall quality perspective. The amount of protein in the seed directly relates to the resulting protein found in the meal following the seed crushing process to obtain canola oil. Seed quality is economically irreplaceable in *B. napus*, and further understanding its genetic control has been the focus of many breeding efforts.

Past breeding strategies have focused on the increase of canola's seed oil content, efforts which have had a detrimental effect on canola's protein quantity and quality (Malabat et al., 2003; Wanasundara, 2011; Rahman et al., 2013). With seed oil content being arguably the most important quantitative trait in canola, its genetic control is very well documented (Delourme et al., 2006; Wang et al., 2013; Javed et al., 2015; Teh and Mollers, 2016; Chao et al., 2017). Extensive efforts in quantitative trait loci (QTL) mapping studies using a wide range of molecular marker types have revealed genomic regions controlling seed content on 17 of the 19 chromosomes (Delourme et al., 2006; Rahman, 2013). The genetic control of seed oil content is highly complex, the same of which can be said for seed protein content (Würschum et al., 2012; Wu et al., 2013; Wen et al., 2017; Chao et al., 2017). Control of the two traits becomes even more complex when considering the well-established and negative correlation they share with one another (Asare and

Scarsbrick, 1995; Andersen et al., 1996; Brennan et al., 2000; Hao et al., 2004; Ahmad et al., 2007; Hu et al., 2013).

Broadening the focus of breeding efforts in canola to include and improve protein quantity and quality can significantly impact the overall value of the crop and its end-use products. In recent years, QTL mapping studies have begun to focus on seed protein content and regions of genetic control have been reported on chromosomes A7 (Würschum et al., 2012), A4 (Wu et al., 2013), A3 and C5 (Chao et al., 2017) in *B. napus*. Even further, Chao et al. (2017) conducted the first study documenting the correlation between genomic locations controlling seed oil and seed protein content using high density linkage maps developed with the Illumina® 60K *Brassica* SNP array.

Beyond the amount of protein present in the seed, the nutritional value of canola's meal is crucially dependent on the protein quality. At a molecular level, amino acids (AAs) are the "building blocks" that constitute all proteins and as such, the amino acid profile is reflective of the overall quality (Parker and Pratt, 2010). From a nutritional standpoint for both humans and animals, the amino acid profile of *B. napus* meal contains a lower content of almost all amino acids when compared to soybean (*Glycine max* L. Merr) meal, the most widely traded plant-based protein (Wanasundara, 2011; Canola Council of Canada, 2015). While understanding the genetic control of protein in the seed is important, understanding the genetic control of amino acid content in the meal is also a tool that can be used to improve the overall quality of canola. Limited research is currently available on the genetic control of amino acid content in the meal of *B. napus* with the exception of a study by Wen et al. (2015) discovering several QTL related to arginine, histidine, glutamine, glycine, proline, alanine, and asparagine. Further validation and identification of regions controlling specific amino acids in the meal of *B. napus* is key for the development of breeding strategies with a focus on protein improvement.

The objectives of this research were to identify chromosomal regions associated with seed oil content, seed protein content, and amino acid contents of *B. napus* meal by conducting a QTL analysis using the Illumina® 60K *Brassica* SNP array on a diverse set of 124 doubled haploid (DH) genotypes developed for seed quality and derived from ethyl methanesulfonate (EMS) treatment. The discovery of these novel regions will aid in future breeding efforts to improve the overall seed and meal quality of *B. napus*, adding potential and value to the industry.

2.0 LITERATURE REVIEW

2.1 Background

2.1.1 Introduction to *Brassica* Species

The *Brassica* genus resides in the Brassicaceae tribe of the crucifer family, containing a total of 3709 species and 338 genera (Warwick et al., 2009). The *Brassica* genus itself contains 39 species, exhibiting great diversity in their use as oils, vegetables, condiments and animal feed (Warwick et al., 2009). Many different *Brassica* species are of economic importance in Canada, including *B. napus* L., *B. rapa* L., *B. juncea* L., and *B. oleracea* L. (Rakow, 2004). *Brassica napus*, *B. rapa*, and *B. juncea* are extensively used for the production of oils and other food products, while cultivated *B. oleracea* is used in the production of kale, cabbage, brussel sprouts, cauliflower, and broccoli (Rakow, 2004). Wild forms of *B. oleracea* exist across many regions including Greece, Turkey, Italy, Yugoslavia, Spain and France and all exhibit very diverse phenotypes (Snogerup, 1980). *B. rapa* is cultivated as a rapeseed oil crop and is believed to have originated near the Mediterranean Sea (Tsunoda, 1980). Interspecific hybridizations have been successful within the *Brassica* genus, including the economically important breeding between *B. rapa* (n = 10) and *B. oleracea* (n = 9) to create *B. napus* (n = 19) (Nagaharu, 1935). Many other hybridization events have been successful as well, including the derivation of *B. juncea* (n = 18) from *B. rapa* and *B. nigra* (n = 8) and the origin of *B. carinata* A. Braun (n = 17) from *B. nigra* and *B. oleracea* (Nagaharu, 1935).

2.1.2 Development of *Brassica napus* L.

Brassica napus L. is known more commonly as oilseed rape (Downey, 1983). Canola is a term that was trademarked in 1978 derived from the words "Canadian oil, low acid" describing a specific seed quality type of oilseed rape low in erucic acid (less than 2%) and glucosinolate content (less than 3 µmol/g) (Eskin, 2012). The legacy of the *Brassica* species was forever changed when two plant breeders, Drs. Baldur Stefansson and Keith Downey, set their focus on creating rapeseed varieties containing low amounts of erucic acid and glucosinolates. Erucic acid is a monounsaturated fatty acid characterized by its 22-carbon fatty acid chain (Lobb and Chow, 2008). Rapeseed and other cruciferous oils contain a much higher fraction of erucic acid than those of their vegetable oil counterparts (Harvey and Downey, 1964). The presence of erucic acid in food products for human and animal consumption has been a controversial issue. Beare (1957) and Vles (1978) were the first to question the presence of erucic acid in rapeseed following the discovery of potential links between erucic acid and myocardial lesions in animals. These discoveries led to a shift in breeding focus to reducing erucic acid content in *B. napus* for its use as a food oil for human consumption. One of canola's major breakthroughs occurred in 1963, when Keith Downey and his colleague B. L. Harvey discovered the underlying genetic inheritance of erucic acid content in rapeseed (Harvey and Downey, 1964). They observed that the trait is controlled by the embryo's genotype and not that of the mother plant, meaning seeds from the same plant could exhibit variation in erucic acid content (Harvey and Downey, 1964). This led to the identification of Liho, a European rapeseed with only 10 % erucic acid, by Baldur Stefansson in 1960 (Eskin, 2012). Downey later bred this low erucic trait into *B. napus*, resulting in the first "low erucic acid rapeseed" (LEAR) variety, "ORO" (Downey and Craig, 1964; Harvey and Downey, 1964).

Glucosinolates are organic compounds in the seed and meal of *Brassica* species that give rise to the pungent nature of the mustards and related plants (Ishida et al., 2014). The presence of glucosinolates not only leads to unpalatability of the meal, but, has been linked to health risks in animals depending on the characteristics of their gastrointestinal tract (Bell, 1993). Gas liquid chromatography (GLC) was used for the widespread screening of germplasm for glucosinolate content, allowing for the identification of materials with low erucic acid as well as material with low glucosinolate content. Breeding of these newly discovered genotypes led to the development of the first *B. napus* genotypes low in both erucic acid and glucosinolates (Eskin, 2012). These materials were heavily utilized in the breeding programs of both Stefansson and Downey (Stefansson et al., 1961; Harvey and Downey, 1964; Eskin, 2012). Stefansson released the first *B. napus* variety to have zero erucic acid and low glucosinolate content (less than 3 µmol/g), called Tower, in 1974 (Stefansson and Kondra, 1974). This was the birth of "canola" as we recognize it today, a crop that has since undergone incredibly rapid growth, adoption and genetic improvement.

2.2 End-Uses of *Brassica napus* L.

The value of the Canola industry to the Canadian economy is valued at $26.7 billion annually (LMC International, 2016). Contributing to this value are various end-products of canola production that include the oil, the meal, and the seed. Canada is a large exporter of these goods, including 1.6 million tonnes of oil, 2.2 million tonnes of meal, and 5.5 million tonnes of seed in the 2015-2016 crop year (Statistics Canada, 2016). Domestic canola production continues to grow with approximately 20 million acres produced in Canada each year (Statistics Canada, 2016). Canada is also home to 14 crushing facilities where canola seed is processed to produce canola oil and canola meal, producing an annual 3 million and 4 million tonnes of each, respectively (Canola

Council of Canada, 2017a). Canola oil contains a lower amount of saturated fat than any edible oil on the market and has been proven to provide many benefits to human health (Canola Council of Canada, 2017b). The canola meal, a natural by-product of the crushing process, is often used as a feed additive in animal production (Bell, 1984; Wanasundara, 2011).

2.2.1 Oil Properties of *Brassica napus*

Breeding efforts to remove erucic acid and glucosinolates resulted in a shift in the physical and chemical properties of canola oil compared to that of its predecessor rapeseed (Malabat et al., 2003). An important characteristic of a frying oil is its smoke point, the temperature at which the oil produces smoke (Pryzbylski, 2011). For canola oil, this smoke point is approximately 220 – 230 C while it can be as low as 190 C in rapeseed (Przybylski et al., 2005). Chemically, canola oil is made up of triglycerides, phospholipids, and various fatty acids (Canola Council of Canada, 2017b). The types of fatty acids present vary from saturated to monounsaturated and polyunsaturated (Eskin, 1989). Removing erucic acid (C22) led to an increase in C18 fatty acids, which now make up approximately 95 % of canola's oil (Przybylski et al., 2005). Common C18 fatty acids include stearic acid, oleic acid, linoleic, and linolenic acids (Lobb and Chow, 2008).

2.2.2 Specialty Oil Production

A lower fraction of linolenic acid (C18:3) provides greater frying stability to canola oil, allowing for frying of foods without the production of *trans* fats that are produced when using traditional canola oil (Liu and Iassonova, 2012). Breeding focus to create *B. napus* varieties with low levels of linolenic acid in the seed oil began in the 1980s and it was observed that efforts to lower linolenic acid resulted in an increase in oleic acid (C18:1) (Downey and Craig, 1964). One of the most

successful specialty *B. napus* oils featuring low linolenic acid and high oleic acid comes from Dow AgroSciences' Nexera™ canola (Dow AgroSciences, 2016). In addition, Cargill also markets Victory® Hybrid canola (Cargill, 2015). Cultivars in the specialty oil sector are available as high-yielding, herbicide-tolerant hybrids, providing financial premiums to the producers who grow them while solidifying value-added chains in the canola industry (Cargill, 2015; Dow AgroSciences, 2016).

Another faction of specialty *B. napus* oils is high erucic acid rapeseed (HEAR) (McVetty et al., 2016). Unlike the traditional canola varieties, which were intentionally bred to contain little to no erucic acid, HEAR varieties are intentionally bred to contain upwards of 50 % erucic acid in the seed (Daun, 1983). HEAR oil is used to produce an additive called erucamide that is used in polythene and polypropylene film in the oleochemical industry (Bunge, 2016). HEAR oil is used in the manufacturing of plastics, rubbers, lubricants, inks, and cosmetics (Nieschlag and Wolff, 1971). The University of Manitoba has released many successful HEAR cultivars, including MillenniUM 03, a traditionally bred *B. napus* variety containing 55.2 % erucic acid (McVetty et al., 2000). Since MilleniUM 03, a series of Roundup Ready® HEAR cultivars have been released, such as Red River 1861 containing 52.6 % erucic acid (McVetty et al., 2012). In 2013, the world's first hybrid Roundup Ready® HEAR variety was released, called HYHEAR 1, containing 52.2 % erucic acid (McVetty et al., 2014). Subsequently, HYHEAR 2 was released in 2016 containing 51.6 % erucic acid (Duncan et al., 2016) and HYHEAR 3 in 2017 with 50.9 % (Duncan et al., 2017). Breeding efforts to alter the fatty acid profiles of *B. napus* to create specialty markets have proven to be successful not only in their adoption by farmers but in their specialty end-use markets as well. With the Canola Council of Canada's target initiative "Keep it Coming – 2025" ambitious goals have been set to increase value-added components of the canola industry through focus on

both seed and meal. There remains large untapped potential in these sectors to create a strong future for canola in Canada.

2.2.3 Canola Meal Production in Canada

While oil is the most valuable component of *B. napus* production, canola and rapeseed meal is the second most widely traded plant-based protein ingredient behind that of soybean meal (Canadian Council of Canada, 2015). In 2015, a total of 4.7 million tonnes of product were crushed in Canada resulting in the production of 2.7 million tonnes of canola meal (Statistics Canada, 2016). The United States of America account for 95 % of total Canadian exports, importing 2.1 million tonnes of the 2.2 million tonnes that were exported in 2015 (Statistics Canada, 2016).

Canola and rapeseed meals are utilized for a variety of functions in different markets around the world (Canola Council of Canada, 2015). Its nutritional composition and amino acid profile differs from that of soybean meal (Table 2.4) and it is often used as an added source of protein in feed rations (Wanasundara, 2011). Canola meal exported to the USA is largely used for feed in dairy cattle operations, though global uses for canola meal extend to swine, poultry, and fish industries (Canola Council of Canada, 2015). In addition to its use in livestock industries, canola meal also contains significant potential as a plant-based protein source in human diets.

Though past breeding efforts have largely focused on increasing seed oil content and improving other oil-related traits (Malabat et al., 2003; Wanasundara, 2011), there is an opportunity to continue to add value to canola by simultaneously focusing on protein related-traits. By enhancing the nutritional value of seed protein and meal-related traits, there is the potential to further increase

the value of canola, not only for producers and growers, but also in its use in animal and human diets.

2.3 Nutritional Quality and Functionality of Canola Meal

2.3.1 Introduction to Canola Meal

The basic chemical composition of canola meal can include moisture content, crude protein content, rumen escape protein (see Section 2.4.1. Ruminant Diets), linoleic acid content, linolenic acid content, crude fibre, total dietary fibre, and glucosinolates, amongst many other components (Canola Council of Canada, 2015). Based on a three-year study conducted by Slominski (2015) and a four-year study conducted by Broderick (2015), canola meal was found to typically have 12 % moisture, 36.7 % crude protein, 43.5 % rumen escape protein, 0.67 % linoleic acid, 0.32 % linolenic acid, 11.2 % crude fibre, 32.4 % total dietary fibre, and 4.2 μmol/g of glucosinolates.

Looking closer, the protein fraction of canola is primarily composed of two major storage protein constituents; 11 or 12S cruciferin (Schewenke et al., 1981; Sjödahl et al., 1991) and 1.7 or 2S napin (Lönnerdal and Janson, 1972; Crouch et al., 1982). In addition, a structural protein called oleosin is also present in a small fraction and is associated with the oil content of the seed (Uppstrom, 1995). Cruciferin constitutes the majority of canola's protein fraction, making up approximately 60 % of the seed storage proteins (Crouch and Sussex, 1981). Napin contributes less to the total seed protein, constituting only 20 % (Ericson et al., 1986; Höglund et al., 1992). The combination of cruciferin and napin gives rise to a unique essential amino acid profile, biological value, and net protein utilization value for canola meal in its use as a food protein (Wanasundara, 2011). Functionally, the properties of canola meal can differ depending on the cruciferin to napin ratio

(Wanasundara, 2011; Wanasundara et al., 2016). This is due to the fact that the two proteins differ in their amino acid composition, size, structure, and physical and chemical properties (Wansundara, 2011; Campbell et al., 2016; Wanasundara et al., 2016). When considering the solubility of canola protein products, napin-rich products tend to be more soluble than those rich in cruciferin (Wanasundara, 2011). The protein ratio also has an effect on the emulsifying and processing properties of canola meal. Many studies have shown that napin exhibits poor emulsifying capabilities in comparison to cruciferin, posing a detrimental effect on the overall emulsifying properties of canola meal that contains both proteins (Wu and Muir, 2008; Wanasundara, 2011; Tan et al., 2011; Aachary and Thiyam, 2012; Tan et al., 2014). Amino acids are considered to be the building blocks of proteins and are just as diverse in their functions as the proteins that they form.

2.3.2 Introduction to Amino Acids

By definition, amino acids (AAs) are organic substances that contain both a primary amino group (NH_2) and a carboxyl group (COOH) with varying side chains that determine their overall structure and function (Wu, 2009). Amino acids serve vital functions in growth, maintenance, and reproduction at various life stages either directly or in the form of substrates used to produce numerous key substrates in the body (Wu, 2009). Amino acids are quite often distinguished as either being “essential” or “non-essential” based on whether or not they can be synthesized de novo within an organism. Wu (2009; 2010; 2013a; 2013b) has done extensive work to characterize the metabolic functions of amino acids in growth, reproduction, and health. In addition, Wu (2010; 2013b) has introduced the concept of “functional amino acids”, which are considered to be regulators of key metabolic pathways, regardless of whether they are essential amino acids or not.

2.3.2.1 Essential Amino Acids (EAA)

Essential amino acids are defined by Wu (2009) as "amino acids whose carbon skeletons cannot be synthesized or adequately synthesized de novo by the body relative to needs and which must be provided from the diet to meet optimal requirements". In relation to humans, there are nine essential amino acids – histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine (Institute of Medicine, 2006) and can be found in Table 2.1.

2.3.2.2 Non-Essential Amino Acids (NEAA)

Non-essential amino acids are defined by Wu (2009) as amino acids "which can be synthesized de novo in adequate amounts by the body to meet optimal requirements". Non-essential amino acids include alanine, aspartic acid, asparagine, glutamic acid, and serine (Institute of Medicine, 2006) and can be found in Table 2.2.

2.3.2.3 Conditionally-Essential Amino Acids (CEAA)

Some amino acids fall into the category of "conditionally-essential" because of the complex nature of their synthesis, resulting at times in the body being limited by the amount of that specific AA that it is able to produce (Reeds, 2000). Amino acids that are considered conditionally-essential, or limited under certain conditions, include arginine, cysteine, glycine, glutamine, proline, and tyrosine (Institute of Medicine, 2006) and can be found in Table 2.3.

Table 2.1. Direct and indirect functions of essential amino acids (adapted from Wu, 2009).

Amino Acid	Function
Histidine	Involved in forming antioxidative dipeptides; precursor of histamine (immune response, allergic reaction)
Isoleucine	Branched-chain amino acid; involved in synthesis of glutamine and alanine; involved in muscle endurance and healing
Leucine	Branched-chain amino acid; involved in muscle building and weight loss; involved in glutamine and alanine synthesis
Lysine	Regulation of signaling molecule nitric oxide (NO); antiviral properties; precursor of OH-lysine involved in structure and function of collagen
Methionine	Sulfur-containing; precursor of homocysteine, betaine, choline, cysteine, DCSAM*, taurine, and phospholipids; resulting functions include oxidant production, cellular metabolism and nutrition, methylation of proteins and DNA, vascular, muscular, cardiac, and retinal care, anti-inflammation, and more
Phenylalanine	Involved in neurological development and function; synthesis of tyrosine
Threonine	Involved in maintaining intestinal and gut function; immune function; synthesis of glycine

Table 2.1 continued. Direct and indirect functions of essential amino acids (adapted from Wu, 2009) continued.

Amino Acid	Function
Tryptophan	Precursor of serotonin, *N*-acetylserotonin, melatonin, anthanilic acid, and niacin; resulting functions include neurotransmission, antioxidant production, immune function, production of NAD and NADPH, and more
Valine	Branched-chain amino acid; involved in synthesis of glutamine and alanine; muscle building; nervous system functioning

*DCSAM = decarboxylated *S*-adenosyl-L-methionine (Junker et al., 2013). **NAD = nicotinamide adenine dinucloeitde, NADPH = nicotinamide adenine dinucleotide phosphate.

Table 2.2. Direct and indirect functions of non-essential amino acids (adapted from Wu, 2009).

Amino Acid	Function
Alanine	Involved in gluconeogenesis (formation of glucose); glucose-alanine cycle
Aspartic Acid	Involved in synthesis of asparagine and arginine; purine and pyrimidine synthesis (nucleotides forming DNA); urea cycle
Asparagine	Involved in cell metabolism and physiology; gene expression; immune function; nervous system function; ammonia detoxification; precursor of acrylamide
Glutamic Acid	Involved in synthesis of glutamine and arginine; precursor of GABA* (excitatory neurotransmitter)
Serine	Involved in synthesis of cysteine and glycine; synthesis of tryptophan in bacteria; gluconeogenesis in ruminants

*GABA = γ-aminobutyrate

Table 2.3. Direct and indirect functions of conditionally-essential amino acids (adapted from Wu, 2009).

Amino Acid	Function
Arginine	Involved in regulation of gene and immune function; precursor of nitric oxide (NO) involved in nutrition, metabolism, embryogenesis, fertility, immune function, hormone secretion, wound healing, neurotransmission and more; antioxidative function; methylation of proteins; deamination of proteins; precursor of ornithine involved in wound healing
Cysteine	Sulfur-containing; involved in synthesis of taurine and H_2S (signaling molecule; antioxidative function
Glycine	Involved in calcium influx; synthesis of serine and purines; inhibitory neurotransmitter; precursor of heme groups including hemoproteins (hemoglobin, myglobin, etc.) and carbon monoxide (signaling molecule)
Glutamine	Involved in gene function and immune function; synthesis of arginine, proline, and asparagine; synthesis of purines and pyrimidines; precursor of glutamine and aspartic acid; neurotransmission; ammonia detoxification; precursor of ammonia involved in acid-base balance and synthesis of glutamate
Proline	Involved in collagen structure; neurological function; precursor of H_2O_2 (signaling molecule) involved in immunity, intestinal function, and pathogen control; precursor of P5C* involved in DNA synthesis, gene expression, and stress response

Table 2.3 continued. Direct and indirect functions of conditionally-essential amino acids (adapted from Wu, 2009).

Amino Acid	Function
Tyrosine	Involved in protein phosphorylation; precursor of dopamine (neurotransmitter) involved in immune response; precursor of epinephrine and norepinephrine (neurotransmitters) involved in cell metabolism; precursor of melanin (antioxidant)

*P5C = pyrroline-5-carboxylate

2.3.2.4 Functional Amino Acids (FAA)

In addition to the above classification, the concept of "functional amino acids" has also been introduced. According to Wu (2009), functional amino acids are those that are "important regulators of key metabolic pathways that are necessary for maintenance, growth, reproduction and immunity in organisms". Their functionality separates them from previously defined categories (essential, non-essential, and conditionally-essential) based on their metabolic purposes rather than the body's ability to synthesize them. Functional amino acids include arginine, cysteine, glutamine, leucine, proline, and tryptophan (Wu, 2009).

2.3.2.5 Branched-Chain Amino Acids (BCAA)

As previously mentioned, the structure of each AA is completed with a unique side chain that determines the overall structure and function of that amino acid. Amino acids with a side chain containing a branch are those referred to as "branched-chain amino acids". The amino acids that fall into this category are leucine, isoleucine, and valine (Chuang, 2013). As Wu (2009) described the branched-chain amino acids play a key role in muscle development and function. Supplementation of these amino acids has become increasingly popular in active individuals seeking muscle growth and/or weight loss. Two separate studies conducted by Louard et al., (1990) and Nair et al., (1992) suggested that BCAAs were capable of simultaneously increasing protein synthesis while inhibiting protein degradation in skeletal muscle tissue. It has since been discovered that BCAAs are involved in the activation of key enzymes related to the synthesis of proteins, most notably during the recovery period after exercise (Blomstrand et al., 2006).

2.4 Amino Acid Profiles in Canola Meal

The amino acid profiles of canola meal are considered to be well-balanced (Wanasundara, 2011; Tan et al., 2011), but its adequacy as a plant-based source of protein varies depending on the specific dietary requirements of the human or animal in question. Amino acid requirements of a given organism can be as variable as the amino acids themselves. The species of the organism, it's developmental stage, physiological status, what species and quantity of microflora exist in the small intestine, environmental influences, and pathological states can all cause changes in dietary AA requirements (Dai et al., 2011; 2012a; 2012b; 2013). In general, Canadian canola meal contains a lower amount of all of the amino acids described with the exception of methionine and cysteine when compared to *Glycine max* as shown below in Table 2.4.

Table 2.4. Comparison of mean amino acid contents in Canadian canola meal versus USA soybean meal on a percent as received basis.

Amino Acid	Canadian Canola Meal (2015)*	USA Soybean Meal (2015)**
Alanine	1.57	2.05
Arginine	2.08	3.48
Asparagine	2.61	5.52
Cysteine	0.86	0.79
Glutamine	6.53	8.62
Glycine	1.77	1.97
Histidine	1.12	1.21
Isoleucine	1.56	2.17
Leucine	2.54	3.60
Lysine	2.00	2.89
Methionine	0.74	0.63
Met + Cys	1.60	1.36
Phenylalanine	1.38	2.37
Proline	2.15	2.37
Serine	1.44	2.38
Threonine	1.58	1.84
Tryptophan	0.48	0.63
Tyrosine	1.16	1.68
Valine	1.97	2.30

* Canola Council of Canada (2015); ** American Soybean Association (2015)

2.4.1 Ruminant Diets

Ruminants are animals whose digestive systems contain four stomachs, the largest being the rumen (Hungate, 1966). The rumen itself contains micro-organisms that are able to break down food ingested by the animal (Ross et al., 2013). This can make the development of feed rations for ruminant animals difficult as a large portion of the dietary protein can be degraded in the rumen as a result of microbial activity (Broderick, 1987). Only some of the protein that enters the rumen is able to pass through without degradation, creating the classification of ruminally-degraded protein (RDP) and ruminally-undegraded protein (RUP) (Hedqvist and Udén, 2006). Both RDP and RUP serve an important function in ruminant nutrition, RDP serving as a food source for the microbial flora that flourish in the rumen and RUP serving as a source of proteins and amino acids that are actually digested by the animal itself (Schwab, 2012). Many studies have been conducted looking at the value of canola meal in ruminant diets, particularly in dairy and beef cattle.

A series of studies conducted over almost ten years set out to look at the amount of rumen un-degraded protein (RUP) present in animal diets when fed canola meal versus soybean meal (Hedqvist and Udén, 2006; Tylutki et al., 2008; Maxin et al., 2013; Jayasinghe et al., 2014; Ross, 2015; Broderick, 2015). Though the method of meal extraction has an effect on the availability of nutrients to the animal, canola meal has been proven to consistently have a higher amount of RUP as a percent of total meal protein (Hedqvist and Udén, 2006; Tylutki et al., 2008; Maxin et al., 2013; Jayasinghe et al., 2014; Ross, 2015; Broderick, 2015). When considering the RUP fraction as a canola to soybean ratio, this series of studies found that the ratio ranged from 1.09 to 2.07, showing the superiority of canola meal over soybean meal (Hedqvist and Udén, 2006; Tylutki et al., 2008; Maxin et al., 2013; Jayasinghe et al., 2014; Ross, 2015; Broderick, 2015). The amino

acid profile of canola is suited for the nutrient requirements of ruminant animals (Hedqvist and Udén, 2006; Tylutki et al., 2008; Maxin et al., 2013; Jayasinghe et al., 2014; Ross, 2015; Broderick, 2015) and its ability to pass through the rumen un-degraded enhances its potential. A study conducted by Ross et al. (2013) set out to look at the overall composition of essential amino acids in the RUP fraction of canola meal and found that canola meal delivers a significant amount of methionine. This provides another edge for canola meal as ruminant feed, as methionine is usually one of the most limiting essential amino acids (Miller, 2004).

A group of three meta-analyses were conducted to specifically look at the link between canola meal and increased milk production in dairy cattle. The first, a study conducted by Huhtanen et al. (2011), included 292 treatments across 122 studies and looked at canola meal's effect on milk production in comparison to that of soybean. Two separate studies conducted by Martineau et al. (2013; 2014) looked at the effect of replacing a series of other meals with canola in relation to milk production. The first meta-analysis showed that feeding canola meal results in a 3.4 kg increase in milk production while feeding soybean meal only results in a 2.4 kg increase (Huhtanen et al., 2011). The second study showed that canola meal provides an overall 1.4 kg milk increase compared to all other vegetable proteins, but only a 0.70 kg increase when compared to soy alone (Martineau et al., 2013). The last meta-analysis looked at the presence of essential amino acids in the plasma of lactating dairy cattle and found that canola meal resulted in an overall increase of total amino acids, total essential amino acids, and individual essential amino acids (Martineau et al., 2014). These meta-analyses have provided knowledge and insight into feed requirements for lactating dairy cattle and have shed light on the potential canola meal holds.

When considering specific amino acids, methionine and lysine have been proven to be the two most limiting, perhaps "co-limiting", amino acids in lactating dairy cattle (Broderick et al., 1974; Schwab et al., 1976). In addition, a study by Broderick et al. (1974) identified valine as a limiting amino acid while separate studies by Chandler and Polan (1972) and Derrig et al. (1974) also identified phenylalanine as limiting. While the value of canola meal as dairy feed is reflected by its higher methionine content, the content of lysine and other essential amino acids in comparison to soybean meal is reduced.

Canola meal has also shown promise as a feed source in beef cattle. Different studies from Li et al. (2013) and Yang et al. (2013) linked canola meal to increased weight gain in heifers and steers, respectively. Cattle in the beef industry have a very wide range of nutritional requirements depending on whether heifers, steers, or calves are being discussed. Canola meal has shown promise in each of these areas (Petit et al., 1994; Anderson and Schoonmaker, 2004; He and Armentano, 2011; Miller-Cushon et al., 2014), but more research is required.

2.4.2 Swine Diets

When considering a vegetable meal source as a feedstuff in swine production, many factors have to be considered. Nyachoti et al. (2004) explained that palatability, dietary inclusion, feed energy, fibre content, and mineral balance all have an effect on the actual feed intake of the animal. Canola meal is often considered a poor energy source for swine, in conjunction with the fact that hogs themselves have less of an affinity for it compared to other vegetable protein sources in terms of palatability (Landero et al., 2011). Another large issue with feeding swine canola meal comes from the presence of glucosinolates. The reduction of glucosinolates in canola meal has been a breeding objective since the early days of the crop's history. Still, glucosinolates have been linked to poor

feed intake, growth, and thyroid function at levels at or above 2.5 µmol/g (Schöne et al., 1997). Bell (1993) made the recommendation that canola meal for swine diets should contain no more than 2.4 µmol/g of glucosinolates, which was later confirmed by the studies performed by Schöne et al., (1997). Recent work by Mejicanos et al. (2016) has shown that breeding efforts in canola have resulted in an increase in the protein and fibre content while decreasing anti-nutritional factors, opening the door for canola meal as a cost-effective alternative to soybean meal.

Additionally, the amino acid digestibility of canola meal is also in question when it comes to swine (Sauer et al., 1986; Stein et al., 2007). It is important to consider the total amount of amino acids that are actually digestible in the swine stomach than to consider the total amount of amino acids present in the meal. The ileum, part of the small intestine, is where a large amount of nutrient absorption takes place (Yen, 2000). Based on this, a unit of measurement for the digestibility of amino acids has been established as the standardized ileal digestibility (SID) (Jondreville et al., 1995; Mosenthin et al., 2000; Jansman et al., 2002). Having this standardized unit of measure allows for easier formulation of feed diets and direct comparison of the availability of particular amino acids amongst different meal types (Stein et al., 2007). Studies looking at SID of amino acids in swine when soybean meal is replaced with canola meal have shown that both digestible lysine and threonine become very limiting, affecting the overall growth rate (Bell et al., 1988; Bell and Keith, 1989). Hickling (1994) conducted feeding trials across Manitoba, Saskatchewan, and Alberta to assess lysine and threonine as limiting amino acids in swine diets and confirmed that they are in fact the first and second-most limiting amino acids.

2.4.3 Poultry Diets

The nutritional requirements of poultry diets are equally as diverse as the previous examples depending on whether the meal is being considered as feed for broilers, breeding hens, or layers (Newkirk and Classen, 2002; Ramesh et al., 2006; Naseem et al., 2006; Oryschak and Beltranena, 2013; Rogiewicz et al., 2015). Overall, one of the largest limitations canola meal has in comparison to soybean meal is the amount of energy available (Wanasundara, 2011; Tan et al., 2011). Many studies have shown the requirements for supplemented dietary enzymes to aid in nutrition uptake from canola meal, recently including one conducted by Jia et al. (2012). Canola is currently used as a feed ingredient in laying hens, though many limitations exist (Canola Council of Canada, 2015). Work done by Novak et al. (2004) has shown a link between insufficient lysine content and egg weight reduction while Kaminska (2003) found evidence of insufficient lysine content in poultry fed canola meal diets. The actual amount of canola meal that should be included in the diet of laying hens has been suggested at 15 to 20 % in separate studies by Oryshack and Beltranena (2013) and Rogiewicz et al. (2015). In the case of both breeding hens and broilers, canola meal provides little to no advantage over other plant-based proteins (Nasser et al., 1985; Ahmadi et al., 2007; Newkirk and Classen, 2002; Ramesh et al., 2006). Therefore, canola meal remains a protein additive fraction rather than a primary component in poultry nutrition.

2.4.4 Aquaculture Diets

Though smaller than those previously discussed, Canada does have a relatively large aquaculture industry consisting of the production of finfish, shellfish, and Atlantic salmon (Canola Council of Canada, 2015). The Canadian aquaculture industry contributes an estimated $244 million to the economy, with 88 % of aquaculture sales coming from finfish production (Statistics Canada,

2014). Studies have shown the value of canola meal as feed in the aquaculture industry, with Fangfang et al. (2014) observing an increase in feed intake when fish were fed canola meal over soybean meal. In general, the digestibility of canola proteins and its constituent amino acids is quite high when fish are fed canola meal. According to the National Research Council (2011), protein digestibility for canola meal is 91 %, 85 %, and 89 % for rainbow trout, blue tilapia, and cobia, respectively.

When compared to other vegetable protein sources, Friedman (1996) stated that the amino acid balance of canola proteins is the best commercially and currently available for fish production. In aquaculture, the protein efficiency ratio (PER) is often considered when comparing different feed meals (Canadian Food Inspection Agency, 2017). The PER corresponds to the amount of weight gained per gram of protein fed. Canola proteins perform well in this regard, having a PER of 3.29 while soybean proteins only have a PER of 1.60 (Friedman, 1996). When looking at improving the nutritional qualities of canola meal it is important to consider its use in all industries. Though the amino acid profile of canola meal is already well-suited for aquaculture diets, the industry could no doubt benefit from an overall increase in available meal protein.

2.4.5 Human Consumption

A basic requirement for human life is an intake of protein in our diets (World Health Organization, 2007). Plant-based proteins make up roughly 65 % of the world's supply of edible protein (Young and Pellett, 1994) and are no doubt an excellent source of the required dietary protein we as humans need. Unfortunately, there are still roughly one billion people in our world that have inadequate protein intake (FAO, 2013). The intensive growth of the world's population has put heavy pressure on global sources of animal-based proteins. As a result, there has never been a higher demand for

not only an increase in quantity but also an increase in quality of plant-based proteins (Wu et al., 2014). Specific human dietary requirements for proteins and amino acids vary by gender and age (World Health Organization, 2007). If the goal is to isolate particular amino acids that are the most important overall, studies vary. However, it is generally agreed upon that lysine, threonine, and tryptophan are easily limiting and indispensable in the human diets (Young and Pellet, 1994; Reeds, 2000). The review conducted by Tan et al. (2011) once again confirms that the amino acid contents in canola meal are balanced in comparison to one another but are lacking in comparison to other vegetable protein sources like soybean. Nevertheless, there have been successful initiatives to move canola into the plant-based protein market with companies like Burcon NutraScience. Burcon released the first canola-based protein isolate powders under the trade names "Puratein®" and "Supertein™" in 2010 (Burcon NutraScience Corporation, 2016). While Puratein® holds more potential as a protein-added baking and emulsifying agent, the recommended use for Supertein™ is as a protein beverage (Burcon NutraScience Corporation, 2016). The introduction of canola protein-based products such as these, especially Supertein™ with its balanced amino acid profile and high digestibility, provide huge potential for canola to compete in a market where soybean-based protein isolates are dominant. After reviewing the dietary requirements of many different animals, including humans, it is easy to see that there is a large amount of untapped potential for canola meal to become an important contributor as a highly nutritious plant-based protein.

It is worth noting as well that values expressed for crude protein or amino acid contents are done so using seed obtained from bulked seed composites from a mixture of cultivars and locations. This is the case with available data from the industry (Canola Council of Canada, 2015) as well as in the case of each of the individual studies of canola meal as a feed in animal diets that has been

previously discussed in this review. The approach to view and understand individual amino acid contents at a genotype-specific level is currently extremely limited. Studies by Wen et al. (2015) in *B. napus* and Warrington et al., (2016) in *G. max* are some of the only studies available and will be discussed in more detail later in this review.

2.5 Comparison of Canola Meal to Soybean Meal

As has been established, canola's largest competitor when considering the meal as a food source is soybean. Soybean is the world's top oilseed with a projected production of 3880 million bushels in the 2016 – 2017 season (United States Department of Agriculture, 2016). Table 2.4 provides a direct amino acid comparison between canola and soybean meals. With dairy cattle being the main exception, soybean meal has many nutritional advantages over that of canola (Baker, 2000; Stein et al., 2008). If strictly considering amino acids, the only advantage canola meal has over soybean meal is that it is naturally higher in the sulfur containing amino acids methionine and cysteine (Canola Council of Canada, 2015; American Soybean Association, 2015). When considering total protein content, the USDA compares canola meal at 34 – 38 % to soybean meal at 44 – 49 % (USDA, 2016a).

Overall, focused and ambitious goals need to be set for the improvement of canola meal if it is to ever compete economically with soybean. The potential for canola meal to surpass soybean does exist in many different markets. As outlined in this review, extensive research has been conducted to better understand the nutritive role of canola meal in the diets of both animals and humans. With an in-depth understanding of where canola's limitations lie, it is perhaps time to focus on meal improvement from a genetic standpoint just as Drs. Keith Downey and Baldur Stefansson did with oil improvement when they first created canola as we know it today. Breeding efforts in canola

have heavily focused on understanding the genetic inheritance behind seed oil content and quality and using this understanding to increase the value of the traits. There is no reason that the same approach cannot be taken to improve the protein content of the seed as well. It is possible to further expand this type of research to look at the contents of particular amino acid values in the meal. Identification of the underlying genetics of these traits may allow plant breeders to oversee the improvement of protein-related traits, enhancing the value of canola on a global scale.

2.6 Introduction to Breeding Methods

2.6.1 Importance of Germplasm Resources for Genetic Diversity

Since the development and commercialization of *B. napus* as "canola" in the 1970s, the crop has been exposed to intensive breeding pressure (Malabat et al., 2003; Wanasundara, 2011; Rahman, 2013; Rahman et al., 2013; Fu and Dong, 2015). The value of the canola industry on both a national and global scale has led to the desire for rapid cultivar development, in many cases creating further uniformity within *B. napus* gene pools (Rahman, 2013). Numerous studies have visually displayed the lack of genetic diversity in both winter and spring canola types through the use of molecular marker analysis. In a study conducted by Hasan et al. (2006), 96 different accessions were assessed, and the group concluded that the spring oilseed types had the least number of unique alleles compared to winter and vegetable types. Qian et al. (2006) compared spring and winter accession to Chinese genotypes derived from interspecies crosses and also found very low levels of diversity among the spring lines. Chen et al. (2007) analyzed a set of Australian cultivars released between the 1970s and 2000s and also reported an observed loss of genetic diversity. Similarly, Fu and Gugel (2010) observed a decline of allelic variation in Canadian canola cultivars developed between the 1940s and 1990s. Bus et al., (2011) assessed the amount of genetic diversity

in winter types and once again confirmed the lack of genetic diversity. It has therefore been well documented that *B. napus* gene pools are in dire need of allelic and genetic diversity. Rahman (2013) stated that breeding efforts based on currently restricted gene pools pose a risk for the future of canola cultivars. Without allelic diversity, the amount of available recombination events may become exhausted, limiting the potential for any further genetic improvement. Rahman (2013) then proposed that the genetic base of breeding programs can be broadened through the introduction of new alleles.

There are numerous approaches that have been successfully applied in breeding programs to enhance genetic diversity through the introduction of unique alleles. Some of these approaches include re-synthesis breeding, utilization of exotic germplasm, and mutagenesis. Re-synthesis breeding refers to the re-creation of the original interspecific hybridization that occurred between *B. rapa* and *B. oleracea* to create *B. napus* (Rahman, 2013). Early generation progeny from re-synthesis crosses contain greater diversity due to lower exposure to intense breeding pressure, as has been applied in the creation of our current low erucic, low glucosinolate varieties (Becker et al., 1995; Jesske et al., 2011; Rahman, 2013). These experimental genotypes can serve as sources of unique alleles for introgression into breeding lines with low diversity. Utilizing the gene pools of related *Brassica* species such as *B. juncea* and *B. nigra* has proven to be useful in the identification of novel alleles and traits in many studies (Rahman, 2013). One of the greatest successes has come from the introgression of disease resistance, including blackleg caused by *Leptosphaeria maculans* (Desm. Ces. and de Not. anamorph Phoma lingam (Tode ex. Fr.) (Desm.). Resistance to the blackleg infection was originally identified in *B. juncea* and has since been introgressed into *B. napus* (Roy, 1984). While these two approaches capitalize on the discovery of alleles that naturally exist in alternative or re-created gene pools, plant breeders have also used

tools to manipulate the current *B. napus* genome to artificially create the unique alleles they seek (Pérez-de-Castro et al., 2012). This has been accomplished through the use of induced genetic mutations (Sikora et al., 2012).

2.6.2 Introduction to Mutagenesis Breeding

Mutation breeding, or mutagenesis, is the act of inducing controlled genetic mutations to gain favorable alleles or traits within a breeding population (Parry et al., 2009; Wilde, 2015). This has been in practice since the early 20th century and has been achieved through use of a variety of mutagenizing agents (Sikora et al., 2012). One of the first documented studies involving induced mutations was conducted by Muller (1927) where he looked at the effect of X-ray treatments on gene mutations in *Drosophila melanogaster* Meigen, 1830. He concluded that the X-ray treatment was capable of increasing the level of mutation in the genome by up to 15 000 % (Muller, 1927). At the same time, Stadler (1928) was studying the effects of mutagenizing agents such as X-ray and radium treatments on barley and maize populations. He discovered that exposure to mutation treatments resulted in an increase in phenotypic variation amongst the plant populations (Stadler, 1928). Though radiation-based techniques have proven to be successful, an overall shift to the use of chemical mutagens has been observed due to the highly destructive nature of radiation elements. In general, chemical mutagens tend to cause single changes in the base pair (bp) sequence of an individual (Sikora et al., 2012). This single base pair change is also known as a single nucleotide polymorphism (SNP) (Brookes, 1999; Mohler and Schwarz, 2004). Though many chemical mutagens exist, one of the most widely and commonly used today is ethyl methanesulfonate (EMS). EMS introduces alkyl groups into the purine compound guanine, one of the nitrogenous bases that make up DNA (Sega, 1984). This alkylation results in DNA-polymerase, the enzyme

responsible for DNA synthesis, to selectively replace cytosine with thymine causing a random point mutation (Sikora et al., 2012). The majority (70 – 99 %) of DNA changes in EMS-mutated populations are GC to AT base pair transitions (Till et al., 2004; 2007).

The introduction of mutagenizing agents into plant breeding programs has been an indispensable tool when it comes to improving existing traits or discovering novel ones. As of 2016, the FAO/IAEA reported 3234 crop cultivars containing one or more desirable traits achieved from mutagenesis (FAO/IAEA Mutant Variety Database, 2016). With the success of EMS and other mutagenizing agents, a new breeding methodology calling TILLING (Targeted Induced Local Lesions in Genomes) has been developed (McCallum et al., 2000). TILLING involves the mutation of individual lines within a population followed by the extraction and characterization of each individual's DNA. These TILLING populations are then screened using next-generation technology to assess both the location and the effect of the mutations. Focus is often applied to a targeted region of the genome previously believed to have importance to a particular trait or breeding objective. In recent years, many TILLING studies have been successful in *B. napus*. Wang et al., (2008) used TILLING to identify mutations within FAE1 (fatty acid elongase 1), a gene in control of seed erucic acid in *B. napus*. They were able to identify mutations resulting in a reduction in seed erucic acid content, a trait vital in the production of low erucic acid "canola" type varieties. A study conducted by Gilchrist et al., (2013) aimed to maximize the potential of TILLING populations by looking at favorable mutations across the entire genome of *B. napus*. They were able to identify 432 unique mutations across 26 genes, a study truly showcasing the potential for mutagenesis studies to benefit breeding programs and future cultivars. There remains an extreme pressure on the agriculture and breeding industries to introduce diversity into their

current and future cultivars to ensure longevity, and mutagenesis is one plant breeding tool that is helping to achieve this.

2.7 Introduction to Molecular Markers

Another tool that has become invaluable to plant breeders and their breeding programs is the use of molecular markers (Jonah et al., 2011). They are key in gaining a deeper understanding of the underlying genetics of breeding materials. Markers allow breeders to identify and track specific loci that are related to traits of interest within their experimental material, the process of which is referred to as marker-assisted selection (MAS) (Langridge and Chalmers, 2004). In addition to understanding gene location, molecular markers assist in the evaluation of polymorphisms present at particular loci across amongst populations (Staub et al., 1996; Langridge and Chalmers, 2004). Applications of molecular markers extend to cultivar identification, studies of genetic diversity and gene mapping (Langridge and Chalmbers, 2004; Mohler and Schwarz, 2004; Jonah et al., 2011). Since the development of the first molecular markers in the 1980s, a wide variety of marker types have been developed, each having unique characterizations and utilizations.

2.7.1 Restriction Fragment Length Polymorphisms (RFLPs)

Restriction fragment length polymorphisms (RFLPs) were discovered by Botstein et al. (1980) through an understanding that DNA restriction enzymes (REs) function through recognition and cleavage at specific DNA sequences. The cutting of DNA by restriction enzymes at such specific sites results in DNA fragments of specific lengths accumulating after RE treatment. Therefore, differences in fragment lengths amongst individuals within a population after exposure to restriction enzymes could provide insight to larger genotypic differences (Botstein et al., 1980).

RFLP markers were used in the construction of the first whole-genome linkage maps in plants (Bernatzky and Tanksley, 1986; Helentjaris et al., 1986). RFLPs allow for visualization of co-dominant alleles at a single locus, though they require large amounts of good quality DNA. Molecular marker analysis using RFLPs is considered very time-consuming and expensive, and the requirement of radioactive agents to detect DNA fragments have all contributed to the decrease in their use in breeding programs (Mohler and Schwarz, 2004). Nonetheless, they opened the door to the development of other molecular markers that have since greatly improved breeding efforts.

2.7.2 Simple Sequence Repeats (SSRs)

Simple Sequence Repeats (SSRs) are part of a classification of molecular markers known as "microsatellites". These markers are tandem repeats of short nucleotide sequences (Mohler and Schwarz, 2004). Tautz and Renz (1984) realized that these short sequence repeats were abundant and evenly spaced throughout the genomes of eukaryotes. SSR markers are able to identify single, co-dominant loci using a small amount of good quality DNA. The number of SSR loci are extremely variable among individuals of the same species, allowing the markers to represent diversity within breeding populations (Goldstein and Pollock, 1997). The simplicity of using SSRs had led to the development of large marker databases across a variety of agriculturally significant plant species including *Solanum tuberosum* L. (Milbourne et al., 1998), *Triticum aestivum* L. (Röder et al., 1998), *Glycine max* L. (Cregan et al., 1999), *Hordeum vulgare* L. (Ramsay et al., 2000), *Oryza sativa* L. (Temnykh et al., 2000), *Zea mays* L. (Sharopova et al., 2002), and *B. napus* (Uzunova and Ecke, 1999). SSR markers have greatly contributed to a further genetic understanding of both crop and other plant species.

2.7.3 Random Amplified Polymorphic DNAs (RAPDs)

Discovered by Williams et al. (1990) and Welsh and McClelland (1990), random amplified polymorphic DNA (RAPD) markers are indiscriminate DNA fragments that are amplified through the use of a single oligonucleotide primer. The oligonucleotide primers used to visualize RAPDs typically contain ten base pairs with varying GC content (Mohler and Schwarz, 2004). The assumption in the use of these markers is that the regions amplified by the primer used are randomly spaced throughout the genome, and genetic differences between individuals will be revealed based on the DNA fragments that are amplified. Though the establishment of universal primers for DNA amplification is a large benefit to this technology, there have been many documented issues with its reproducibility (Ellsworth et al., 1993; Muralidharan and Wakeland, 1993; Schierwater and Ender, 1993). For this reason, the use of RAPD markers has greatly declined.

2.7.4 Amplified Fragment Length Polymorphisms (AFLPs)

Amplified fragment length polymorphism (AFLP) markers (Vos et al., 1995) operate under a very similar understanding to RAPDs. AFLPs are amplified fragments resulting from a double-digestion of genomic DNA using oliger primers (Mohler and Schwarz, 2004). In comparison to RAPDs, AFLPs are amplified using a much higher resolution and more stringent techniques (Mohler and Schwarz, 2004). AFLP markers are considered both reliable and reproducible and have been extensively used in both whole genome and high-resolution genetic maps (Thomas et al., 1995; Simons et al., 1998; Schwarz et al., 1999; Ballvora et al., 2002).

2.7.5 Single Nucleotide Polymorphisms (SNPs)

Considered to be the next generation of molecular markers, single nucleotide polymorphisms (SNPs) are defined as single base pair positions within the genome where different alleles exist between individuals within a population or species (Brookes, 1999; Mohler and Schwarz, 2004). SNPs are one of the most common types of genetic variation and are considered accurate representations of genomic polymorphisms even when present in low numbers (Ganal et al., 2009). They are also well suited for next generation sequencing technologies, having shown great success when used in automated and high-throughput systems. Because SNPs are present in such large numbers in plant genomes, they have been extensively used in the construction of genetic linkage maps, population structure analyses, nested and genome-wide association mapping, map-based gene discovery, and many other breeding applications (Kumar et al., 2012). In *B. napus,* studies have revealed that there are SNPs every 600 base pairs, leading to roughly 1.7 million SNPs based on the ~1 gb genome of *B. napus* (Fourmann et al., 2002; Batley and Edwards, 2007; Hayward et al., 2012). This high SNP density is important when dealing with the allopolyploid nature of *B. napus* as it can become difficult to distinguish which SNPs are occurring within the A or C genome versus between them (Ganal et al., 2009). True SNPs can be discovered by direct sequencing of each of the genomes simultaneously or through genome comparison to the progenitor species, *B. rapa* and *B. oleracea* (Ganal et al., 2009). A study conducted by Huang et al., (2013) used the Illumina® GoldenGate Assay to discover 892,536 SNPs across the entire genome of *B. napus*. This confirmed that there was potential for development of a large genotyping array to be used in future *B. napus* genetic studies.

2.8 Creation of the *Brassica napus* Reference Genome

In 2014, a worldwide effort to sequence the genome of *B. napus* was conducted through the collaborative efforts of over seventy different scientists (Chalhoub et al., 2014). The cultivar that was used as the focus of this initiative was a European winter-type *B. napus* called 'Darmor-*bzh*'. Various types of sequencing methods were utilized by the scientists involved including 454 GS FLX+ Titanium pyrosequencing (Roche Diagnostics Corporation, Indiana, USA), Sanger sequencing (Sanger et al., 1977), and Illumina® HiSeq high-throughput sequencing system (Illumina, California, USA). These separate sequencing initiatives were assembled together to obtain 849.7 Mb of genetic data. "Scaffolding" data from *B. rapa* and *B. oleracea* (20,702 markers total) were assigned to either the A or C genomes and a SNP map was created, physically anchoring 712.3 Mb of the whole genome assembly. The applications of this open-source knowledge and technology are endless, and it has greatly contributed to *B. napus* breeding programs around the world. The authors of the *Brassica* consortium also highlight the loss of genetic diversity that *B. napus* exhibits when compared to its progenitor species, particularly in the areas associated with traits that have been the intense focus of past breeding efforts (ie. seed quality) (Chalhoub et al., 2014; Clarke, 2016). Therefore, this vast scientific effort is also able to contribute to the identification of diversity within exotic germplasm as well as its introgression into more elite breeding lines. For a complete outline of the methods and results of the *Brassica* consortium as well as access to all supplemental materials, please refer to Chalhoub et al., (2014).

2.8.1 Development and Use of the Illumina® 60K *Brassica* SNP Chip

The release of the *B. napus* reference genome greatly contributed to the advancement of other molecular techniques that are being applied to provide a greater understanding of *B. napus*

genetics. One such technology is the Illumina® Infinium 60K SNP Array (Clarke, 2016). Illumina® describes the array as "SNPs extracted and validated in 80 lines from *B. napus*, 4 lines from *B. oleracea*, and 4 lines from *B. rapa*, designed to work in any *Brassica* A or C genome species" (Illumina®, 2016). The array uses Illumina®'s BeadChip technology. The BeadChip works through whole genome amplification followed by an allele-specific hybridization step based on the SNPs in use (Hayward et al., 2012). These hybridization reactions are carried out on the BeadChip and the fluorescence of specific alleles are typically completed using BeadArray Readers, also a product of Illumina® (Hayward et al., 2012). The capacity of these BeadChips in terms of the number of SNPs they assay for and the number of samples that can be accommodated vary widely depending on the species in question and genetic information available for that species. The Infinium *Brassica* array itself contains 28,044 SNPs across the A1 – A10 chromosomes and 30,420 SNPs across the C1 – C19 chromosomes for a total of 58,464 SNPs, hence giving it the "60K" designation (Sharpe, 2012; Clarke, 2016; Mason et al., 2017). Of the 58,464 SNPs available on the array, 52,127 (89%) of them are considered to be a good, working set (Sharpe, 2012; Mason et al., 2017). Since the development of the Infinium *Brassica* array it has been utilized in a number of studies (Li et al., 2014; Hatzig et al., 2015; Mason et al., 2017; Chao et al., 2017). Because of its distribution of SNPs across the genome, it has been successful in genome-wide association mapping studies (GWAS). Li et al. (2014) used the 60K array to analyze 472 *B. napus* accessions and found two genomic regions on chromosomes A8 and C3 that were associated with seed erucic acid content, demonstrating the SNP chip's use in GWAS studies. Another GWAS study performed by Hatzig et al., (2015) analyzed 248 diverse *B. napus* lines and identified a series of *B. napus* genes orthologous to *Arabidopsis thaliana* genes associated with

seed germination and seedling vigor. In addition to GWAS studies, the 60K *Brassica* chip has also contributed greatly to the study of quantitative trait loci (QTL) analysis.

2.9 Introduction to Quantitative Trait Loci (QTL) Analysis

Many agriculturally significant traits are genetically complex, their expression determined not by a single gene at one locus but rather the interaction of many genes (MacKay et al., 2009). Traits controlled by a single gene are considered qualitative traits, while traits controlled by multiple genes are considered quantitative traits (Miles and Wayne, 2008; Asins et al., 2010). It is through this designation that regions containing a group of loci that are responsible for or related to a particular trait are called quantitative trait loci (QTL) (Asins et al., 2010). QTL analysis itself is a statistical procedure that compares the genetic variation among individuals within a population to link together phenotypic data and genotypic data (Kearsey and Farquhar, 1998; Lynch and Walsh, 1998). QTL analysis statistically compares the phenotypic variation within a population to the genetic differences amongst the individuals within it (Miles and Wayne, 2008). Through this comparison, QTL analysis allows for the identification of specific chromosomal regions that contain loci associated with the phenotype in question (Miles and Wayne, 2008; Asins et al., 2010). The basis of conducting a QTL analysis begins with the design of the population in focus. When creating a QTL mapping population, the two parental lines should be genetically distinct from one another when it comes to the trait of interest (Miles and Wayne, 2008). The breeding of these intentionally distinct individuals allows for the maximum amount of polymorphism with respect to the studied phenotype. Typically, an F_2 population is used as the subject of the mapping study as the maximum amount of phenotypic variation is observed at this time (Miles and Wayne, 2008). Phenotypic evaluation is conducted based on the trait being studied. The second component that

is required is genotypic information that is achieved through a collection of molecular markers (SSRs, RFLPs, SNPs, etc). Once both phenotypic and genotypic datasets are established, QTL analysis is completed by comparing individual markers to the resulting phenotype across individuals within the population (Miles and Wayne, 2008; Asins et al., 2010). Molecular markers that exist in the regions of QTL associated with the phenotype will more frequently segregate with the trait, while molecular markers that aren't associated with the QTL region will have a non-significant relationship with the trait (Miles and Wayne, 2008). QTL have been identified for numerous traits in numerous species, and from a plant breeding perspective have greatly contributed to the selection for and improvement of traits based on a genetic understanding (Mauricio, 2001).

2.9.1 QTL for Seed Quality Traits in Canola

While QTL are well documented in many crop species, work to characterize genomic regions in *B. napus* has been limited to seed oil, erucic acid, and glucosinolate contents as well as for traits such as disease resistance and yield. Recently, Wang et al. (2013) used a doubled haploid (DH) population of 348 individuals and 403 SSR and sequence-related amplified polymorphism (SRAP) molecular markers to identify a total of 63 QTL related to oil content explaining 2.6 – 17.9 % of the phenotypic variation in the population. From these, they were able to identify 24 consensus QTL on A1, A2, A3, A8, A9, A10, C3, C5, C6, C8, and C9, explaining 3.6 – 11.7 % of the observed variation. In addition, Qiu et al. (2006) analyzed 202 doubled haploid *B. napus* lines using RFLP, SSR, AFLP, and SNP markers. It should be noted that this particular population was given the designation of the 'TN DH' mapping population, based on it being derived from winter *B. napus* cultivars Ningyou 7 and Tapidor. This same TN DH population has since been exploited in QTL

analyses for many other traits. In this study, they were able to identify four major QTL on chromosomes A8 and C3 for erucic acid content controlling 45.0 % and 30.0 % of phenotypic variation, respectively. This discovery is consistent with the previously discovered QTL, *eru1* and *eru2* also on A8 and C3, found by Howell et al. (1996). In addition to erucic acid, they found seven QTL for oil content on A1, A8, and C3 describing a cumulative 55.0 % of the variation. Breeding focus has been heavily applied in the discovery of genomic regions controlling oil content with it being the most economically significant trait in *B. napus*. However, not until very recently have breeders began to expand QTL analyses to look at protein-related traits. One of the first and only published studies looking at QTL for amino acid related traits was conducted by Wen et al. (2016) using the 202 'TN DH' mapping population. A varied set of 786 molecular markers (RFLP, AFLP, MS-AFLP, SSR, SNP, STS, SSCP, CAPS) were used. They discovered 28 QTL associated with arginine, histidine, glutamic acid, glycine, proline, alanine, and aspartic acid. They considered 12 of these QTL to be major, explaining 13.3 to 35.7 % of the phenotypic variation. The largest clusters of QTLs they discovered in relation to the amino acids mentioned were located on chromosomes A1, A4, A5, A7, and C2. A study conducted by Xu et al., (2013) is one of the only published studies to compare phenotypic and genotypic data to look for regions controlling seed protein content. They too used the TN DH mapping population and discovered 3 QTL for protein content on chromosomes A4, A7, and C5 cumulatively explaining 59.6 % of phenotypic variation. They consider the QTL located on A4 to be the only major QTL, explaining 37.9 % of the variation on its own. More recently, a study conducted by Chao et al. (2017) discovered QTL for both seed protein and seed oil content while documenting the genetic relatedness of the two traits. Chao et al. (2017) discovered 67 QTL for seed oil content and 38 QTL for seed protein content across 12 of the 19 chromosomes in *B. napus*, explaining up to 22.2 % and 27.5 %, respectively. These

studies looking at oil, protein, and amino acid contents have shown the potential for QTL discovery in relation to these quality traits. However, conducting larger QTL mapping studies in more desirable populations is required for further confirmation of the genomic regions if they are to be used in future breeding efforts to improve protein-related traits.

2.9.2 QTL for Protein Related Traits in Soybean

Breeding efforts to identify genomic regions associated with protein and amino acid contents have been successfully implemented in soybean, a crop in which these traits are currently of higher importance. A study conducted by Fallen et al., (2013) used 282 recombinant inbred lines (RILs) and 1,536 SNP markers in the discovery of 10 QTLs on chromosomes 5, 7, 9, 10, 13, and 20 in *G. max*. These QTL explained between 5 and 14 % of the trait variation while considering a variety of the eighteen amino acids that were analyzed. A more recent study conducted by Khandaker et al. (2015) addresses the fact that QTL analysis for protein-related traits is also still limited in soybean in relation to other traits. They used 92 recombinant inbred lines and 537 SNP markers from the Illumina® Infinium SoySNP6K BeadChip array. They discovered 13 QTL related to threonine, proline, serine, tryptophan, cysteine, arginine, and histidine. The amount of variation ranges from 23.0 % for threonine content to 93.0 % for serine content. Finally, Warrington et al. (2015) recently looked for both protein and amino acid content QTL in 140 RILs that were grown across 5 different environmental locations. They discovered 4 QTL for crude protein content, one being major and explaining 55.0 % of phenotypic variation. They analyzed individual amino acids in relation to the crude protein content and found QTL for lysine, threonine, methionine, cysteine, and methionine/cysteine together that explained between 6.0 and 52.0 % of the phenotypic variation.

In closing, literature on protein and amino acid focused QTL discovery is currently limited in *B. napus* at this time. Only recently has there been research focused on breeding these traits in a molecular and genomic-based manner. Research initiatives on this topic have proven to be limited even in soybean, whose meal is more economically important on a global scale than that of canola. With the number of molecular tools on hand and a growing demand for improved protein sources in plant-based foods, discovery of protein-related QTL could potentially have a lasting impact on future breeding initiatives in *B. napus*.

The objectives of this research were to use a molecular and genomic-based approach to identify chromosomal regions within *B. napus* that are associated with traits contributing to improved protein quality and quantity. The use of breeding materials treated with EMS gave rise to the allelic variation necessary to identify genomic regions associated with protein-related traits such as seed protein content and specific amino acid contents within the meal. The goal of this research was to provide a greater understanding of the genetic control of these traits by applying modern breeding techniques for characterization and use in future breeding efforts. In two separate studies, QTL mapping of seed protein content and meal amino acid contents were performed on the same population of *B. napus*. This research aims to provide information that will advance breeding for protein and amino acid content in *B. napus*. The novel discovery of these genetic regions will provide another tool for future *B. napus* improvement in a market where improvement is needed for canola meal to become a true competitor with that of soybean. Even within current markets where canola meal is utilized, there exists room for improvement in both overall protein quantity and quality. There is a growing global need for plant-based proteins for both human and animal consumption. The results of this research provide information that could make this value-added initiative a reality.

3.0 Discovery and Mapping of Quantitative Trait Loci Contributing to Seed Oil Content and Seed Protein Content in *Brassica napus* L.

3.1 Abstract

Canola (*Brassica napus* L.) has many economical by-products that have been a key to its success in the agricultural sector. Seed oil and fatty acid content have been the focus of breeding efforts in *B. napus* since the development of the crop in the 1970s. During this same time period, little research effort has been directed to modifying seed protein content. The total protein fraction of canola seed should also be considered an important trait when it comes to improving the nutritional potential of canola as a plant-based protein for human and animal consumption. The current research was initiated to determine the underlying genetic control of seed protein and seed oil content, as well as to study the relationship between the two. Through a quantitative trait loci (QTL) analysis, we revealed regions on chromosome A3 controlling 4.6 – 7.2 % of the phenotypic variation for seed oil content and 11.6 – 12.3 % of the phenotypic variation for seed protein content. The negative correlation between seed oil and seed protein contents was documented as well as the co-localization of QTL for the two traits in both years of the study. These discoveries provide useful information for future breeding efforts in *B. napus* to enhance seed quality from a protein and protein-related trait perspective. The diversification of seed quality beyond oil is important to ensure a strong future for the canola industry in Canada and around the world.

3.2 Introduction

Canola (*Brassica napus* L.) is highly significant to Canada's agricultural industry, contributing $26.7 billion to the Canadian economy (LMC International, 2016). The seed quality characteristics

of canola have been a large focus of breeding efforts in the past and have also largely contributed to its rapid adoption over the last several decades (Morrison et al., 2016). Canola's oil products are considered one of the healthiest edible oils available in the market for human consumption (Przybylski et al., 2005). Selection pressure to increase seed oil content through breeding has been applied intensely in *B. napus* (Malabat et al., 2003; Rahman et al., 2013). While this has been successful in increasing seed oil content to current averages of 45 % compared to 40 % in the first low erucic acid rapeseed cultivar Tower (Stefansson and Kondra, 1975; Rahman, 2013), there has also been an observed decrease in both the quantity and quality of canola's proteins (Malabat et al., 2003; Wanasundara, 2011; Rahman et al., 2013; Wanasundara et al., 2016).

This negative relationship between seed oil and seed protein content is well documented (Asare and Scarisbrick, 1995; Andersen et al., 1996; Brennan et al., 2000; Hao et al., 2004; Ahmad et al., 2007; Hu et al., 2013) and has provided challenges when it comes to preserving the functionality of canola's protein as a nutritional food source. In addition to the negative correlation between their expression of these two seed quality traits, Chao et al (2017) discovered that similar genetic regions are responsible for their control. Understanding the interrelatedness of seed oil and seed protein is important for future efforts to further improve canola oil while minimizing the detrimental effect that such a narrow focus can have on the resulting protein.

Knowledge of the genetic control of oil content in *B. napus* is well documented and is actively utilized in breeding programs. Various molecular markers have been used to execute quantitative trait loci (QTL) mapping studies and have revealed regions associated with seed oil content on 17 of the 19 chromosomes in *B. napus* (Delourme et al., 2006; Rahman et al., 2013). Molecular breeding approaches have been somewhat limited in canola with respect to seed protein content. To date, QTL controlling seed protein have been reported on chromosome A7 by Würschum et al.

(2012) explaining 10.2 % of variation and on chromosome A4 by Wu et al. (2013) explaining 37.9 % of variation. In a study conducted by Chao et al. (2017), 68 QTL for seed protein content were discovered on 12 different chromosomes in *B. napus* controlling 2.2 to 27.5 % of observed phenotypic variation. QTL for seed protein content have also been reported in *Glycine max* L. Merr (Warrington et al., 2015) as well as *Arabidopsis thaliana* L. Heynh (Jasinski et al., 2016). Warrington et al. (2015) reported the discovery of four crude protein QTL in *G. max* with the strongest controlling 55 % of phenotypic variation, located on chromosome 20. Similarly, Jasinski et al. (2016) also discovered four seed protein content QTL in *A. thaliana* with the strongest located on chromosome 4 and explaining more than 10 % of the observed phenotypic variation in the mapping population. Gaining a better understanding of not only the control of seed protein content, but the relationship it shares with seed oil content, is crucial for enhancing diversity in *B. napus* with respect to overall seed quality.

The objectives of this research were to further understand the genetic control of seed protein content and to study the relationship between seed oil content and seed protein content. A DH mapping population of 124 individuals designed for diversity in seed quality was evaluated through a quantitative trait loci (QTL) mapping analysis using genotypic datasets generated with the Illumina® 60K *Brassica* SNP array. The identification of genomic regions controlling these traits, as well as a closer look at the phenotypic and genotypic relationship they share will provide valuable information for future efforts to improve canola's overall seed quality. The research hypothesizes that linkage mapping and QTL analyses will result in identification chromosomal regions associated with phenotypic expression of seed oil content and seed protein content in *Brassica napus*.

3.3 Materials and Methods

3.3.1 Plant Material and Field Evaluation Details

A diverse set of 343 *B. napus* genotypes was evaluated in field nursery experiments at the University of Manitoba in Winnipeg, Manitoba from 2014 – 2016. The genotypes were derived from two *B. napus* genotypes; ZSDH5225 and ZSDH5825. ZSDH5225 was developed from a cross of Millenium03 / 1123 (McVetty et al., 2000). F_1 seeds were subject to ethyl methanesulfonate (EMS) treatment (Sega, 1984). These plants were pollen donors for standard doubled haploid (DH) procedures (Chen et al., 1994) and ZSDH5225 was selected from this DH population with a focus on fatty acid content. ZSDH5825 is a DH line developed from a backcross between 04-2026 // 04-2026 / Millenium03 (McVetty et al., 2000). F_1 genotypes from this cross were used as pollen donors for the DH process and ZSDH5825 was selected from this DH population. Seed quality characteristics of the two selected genotypes are quite similar with ZSDH5225 exhibiting 42.2% seed oil and 31.5% seed protein and ZSDH5825 exhibiting 39.8% seed oil and 33.0% seed protein. Further, ZSDH5225 was crossed with ZSDH5825 to produce the population that was used as the basis for the research described below.

The genotypes in this study were phenotypically evaluated in single row nursery trials at the University of Manitoba's Winnipeg research field testing station (Latitude 49.80 and Longitude -97.16). Soil at the Winnipeg site is classified as "Riverdale Silty Clay" as described in soil surveys conducted by Ehrlich et al. (1953). For all years of field study (2014, 2015, 2016) 0.50 g of seed was measured for each genotype and planted in 3 m rows with a row spacing of 0.40 m. The spring rapeseed cultivar MilleniUM 03 (McVetty et al., 2000) was inserted every twentieth entry to be used as a check and guard rows were seeded on all outer edges of the experimental rows planted

to serve as a border. In 2014, the genotypes were planted on May 17 in a single 3 m row, single replication field experiment as described above. Prior to initiating this research, Edge® Granular Herbicide (Gowan Canada, Winnipeg, MB) was worked into the soil using a tandem disc in the fall of 2013. Fertilizer, based upon a soil test of the plot area was applied using a broadcast spreader at a rate of 123.3 kg/ha of nitrogen, 44.8 kg/ha of phosphorus, 0 kg/ha of potassium, and 28.0 kg/ha sulfur. Decis® (Bayer CropScience, Leverkusen, Germany) was applied once on May 27 2014 and once on June 4 2014 at a rate of 60 ml per 37.9 litres of water per hectare for control of flea beetles. On June 10 2014, a tank mixture of Lontrel™ 360 (Dow AgroSciences, Indianapolis, IN), Poast® Ultra herbicide (BASF, Ludwigshafen, Germany), and Muster® (DuPont Canada, Mississauga, ON) was applied. Lontrel™ 360 was applied at a rate of 0.67 litres/hectare for thistle control, Poast® Ultra herbicide was applied at a rate of 0.67 litres/hectare with 1 litre/100 litres water of Merge surfactant for grassy weed control, and Muster® was applied at a rate of 29.7 grams/hectare for broadleaf weed control. Determination of physiological maturity is described in canola as the point at which 60 % of the seed on main stem has changed in colour from green to brown (Canola Council of Canada, 2012). A small sample of pods were selected from nursery rows each year to assess for maturity prior to harvest. The nursery trial was harvested between September 9 and 11, 2014. All rows were individually harvested in the field using a Wintersteiger small-plot combine (Nursery Master Classic, Wintersteiger, Salt Lake City, UT). Following harvest, seed samples were manually cleaned using a spiral seed cleaner (Can-Seed Equipment Ltd., Saskatoon, SK) and a 30 g sample was taken from the total yield and analyzed for seed quality.

From the original population of 343 *B. napus* genotypes, selections were conducted based upon quality characteristics such as seed protein and seed oil content. Selections were made with focus on achieving both high seed oil and seed protein contents while maintaining the variation

appropriate for a mapping study. From the original population, 143 genotypes were selected from the ZSDH5225 / ZSDH5825 population and the selected genotypes were carried forward and evaluated in a completely randomized experimental design with three replicates per genotype in 2015 and 2016.

In 2015, the 143 selected genotypes were planted on May 23 in a 3 m row, completely randomized field experiment with three replicates per genotype. Prior to initiating this research, Edge® granular herbicide was worked into the soil using a tandem disc in the fall of 2014. Fertilizer, based upon a soil test, was applied a rate of 151.3 kg/ha of nitrogen, 44.8 kg/ha of phosphorus, 0 kg/ha of potassium, and 33.6 kg/ha sulfur. Decis® was applied as described above on May 30, 2015 and June 10, 2015 for control of flea beetles. On June 9 2015 a tank mixture of Lontrel™ 360, Poast®, and Muster® was applied as described above. Lontrel™ 360 was applied at a rate of 0.84 litres/hectare, Poast® Ultra at a rate of 0.27 litres/hectare with 1 litre/100 litres water of Merge, and Muster® at a rate of 29.7 grams/hectare. The nursery trial was harvested between September 23 and 24, 2015 using the same protocol and equipment as described above.

In 2016, the genotypes were planted on May 7 in a 3 m row, randomized triple replicate field experiment. Prior to initiating this research, Edge® granular herbicide was applied as described above in fall of 2015. Fertilizer, based upon a soil test, was applied with a rate of 134.5 kg/ha of nitrogen, 44.8 kg/ha of phosphorus, 0 kg/ha of potassium, and 28.0 kg/ha sulfur. Decis® was applied as described above on June 14 2016 for control of flea beetles. On June 16 2016 a tank mixture of Lontrel™ 360, Poast®, and Muster® was applied as described above. Lontrel™ 360 was applied at a rate of 0.84 litres/hecate, Poast® Ultra at a rate of 0.67 litres/hectare with 1 litre/100 litres water of Merge, and Muster® at a rate of 29.7 grams/hectare for weed control. Sub-samples

(10 – 20 g) were hand harvested and the nursery rows were bulk-combined using the same protocol and equipment as described above.

3.3.2 Seed Quality Analysis

All genotypes were evaluated for seed oil and seed protein content using near infrared spectroscopy (NIR) analysis. A monochromator FOSS NIR System, Model 6500 (Foss NIRSystems Inc., Maryland, USA) was used to quantify these values using the American Oil Chemist's Society (AOCS) approved guidelines for the determination of seed protein, oil, and moisture (Daun et al., 1994; Kim et al., 2007; Javed et al., 2016). NIR analysis was conducted at the University of Manitoba's seed quality lab, which is annually certified by the Canadian Grain Commission. Briefly, seed was measured into 6 g samples and placed into standard ring cups before being placed into the NIR for scanning. Spectral data was collected based on measurement of the reflectance of the sample at wavelengths ranging from 400 to 2500 nanometers. WinISI II software (Version 1.04a) was used for processing and statistical analysis of the reflectance data. Results are expressed as the amount of seed protein and seed oil expressed as a percentage of the total whole, intact seed on a dry matter basis.

3.3.3 DNA Extraction and Quantification

The original greenhouse-developed seed for each of the genotypes were seeded into 4 x 3 cell packs containing Sunshine Metro Mix potting soil (Sungro® Horticulture, MA, USA), watered daily, and kept for germination inside a growth chamber (day temperature 22 C; night temperature 18 C; light cycle 16 h light, 8 h dark) for procurement of DNA for sequencing analysis. When the plants reached the two-leaf stage, they were transplanted into 15 x 15 cm pots containing Sunshine

Metro Mix potting soil, watered daily, and grown to maturity inside a greenhouse facility (day temperature 25 C; night temperature 22 C; relative humidity 40 – 50 %; light cycle 16 h light, 8 h dark). Tissue from a newly expanded leaf (the youngest tissue) was sampled for each of the genotypes when the plants had reached Growth Stage 3 (stem elongation or "bolting" phase) (Canola Council of Canada, 2017c) and the tissue was subsequently stored at -80 C until DNA extraction could occur.

DNA was extracted from the leaf tissue using a modified version of the cetyl trimethylammonium bromide (CTAB) extraction protocol (Murray and Thompson, 1980). Leaf tissue was ground in a 1.5 ml micro-centrifuge tube. A volume of 400 µl of 2X CTAB buffer was added to the ground leaf tissue before it was gently ground again. The samples were incubated for 60 min at 65 C. An equal volume of chloroform:isoamyl alcohol (24:1) was added to each sample and samples were centrifuged at 10 000 g for 5 min for phase separation. The aqueous phase was transferred to a new 1.5 ml micro-centrifuge tube. For precipitation, 5 M ammonium acetate (0.40 volumes) and isopropanol (2 volumes) were added and the samples were kept overnight at -20 C. DNA was recovered through a 10 min centrifugation period at 12 000 g. The resulting DNA pellet was washed in 70 % ethanol for removal of isopropanol residue before allowing it to air dry. The DNA was resuspended in 400 µl of Tris-EDTA (TE) buffer and 1.0 µl of 10 mg per ml heat-treated RNAse. The samples were then incubated for 60 min at 37 C. Following incubation, an equal volume of phenol was added and samples were centrifuged at 5 000 g for 5 min for separation of the aqueous and organic layers. The aqueous layer was transferred to a new 1.5 ml micro-centrifuge tube where an equal volume of phenol:chloroform (1:1) was added before another round of centrifugation. The process was repeated once more using an equal volume of chloroform. The DNA was precipitated for a second time through the addition of 3 M sodium acetate (0.10 volumes)

and ethanol (2.0 volumes). Samples were left to precipitate overnight at -20 C before being centrifuged at 10 000 g for 15 min for DNA recovery. The resulting pellet was rinsed with 70 % ethanol and left to air dry before being re-suspended in TE buffer for storage. DNA was quantified using a NanoDrop™ 2000 UV-Vis Spectrophotometer (Thermo Fisher Scientific, Massachusetts, USA). A volume of 2 µl of DNA representing each of the *B. napus* genotypes was loaded on to the pedestal using a micro-pipette. The quantity of DNA (ng/µl) was obtained by taking a mean of three samples for DNA quantification. The absorbance of the sample at 260 nm and 280 nm was obtained to determine a A260:A280 ratio. The ratio was used as a determinate of sample quality and purity. All results were recorded and used as a representation of DNA quantity and quality for SNP array analysis.

3.3.4 Illumina® 60K *Brassica* Array Analysis

Sequencing of the *B. napus* population using the Illumina® 60K *Brassica* array was conducted as described by Clarke et al. (2016). Visualization of SNP marker data was completed using the GenomeStudio V2011.1 software available from Illumina Inc. (California, USA). The working dataset from the array contained 52,157 SNPs (Clarke et al., 2016). Criteria for selection of polymorphic SNP data within Genome Studio was conducted by filtering of all SNPs with a GenTrain score of zero. This resulted in the removal of 4,854 SNPs, leaving 47,303 SNPs. Based on visual evaluation of the dataset, SNPs with a call rate of less than 28 % were filtered out. This resulted in the removal of 5,339 SNPs leaving a total of 41,964 in the working dataset. Based on visual evaluation of the dataset, SNPs with a minor allele frequency of less than 10 % were filtered. This resulted in the removal of 36,741 SNPs to leave a working dataset of 5,223 SNPs. Given the doubled haploid nature of the populations in this study, any SNP with an AB frequency of more

than 10 % were filtered. This removed 396 SNPs, leaving 4,827. The filtering of any SNPs with "no calls" greater than 60 removed 18 more SNPs, leaving 4,809 in the working dataset. Finally, three SNPs were manually removed (Bn11-p6742595, Bn11-p454442, Bn11-p449385) on the basis of too many missing data points. This resulted in a final working dataset of 4,806 SNP markers. This working dataset was used for linkage mapping and QTL analysis.

3.3.5 Linkage Mapping Analysis

Development of linkage groups was completed using MSTMap (Wu et al., 2008) and MapDisto 1.7.7 (Lorieux, 2012). The curated marker dataset of 4,806 SNPs was run in MTSMap using the Kosambi mapping function (Vinod, 2011). The population was designated as a DH population with mapping size threshold of 2 cM, a mapping distance threshold of 15 cM, a p-value threshold of 1×10^{-9} and a missing data threshold of 50%.

The linkage groups developed from MSTMap were separated by chromosome and manually curated. The curation process involved a visual and manual manipulation of marker data based on a color-blocking scheme. The SNP markers were manually sorted and organized based on both the Darmor-ID, indicating the SNP's physical position, and the visual fit of the marker across all genotypes in question. In total, 11 genotypes were culled across all linkage groups as a result of missing data. Following manual organization, the markers were placed into marker "bins" through the selection of one marker to represent a series of coincidence markers. The representative markers were then used to create a smaller overall mapping file run in MapDisto 1.7.7 (Lorieux, 2012).

The curated mapping files were run per chromosome using MapDisto 1.7.7 using the Kosambi mapping function, with SARF (sum of adjacent recombination fractions) ordering criteria, seriation ordering method, and a classical recombination frequency (RF) estimate. A LOD score minimum of 3.0 and r value maximum of 0.30 were used to separate loci data into linkage groups. Loci with more than 10 % missing data were also sorted using the MapDisto 1.7.7 program. The marker order and distance developed from these maps was used to create the data files run in QGene 4.3.10 (Joehanes and Nelson, 2008) for QTL analysis.

3.3.6 Quantitative Trait Loci Mapping Analysis

Quantitative trait loci (QTL) mapping analysis was conducted using the Simple Interval Mapping (SIM) procedure in QGene 4.3.10 (Joehanes and Nelson, 2008). Linkage mapping data from MSTMap and MapDisto were prepared and run in QGene using the "QTL mapping" function. QGene was used to obtain LOD scores and R^2 values for seed protein and seed oil content. The R^2 value represents the percentage of observed phenotypic variation that is explained by the marker-trait association represented by the QTL peak. The larger mapping population (ZSDH5225 / ZSDH5825) of 143 individuals was reduced initially during the development of the mapping dataset through manual removal of 11 genotypes with missing data. The resulting population of 132 individuals was furthered reduced through the removal of 8 additional genotypes that had issues during amino acid determination. Therefore, QTL analysis was performed on a working mapping dataset of 124 *B. napus* genotypes.

3.3.7 Statistical Analysis

Statistical analysis was performed using Statistical Analysis System (SAS) version 9.3 (SAS Institute, North Carolina, USA) to conduct an analysis of variance (ANOVA) using PROC MIXED to determine the individual effects of genotype and year on the phenotypic expression of seed oil content and seed protein content. Phenotypic correlations between seed protein and seed oil content were determined through the "Trait Correlation" function in QGene 4.3.10.

3.4 Results

3.4.1 Phenotypic Variation

Individuals selected within the ZSDH5225 / ZSDH5825 population following the 2014 field analysis displayed seed oil content ranging 41.2 to 51.8 % and seed protein content ranging from 22.7 to 30.7 % and were selected based on this observed variation in seed quality with focus on maintenance of both oil and protein content. In 2015, the distribution of seed oil content ranged from 37.4 to 46.9 % and had a mean of 43.2 % (Table 3.1, Figure 3.1). In 2016, seed oil content ranged from 37.2 to 43.0 % with a mean oil content of 40.1 %. Seed oil content differed by 3.1 % between 2015 to 2016. Phenotypic distribution of seed protein content in 2015 for the 124 B. napus genotypes ranged from 28.2 to 34.1 % (Table 3.1, Figure 3.2). The mean seed protein content in 2015 was 30.8 %. Seed protein content of the same individuals following the 2016 analysis ranged from 29.9 to 34.9 %. The mean seed protein content was 32.2 % in 2016. Though no selection pressure was applied on the mapping population following the 2015 season, there was an observed 1.4 % increase in mean seed protein in 2016.

Within the population, specific genotypes exhibited exceptional performance for both seed oil and seed protein content, simultaneously (Table 3.2). Genotype 13ZSDH182 expressed 48.1 % seed oil and 28.1% seed protein in 2015, and 41.5 % seed oil and 31.4 % seed protein in 2016. Genotype 13ZSDH130 expressed 47.4 % seed oil and 28.9 % seed protein in 2015, and 43.0 % seed oil and 31.3 % seed protein in 2016. Both of these genotypes expressed high oil content, exceeding the industry mean in 2015 while displaying exceptional protein values in both years. Genotype 13ZSDH348 expressed 44.5 % seed oil and 31.1 % seed protein content in 2015, and 41.1 % seed oil and 31.6 % seed protein in 2016. For this genotype, there was less of an observed difference in trait value between the two years of data, maintaining acceptable oil content while exceeding industry standards for protein content for both years. In addition, genotype 13ZSDH386 expressed 42.7 % seed oil and 33.3 % seed protein content in 2015, and 40.1 % seed oil and 33.1 % seed protein content in 2016. These genotypes are representative examples of the acceptable performance of the population in terms of both seed oil and seed protein content, simultaneously.

Table 3.1 Phenotypic values and distribution for seed oil content and seed protein content showing value ranges and means for 124 *Brassica napus* genotypes determined using near infrared spectroscopy (NIR) in 2015 and 2016.

	Seed Oil Content (%)	**Seed Protein Content (%)**
2015 Range	37.4 – 46.9	28.2 – 34.1
2015 Mean	43.2	30.8
2016 Range	37.2 – 43.0	29.9 – 34.9
2016 Mean	40.1	32.2

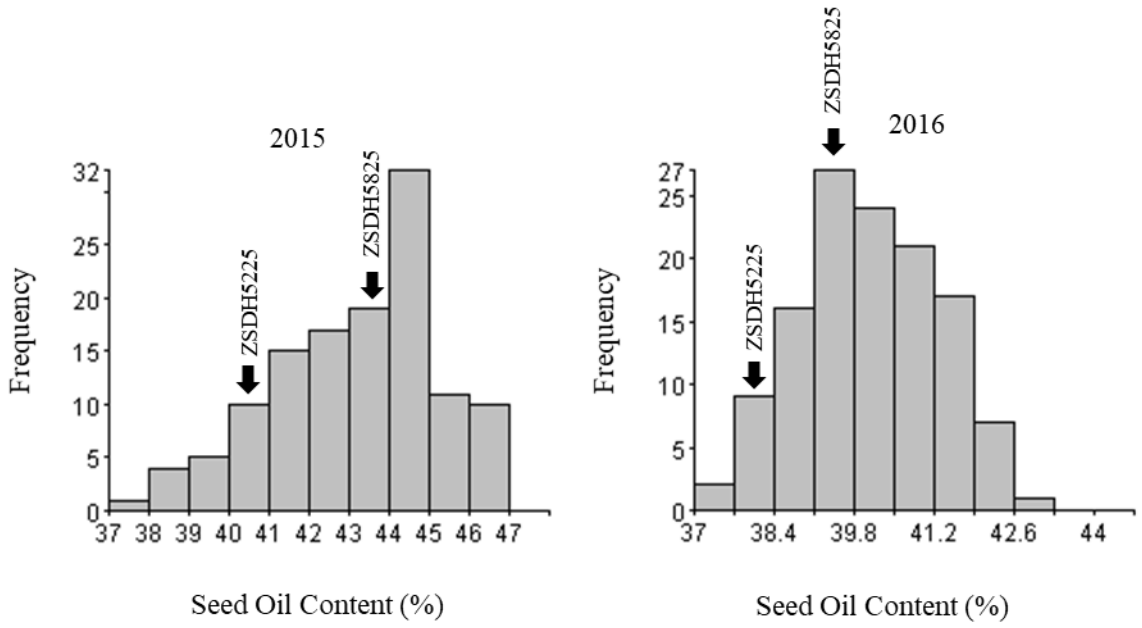

Figure 3.1. Phenotypic distribution of seed oil content determined using near-infrared spectroscopy (NIR) amongst 124 *Brassica napus* genotypes in a DH mapping population following field analysis at the University of Manitoba in 2015 (left) and 2016 (right).

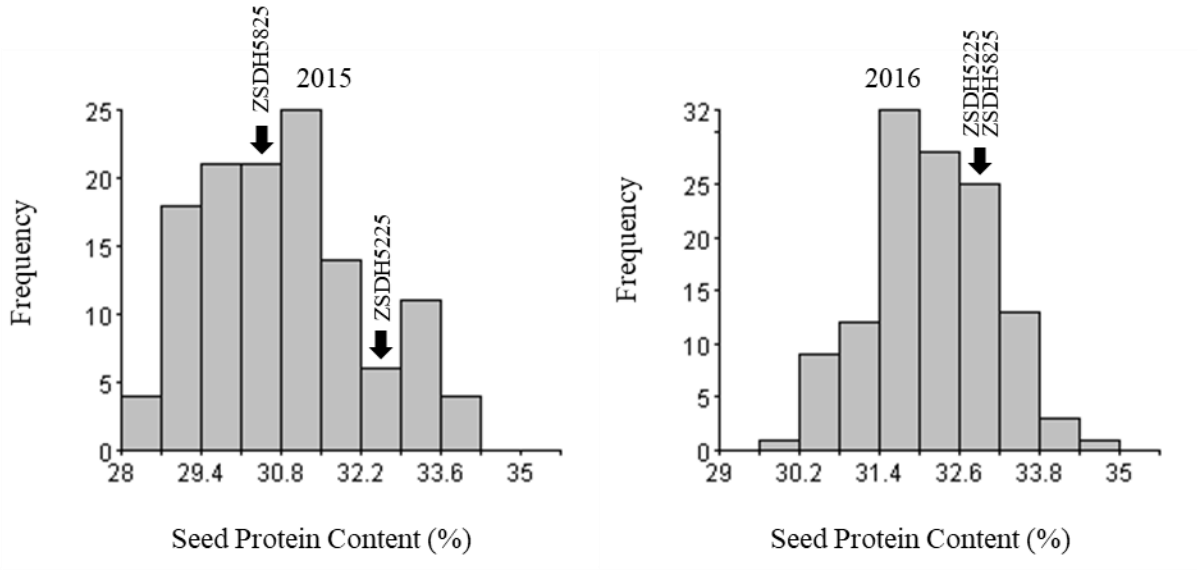

Figure 3.2. Phenotypic distribution of seed protein content determined using near-infrared spectroscopy (NIR) amongst 124 *Brassica napus* genotypes in a DH mapping population following field analysis at the University of Manitoba in 2015 (left) and 2016 (right).

Table 3.2. Phenotypic data highlighting performance amongst the 124 *Brassica napus* genotypes included in the mapping study, showing seed quality values for both seed oil and seed protein content determined using near infrared spectroscopy in 2015 and 2016.

Genotype	Seed Oil 2015 (%)	Seed Oil 2016 (%)	Seed Protein 2015 (%)	Seed Protein 2016 (%)
13ZSDH182	48.1	41.5	28.1	31.4
13ZSDH130	47.4	43.0	28.9	31.3
13ZSDH348	44.5	41.1	31.1	31.6
13ZSDH386	42.7	40.1	33.3	33.1

3.4.2. Statistical Analysis

An ANOVA was used to evaluate the effect of both genotype and year on the phenotypic expression of seed oil and seed protein content. Genotype and year were found to have a significant effect on the expression of both traits (Tables 3.3 and 3.4). In all cases, the reported p-values were less than 0.0001 indicating the significance of both plant genotype and year on seed oil and seed protein contents. As a result, QTL discovered in the subsequent mapping study were presented separately for 2015 and 2016. In both 2015 and 2016, a negative correlation was observed between seed protein content and seed oil content. The correlation coefficient was -0.85 between the two traits in 2015 and -0.73 between the two traits in 2016.

3.4.3 Linkage Mapping Analysis

Initial analysis of the ZSDH5225 / ZSDH5825 population in MSTMap gave rise to a total of 31 linkage groups covering 2595.64 cM. Prior to analysis in MSTMap, the marker dataset was curated using the filtration parameters in GenomeStudio. The resulting 4,806 markers were distributed across these 31 linkage groups. Negligible coverage was discovered within many of these linkage groups and following further curation, several small linkage groups were removed. Marker coverage was poor on chromosomes A2 and A8, forcing them to be absent from this study. In many cases, two or more smaller linkage groups were discovered on the same chromosome due to gaps in chromosome coverage between the groups themselves, preventing the development of one large linkage group. Following data curation, prior to analysis in MapDisto, the number of linkage groups was reduced from 31 linkage groups to 25 linkage groups. The total map distance was also reduced from 2595.64 cM to 650.56 cM. A complete outline of chromosomal coverage of the discovered linkage groups, following marker data curation, can be found in Table 3.5.

The linkage group that was of interest to the development of QTL for seed oil and seed protein content was discovered on chromosome A3. The raw linkage group data output by MSTMap following the initial stage of analysis revealed a total map distance of 347.7 cM for chromosome A3. Following data curation and removal of poor marker data, the total map distance of A3 was reduced to 97.1 cM. Using an estimate of genomic coverage in megabase pair (Mbp) from the visualization data and Darmor-ID from GenomeStudio, coverage for A3 ranged from 1.7 to 35 Mbp. The total length of A3 based on Darmor-*bzh* sequencing is 35 Mbp (Chalhoub et al., 2014).

3.4.4 QTL Mapping Analysis

Two different QTL were discovered on chromosome A3 controlling seed oil content. The physical position of these QTL differed between the 2015 and 2016 field analyses (Table 3.6). The first oil content QTL, *qOIL-A3a,* was located at a physical position of 81 cM bordered by SNP markers Bn-N3-p24842485 (80.3 cM) and Bn-N3-p25516345 (81.1 cM). The QTL had a LOD score of 2.02 and explained 7.2 % of observed phenotypic variation. The second oil QTL, *qOIL-A3b*, was found following 2016 field analysis. This QTL was at a physical position of 34.0 – 34.7 cM bordered by SNP markers Bn-N3-p11247217 (34.0 cM) and Bn-N3-p11539306 (34.7 cM). The second oil QTL had a LOD score of 1.3 and explained 4.6 % of observed phenotypic variation.

Two different QTL for seed protein content were also found on chromosome A3, differing in physical position between the 2015 and 2016 field years (Table 3.6). The first seed protein QTL, *qPROT-A3a*, was detected within the A3 chromosome of *B. napus* in 2015. The QTL is located at a physical position of 78.1 cM bordered by SNP markers Bn-N3-p23695516 (78.1 cM) and Bn-N3-p24614035 (78.8 cM). With a LOD score of 3.54, *qPROT-A3a* explained 12.3 % of the observed phenotypic variation. The second seed protein QTL, *qPROT-A3b*, was also detected on

the A3 chromosome of *B. napus*, found in 2016. The QTL is located at a physical position of 27.1 cM, bordered by SNP markers Bn-N3-p7687041 (27.1 cM) and Bn-N3-p7932668 (27.9 cM). With a LOD score of 3.33, *qPROT-A3b* explained 11.6 % of observed phenotypic variation. The linkage map shown in Figure 3.3 was developed using 28 representative markers from the total 107 markers on linkage group A3. Representative markers were selected every 5 to 10 cM from the original group, ensuring that SNP markers bordering the discovered QTL were included.

Figure 3.4 displays the relationship between *qPROT-A3a* and *qOIL-A3a* in 2015 and the relationship between *qPROT-A3b* and *qOIL-A3b* in 2016. In comparison of physical peak location, *qPROT-A3a* had the highest LOD value at 78.1 cM while *qOIL-A3a* was close in proximity at 81 cM. The same was observed in 2016 with *qPROT-A3b* peaking at 27.1 cM and *pOIL-A3b* at approximately 34 cM. In both years, the observed protein-related QTL represented a higher explanation of the phenotypic variation than those discovered relating to oil content.

Table 3.3. Analysis of variance of the effect of genotype and year on seed oil content for 124 *Brassica napus* genotypes in 2015 and 2016.

Source	DF	Type I Sum of Squares	Mean Square	F Value	Pr > F
Genotype	123	346.0048271	2.8130474	2.71	< 0.0001
Year	1	68.6284578	68.6284578	66.15	< 0.0001

Table 3.4. Analysis of variance of the effect of genotype and year on seed protein content for 124 *Brassica napus* genotypes in 2015 and 2016.

Source	DF	Type I Sum of Squares	Mean Square	F Value	Pr > F
Genotype	123	660.7458345	5.3719174	81.33	< 0.0001
Year	1	543.8313902	543.8313902	232.02	< 0.0001

Table 3.5. Chromosome location and map coverage (cM) of linkage groups developed in MapDisto 1.7.7 and MSTmap using a *Brassica napus* DH mapping population of 124 individuals in 2015 and 2016.

Linkage Group (LG)	Chromosome (*B. napus*)	Total Map Coverage (cM)
N1.1	A1	2.27
N1.2	A1	2.27
N3	A3	97.12
N4	A4	16.31
N5	A5	6.09
N6.1	A6	1.51
N6.2	A6	40.19
N7.1	A7	0.77
N7.2	A7	3.79
N7.3	A7	52.71
N9	A9	18.19
N10	A10	10.80
N11	C1	75.29
N12.1	C2	2.27
N12.2	C2	61.08
N13.1	C3	6.83
N13.2	C3	9.88
N14.1	C4	17.46

Table 3.5 continued. Chromosome location and map coverage (cM) of linkage groups developed in MapDisto 1.7.7 and MSTmap using a *Brassica napus* DH mapping population of 124 individuals in 2015 and 2016.

Linkage Group (LG)	Chromosome (*B. napus*)	Total Map Coverage (cM)
N14.2	C4	22.00
N15	C5	73.09
N16	C6	46.25
N17	C7	46.25
N18.1	C8	24.48
N18.2	C8	8.35
N19	C9	5.31

Table 3.6. Summarized list of QTL detected for seed oil and seed protein content determined from simple interval mapping (SIM) in MapDisto 1.7.7 using 124 DH individuals from a *Brassica napus* mapping population assessed in 2015 (*qOIL-A3a; qPROT-A3a;*) and 2016 (*qOIL-A3b; qPROT-A3b*).

QTL	LG	Position	LOD	R^2	Flanking Markers
qOIL-A3a	A3	81 cM	2.02	7.2	Bn-N3-p24842485 / Bn-N3-p25516345
qOIL-A3b	A3	34.0 – 34.7 cM	1.28	4.6	Bn-N3-p11247217 / Bn-N3-p11539306
qPROT-A3a	A3	78.1 cM	3.54	12.3	Bn-N3-p23695516 / Bn-N3-p24614035
qPROT-A3b	A3	27.1 cM	3.33	11.6	Bn-N3-p7687041 / Bn-N3-p7932668

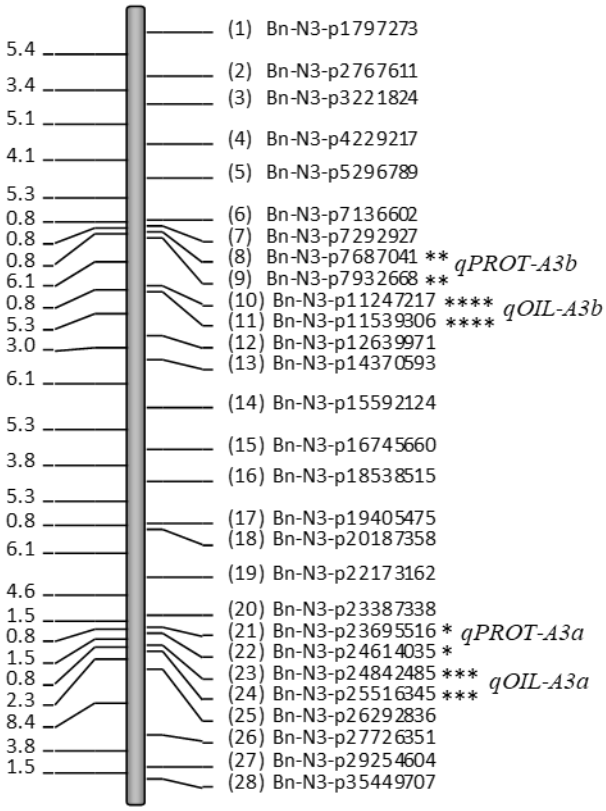

Figure 3.3. Refined linkage map depicting chromosome A3 in *Brassica napus* showing the location and bordering SNP markers for two discovered seed quality QTL in 2015 (*qPROT-A3a**; *qOIL-A3a****) and two discovered seed quality QTL in 2016 (*qPROT-A3b***; *qOIL-A3b*****) in a mapping population of 124 individuals.

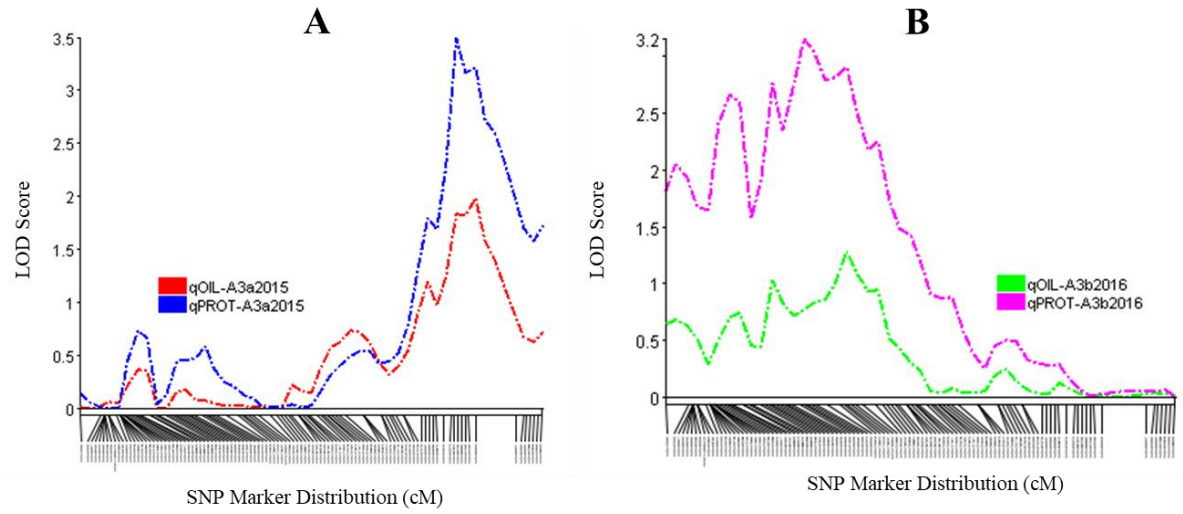

Figure 3.4 QTL controlling seed oil content (*qOIL-A3a*) and seed protein content (*qPROT-A3a*) on the A3 chromosome of *Brassica napus* detected within a doubled haploid mapping population analyzed in 2015 (A) and QTL controlling seed oil content (*qOIL-A3b*) and seed protein content (*qPROT-A3b*) on the A3 chromosome of *Brassica napus* detected within a doubled haploid mapping population analyzed in 2016 (B).

3.5 Discussion

The crude protein content of Canadian canola seed has a mean value of 20.3 % based on representative samples analyzed from 2010 to 2016 (Canadian Grain Commission, 2017a) while the minimum protein content value observed in the current mapping population was 28.2 % in 2015 and 29.9 % in 2016. All genotypes in both years exhibited seed protein values that exceed the national average for current commercially available canola cultivars. Even with these high levels of seed protein content, values for seed oil content were acceptable (34.0 to 49.2 % in 2015 and 36.0 to 43.3 % in 2016). The oil content of Canadian canola seed has a mean value of 44.2 % based on data collected from 2010 to 2016 (Canadian Grain Commission, 2017b). Based on these national averages, current canola quality standards reflect approximately 44 % oil and 20 % protein. The current population exhibited a mean value of 43 % oil and 31 % protein in 2015 and 40 % oil and 32 % in 2016, respectively. The range of observed oil and protein contents were 37.4 to 46.9 % and 28.2 to 34.1 % in 2015, and 37.2 to 43.0 % and 29.9 to 34.9 % in 2016, respectively. The population as a whole displayed seed quality values that are highly competitive with current market standards.

Many genotypes included in this study, including 13ZSDH182, 13ZSDH130, 13ZSDH348, and 13ZSDH386 achieved high protein content while not sacrificing oil content. Therefore, the population has proven to contain genotypes that display acceptable phenotypic performance for both seed oil and seed protein content. This reflects the effect diversity and mutagenesis can have on seed quality while at the same time highlights the potential negative effect that material highly bred for oil content can have on protein content (Malabat et al., 2003; Wanasundara, 2011; Rahman et al., 2013). This research also demonstrates that it is possible to achieve acceptable seed oil levels

with increased levels of seed protein. In this particular study, check cultivar Millenium03 (McVetty et al., 2000) displayed mean seed oil content values of 42% and mean seed protein values of 30%.

A recent study conducted by Chao et al. (2017) looked to dissect the association between seed oil and seed protein content by looking at the genetic control of each. The negative relationship between these two traits has been shown and referred to throughout the history of *B. napus* (Asare and Scarisbrick, 1995; Andersen et al., 1996; Brennan et al., 2000; Hao et al., 2004; Ahmad et al., 2007; Hu et al., 2013). Chao et al. (2017) discovered 67 QTL for seed oil and 38 QTL for seed protein content and performed bulk segregate analysis (BSA) sequencing within a mapping population grown in a total of three macro-environments and twelve micro-environments. They observed a mean negative correlation coefficient of -0.65 between seed oil and seed protein content across all environments in study. A similar relationship was evident in the ZSDH5225 / ZSDH5825 population with a correlation coefficient of -0.85 in 2015 and -0.73 in 2016. Negative correlations between seed and oil content have also been documented in *G. max* (soybean), with values ranging from -0.34 to -0.88 as reported by Eskendari et al. (2013). The negative relationship between the genetic control of seed oil and seed protein has also been documented through QTL mapping studies in *A. thaliana* (Chardon et al., 2014; Jasinski et al., 2016). With a greater understanding of this relationship, current breeding strategies need to be conscientious of the effect that selection for oil content has on protein content. Future breeding strategies should focus on the overall improvement of seed quality in *B. napus*, extending beyond oil content and include protein content and nutritive value.

In addition to the relationship between seed oil and seed protein content, genetic control of the two traits has been separately studied in *B. napus* (Würschum et al., 2012; Wang et al., 2013; Wu et al., 2013; Javed et al., 2015; Teh and Mollers, 2016; Wen at al., 2016). Documentation is much more readily available for seed oil content, given its greater importance to past breeding efforts compared with that of seed protein content. A wide variety of mapping studies have been performed for seed oil content in *B. napus*, implementing the use of various molecular markers, and leading to the discovery of various QTL on 17 of the 19 chromosomes in *B. napus* (Delourme et al., 2006; Rahman, 2013). Even in recent studies, seed oil content QTL have been reported on chromosome A10 (Javed et al., 2015); chromosomes A1, A2, A7, C3, C8 (Teh and Mollers, 2016); and chromosomes A4, A5, A6, A7, C1, C4, and C7 (Chao et al., 2017). Our results depict two oil content QTL located on A3, differing in physical location from the above studies. For closer comparison, Chao et al. (2017) report five QTL controlling seed oil content present in the 3.36 to 23.4 Mbp range of the A3 chromosome of *B. napus*. Our research shows that *qOIL-A3a* was present at 24.6 to 25.3 Mbp and *qOIL-A3b* was present at 11.1 to 11.3 Mbp. There are similarities between the location of the discovered QTL on the overall coverage of chromosome A3 by the SNP markers that make up the Illumina *Brassica* 60K array. The QTL discovered in this study for oil content fall within the range of the discovered QTL from the study by Chao et al. (2017) but none of the QTL are present at the exact same location. However, it is worth noting that oil content was not the main trait of interest in our study; it was more thoroughly studied as it relates to seed protein.

The primary objective of this research was to better understand the control of seed protein content in *B. napus*. Published data from genetic mapping studies for seed protein content are limited to date. As described, the relationship between seed oil and seed protein is well documented but the

actual control of seed protein content has not been as important to breeding efforts in the past. Würschum et al. (2012) discovered a QTL explaining 10.2 % of genetic variation for protein content on chromosome A7 while Xu et al. (2013) discovered a QTL explaining 37.9 % of phenotypic variation in protein content on chromosome A4. In addition, Chao et al. (2017) discovered regions containing QTL on chromosomes A3 and C5. This particular study is the most similar in nature to ours with use of the *Brassica* 60K Illumina Infinium™ SNP genotyping array for development of linkage maps for QTL analysis in both cases. Our reported QTL for protein content, *qPROT-A3a* and *qPROT-A3b* were both similarly located on chromosome A3. The study conducted by Chao et al. (2017) involved the production of one of the highest density SNP maps published to date looking at the control of seed oil and seed protein in *B. napus*. The similarities in QTL location between these two studies is worth noting. Chao et al. (2017) report four QTL controlling seed protein present in the 16.1 to 19.8 Mbp range of the A3 chromosome of *B. napus*. Our discovered QTL, *qPRO-A3a* and *qPRO-A3b* were discovered at 23.4 to 24.4 Mbp and 7.63 to 7.87 Mbp, respectively. The location of these QTL are similar, with *qPRO-A3a* falling within the range of discovered QTL by Chao et al. (2017). However, *qPRO-A3b* falls approximately 4 Mbp further into the A3 chromosome than their reported QTL. Looking at both traits simultaneously, Chao et al. (2017) also discovered QTL clusters controlling both seed oil and seed protein content on the A9, C3, and C5 chromosomes of *B. napus*. QTL controlling seed protein content discovered by Würschum et al. (2012) and Xu et al. (2013) were reported on the same chromosomes that have been previously associated with QTL for seed oil content by Teh and Mollers (2016) and Chao et al. (2017).

Overall, the goal of this research was to provide information on the genetic control of seed protein content with focus on its relationship with the control of seed oil content. The discovered

correlation between the physical location of QTL controlling both traits and the association between their phenotypic expression is useful in explaining the interrelatedness of the two traits. Seed oil and seed protein content are related both genotypically and phenotypically, which has been previously documented and further confirmed with this research. Exceptional seed oil content is no doubt essential for the economic success of the *B. napus* industry in Canada, but this research is unique in that it has identified a vast number of genotypes containing acceptable values for both seed oil and seed protein content. In conjunction with the observed phenotypic values, the resulting genotypic data provides both information and opportunity. The discovery of four QTL controlling seed oil and seed protein content on chromosome A3 provides insight to a potential area of focus for future breeding efforts to add value to the seed quality of *B. napus* as it extends beyond seed oil content. The introgression of improved seed protein quality traits provides not only economical but nutritional opportunities for the industry as well. The exploration of enhanced seed quality can extend beyond seed protein content. The next chapter will explore the variation in amino acid profile and the regions associated with amino acid control.

4.0 Discovery and Mapping of Quantitative Trait Loci Contributing to Amino Acid Contents in *Brassica napus* L. Meal

4.1 Abstract

The meal of *Brassica napus* L. is an important by-product in the Canadian canola industry with applications extending to ruminant, swine, poultry and aquaculture industries. The nutritional profile of Canadian canola meal is lower in sixteen of the eighteen amino acids (AA) including important amino acids such as lysine, isoleucine and leucine when compared to soybean meal produced in the United States. Breeding efforts to improve canola's protein content requires a better understanding of the underlying genetic control of these traits. Past breeding efforts for seed quality have focused on seed oil and fatty acid content, and as a result little breeding effort has been directed to the quantity and quality of the protein fraction. This research was able to provide new information identifying genomic regions controlling ten different amino acids in *B. napus*. Two major QTL were discovered controlling cysteine content on chromosomes C1 and C2, explaining 12.1 and 15.0 % of observed phenotypic variation, respectively. Chromosome C5 revealed major control over arginine, glutamine, isoleucine, leucine, and lysine content explaining 11.0 – 14.0 % of observed phenotypic variation and minor control over asparagine, phenylalanine, serine, and valine content explaining 7.0 – 10.0 % of observed phenotypic variation. Future breeding efforts to improve the nutritional quality of protein-related traits in *B. napus* will benefit from this knowledge, which should lead to genomic-based breeding approaches for AA content.

4.2 Introduction

The key to ensuring future growth of the Canola industry is to continue to diversify the crop for increased value from an end-use standpoint. From a breeding standpoint, efforts in the past have been highly intensive in nature and have heavily focused on an improvement in oil and fatty acid content, narrowing available *B. napus* gene pools (Malabat et al., 2003; Wanasundara, 2011; Rahman, 2013). Diversification for the future of a strong canola industry relies on the introduction of new alleles into current gene pools, as well as the development and characterization of new traits that provide additional end-use value to current canola products. One such area that could benefit from increased breeding effort is the improvement of protein-related traits in the seed and the resulting meal.

On a global scale, the meal of *B. napus* is the second most widely traded plant-based protein behind soybean meal (Canola Council of Canada, 2015). The nutritional composition and amino acid profile of canola meal differs from soybean meal in such a way that it is often used only as an additive in feed rations for animals (Wanasundara, 2011). Data from the Canola Council of Canada (2015) and the American Soybean Association (2015) outlining the nutritional profiles of both meal products highlight the fact that American soybean meal is naturally higher in the content of all amino acids except for methionine and cysteine. Lysine is considered one of the most limiting amino acids in both human (Young and Pellet, 1994; Reeds, 2000) and animal diets extending to dairy cattle (Broderick et al., 1974; Schwab et al., 1976), swine (Bell and Keith, 1989; Hickling, 1994; Bell et al., 1998), and poultry (Kaminska, 2003; Novak et al., 2004). Bulk canola meal contains higher contents of the sulfur-containing amino acids methionine and cysteine, which has proven useful for its use as feed for dairy cattle where several meta-analyses have shown its efficacy in increasing milk production (Huhtanen et al., 2011; Martineau et al., 2013; Marinteau

et al., 2014). While the dairy industry is currently benefiting from the available nutrition in canola meal, the overall profile requires improvement to be truly competitive with soybean meal as a plant-based protein source in numerous other industries. It is worth noting, that representative industry standards for canola meal values are based on bulked composite samples that are typically acquired from numerous crushing facilities (Canola Council of Canada, 2015). Currently, the nutritional value of canola meal is not being compared or analyzed at a genotype-specific level, which may be the key to unlocking and developing breeding tools for overall nutritional improvement.

Molecular breeding efforts to improve protein and protein-related traits in *B. napus* have been limited to date with a historical focus on oil content. Quantitative trait loci (QTL) mapping studies have been conducted in recent years, showing evidence of genetic control of seed protein content on chromosomes A7 (Würschum et al., 2012), A4 (Wu et al., 2013), A3 and C5 (Chao et al., 2017). As previously stated, this research has also found evidence of genetic control of seed protein content on chromosome A3 (Chapter 3). Studies to better understand regions of the *B. napus* genome associated with amino acid contents have been even more limited. Wen et al. (2016) discovered QTL's for arginine content on chromosomes A1, A4, A5, A7, and A8 explaining a cumulative 37.3 % of observed phenotypic variation and QTL's for glutamine content on chromosome A1, A4, and A7 explaining a cumulative 54.2 % of observed phenotypic variation. This work by Wen et al. utilized a group of RFLP, AFLP, SSR, and SNP markers and obtained amino acid content data for arginine, histidine, glutamine, glycine, proline, alanine, and aspartic acid using NIR.

In comparison, the current study utilized the Illumina® *Brassica* 60K Infinium SNP array to look at the content of amino acids in the dried, defatted meal using acid and oxidized hydrolysis

procedures. The objectives of this research were to provide information on the underlying genetic architecture controlling specific amino acid contents in *B. napus* meal. A QTL mapping analysis was performed on a diverse set of 124 doubled haploid (DH) genotypes developed for seed quality and derived from ethyl methanesulfonate (EMS) treatment. The DH population was genotyped using the Illumina® *Brassica* 60K Infinium SNP array. Genomic regions associated with alanine, arginine, asparagine, cysteine, glutamine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tyrosine, and valine were examined. The discovery of these regions will aid in future breeding efforts with a focus on improving protein-related traits in *B. napus* from a nutritional and value-added perspective. The research hypothesizes that linkage mapping and QTL analyses will result in the identification of chromosomal regions associated with phenotypic expression of amino acid contents in the meal of *Brassica napus*.

4.3 Materials and Methods

4.3.1 Plant Material and Field Evaluation Details

A diverse set of 343 *B. napus* genotypes was evaluated in field nursery experiments at the University of Manitoba in Winnipeg, Manitoba from 2014 – 2016. The genotypes were derived from two *B. napus* genotypes; ZSDH5225 and ZSDH5825. ZSDH5225 was developed from a cross of Millenium03 / 1123 (McVetty et al., 2000). F_1 seeds were subject to ethyl methanesulfonate (EMS) treatment (Sega, 1984). These plants were pollen donors for standard doubled haploid (DH) procedures (Chen et al., 1994) and ZSDH5225 was selected from this DH population with a focus on fatty acid content. ZSDH5825 is a DH line developed from a backcross between 04-2026 // 04-2026 / Millenium03 (McVetty et al., 2000). F_1 genotypes from this cross were used as pollen donors for the DH process and ZSDH5825 was selected from this DH

population. Seed quality characteristics of the two selected genotypes are quite similar with ZSDH5225 exhibiting 42.2% seed oil and 31.5% seed protein and ZSDH5825 exhibiting 39.8% seed oil and 33.0% seed protein. Further, ZSDH5225 was crossed with ZSDH5825 to produce the population that was used as the basis for the research described below.

The genotypes in this study were phenotypically evaluated in single row nursery trials at the University of Manitoba's Winnipeg research field testing station (Latitude 49.80 and Longitude -97.16). Soil at the Winnipeg site is classified as "Riverdale Silty Clay" as described in soil surveys conducted by Ehrlich et al. (1953). For all years of field study (2014, 2015, 2016) 0.50 g of seed was measured for each genotype and planted in 3 m rows with a row spacing of 0.40 m. The spring rapeseed cultivar MilleniUM 03 (McVetty et al., 2000) was inserted every twentieth entry to be used as a check and guard rows were seeded on all outer edges of the experimental rows planted to serve as a border. In 2014, the genotypes were planted on May 17 in a single 3 m row, single replication field experiment as described above. Prior to initiating this research, Edge® Granular Herbicide (Gowan Canada, Winnipeg, MB) was worked into the soil using a tandem disc in the fall of 2013. Fertilizer, based upon a soil test of the plot area, was applied using a broadcast spreader at a rate of 123.3 kg/ha of nitrogen, 44.8 kg/ha of phosphorus, 0 kg/ha of potassium, and 28.0 kg/ha sulfur. Decis® (Bayer CropScience, Leverkusen, Germany) was applied once on May 27 2014 and once on June 4 2014 at a rate of 60 ml per 37.9 litres of water per hectare for control of flea beetles. On June 10 2014, a tank mixture of Lontrel™ 360 (Dow AgroSciences, Indianapolis, IN), Poast® Ultra herbicide (BASF, Ludwigshafen, Germany), and Muster® (DuPont Canada, Mississauga, ON) was applied. Lontrel™ 360 was applied at a rate of 0.67 litres/hectare for thistle control, Poast® Ultra herbicide was applied at a rate of 0.67 litres/hectare with 1 litre/100 litres water of Merge surfactant for grassy weed control, and Muster® was applied at a rate of 29.7

grams/hectare for broadleaf weed control. Determination of physiological maturity is described in canola as the point at which 60 % of the seed on main stem has changed in colour from green to brown (Canola Council of Canada, 2012). A small sample of pods were selected from nursery rows each year to assess for maturity prior to harvest. The nursery trial was harvested between September 9 and 11, 2014. All rows were individually harvested in the field using a Wintersteiger small-plot combine (Nursery Master Classic, Wintersteiger, Salt Lake City, UT). Following harvest, seed samples were manually cleaned using a spiral seed cleaner (Can-Seed Equipment Ltd., Saskatoon, SK) and a 30 g sample was taken from the total yield and analyzed for seed quality.

From the original population of 343 *B. napus* genotypes, selections were conducted based upon quality characteristics such as seed protein and seed oil content. Selections were made with focus on achieving both high seed oil and seed protein contents while maintaining the variation appropriate for a mapping study. From the original population, 143 genotypes were selected from the ZSDH5225 / ZSDH5825 population and the selected genotypes were carried forward and evaluated in a completely randomized experimental design with three replicates per genotype in 2015 and 2016.

In 2015, the 143 selected genotypes were planted on May 23 in a 3 m row, completely randomized field experiment with three replicates per genotype. Prior to initiating this research, Edge® granular herbicide was worked into the soil using a tandem disc in the fall of 2014. Fertilizer, based upon a soil test, was applied a rate of 151.3 kg/ha of nitrogen, 44.8 kg/ha of phosphorus, 0 kg/ha of potassium, and 33.6 kg/ha sulfur. Decis® was applied as described above on May 30, 2015 and June 10, 2015 for control of flea beetles. On June 9 2015 a tank mixture of Lontrel™ 360, Poast®,

and Muster® was applied as described above. Lontrel™ 360 was applied at a rate of 0.84 litres/hectare, Poast® Ultra at a rate of 0.27 litres/hectare with 1 litre/100 litres water of Merge, and Muster® at a rate of 29.7 grams/hectare. The nursery trial was harvested between September 23 and 24, 2015 using the same protocol and equipment as described above.

In 2016, the genotypes were planted on May 7 in a 3 m row, randomized triple replicate field experiment. Prior to initiating this research, Edge® granular herbicide was applied as described above in fall of 2015. Fertilizer, based upon a soil test, was applied with a rate of 134.5 kg/ha of nitrogen, 44.8 kg/ha of phosphorus, 0 kg/ha of potassium, and 28.0 kg/ha sulfur. Decis® was applied as described above on June 14 2016 for control of flea beetles. On June 16 2016 a tank mixture of Lontrel™ 360, Poast®, and Muster® was applied as described above. Lontrel™ 360 was applied at a rate of 0.84 litres/hecate, Poast® Ultra at a rate of 0.67 litres/hectare with 1 litre/100 litres water of Merge, and Muster® at a rate of 29.7 grams/hectare for weed control. Sub-samples (10 – 20 g) were hand harvested and the nursery rows were bulk-combined using the same protocol and equipment as described above.

4.3.2 Amino Acid Analysis

4.3.2.1 Fat Extraction

To allow for the quantification of specific amino acid content within the protein fraction, meal samples for each genotype were de-fatted and dried prior to amino acid analysis as described by Klockleman et al. (1997). Seed samples were collected from harvested plants grown in the replicated nursery experiment located at the Point on the University of Manitoba campus from May to September of 2015. Seed samples of 1.2 g were weighed for each genotype, and the seed was heat-sealed inside of an ANKOM XT-4.0 filter bag. The seed samples were manually crushed

within the filter bags, using a steel rolling pin to obtain the meal sample. The crushed meal was dried at 102 C for 3 hours to remove moisture prior to fat extraction. A hexane fat extraction process was performed as described by Tzeng et al. (1988) and Wu and Muir (2008) using an ANKOM^{XT10} Extraction System (ANKOM Technology, New York, USA) for small batches of samples and a Kontes Macro Soxhlet apparatus (Fisher Scientific, Edmonton, AB) for larger batches of samples. Following fat extraction, samples were weighed and dried at room temperature in a fume hood.

4.3.2.2 Acid Hydrolysis Procedure

The acid hydrolysis procedure was performed to quantify "regular" amino acids (non-sulfur containing) including alanine, arginine, asparagine, glutamine, glycine, histidine, isoleucine, leucine, lysine, phenylalanine, proline, serine, threonine, tyrosine and valine following the official method of analysis as described by the Association of Official Analytical Chemists (AOAC) (1995). Quantification of tryptophan requires a separate hydrolysis method and was not included in the analysis based on resource constraints. For each genotype, 100 mg of dried and de-fatted meal were weighed into individual hydrolysis tubes. Phenolic hydrochloric acid and 2-Octanol were added to each tube before air removal using a vacuum pump. The samples were digested within the tubes at a temperature of 110 C for 24 hr. Following digestion, 25 % w/v sodium hydroxide was added for neutralization. The solutions were transferred into volumetric flasks and rinsed using sodium citrate (pH 2.2) before filtration into scintillation vials. The samples were frozen at this stage for subsequent analysis.

4.3.2.3 Oxidized Hydrolysis Procedure

The oxidized hydrolysis procedure was performed to quantify the sulfur-containing amino acids, including methionine and cysteine following the official method of analysis as described by the Association of Official Analytical Chemists (AOAC) (1995). For each genotype, 100 mg of dried and de-fatted meal were weighed into individual glass stopper tubes. Performic acid and 2-Octanol were added to each tube. The tubes were placed in a refrigerator at 4 C and left for 18 to 20 hours. Performic acid was prepared prior to each extraction by combining formic acid (88%) and hydrogen peroxide (35%) at a ratio of 9:1, respectively and allowing the mixture to stand for one hour at 4 C before use. Following the 18 – 20 hour standing period, 0.35 mg of sodium meta bisulfite was added to each sample and left to stand in a fume hood for 2 hours. Concentrated hydrochloric acid was added to each tube and the samples were digested at 110 C for 16 hours. Following digestion, 25% w/v sodium hydroxide was added for neutralization. The solutions were transferred into volumetric flasks and rinsed using sodium citrate (pH 2.2) before filtration into scintillation vials. Amino acid quantification was performed using ion-exchange chromatography with a Sykam Amino Acid Analyzer, Model S2100/S4300 following the procedures outlined by the manufacturer (Skyam GmbH, Eresing, Germany).

4.3.3 DNA Extraction and Quantification

The original greenhouse-developed seed for each of the genotypes were seeded into 4 x 3 cell packs containing Sunshine Metro Mix potting soil (Sungro® Horticulture, MA, USA), watered daily, and kept for germination inside a growth chamber (day temperature 22 C; night temperature 18 C; light cycle 16 h light, 8 h dark) for procurement of DNA for sequencing analysis. When the plants reached the two-leaf stage, they were transplanted into 15 x 15 cm pots containing Sunshine

Metro Mix potting soil, watered daily, and grown to maturity inside a greenhouse facility (day temperature 25 C; night temperature 22 C; relative humidity 40 – 50 %; light cycle 16 h light, 8 h dark). Tissue from a newly expanded leaf (the youngest tissue) was sampled for each of the genotypes when the plants had reached Growth Stage 3 (stem elongation or "bolting" phase) (Canola Council of Canada, 2017c) and the tissue was subsequently stored at -80 C until DNA extraction could occur.

DNA was extracted from the leaf tissue using a modified version of the cetyl trimethylammonium bromide (CTAB) extraction protocol (Murray and Thompson, 1980). Leaf tissue was ground in a 1.5 ml micro-centrifuge tube. A volume of 400 µl of 2X CTAB buffer was added to the ground leaf tissue before it was gently ground again. The samples were incubated for 60 min at 65 C. An equal volume of chloroform:isoamyl alcohol (24:1) was added to each sample and samples were centrifuged at 10 000 g for 5 min for phase separation. The aqueous phase was transferred to a new 1.5 ml micro-centrifuge tube. For precipitation, 5 M ammonium acetate (0.40 volumes) and isopropanol (2 volumes) were added and the samples were kept overnight at -20 C. DNA was recovered through a 10 min centrifugation period at 12 000 g. The resulting DNA pellet was washed in 70 % ethanol for removal of isopropanol residue before allowing it to air dry. The DNA was resuspended in 400 µl of Tris-EDTA (TE) buffer and 1.0 µl of 10 mg per ml heat-treated RNAse. The samples were then incubated for 60 min at 37 C. Following incubation, an equal volume of phenol was added and samples were centrifuged at 5 000 g for 5 min for separation of the aqueous and organic layers. The aqueous layer was transferred to a new 1.5 ml micro-centrifuge tube where an equal volume of phenol:chloroform (1:1) was added before another round of centrifugation. The process was repeated once more using an equal volume of chloroform. The DNA was precipitated for a second time through the addition of 3 M sodium acetate (0.10 volumes)

and ethanol (2.0 volumes). Samples were left to precipitate overnight at -20 C before being centrifuged at 10 000 g for 15 min for DNA recovery. The resulting pellet was rinsed with 70 % ethanol and left to air dry before being re-suspended in TE buffer for storage. DNA was quantified using a NanoDrop™ 2000 UV-Vis Spectrophotometer (Thermo Fisher Scientific, Massachusetts, USA). A volume of 2 µl of DNA representing each of the *B. napus* genotypes was loaded on to the pedestal using a micro-pipette. The quantity of DNA (ng/µl) was obtained by taking a mean of three samples for DNA quantification. The absorbance of the sample at 260 nm and 280 nm was obtained to determine a A260:A280 ratio. The ratio was used as a determinate of sample quality and purity. All results were recorded and used as a representation of DNA quantity and quality for SNP array analysis.

4.3.4 Illumina® 60K *Brassica* Array Analysis

Sequencing of the *B. napus* population using the Illumina® 60K *Brassica* array was conducted as described by Clarke et al. (2016). Visualization of SNP marker data was completed using the GenomeStudio V2011.1 software available from Illumina Inc. (California, USA). The working dataset from the array contained 52,157 SNPs (Clarke et al., 2016). Criteria for selection of polymorphic SNP data within Genome Studio was conducted by filtering of all SNPs with a GenTrain score of zero. This resulted in the removal of 4,854 SNPs, leaving 47,303 SNPs. Based on visual evaluation of the dataset, SNPs with a call rate of less than 28 % were filtered out. This resulted in the removal of 5,339 SNPs leaving a total of 41,964 in the working dataset. Based on visual evaluation of the dataset, SNPs with a minor allele frequency of less than 10 % were filtered. This resulted in the removal of 36,741 SNPs to leave a working dataset of 5,223 SNPs. Given the doubled haploid nature of the populations in this study, any SNP with an AB frequency of more

than 10 % were filtered. This removed 396 SNPs, leaving 4,827. The filtering of any SNPs with "no calls" greater than 60 removed 18 more SNPs, leaving 4,809 in the working dataset. Finally, three SNPs were manually removed (Bn11-p6742595, Bn11-p454442, Bn11-p449385) on the basis of too many missing data points. This resulted in a final working dataset of 4,806 SNP markers. This working dataset was used for linkage mapping and QTL analysis.

4.3.5. Linkage Mapping Analysis

Development of linkage groups was completed using MSTMap (Wu et al., 2008) and MapDisto 1.7.7 (Lorieux, 2012). The curated marker dataset of 4,806 SNPs was run in MTSMap using the Kosambi mapping function (Vinod, 2011). The population was designated as a DH population with mapping size threshold of 2 cM, a mapping distance threshold of 15 cM, a p-value threshold of 1×10^{-9} and a missing data threshold of 50%.

The linkage groups developed from MSTMap were separated by chromosome and manually curated. The curation process involved a visual and manual manipulation of marker data based on a color-blocking scheme. The SNP markers were manually sorted and organized based on both the Darmor-ID, indicating the SNP's physical position, and the visual fit of the marker across all genotypes in question. In total, 11 genotypes were culled across all linkage groups as a result of missing data. Following manual organization, the markers were placed into marker "bins" through the selection of one marker to represent a series of co-localized markers. The representative markers were then used to create a smaller overall mapping file run in MapDisto 1.7.7 (Lorieux, 2012).

The curated mapping files were run per chromosome using MapDisto 1.7.7 using the Kosambi mapping function, with SARF (sum of adjacent recombination fractions) ordering criteria, seriation ordering method, and a classical recombination frequency (RF) estimate. A LOD score minimum of 3.0 and r value maximum of 0.30 were used to separate loci data into linkage groups. Loci with more than 10 % missing data were also sorted using the MapDisto 1.7.7 program. The marker order and distance developed from these maps was used to create the data files run in QGene 4.3.10 (Joehanes and Nelson, 2008) for QTL analysis.

4.3.6. Quantitative Trait Loci Mapping Analysis

Quantitative trait loci (QTL) mapping analysis was conducted using the Simple Interval Mapping (SIM) procedure in QGene 4.3.10 (Joehanes and Nelson, 2008). Linkage mapping data from MSTMap and MapDisto were prepared and run in QGene using the "QTL mapping" function. QGene was used to obtain LOD scores and R^2 values for seed protein and seed oil content. The R^2 value represents the percentage of observed phenotypic variation that is explained by the marker-trait association represented by the QTL peak. The larger mapping population (ZSDH5225 / ZSDH5825) of 143 individuals was reduced initially during the development of the mapping dataset through manual removal of 11 genotypes with missing data. The resulting population of 132 individuals was furthered reduced through the removal of 8 additional genotypes that had issues during amino acid determination. Therefore, QTL analysis was performed on a working mapping dataset of 124 *B. napus* genotypes.

4.3.7 Statistical Analysis

Correlation analysis was performed using the "Trait Correlation" function in QGene 4.3.10 to look at the interrelatedness of phenotypic values for all of the amino acids tabulated in this study determined from one technical replicate of hydrolysis analysis.

4.4 Results

4.4.1 Phenotypic Variation

The genotypes within the *B. napus* population displayed phenotypic variation for meal amino acid content. When compared to industry standards for Canadian canola meal, the mean values of the mapping population surpassed national mean values for all seventeen amino acids that were included in this study (Table 4.1). When compared to industry means for USA soybean meal in 2015, the largest competitor of canola meal as a feed source, mean values of the mapping population surpassed the mean for soybean meal in the case of alanine, cysteine, glycine, histidine, methionine, proline, and threonine. Aside from the mean values, genotypes within this *B. napus* population expressed maximum amino acid values that exceeded mean soybean values in the case of arginine, glutamine, leucine, lysine, serine, and valine (Table 4.1). The population exhibited phenotypic values for amino acid content of the meal that are competitive when compared to current standards for both canola and soybean meal.

The phenotypic distribution for amino acid content values across the 124 mapping individuals for all amino acids analyzed can be found in Figure 4.1. In addition, p-values corresponding to the Kolmogorov-Smirnov normality test can be found in Table 4.2. The test determined that all amino acids exhibited a normal phenotypic distribution except for alanine, tyrosine, and histidine at p =

0.05. Distribution for alanine values appear to skew left while distribution for tyrosine and histidine values appear to skew right (Figure 4.1).

Correlation coefficients displaying the interrelatedness of the observed phenotypic values for all of the amino acids in this study can be found in Table 4.3. Of all the coefficients determined, 30.1 % of the relationships reported exceeded correlation values of 0.75 indicating strong correlation among many of the amino acid contents in the meal of *B. napus*. While this data reflects the many strong and positive associations that these amino acids have amongst themselves, methionine displays little correlation with any of the other amino acids except for cysteine with a correlation coefficient of 0.43. Methionine and cysteine are more similar with one another than they are the other amino acids because their side chains contain sulfur. The branched-chain amino acids (BCAAs) including leucine, isoleucine, and valine all displayed a strong correlation with one another. The correlation coefficients were 0.81, 0.92, and 0.81 for the relationships between leucine and isoleucine, leucine and valine, and isoleucine and valine, respectively.

4.4.2. Linkage Mapping Analysis

Initial analysis of the ZSDH5225 / ZSDH5825 population in MSTMap gave rise to a total of 31 linkage groups covering 2595.64 cM. Prior to analysis in MSTMap, the marker dataset was curated using the filtration parameters in GenomeStudio. The resulting 4,806 markers were distributed across these 31 linkage groups. Negligible coverage was discovered within many of these linkage groups and following further curation, several small linkage groups were removed. Marker coverage was poor on chromosomes A2 and A8, forcing them to be absent from this study. In many cases, two or more smaller linkage groups were discovered on the same chromosome due to gaps in chromosome coverage between the groups themselves, preventing the development of one

large linkage group. Following data curation, prior to analysis in MapDisto, the number of linkage groups was reduced from 31 linkage groups to 25 linkage groups. The total map distance was also reduced from 2595.64 cM to 650.56 cM. A complete outline of chromosomal coverage of the discovered linkage groups, following marker data curation, can be found in Table 4.4.

Table 4.1. Minimum, maximum and mean amino acid contents among 124 *B. napus* genotypes in a DH mapping population as well as national mean amino acid values for Canadian canola and USA soybean meal determined from one replicate of hydrolysis analysis.

Amino Acid	Min	Max	Mean	Canola Mean (2015)[A]	Soybean Mean (2015)[B]
Alanine	1.16	3.33	2.18	1.57	2.05
Arginine	1.83	3.80	2.99	2.08	3.48
Asparagine	2.22	4.87	3.92	2.61	5.52
Cysteine	0.62	1.08	0.93	0.86	0.79
Glutamine	5.16	10.74	8.54	6.53	8.62
Glycine	1.35	2.73	2.19	1.77	1.97
Histidine	0.01	1.71	1.30	1.12	1.21
Isoleucine	1.00	2.04	1.65	1.12	2.17
Leucine	1.89	4.11	3.19	2.54	3.60
Lysine	1.50	3.15	2.49	2.00	2.89
Methionine	0.48	1.06	0.79	0.74	0.63

Table 4.1 continued. Minimum, maximum and mean amino acid contents among 124 *B. napus* genotypes in a DH mapping population as well as national mean amino acid values for Canadian canola and USA soybean meal determined from one replicate of hydrolysis analysis.

Amino Acid	Min	Max	Mean	Canola Mean (2015)[A]	Soybean Mean (2015)[B]
Phenylalanine	1.11	2.33	1.87	1.38	2.37
Proline	1.67	4.63	2.92	2.15	2.37
Serine	1.22	2.78	2.11	1.55	2.38
Threonine	1.14	2.47	1.89	1.58	1.84
Tyrosine	0.01	1.68	1.24	1.16	1.68
Valine	1.29	2.68	2.05	1.97	2.30

[A] Canola Council of Canada (2015); [B] American Soybean Association (2015)

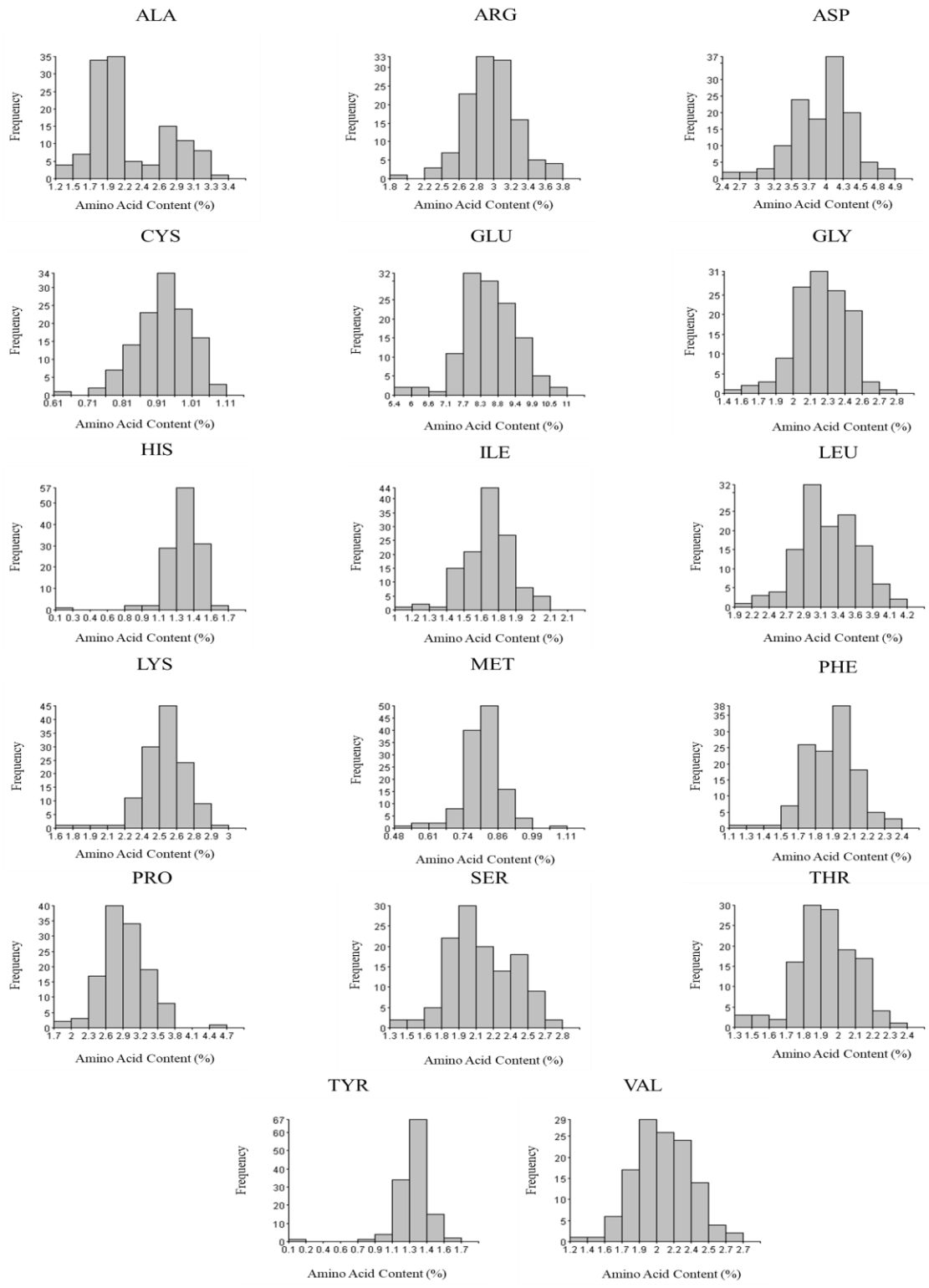

Figure 4.1. Phenotypic distribution of meal content values for alanine (ALA), arginine (ARG), asparagine (ASP), cysteine (CYS), glutamine (GLU), glycine (GLY), histidine (HIS), isoleucine

(ILE), leucine (LEU), lysine (LYS), methionine (MET), phenylalanine (PHE), proline (PRO), serine (SER), threonine (THR), tyrosine (TYR), and valine (VAL) among 124 *Brassica napus* genotypes in a DH mapping population in 2015.

Table 4.2. Resulting p-values for the Kolmogorov-Smirnov test for normal distribution following phenotypic analysis of amino acid content in the meal of 124 *Brassica napus* genotypes using trait analysis in QGene 4.3.1.

Amino Acid	p-value
Asparagine (ASP)	0.3148
Threonine (THR)	0.9707
Serine (SER)	0.4507
Glutamine (GLU)	0.8830
Proline (PRO)	0.6078
Glycine (GLY)	0.9945
Alanine (ALA)	**0.0002**
Cysteine (CYS)	0.5769
Valine (VAL)	0.9652
Methionine (MET)	0.0678
Isoleucine (ILE)	0.2778
Leucine (LEU)	0.7307
Tyrosine (TYR)	**0.0075**
Phenylalanine (PHE)	0.7961
Histidine (HIS)	**0.0245**
Lysine (LYS)	0.3578
Arginine (ARG)	0.9619

Table 4.3. Correlation coefficients displaying the relationships between 17 amino acids* in the meal of 124 *Brassica napus* mapping individuals determined using trait correlation analysis in QGene 4.3.1.

	THR	SER	GLU	PRO	GLY	ALA	CYS	VAL	MET	ILE	LEU	TYR	PHE	HIS	LYS	ARG
ASP	0.40	**0.91**	**0.88**	0.40	0.43	**0.81**	-0.02	**0.87**	0.04	**0.77**	**0.89**	0.43	**0.85**	**0.80**	0.72	**0.86**
THR		0.15	0.28	-0.13	**0.94**	-0.11	0.00	0.19	0.01	0.44	0.17	0.27	0.26	0.30	0.28	0.29
SER			**0.94**	0.54	0.15	**0.94**	0.09	**0.91**	-0.01	0.68	**0.94**	0.40	**0.85**	**0.78**	**0.79**	**0.86**
GLU				0.62	0.30	**0.86**	0.22	**0.88**	0.10	**0.75**	**0.92**	0.46	**0.89**	**0.75**	**0.89**	**0.89**
PRO					-0.10	0.54	0.42	0.42	0.14	0.30	0.50	0.25	0.59	0.29	0.67	0.51
GLY						-0.09	-0.01	0.24	0.04	0.52	0.19	0.35	0.32	0.31	0.32	0.37
ALA							0.09	0.87	0.05	0.60	**0.91**	0.32	**0.80**	0.74	0.71	**0.77**
CYS								0.04	**0.43**	0.04	0.08	0.13	0.16	0.10	0.37	0.11
VAL									0.12	**0.80**	**0.92**	0.45	**0.84**	**0.74**	**0.77**	**0.87**
MET										0.14	0.07	0.08	0.10	0.00	0.10	0.11
ILE											**0.81**	0.47	**0.81**	0.61	0.71	0.73
LEU												0.49	**0.90**	**0.77**	**0.78**	**0.85**
TYR													0.55	0.39	0.43	0.49
PHE														0.71	**0.83**	**0.84**
HIS															0.67	0.70
LYS																**0.81**

*ASP: Asparagine, THR: Threonine, SER: Serine, GLU: Glutamine, PRO: Proline, GLY: Glycine, ALA: Alanine, CYS: Cysteine, VAL: Valine, MET: Methionine, ILE: Isoleucine, LEU: Leucine, TYR: Tyrosine, PHE: Phenylalanine, HIS: Histidine, LYS: Lysine, ARG: Arginine.

Table 4.4. Chromosome location and map coverage (cM) of linkage groups developed in MapDisto 1.7.7 and MSTmap using a *Brassica napus* DH mapping population of 124 individuals in 2015 and 2016.

Linkage Group (LG)	Chromosome (*B. napus*)	Total Map Coverage (cM)
N1.1	A1	2.27
N1.2	A1	2.27
N3	A3	97.12
N4	A4	16.31
N5	A5	6.09
N6.1	A6	1.51
N6.2	A6	40.19
N7.1	A7	0.77
N7.2	A7	3.79
N7.3	A7	52.71
N9	A9	18.19
N10	A10	10.80
N11	C1	75.29
N12.1	C2	2.27
N12.2	C2	61.08
N13.1	C3	6.83
N13.2	C3	9.88
N14.1	C4	17.46

Table 4.4 continued. Chromosome location and map coverage (cM) of linkage groups developed in MapDisto 1.7.7 and MSTmap using a *Brassica napus* DH mapping population of 124 individuals in 2015 and 2016.

Linkage Group (LG)	Chromosome (*B. napus*)	Total Map Coverage (cM)
N14.2	C4	22.00
N15	C5	73.09
N16	C6	46.25
N17	C7	46.25
N18.1	C8	24.48
N18.2	C8	8.35
N19	C9	5.31

4.4.3 QTL Mapping Analysis

The C5 chromosome of *B. napus* was found to contain QTLs controlling the meal content of numerous amino acids. The QTL on C5 considered to be "major" with a LOD score greater than 3.0 included QTL controlling the contents of arginine, glutamine, isoleucine, leucine and lysine in *B. napus* meal (Table 4.5). The QTL on C5 considered to be "minor" with a LOD score below 3.0 included QTL controlling asparagine, phenylalanine, serine, and valine (Table 4.6).

There were five major QTL detected on chromosome C5 (Table 4.5). *qARG-C5a*, controlling arginine content, was located at a physical position of 31.9 cM. The arginine QTL explained 13.0 % of the observed phenotypic variation with a LOD score of 3.75 and was bordered by SNP markers Bn-NI5-p4544613 (31.9 cM) and Bn-N15-p7646696 (42.6 cM). *qGLU-C5b*, controlling glutamine content, was located at a physical position of 32.7 – 33.0 cM bordered by SNP markers Bn-NI5-p4544613 (31.9 cM) and Bn-N15-p7646696 (42.6 cM). The glutamine QTL had a LOD score of 4.06 and explained 14.0 % of observed phenotypic variation. *qILE-C5c*, controlling isoleucine content, was located at a physical position of 33.3 – 33.8 cM and bordered by SNP markers Bn-NI5-p4544613 (31.9 cM) and Bn-N15-p7646696 (42.6 cM). The isoleucine QTL explained 11.0 % of the variation with a LOD score of 3.25. *qLEU-C5d*, controlling leucine content, explained 11.0 % of phenotypic variation with a LOD score of 3.23. The leucine QTL was located at a physical position of 33.1 – 33.4 cM, bordered by SNP markers Bn-NI5-p4544613 (31.9 cM) and Bn-N15-p7646696 (42.6 cM). *qLYS-C5e*, controlling lysine content, was located at a physical position of 34.7 cM with a LOD score of 3.09. The lysine QTL controlled 11.0 % of variation and was bordered by SNP markers Bn-NI5-p4544613 (31.9 cM) and Bn-N15-p7646696 (42.6 cM). The location of all major QTL are depicted on the linkage map of chromosome C5 in Figure 4.2.

Table 4.5. Summarized list of major quantitative trait loci (QTL) for meal amino acid contents using simple interval mapping in a *Brassica napus* doubled haploid population in 2015.

QTL	LG*	Position	LOD*	R^2*	Flanking Markers
qCYS-C1	C1	61.6 – 61.7 cM	3.47	12.1	Bn-N11-p5327030 / Bn-N11-p5126350
qCYS-C2	C2	29.4 cM	4.38	15.0	Bn-N12-p26149924 / Bn-N12-p31488960
qARG-C5a	C5	31.9 cM	3.75	13.0	Bn-NI5-p4544613 / Bn-N15-p7646696
qGLU-C5b	C5	32.7 – 33.0 cM	4.06	14.0	Bn-NI5-p4544613 / Bn-N15-p7646696
qILE-C5c	C5	33.3 – 33.8 cM	3.25	11.0	Bn-NI5-p4544613 / Bn-N15-p7646696
qLEU-C5d	C5	33.1 – 33.4 cM	3.23	11.0	Bn-NI5-p4544613 / Bn-N15-p7646696
qLYS-C5e	C5	34.7 cM	3.09	11.0	Bn-NI5-p4544613 / Bn-N15-p7646696

*LG: linkage group; LOD: logarithm of odds; R^2: coefficient of multiple regression.

Table 4.6. Summarized list of minor QTLs detected for meal amino acid contents using simple interval mapping in a *Brassica napus* doubled haploid population in 2015.

QTL	LG*	Position	LOD*	R^2*	Flanking Markers
qASP-C5f	C5	33.7 – 34.3 cM	2.04	7.0	Bn-NI5-p4544613 / Bn-N15-p7646696
qPHE-C5g	C5	33.8 – 34.2 cM	2.43	9.0	Bn-NI5-p4544613 / Bn-N15-p7646696
qSER-C5h	C5	32.9 – 33.4 cM	2.78	10.0	Bn-NI5-p4544613 / Bn-N15-p7646696
qVAL-C5i	C5	32.5 – 33.0 cM	2.88	10.0	Bn-NI5-p4544613 / Bn-N15-p7646696

*LG: linkage group; LOD: logarithm of odds; R^2: coefficient of multiple regression.

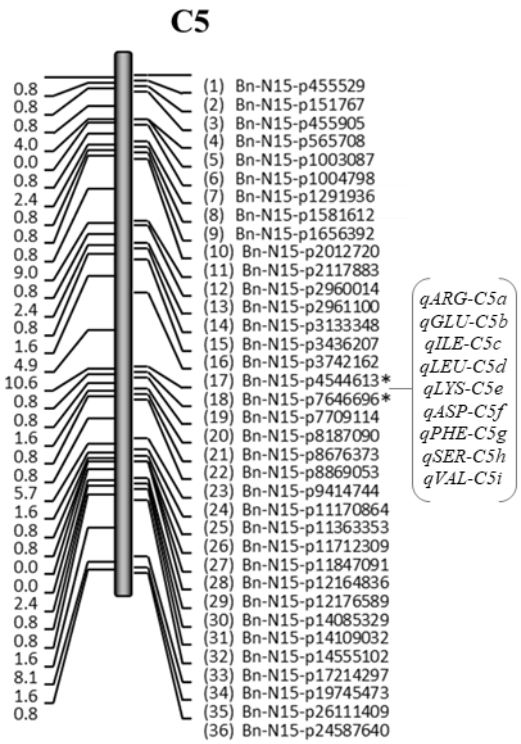

Figure 4.2. Refined linkage map depicting chromosome C5 in *Brassica napus* showing the location and bordering SNP markers for five major AA QTL (*qARG-C5a*; *qGLU-C5b*; *qILE-C5c*; *qLEU-C5d*; *qLYS-C5e*) and four minor QTL (*qASP-C5f*; *qPHE-C5g*; *qSER-C5h*; *qVAL-C5i*).

Four minor QTL were also detected on chromosome C5 (Table 4.6). *qASP-C5f*, controlling asparagine content, was detected at 33.7 – 34.3 cM bordered by SNP markers Bn-NI5-p4544613 (31.9 cM) and Bn-N15-p7646696 (42.6 cM). The asparagine QTL controlled 7.0 % of variation with a LOD score of 2.04. *qPHE-C5g*, controlling phenylalanine content, was detected at 33.8 – 34.2 cM bordered by SNP markers Bn-NI5-p4544613 (31.9 cM) and Bn-N15-p7646696 (42.6 cM). The phenylalanine QTL had a LOD score of 2.43 and controlled 9.0 % of observed variation. The serine QTL, *qSER-C5h*, was located at 32.9 – 33.4 cM bordered by SNP markers Bn-NI5-p4544613 (31.9 cM) and Bn-N15-p7646696 (42.6 cM). *qSER-C5h* controlled 10.0 % of observed phenotypic variation with a LOD score of 2.78. The valine QTL, *qVAL-C5i*, controlled 10.0 % of phenotypic variation with a LOD score of 2.88. The valine QTL was located at 32.5 – 33.0 cM, bordered by SNP markers Bn-NI5-p4544613 (31.9 cM) and Bn-N15-p7646696 (42.6 cM). The location of all minor QTL are depicted on the linkage map of chromosome C5 in Figure 4.2.

Two QTL controlling cysteine content in the meal of *B. napus* were detected on the C1 and C2 chromosomes of *B. napus* (Table 4.5). The first, *qCYS-C1*, was located at a physical position of 61.6 to 61.7 cM on the C1 chromosome, bordered by SNP markers Bn-N11-p5327030 (61.5 cM) and Bn-N11-p5126350 (63.1 cM). *qCYS-C1* had a LOD score of 3.47 and explained 12.1 % of the observed variation in cysteine content (Figure 4.3). The second QTL, *qCYS-C2*, was located at a physical position of 29.4 cM on the C2 chromosome, bordered by SNP markers Bn-N12-p26149924 (28.8 cM) and Bn-N12-p31488960 (29.5 cM). With a LOD score of 4.38, *qCYS-C2* explained 15.0 % of the observed variation in cysteine content (Figure 4.3). The *qCYS-C2* QTL had the highest LOD score and explained the most phenotypic variation out of all QTL considered to be “significant”, ie. having a LOD score exceeding 3.0.

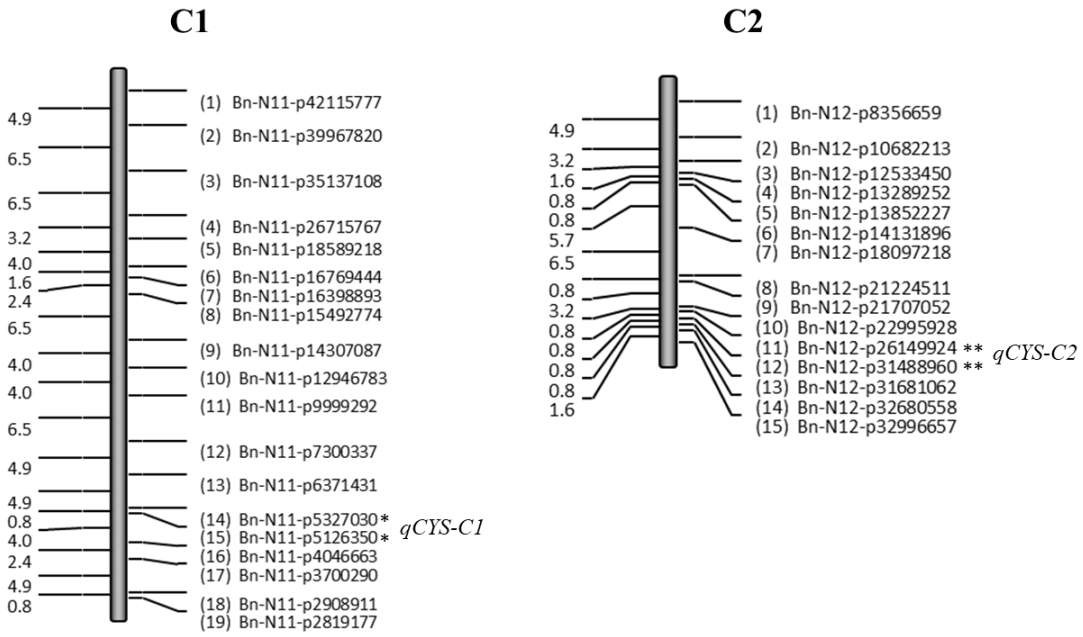

Figure 4.3. Refined linkage maps depicting chromosomes C1 and C2 in *Brassica napus* showing the location and bordering SNP markers for two QTL discovered to control cysteine content (*qCYS-C1** and *qCYS-C2***).

4.5 Discussion

This research determined that the genetic control of amino acid content in the meal of *B. napus* varies depending on the amino acid in question. Of the seventeen amino acids included in this study, nine of them were found to have genetic association with the C5 chromosome in both major and minor ways. Nine QTL were discovered between 31.9 and 34.7 cM on chromosome C5, including 5 major QTL for arginine, glutamine, isoleucine, leucine, and lysine with LOD score ranges from 3.06 to 4.06 and phenotypic explanation ranging from 11.0 to 14.0 %. Of the major QTL discovered, significance may lie in further understanding the genetic control of isoleucine and leucine, which classified as branched-chain amino acids (BCAAs). Inclusion of BCAAs in human nutrition has become a focus due to their key role in muscle development and function as well as their contribution to protein synthesis during exercise recovery periods (Blomstrand et al., 2006; Chuang, 2013; Wu, 2009). Other information on the genetic control of these amino acids is not currently available and therefore this novel discovery of QTLs *qILE-C5a* (isoleucine) and *qLEU-C5d* (leucine) may prove useful in improving the nutritional value of canola meal for an end-use in human nutrition.

This region of C5 was also found to control lysine content (*qLYS-C5e*), arguably one of the most important amino acids. Lysine is the most limiting amino acid in both human and animal diets (Young and Pellet, 1994; Reeds, 2000; Broderick et al., 1974; Schwab et al., 1976; Bell and Keith, 1989; Hickling, 1994; Bell et al., 1998; Kaminska, 2003; Novak et al., 2004). Because lysine is an essential amino acid (EAA) it cannot be synthesized sufficiently by the body itself and therefore must be provided through dietary intake. Better understanding the genetic control of lysine in plant-based protein sources should be a key focus in understanding the genetic control of amino acids from a basic nutritional standpoint, within and beyond *B. napus*. Warrington et al. (2015)

discovered QTLs controlling lysine content on chromosomes 08 and 20 in *G. max*. These understandings in both soybean and canola can serve to further enhance the nutritional value of the meal. Canola falls behind soybean in terms of the amount of lysine content available in the meal and using a genome-based breeding approach to increase lysine through further understanding of its genetic control could provide canola the edge it needs to be competitive.

Other major QTL on C5 include two controlling arginine (*qARG-C5a*) and glutamine (*qGLU-C5b*), both classified as conditionally-essential amino acids (CEAAs). Though only considered limiting amino acids under certain conditions they play an important role in both gene and immune system regulation (Reeds, 2000; Wu, 2009). While the research available on the genetic control of amino acids in *B. napus* is currently limited, a study conducted by Wen et al. (2016) found QTL controlling variation in both arginine and glutamine in DH mapping population. In their study, genetic regions associated with arginine content were discovered on chromosomes A1, A4, A5, A7, and A8 and genetic regions associated with glutamine content were found on A1, A4, and A7 (Wen et al., 2016). Control of these amino acids within the A genome of *B. napus* differs from the results of our study where control was discovered within the C genome. Our research looked at the genetic control of amino acids in the dried and defatted meal of *B. napus* using the oxidized and acid hydrolysis wet chemistry method in comparison to amino acid content data collection through NIR as performed by Wen et al. (2016). The results of both studies can be utilized in future efforts to understand amino acid content control through linkage and QTL mapping.

The major QTL discovered provided evidence of co-localization of QTL controlling amino acid content. This co-localization was further documented with a second set of minor QTLs discovered in the same region of C5. Minor QTL controlling asparagine (*qASP-C5f*), phenylalanine (*qPHE-C5g*), serine (*qSER-C5h*), and valine (*qVAL-C5i*) were discovered with LOD scores ranging from

2.04 to 2.78 and phenotypic explanation ranging from 7.0 to 10.0 %. Valine is considered a branched-chain amino acid along with isoleucine and leucine and serves the same function in human nutrition (Wu, 2009). Phenylalanine is an essential amino acid involved in neurological development while serine is considered a non-essential amino acid involved in the synthesis of other amino acids such as cysteine and glycine (Wu, 2009). Other literature for genetic control of these amino acids is not currently available for comparison in *B. napus*, and so these discoveries are also considered novel. Asparagine is involved in cell metabolism and physiology (Wu, 2009) and is a precursor for aspartic acid, for which QTL were also discovered by Wen at al. (2015). Similar to QTL discovered by Wen et al. (2015) for arginine and glutamine, those discovered for aspartic acid differed in physical location from our research. QTL controlling aspartic acid were found on A1, A4, A6, A9, C2, and C4 (Wen et al., 2015). Again, these separate discoveries may be useful in conjunction with one another to provide an overall understanding of genetic control of amino acid content in future studies.

Separate from the QTL discovered on C5, two major QTL controlling cysteine content were discovered on chromosomes C1 and C2. The two QTLs, *qCYS-C1* and *qCYS-C2*, were amongst the highest of all QTL discovered in this study explaining 12.1 and 15.0 % of the observed phenotypic variation, respectively. Cysteine is classified as a functional amino acid (FAA) and has important signaling and anti-oxidative properties (Wu, 2009). Cysteine and methionine are the only amino acids that are sulfur-containing. Data reflecting the average amino acid contents of both canola and soybean meal show that *B. napus* already naturally contains higher amounts of methionine and cysteine, amino acids which have proven to increase milk production in lactating dairy cattle (Huhtanen et al., 2011; Martineau et al., 2013; Martineau et al., 2014). Therefore, canola meal is already considered to have an advantage as a feed for dairy cattle and this research

may serve to advance that advantage. Though a QTL for methionine was not discovered in this study, the development of individuals for a mapping study with specific focus on amino acid contents in the future may provide opportunity to take a closer look at the genotypic relatedness of these two sulfur-containing amino acids. It is also worth noting that literature for the genetic control of cysteine content in *B. napus* is not currently available therefore these QTL are considered novel.

The genetic control of amino acid content is complex, varying just as their essentiality and functions do. This research provides information on the genetic control of arginine, glutamine, isoleucine, leucine, lysine, asparagine, phenylalanine, serine, valine, and cysteine. QTLs for these amino acids were located on C5 with the exception of *qCYS-C1* and *qCYS-C2* controlling cysteine. This study has displayed the interrelatedness of the genetic control of these amino acids through the co-localization of their control on C5, a chromosome that should be studied in more detail in future studies looking at amino acid content in the meal of *B. napus*. While the QTL location controlling cysteine differs from the other amino acids, this may be consistent with its structural differences from the rest of the amino acids as well. In future studies, C1 and C2 should also be further explored to better understand the control of not just cysteine content but methionine content as well. Many of the genetic regions discovered in this study are considered novel and further research is required. A mapping population designed with the intention of looking at the genetic control of seed and meal quality in terms of protein and amino acids will provide the framework required for the type of mapping study needed to confirm our results. Imposing selection pressure for the overall nutritional quality of canola seed and meal are made easier through a genetic understanding of these traits and eventually the development of molecular markers that may be used in a marker-assisted selection breeding approach. The implementation of these genome-based

breeding strategies can give rise to new genetic material in *B. napus*, allowing for it to be nutritionally competitive with that of soybean.

5.0 GENERAL DISCUSSION AND CONCLUSIONS

The QTL mapping conducted in this research was able to provide information both novel and comparative to previously discovered genomic regions associated with seed oil content, seed protein content, and amino acid content in the meal in *Brassica napus*. While the primary goal was to analyze seed protein and meal amino acid contents in more detail than currently available in the literature, the research also gave rise to information about seed oil content and its relatedness to seed protein content on a phenotypic and genotypic level.

The economic success of canola would not be possible without its seed quality characteristics. This study was successful in displaying numerous *B. napus* genotypes containing exceptional overall quality profiles. This research has proven it is possible to identify and perhaps focus on an increase in seed protein content, extended even to specific amino acid contents within the meal, without sacrificing seed oil content. However, this study also displayed the negative correlation that these seed quality traits have with one another. Seed oil and seed protein content are closely related, and this cannot be overlooked in breeding efforts to improve one, the other, or both traits simultaneously. In addition, individual amino acid contents within the meal of *B. napus* have varying degrees of interaction when it comes to the final constitution of the AA profile.

The amino acid profile of *B. napus* is an area of quality that currently is not being utilized on a selection level in canola breeding programs, largely due to an absence of understanding of the traits on a genotype-specific level. This research was one of the first to look at the genetic control of specific amino acids in the meal and was able to reveal regions associated with ten of the seventeen amino acids in the study. Canola meal is currently utilized to some extent in the feeding regimes of various animal diets, but still suffers largely in comparison to soybean in terms of overall protein quantity and quality. A focused approach to improve the quality of canola's protein

requires a better understanding of the genetic control of protein-related traits and this research was able to provide valuable information.

One of the shortfalls of this research is the lack of variation that was observed among individuals in the mapping population, given that it was not developed specifically for variation of the traits of focus in this study. Linkage mapping analysis was still able to provide good chromosomal coverage on most of the *B. napus* chromosomes, but there were definitely areas where a lack of variation resulted in poor marker coverage. Maximizing both phenotypic and genotypic variation in the mapping population is key to the development of good chromosome coverage, and this should be considered in future studies.

To conclude, this research provides further evidence of genetic control of seed quality traits such as seed oil and seed protein within the A3 chromosome of *B. napus*. In addition, it displayed the interrelatedness of genomic regions controlling seed quality through the association observed between *qPROT-A3a* and *qOIL-A3a* as well as *qPROT-A3b* and *qOIL-A3b* across two years of field and mapping analyses. The research also demonstrated strong amino acid control on the C5 chromosome in *B. napus*. The C5 chromosome was found to have major control of arginine (*qARG-C5a*), glutamine (*qGLU-C5b*), isoleucine (*qILE-C5c*), leucine (*qLEU-C5d*), and lysine (*qLYS-C5e*) while maintaining minor control over asparagine (*qASP-C5f*), phenylalanine (*qPHE-C5g*), serine (*qSER-C5h*), and valine (*qVAL-C5i*). The will contribute greatly to future breeding efforts to characterize and incorporate protein and amino acid related traits into value-added products of *B. napus.*

6.0 FUTURE RESEARCH RECCOMENDATIONS

Future research is required given the novel nature of some of the discoveries in this study. One similarity between this research and other studies is that the subject mapping population was not specifically designed for seed protein or meal amino acid contents. To conduct a thorough analysis on the genetic control of these traits, a linkage mapping and QTL analysis study should be conducted on a population designed with the intent of maximizing variation in seed protein content and in individual amino acid contents of the meal in *B. napus*.

The nature of the methodology for obtaining amino acid contents of the meal through the full oxidized and acid hydrolysis process is time-consuming but increasing the number of technical replicates in future studies should allow for a more thorough understanding. In time, with further information, the development of populations to focus on specific amino acids could be very useful to *B. napus* breeding programs. Creating mapping populations to maximize variation in lysine content may prove useful in further understanding its control as it is one of the most limiting amino acids in both human and animal nutrition.

A more in-depth correlation analysis could also be conducted in the future to better understand the relationship between protein content in the seed to the crude protein content in the meal, and even further how these values relate to the overall amino acid profile. In addition, looking further into the relationship between the total protein value of canola meal with the contents of specific amino acids may be very useful from an animal feed perspective.

The data produced in this research regarding amino acid content of the 124 *B. napus* individuals could be useful in the development of models for identification of these traits using near-infrared spectroscopy (NIR). More data will be required from additional genotypes, but the dataset

provided from this research is a good start. The ability to screen populations for amino acid content using NIR is more time efficient and cost effective and would allow for rapid development of the datasets required to implement breeding strategies with to focus on protein and amino acids.

Eventually, the development of molecular markers for seed protein and specific amino acids will facilitate breeding efforts to select for these enhanced seed quality traits. The first step in the development of molecular markers is creating a better understanding of the regions of genetic control. The results from this study and recommendations for future research will serve to provide the framework required for molecular marker development in this area. The implementation of this knowledge in future breeding programs will lead to the development of *B. napus* materials low in erucic acid with high oil content, high protein content, and an enhanced amino acid profile.

7.0 LITERATURE CITED

Association of Official Analytical Chemists (AOAC). 1995. Official Methods of Analysis. 16th Edition. Washington, DC, USA. 15-16.

Aachary, A. A., and Thiyam, U. 2012. A pursuit of the functional nutritional and bioactive properties of canola proteins and peptides. Critical Reviews in Food Science and Nutrition. 52(11): 965-979.

Ahmadi, A. S., Shivazad, M., Zaghari, M., and Shahneh, A. Z. 2007. The effect of different levels of rapeseed meal (with or without enzyme) on the broiler breeder flocks performance. Iranian Journal of Agricultural Science. 37*:* 233-240.

American Soybean Association. 2015. *The Nutritional Value of U.S. Soybean Meal*. U.S. Soybean Export Council, Chesterfield, MO, USA. URL: https://ussec.org/wp-content/uploads/2015/10/20120123-Nutritional-value-white-paper-1.pdf. (7 November 2017)

Andersen, M. N., Heidmann, T. and Plauborg, F. 1996. The effects of drought and nitrogen on light interception, growth and yield of winter oilseed rape. Acta Agriculturae Scandinavica. 46(1): 55-67.

Anderson, V. and Schoonmaker, J. 2004. Effect of pulse grains on performance of newly weaned steer calves. NDSU Carrington Research Extension Center Beef Production Field Day Proceedings. 27: 6-8.

Asare, E. and Scarisbrick, D. H. 1995. Rate of nitrogen and sulphur fertilizers on yield, yield components and seed quality of oilseed rape (*Brassica napus* L.). Field Crop Research. 44: 41-46.

Asins, M. J., Bernet, G. P., Villalta, I., and Carbonell, E. A. 2010. QTL analysis in Plant Breeding. In Jain, S. M. and Brar, D. S. (Ed.). Molecular Techniques in Crop Improvement, 2nd Edition. p. 3-21. Science and Business Media LLC, New York, NY, USA.

Baker, D. H. 2000. Nutritional constraints to use of soy products by animals. In Drackley, J. K. (Ed.). Soy in Animal Nutrition. p. 1-12. Federation of Animal Societies. Savoy, IL, USA.

Ballvora. A., Ercolano, M. R., Weib, J., Mesem, K., Bormann, C. A., Oberhagemann, P., Salamini, F., Gebhardt, C. 2002. The R1 gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. The Plant Journal. 30: 361-371.

Batley, J., & Edwards, D. 2007. SNP Applications in Plants. In Oraguzie, N.C., Rikkerink, E. H. A., Gardiner, S. E., and De Silva, H. N. (Ed.). Association Mapping in Plants. p. 95-102. Spring-Verlag Berlin Heidelberg. New York, NY, USA.

Beare, J. 1957. Rapeseed oil as a food. Food Manufacturing. 32: 378-384.

Becker, H., Engqvist, G., and Karlsson, B. 1995. Comparison of rapeseed cultivars and resynthesized lines based on allozyme and RFLP markers. Theoretical and Applied Genetics. 91(1): 62-67.

Beckie, H. J., Harker, K. N., Légère, A., Morrison, M. J., Séguin-Swartz, G., and Falk, K. C. 2011. GM Canola: The Canadian Experience. Farm Policy Journal. 8: 43-49.

Bell, J. 1984. Nutrients and toxicants in rapeseed meal: a review. Journal of Animal Science. 58(4): 996-1010.

Bell, J., and Keith, M. 1989. Factors affecting the digestibility by pigs of energy and protein in wheat, barley and sorghum diets supplemented with canola meal. Animal Feed Science and Technology. 24(3): 253-265.

Bell, J., Keith, M., and Darroch, C. 1988. Lysine supplementation of grower and finisher pig diets based on high protein barley, wheat and soybean meal or canola meal, with observations on thyroid and zinc status. Canadian Journal of Animal Science. 68(3): 931-940.

Bell, J. M. 1993. Factors affecting the nutritional value of canola meal: A review. Canadian Journal of Animal Science. 73(4): 689-697

Bernatzky, R., and Tanksley, S. 1986. Genetics of actin-related sequences in tomato. Theoretical and Applied Genetics. 72(3): 314-321.

Blomstrand, E., Eliasson, J., Karlsson, H. K., and Köhnke, R. 2006. Branched-chain amino acids activate key enzymes in protein synthesis after physical exercise. The Journal of Nutrition. 136(1): 269S-273S.

Botstein, D., White, R. L., Skolnick, M., and Davis, R. W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics. 32(3): 314.

Brennan, R. F., Mason, M. G., and Walton, G. H. 2000. Effect of nitrogen fertilizer on the concentrations of oil and protein in canola (*Brassica napus*) seed. Journal of Plant Nutrition. 23: 339-348.

Broderick, G., Satter, L., and Harper, A. 1974. Use of plasma amino acid concentration to identify limiting amino acids for milk production. Journal of Dairy Science. 57(9): 1015-1023.

Broderick, G. A. 1987. Determination of protein degradation rates using a rumen in vitro system containing inhibitors of microbial nitrogen metabolism. British Journal of Nutrition. 58(3): 463-475.

Broderick, G. A. 2015. Canola science cluster research report. Canola Council of Canada, Winnipeg, MB, Canada.

Brookes, A. J. 1999. The essence of SNPs. Gene. 234(2): 177-186.

Bunge North America. 2016. H.E.A.R. Oil - Crude, Super-Degummed. Bunge North America, St. Louis, MO, USA. URL: https://www.bungenorthamerica.com/products/21-h-e-a-r-oil-crude-super-degummed. (7 November 2017)

Burcon NutraSciences. 2016. *Canola Proteins: Supertein® and Puratein™*. Burcon NutraScience Corporation, Vancouver, BC, Canada. URL: http://www.burcon.ca/products/canada_protein.php. (7 November 2017)

Bus, A., Körber, N., Snowdon, R. J., and Stich, B. 2011. Patterns of molecular variation in a species-wide germplasm set of *Brassica napus*. Theoretical and Applied Genetics. 123(8): 1413-1423.

Campbell, L., Rempel, C. B., and Wanasundara, J. P. 2016. Canola/Rapeseed Protein: Future Opportunities and Directions—Workshop Proceedings of IRC 2015. Plants. 5(2): 17.

Canadian Food Inspection Agency (CFIA). 2017. *Elements within the Nutriton Facts Table: Protein Efficiency Ratios*. Canadian Food Inspection Agency, Ottawa, ON, Canada. URL: http://www.inspection.gc.ca/food/labelling/food-labelling-for-industry/nutrition-labelling/elements-within-the-nutrition-facts-table. (7 November 2017)

Canadian Grain Commission. 2017a. *Quality of Western Canadian Canola 2016 - Oil Content.* Canadian Grain Commission, Winnipeg, MB, Canada. URL: http://www.grainscanada.gc.ca/canola/harvest-recolte/2016/hqc16-qrc16-4-en.htm. (7 November 2017)

Canadian Grain Commission. 2017b. *Quality of Western Canadian Canola 2016 - Protein Content*. Canadian Grain Commission, Winnipeg, MB, Canada. URL: http://www.grainscanada.gc.ca/canola/harvest-recolte/2016/hqc16-qrc16-5-en.htm. (7 November 2017)

Canola Council of Canada. 2015. *Canola Meal Feeding Guide*. 5th Edition. Canola Council of Canada, Winnipeg, MB, Canada. URL: https://www.canolacouncil.org/media/516716/2015_canola_meal_feed_industry_guide.pdf. (7 November 2017)

Canola Council of Canada. 2017a. *Industry Overview*. Canola Council of Canada, Winnipeg, MB, Canada. URL: https://www.canolacouncil.org/markets-stats/industry-overview/. (7 November 2017)

Canola Council of Canada. 2017b. *What is Canola Oil?* Canola Council of Canada, Winnipeg, MB, Canada. URL: https://www.canolacouncil.org/oil-and-meal/canola-oil/. (7 November 2017)

Canola Council of Canada, 2017c. *Growth Stage of the Canola Plant*. Canola Council of Canada, Winnipeg, MB, Canada. URL: https://www.canolacouncil.org/canola-encyclopedia/crop-development/growthstages. (18 April 2018)

Cargill. 2015. *VICTORY® V-Class™ Hybrids*. Cargill Inc., Minneapolis, MN, USA. URL: https://www.cargillag.ca/grow-grain/specialty-seeds-oils/hybrids/victory-v-class-hybrids. (7 November 2017)

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B., Corréa, M., Da Silva, C., Just, J., Falentin, C., Shin Koh, C., Le Clainche, I., Bernard, M., Bento, P., Noel, B., Labadie, K., Alberti, A., Charles, M., Arnaud, D., Guo, H., Daviaud, C., Alamery, S., Jabbari, K., Zhao, M., Edger, P. P., Chelaifa, H., Tack, D., Lasalle, G., Mestiri, I., Schnel, N., Le Paslier, M., Fan, G., Renault, V., Bayer, P. E., Golicz, A. A., Manoli, S., Lee, T., Ha Dinh Thi, V., Chalabi, S., Hu, Q., Fan, C., Tollenaere, R., Lu, Y., Battail, C., Shen, J., Sidebottom, C. H. D., Wang, X., Canaguier, A., Chauveau, A., Bérard, A., Denoit, G., Guan, M., Liu, Z., Sun, F., Pyo Lim, Y., Lyons, E., Town, C. D., Bancroft, I., Wang, X., Meng, J., Ma, J., Pires, J. C., King, G. J., Brunel, D., Delourme, R., Renard, M., Aury, J., Adams, K. L., Batley, J., Snowdown, R. J., Tost, J., Edwards, D., Zhou, Y., Hua, W., Sharpe, A. G., Paterson, A. H., Guan, C., and Wincker, P. 2014. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. Science. 345(6199): 950-953.

Chandler, P., and Polan, C. 1972. Considerations in interpretations of serum amino acids in lactating cows. Journal of Dairy Science. 55: 709.

Chao, H., Wang, H., Wang. X., Guo, L., Gu, J., Zhao, W., Li, B., Chen, D., Rabaonatahiry, N., and Li, M. 2017. Genetic dissection of seed oil and protein content and identification of networks associated with oil content in *Brassica napus*. Scientific Reports. 7: 46295.

Chardon, F., Jasinski, S., Durandet, M., Lecureuil, A., Soulay, F., Bedu, M., Guerche, P., and Masclaux-Daubresse, C. 2014. QTL meta-analysis in Arabidopsis reveals an interaction between leaf senescence and resource allocation to seeds. Journal of Experimental Botany. 65(14): 3949-3962.

Chen, S., Nelson, M., Ghamkhar, K., Fu, T., and Cowling, W. 2007. Divergent patterns of allelic diversity from similar origins: the case of oilseed rape (*Brassica napus* L.) in China and Australia. Genome. 51(1): 1-10.

Chuang, D. T. 2013. Branched-Chain Amino Acids. In Metabolism Vitamins and Hormones. p. 244-249. Elsevier. New York City, NY, USA.

Cregan, P., Jarvik, T., Bush, A., Shoemaker, R. C., Lark, K. G., Kahler, A. L., Kaya, N., vanToai, T. T., Lohnes, D. G., Chung, J., and Specht, J. E. 1999. An integrated genetic linkage map of the soybean genome. Crop Science. 39(5): 1464-1490.

Crouch, M., Tenbarge, K., Simon, A., and Ferl, R. 1982. cDNA clones for *Brassica napus* seed storage proteins: evidence from nucleotide sequence analysis that both subunits of napin are cleaved from a precursor polypeptide. Journal of Molecular and Applied Genetics. 2(3): 273-283.

Crouch, M. L., and Sussex, I. M. 1981. Development and storage-protein synthesis in *Brassica napus* L. embryos in vivo and in vitro. Planta. 153(1): 64-74.

Dai, Z. L., Li, X. L., Xi, P. B., Zhang, J., Wu, G., and Zhu, W. Y. 2012a. Metabolism of select amino acids in bacteria from the pig small intestine. Amino Acids. 42(5): 1597-1608.

Dai, Z. L., Li, X. L., Xi, P. B., Zhang, J., Wu, G., and Zhu, W. Y. 2012b. Regulatory role for L-arginine in the utilization of amino acids by pig small-intestinal bacteria. Amino Acids. 43(1): 233-244.

Dai, Z. L., Li, X. L., Xi, P. B., Zhang, J., Wu, G., and Zhu, W. Y. 2013. L-Glutamine regulates amino acid utilization by intestinal bacteria. Amino Acids. 45(3): 501-512.

Dai, Z. L., Wu, G., and Zhu, W. Y. 2011. Amino acid metabolism in intestinal bacteria: links between gut ecology and host health. Frontiers in Bioscience. 16(1): 1768-1786.

Daun, J. K. 1983. The Introduction of Low Erucic Acid Rapeseed Varieties into Canadian Production. In Kramer, J. K. G., Sauer, F. D., and Pigden, W. J. (Ed.). High and Low Erucic Acid Rapseed Oils: Production, Usage, Chemistry, and Toxilogical Evaluation. p. 162-180. Academic Press. Cambridge, MA, USA.

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).

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

Derrig, R., Clark, J., and Davis, C. 1974. Effect of abomasal infusion of sodium caseinate on milk yield, nitrogen utilization and amino acid nutrition of the dairy cow. The Journal of Nutrition. 104(2): 151-159.

Dill, G. M. 2005. Glyphosate-resistant crops: history, status and future. Pest Management Science. 61(3): 219-224.

Dow AgroSciences. 2016. *Nexera Canola*. Dow AgroSciences Canada Inc., Calgary, AB, Canada. URL: http://nexeracanola.ca/. (7 November 2017)

Downey, R. K. 1983. The Origin and Description of the Brassica Oilseed Crops. In Kramer, J. K. G., Sauer, F. D., and Pigden, W. J. (Ed.). High and Low Erucic Acid Rapseed Oils: Production, Usage, Chemistry, and Toxilogical Evaluation. p. 1-18. Academic Press. Cambridge, MA, USA.

Duncan, R. W., McVetty, P. B. E., Duguid, J. L., Fox, S. L., Paulmann, W., Sass, O., Fernando, W. G. D., and Li, G. 2016. HYHEAR 2 Roundup Ready® high erucic acid, low glucosinolate hybrid summer rape. Canadian Journal of Plant Science. 96(4): 711-713.

Duncan, R. W., McVetty, P. B. E., Duguid, J. L., Fox, S. L., Paulmann, W., Sass, O., Fernando, W. G. D., and Li, G. 2017. HYHEAR 3 Roundup Ready® high erucic acid, low glucosinolate hybrid summer rape. Canadian Journal of Plant Science. 97: 1185-1187.

El Idrissi, A. 2008. Taurine increases mitochondrial buffering of calcium: role in neuroprotection. Amino Acids. 34(2): 321-328.

Ellsworth, D., Rittenhouse, K., and Honeycutt, R. 1993. Artifactual variation in randomly amplified polymorphic DNA banding patterns. Biotechniques. 14(2): 214-217.

Ericson, M., Rödin, J., Lenman, M., Glimelius, K., Josefsson, L.-G., and Rask, L. 1986. Structure of the rapeseed 1.7 S storage protein, napin, and its precursor. Journal of Biological Chemistry. 261(31): 14576-14581.

Eskandari, M., Cober, E. R., and Rajcan, I. 2013. Genetic control of soybean seed oil: II. QTL and genes that increase oil concentration without decreasing protein or with increased seed yield. Theoretical and Applied Genetics. 126(6): 1677-1687.

Eskin, N. A. M. 1989. Chemical and physical properties of canola oil products. p. 16-32. In: Canola Oil Properties and Performance. Canola Council of Canada, Winnipeg, MB, Canada.

Fallen, B. D., Hatcher, C. N., Allen, F. L., Kopsell, D. A., Saxton, A. M., Chen, P., Kantartzi, S. K., Cregan, P. B., Hyten, D. L., and Pantalone, V. R. 2013. Soybean seed amino acid content QTL detected using the Universal Soy Linkage Panel 1.0 with 1,536 SNPs. Journal of Plant Genome Sciences. 1(3): 68-79.

Fangfang, T., Qiping, G., Ruojun, W., and Nernberg, L. 2014. Effects of feeding three kinds of rapeseed meal on growth performance of tilapia and the cost performance of three kinds of rapeseed meal. Feed Industry. 20.

FAO/IEAE Mutant Variety Database. 2016. *Mutant Variety Database*. Food and Agriculture Organization of the United Nations, International Atomic Energy Agency. URL: https://mvd.iaea.org/ (18 April 2018)

Food and Agriculture Organization (FAO). 2013. Dietary protein quality evaluation in human nutrition. FAO Food and Nutrition Paper. 92: 1-66.

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.

Friedman, M. 1996. Nutritional value of proteins from different food sources. A review. Journal of Agricultural and Food Chemistry. 44(1): 6-29.

Fu, Y. B., and Gugel, R. K. 2010. Genetic diversity of Canadian elite summer rape (*Brassica napus* L.) cultivars from the pre-to post-canola quality era. Canadian Journal of Plant Science. 90(1): 23-33.

Ganal, M. W., Altmann, T., and Röder, M. S. 2009. SNP identification in crop plants. Current Opinion in Plant Biology. 12(2): 211-217.

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): e84303.

Goldstein, D., and Pollock, D. 1997. Mutation processes and methods of phylogenetic inference. Journal of Heredity. 88: 335-342.

Gusta, M., Smyth, S. J., Belcher, K., Phillips, P. W., and Castle, D. 2011. Economic benefits of genetically-modified herbicide-tolerant canola for producers. AgBio Forum. 14: 1-13.

Hao, X., Chang, C., and Travis, G. J. 2004. Short communication: effect of long-term cattle manure application on relations between nitrogen and oil content in canola seed. Journal of Plant Nutrition and Soil Science. 167: 214-215.

Harvey, B. L., and Downey, R. K. 1964. The Inheritance of Erucic Acid Content in Rapeseed (*Brassica napus*) Canadian Journal of Plant Science. 44(1): 104-111.

Hasan, M., Seyis, F., Badani, A., Pons-Kühnemann, J., Friedt, W., Lühs, W., and Snowdon, R. 2006. Analysis of genetic diversity in the *Brassica napus* L. gene pool using SSR markers. Genetic Resources and Crop Evolution. 53(4): 793-802.

Hatzig, S. V., Frisch, M., Breuer, F., Nesi, N., Ducournau, S., Wagner, M. H., Leckband, G., Abbadi, A., and Snowdon, R. J. 2015. Genome-wide association mapping unravels the genetic control of seed germination and vigor in *Brassica napus*. Frontiers in Plant Science. 6: 221.

Hayward, A., Mason, A., Dalton-Morgan, J., Zander, M., Edwards, D., and Batley, J. 2012. SNP discovery and applications in *Brassica napus*. Plant Biotechnology. 39(1): 12.

He, M., and Armentano, L. 2011. Effect of fatty acid profile in vegetable oils and antioxidant supplementation on dairy cattle performance and milk fat depression. Journal of Dairy Science. 94(5): 2481-2491.

Hedqvist, H., and Udén, P. 2006. Measurement of soluble protein degradation in the rumen. Animal Feed Science and Technology. 126(1): 1-21.

Helentjaris, T., Slocum, M., Wright, S., Schaefer, A., and Nienhuis, J. 1986. Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. Theoretical and Applied Genetics. 72(6): 761-769.

Hickling, D. 1994. Canola meal hog feeding trials in Western Canada. Canola Council of Canada. Winnipeg, MB, Canada.

Höglund, A.-S., Rödin, J., Larsson, E., and Rask, L. 1992. Distribution of napin and cruciferin in developing rape seed embryos. Plant Physiology. 98(2): 509-515.

Howell, P., Lydiate, D., and Marshall, D. 1996. Towards developing intervarietal substitution lines in *Brassica napus* using marker-assisted selection. Genome. 39(2): 348-358.

Hu, Z. Y., Hua, W., Zhang, L., Deng, L. B., Wang, X. F., Liu, G. H., Hao, W. J., and Wang, H. Z. 2013. Seed structure characteristics to form ultrahigh oil content in rapeseed. PloS one. 8(4): e62099.

Huang, S., Deng, L., Guan, M., Li, J., Lu, K., Wang, H., Fu, D., Mason, A. S., Liu, S., and Hua, W. 2013. Identification of genome-wide single nucleotide polymorphisms in allopolyploid crop *Brassica napus*. BMC Genomics. 14(1): 717.

Huhtanen, P., Hetta, M., and Swensson, C. 2011. Evaluation of canola meal as a protein supplement for dairy cows: A review and a meta-analysis. Canadian Journal of Animal Science. 91(4): 529-543.

Hungate, R. E. 1966. The rumen and its microbes. p. 1-7. Academic Press. New York, USA.

Illumina. 2016. *International Brassica SNP Consortium*. Illumina. San Diego, CA, USA. URL: http://www.illumina.com/ (18 April 2018)

Institute of Medicine. 2006. Dietary Reference Intakes: The Essential Guide to Nutrient Requirements. The National Academies Press. Washington, DC, USA.

Ishida, M., Hara, M., Fukino, N., Kakizaki, T., and Morimitsu, Y. 2014. Glucosinolate metabolism, functionality and breeding for the improvement of *Brassicaceae* vegetables. Breeding Science. 64(1): 48-59.

Jansman, A., Smink, W., Van Leeuwen, P., and Rademacher, M. 2002. Evaluation through literature data of the amount and amino acid composition of basal endogenous crude protein at the terminal ileum of pigs. Animal Feed Science and Technology. 98(1): 49-60.

Jasinski, S., Lecureuil, A., Durandet, M., Bernard-Moulin, P., and Guerche, P. 2016. *Arabidopsis* Seed Content QTL Mapping Using High-Throughput Phenotyping: The Assets of Near Infrared Spectroscopy. Frontiers in Plant Science. 7: 1682.

Javed, N., Geng, J., Tahir, M., McVetty, P. B. E., Li, G., and Duncan, R. W. 2015. Identification of QTL influencing seed oil content, fatty acid profile and days to flowering in *Brassica napus* L. Euphytica. 207(1): 191-211.

Jayasinghe, N. 2014. Ruminal degradability and intestinal digestibility of protein and amino acids in canola meal. Paper presented at the 2014 ADSA-ASAS-CSAS Joint Annual Meeting. Kansas City, MO, USA.

Jesske, T., Olberg, B., and Becker, H. 2011. Resynthesized rapeseed with wild Brassica species as new genetic resource for breeding. Paper presented at the 13th International Rapeseed Congress. Prague, Czech Republic.

Jia, W., Mikulski, D., Rogiewicz, A., Zduńczyk, Z., Jankowski, J., and Slominski, B. A. 2012. Low-fiber Canola. Part 2. Nutritive Value of the Meal. Journal of Agricultural and Food Chemistry. 60(50): 12231-12237.

Joehanes, R., and Nelson, J. C. 2008. QGene 4.0, an extensible Java QTL-analysis platform. Bioinformatics 24(23): 2788-2789.

Jonah, P., Bello, L., Lucky, O., Midau, A., and Moruppa, S. 2011. Review: the importance of molecular markers in plant breeding programmes. Global Journal of Science Frontier Research. 11(5): 5-12.

Jondreville, C., Van den Broecke, J., Gatel, F., and Van Cauwenberghe, S. 1995. Ileal digestibility of amino acids in feedstuffs for pigs. Eurolysine/ITFC Publication. Paris, France.

Junker, A., Fischer, J., Sichhart, Y., Brandt, W., and Dräger, B. 2013. Evolution of the key alkaloid enzyme putrescine N-methyltransferase from spermidine synthase. Frontiers in Plant Science. 4: 260.

Kaminska, B. 2003. Substitution of soyabean meal with 00 rapeseed meal or its high-protein fraction in the nutrition of hens laying brown-shelled eggs. Journal of Animal and Feed Sciences. 12(1): 111-119.

Kearsey, M., and Farquhar, A. 1998. QTL analysis in plants; where are we now? Heredity. 80(2): 137-142.

Khandaker, L., Akond, M., Liu, S., Kantartzi, S. K., Meksem, K., Bellaloui, N., Lightfoot, D. A., and Kassem, M. A. 2015. Mapping of QTL Associated with Seed Amino Acids Content in" MD96-5722" by" Spencer" RIL Population of Soybean Using SNP Markers. Food and Nutrition Sciences. 6(11): 974.

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: 15-20.

Klockeman, D. M., Toledo, R., and Sims, K. A. 1997. Isolation and chacterization of defatted canola meal protein. Journal of Agricultural and Food Chemistry. 45(10): 3867-3870.

Kumar, S., Banks, T. W., and Cloutier, S. 2012. SNP discovery through next-generation sequencing and its applications. International Journal of Plant Genomics. Article ID: 831460.

Landero, J., Beltranena, E., Cervantes, M., Morales, A., and Zijlstra, R. 2011. The effect of feeding solvent-extracted canola meal on growth performance and diet nutrient digestibility in weaned pigs. Animal Feed Science and Technology. 170(1): 136-140.

Langridge, P., and Chalmers, K. 2004. The principle: identification and application of molecular markers. In Molecular Marker Systems in Plant Breeding and Crop Improvement. p. 3-22. Springer-Verlag Berlin Heidelberg. New York, USA.

Li, C., Beauchemin, K., and Yang, W. 2013. Effects of supplemental canola meal and various types of distillers grains on ruminal degradability, duodenal flow, and intestinal digestibility of protein and amino acids in backgrounded heifers. Journal of Animal Science. 91(11): 5399-5409.

Li, F., Chen, B., Xu, K., Wu, J., Song, W., Bancroft, I., Harper, A. L., Trick, M., Liu, S., Gao, G., Wang, N., Yan, G., Qiao, J., Li, J., Li, H., Xiao, X., Zhang, T., and Wu, X. 2014. Genome-wide association study dissects the genetic architecture of seed weight and seed quality in rapeseed (*Brassica napus* L.). DNA Research. 21(4): 355-367.

Liu, L., and Iassonova, D. 2012. High-oleic canola oils and their food applications. American Oil Chemists Society (AOCS) Inform. 9: 9-11.

LMC International. 2013. *The Economic Impact of Canola on the Canadian Economy*. [Online Database] Canola Council of Canada, Winnipeg, MB, Canada. URL: http://www.canolacouncil.org/ (18 April 2018)

LMC International. 2016. *The Economic Impact of Canola on the Canadian Economy*. [Online Database] Canola Council of Canada, Winnipeg, MB, Canada. URL: http://www.canolacouncil.org/ (18 April 2018)

Lobb, K. and Chow, C. K. 2008. Fatty acid classification and nomenclature. In Chow, C. K. (Ed.). Fatty acids in foods and their health implications. p. 1-15. Marcel Dekker Inc. New York City, NY, USA.

Lönnerdal, B., and Janson, J. C. 1972. Studies on *Brassica* seed proteins: I. The low molecular weight proteins in rapeseed. Isolation and characterization. Biochimica et Biophysica Acta. 278(1): 175-183.

Lorieux, M. 2012. MapDisto: fast and efficient computation of genetic linkage maps. Molecular Breeding. 30(2): 1231-1235.

Louard, R. J., Barrett, E. J., and Gelfand, R. A. 1990. Effect of infused branched-chain amino acids on muscle and whole-body amino acid metabolism in man. Clinical Science. 79(5): 457-466.

Lupi, A., Tenni, R., Rossi, A., Cetta, G., and Forlino, A. 2008. Human prolidase and prolidase deficiency: an overview on the characterization of the enzyme involved in proline recycling and on the effects of its mutations. Amino Acids. 35(4): 739-752.

Lynch, M., and Walsh, B. 1998. Genetics and Analysis of Quantitative Traits. Sinauer. Sunderland, MA, USA.

Pèrez-de-Castro, M., Volanova, S., Cañizares, J., Pascual, L., Blanca, J. M., Diez, M. J., Prohens, J., and Picó, B. 2012. Application of Genomic Tools in Plant Breeding. Current Genomics. 13(3): 179-195.

Mackay, T. F., Stone, E. A., and Ayroles, J. F. 2009. The genetics of quantitative traits: challenges and prospects. Nature Reviews Genetics. 10(8): 565-577.

Malabat, C., Atterby, H., Chaudhry, Q., Renard, M., and Guéguen, J. 2003. Genetic variability of rapeseed protein composition. Paper presented at the 11th International Rapeseed Congress. Copenhagen, Denmark.

Martineau, R., Ouellet, D., and Lapierre, H. 2013. Feeding canola meal to dairy cows: A meta-analysis on lactational responses. Journal of Dairy Science. 96(3): 1701-1714.

Martineau, R., Ouellet, D., and Lapierre, H. 2014. The effect of feeding canola meal on concentrations of plasma amino acids. Journal of Dairy Science. 97(3): 1603-1610.

Mason, A. S., Higgins, E. E., Snowdon, R. J., Batley, J., Stein, A., Werner, C., and Parkin, I. A. 2017. A user guide to the *Brassica* 60K Illumina Infinium SNP genotyping array. Theoretical and Applied Genetics. 130(4): 621-633.

Mauricio, R. 2001. Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. Nature Reviews Genetics. 2: 370-381.

Maxin, G., Ouellet, D., and Lapierre, H. 2013. Effect of substitution of soybean meal by canola meal or distillers grains in dairy rations on amino acid and glucose availability. Journal of Dairy Science. 96(12): 7806-7817.

McCallum, C. M., Comai, L., Greene, E. A., and Henikoff, S. 2000. Targeting induced locallesions in genomes (TILLING) for plant functional genomics. Plant Physiology. 123(2): 439-442.

McVetty, P. B. E., Cuthbert, J., Marwede, V., Paulmann, W., Sass, O., Duncan, R. W., Fernando, W. G. D., Li, G., and Zelmer, C. D. 2014. HYHEAR 1 hybrid Roundup Ready™ high erucic acid, low glucosinolate summer rape. Canadian Journal of Plant Science. 94(2): 453-455.

McVetty, P. B. E., Duncan, R. W, Fernando, W., Li, G., and Zelmer, C. 2012. Red River 1861 Roundup Ready™ high erucic acid, low glucosinolate summer rape. Canadian Journal of Plant Science. 92(7): 1407-1409.

McVetty, P. B. E., Scarth, R., and Rimmer, S. 2000. MillenniUM 03 summer rape. Canadian Journal of Plant Science. 80(3): 611-612.

McVetty, P. B. E., Mietkiewska, E., Omonov, T., Curtis, J., Taylor, D. C., Weselake, R. J. 2016. Brassica spp. Oils. In Industrial Oil Crops. p. 113-156. Elsevier. New York City, NY, USA.

Mejicanos, G., Sanjayan, N., Kim, I., and Nyachoti, M. 2016. Recent advances in canola meal utilization in swine nutrition. Journal of Animal Science and Technology. 58(7).

Milbourne, D., Meyer, R., Collins, A., Ramsay, L., Gebhardt, C., and Waugh, R. 1998. Isolation, characterization and mapping of simple sequence repeat loci in potato. Molecular and General Genetics. 259(3): 233-245.

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

Miller, E. (2004). Protein nutrition requirements of farmed livestock and dietary supply. Paper presented at the Protein sources for the animal feed industry. FAO Expert Consultation and Workshop. Bangkok, Thailan

Miller-Cushon, E., Terré, M., DeVries, T., and Bach, A. 2014. The effect of palatability of protein source on dietary selection in dairy calves. Journal of Dairy Science. 97(7): 4444-4454.

Mohler, V. and Schwarz, G. 2004. Genotyping tools in plant breeding: from restriction fragment length polymorphisms to single nucleotide polymorphisms. In Molecular Marker Systems in Plant Breeding and Crop Improvement. p. 23-38. Springer-Verlag Berlin Heidelberg. New York City, NY, USA.

Morrison, M. J., Harker, K. N., Blackshaw, R. E., Holzapfel, C. J., and O'Donovan, J. T. 2016. Canola yield improvement on the Canadian Prairies from 2000 to 2013. Crop and Pasture Science. 67(4): 245.

Mosenthin, R., Sauer, W., Blank, R., Huisman, J., and Fan, M. 2000. The concept of digestible amino acids in diet formulation for pigs. Livestock Production Science. 64(2): 265-280.

Muller, H. J. (1927). Artificial Transmutation of the Gene. Science. 66(1699): 84-87.

Muralidharan, K., and Wakeland, E. K. 1993. Concentration of primer and template qualitatively affects products in random-amplified polymorphic DNA PCR. Biotechniques. 14(3): 362-364.

Murray, M. G., and Thompson, W. F. 1980. Rapid isolation of high moleular weight plant DNA. Nucleic Acids Research. 8(19): 4321-4325.

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

Nair, K. S., Schwartz, R., and Welle, S. 1992. Leucine as a regulator of whole body and skeletal muscle protein metabolism in humans. American Journal of Physiology-Endocrinology and Metabolism. 263(5): E928-E934.

Naseem, M. Z., Khan, S. H., and Yousaf, M. 2006. Effect of feeding various levels of canola meal on the performance of broiler chicks. Journal of Animal and Plant Sciences. 16: 3-4.

Nassar, A., Goeger, M., and Arscott, G. 1985. Effect of canola meal in laying hen diets. Nutrition Reports International. 31(6): 1349-1355.

National Research Council (NRC). 2011. Nutrient requirements of fish and shrimp. National Academies Press, Washington, DC, USA. 800: 624-6242.

Newkirk, R., and Classen, H. 2002. The effects of toasting canola meal on body weight, feed conversion efficiency, and mortality in broiler chickens. Poultry Science. 81(6): 815-825.

Nieschlag, H., and Wolff, I. 1971. Industrial uses of high erucic oils. Journal of the American Oil Chemists Society. 48(11): 723-727.

Novak, C., Yakout, H., and Scheideler, S. 2004. The combined effects of dietary lysine and total sulfur amino acid level on egg production parameters and egg components in Dekalb Delta laying hens. Poultry Science. 83(6): 977-984.

Novelli, A., and Tasker, R. A. R. 2007. Excitatory amino acids in epilepsy: from the clinics to the laboratory. Amino Acids. 32(3): 295-297.

Nyachoti, C., Zijlstra, R., De Lange, C., and Patience, J. 2004. Voluntary feed intake in growing-finishing pigs: A review of the main determining factors and potential approaches for accurate predictions. Canadian Journal of Animal Science. 84(4): 549-566.

Oryshack, M., and Bentranena, E. 2013. Solvent-extracted vs. expeller-pressed *B. napus* and *B. juncea* fed to layers: Effects on feed intake, egg production and physical egg quality. Poultry Science. 92(E-Suppl. 1): 80.

Parker, E. J., and Pratt, A. J. 2010. Part One Amino Acids as Building Blocs: Amino Acid Biosynthesis. In Hughes, A. B. (Ed) Amino Acids, Peptides, and Proteins in Organic Chemistry, Volume 3, Building Blocks, Catalysis, and Coupling Chemistry. p. 1-82. John Wiley & Sons. Hoboken, NJ, USA.

Parry, M. A., Madgwick, P. J., Bayon, C., Tearall, K., Hernandez-Lopez, A., Baudo, M., Rakszegi, M., Hamada, W., Al-Yassin, A., Ouabbou, H., Labhilili, M., and Phillips, A. L. 2009. Mutation discovery for crop improvement. Journal of Experimental Botany. 60(10): 2817-2825.

Petit, H., Veira, D., and Yu, Y. 1994. Growth and carcass characteristics of beef steers fed silage and different levels of energy with or without protein supplementation. Journal of Animal Science. 72(12): 3221-3229.

Phang, J. M., Donald, S. P., Pandhare, J., and Liu, Y. 2008. The metabolism of proline, a stress substrate, modulates carcinogenic pathways. Amino Acids. 35(4): 681-690.

Przybylski, R., Mag, T., Eskin, N., and McDonald, B. 2005. Canola Oil. In Bailey's Industrial Oil and Fat Products, Sixth Edition. John Wiley & Sons. Hoboken, NJ, USA.

Przybylski, R. 2011. Canola/Rapeseed Oil. In Gunstone, F. (Ed) Vegetable Oils in Food Technology, Composition, Properties, and Uses, Second Addition. p. 107-133. Wiley-Blackwell. Hoboken, NJ, USA.

Qian, W., Meng, J., Li, M., Frauen, M., Sass, O., Noack, J., and Jung, C. 2006. Introgression of genomic components from Chinese Brassica rapa contributes to widening the genetic diversity in rapeseed (*B. napus* L.), with emphasis on the evolution of Chinese rapeseed. Theoretical and Applied Genetics. 113(1): 49-54.

Qiu, D., Morgan, C., Shi, J., Long, Y., Liu, J., Li, R., Zhuang, X., Wang, Y., Tan, X., Dietrich, E., Weihmann, T., Everett, C., Vanstraelen, S., Beckett, P., Fraser, F., Trick, M., Barnes, S., Wilmer, J., Schmidt, R., Li, J., Li, D., Meng, J., and Bancroft, I. 2006. A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. Theoretical and Applied Genetics. 114(1): 67-80.

Rahman, H. 2013. Review: Breeding spring canola (*Brassica napus* L.) by the use of exotic germplasm. Canadian Journal of Plant Science. 93(3): 363-373.

Rahman, H., Harwood, J., and Weselake, R. 2013. Increasing seed oil content in *Brassica* species through breeding and biotechnology. Lipid Technology. 25(8): 182-185.

Rakow, G. (2004). Species Origin and Economic Importance of *Brassica*. In Pua, E. C. and Douglas, C. J. (Eds.) Brassica. Biotechnology in Agriculture and Forestry. p. 3-11. Springer-Verlag Berlin Hiedelberg. New York City, NY, USA.

Ramesh, K., Devegowda, G., and Khosravinia, H. 2006. Effects of enzyme addition to broiler diets containing varying levels of double zero rapeseed meal. Asian Australasian Journal of Animal Sciences. 19(9): 1354.

Ramsay, L., Macaulay, M., Degli Ivanissevich, S., MacLean, K., Cardle, L., Fuller, J., Edwardsm K. J., Tuvesson, S., Morgante, M., Massari, A., Maestri, E., Marmiroli, N., Sjakste, T., Ganal, M., Powell, W., and Waugh, R. 2000. A simple sequence repeat-based linkage map of barley. Genetics. 156(4): 1997-2005.

Reeds, P. J. 2000. Dispensable and indispensable amino acids for humans. The Journal of Nutrition. 130(7): 1835S-1840S.

Röder, M. S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M. H., Leroy, P., and Ganal, M. W. 1998. A microsatellite map of wheat. Genetics. 149(4): 2007-2023.

Rogiewicz, A., Dyck, B.; Slominski, B.A. 2015. High-Inclusion level of canola meal in laying-hen diets. Paper presented at the International Rapeseed Congress. Saskatoon, SK, Canada.

Ross, D., Gutierrez-Botero, M., and Van Amburgh, M. 2013. Development of an in vitro intestinal digestibility assay for ruminant feeds. p. 190-202. Proceedings of Cornell Nutrition Conference.

Roy, N. 1984. Interspecific transfer of *Brassica juncea*-type high blackleg resistance to *Brassica napus*. Euphytica. 33(2): 295-303.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences. 74(12): 5463-5467.

Sauer, W. C., and Ozimek, L. 1986. Digestibility of amino acids in swine: Results and their practical applications. A review. Livestock Production Science. 15(4): 367-388.

Schierwater, B., and Ender, A. 1993. Different thermostable DNA polymerases may amplify different RAPD products. Nucleic Acids Research. 21(19): 4647-4648.

Schöne, F., Rudolph, B., Kirchheim, U., and Knapp, G. 1997. Counteracting the negative effects of rapeseed and rapeseed press cake in pig diets. British Journal of Nutrition. 78(06): 947-962.

Schwab, C. (2012). The principles of balancing diets for amino acids and their impact on N utilization efficiency. Paper presented at the Proceedings of the 23rd Ruminant Nutrition Symposium. Gainsville, FL, USA.

Schwab, C. G., Satter, L., and Clay, A. 1976. Response of lactating dairy cows to abomasal infusion of amino acids. Journal of Dairy Science. 59(7): 1254-1270.

Schwarz, G., Michalek, W., Mohler, V., Wenzel, G., and Jahoor, A. 1999. Chromosome landing at the Mla locus in barley (*Hordeum vulgare* L.) by means of high-resolution mapping with AFLP markers. Theoretical and Applied Genetics. 98: 521-530.

Schwenke, K., Raab, B., Linow, K. J., Pahtz, W., and Uhlig, J. 1981. Isolation of the 12S globulin from Rapeseed (*Brassica napus* L.) and characterization as a "neutral" protein on seed proteins. Food/Nahrung. 25(3): 271-280.

Sega, G. A. 1984. A review of the genetic effects of ethyl methanesulfonate. Mutation Research/Reviews in Genetic Toxicology. 134(2-3): 113-142.

Sharopova, N., McMullen, M. D., Schultz, L., Schroeder, S., Sanchez-Villeda, H., Gardiner, J., Bergstrom, D., Houchins, K., Melia-Hancock, S., Musket, T., Duru, N., Polacco, M., Edwards, K., Ruff, T., Register, J. C., Brouwer, C., Thompson, R., Velasco, R., Chin, E., Lee, M., Woodman-Clickeman, W., Long, M. J., Liscum, E., Cone, K., Davis, G., and Coe Jr, E. H. 2002. Development and mapping of SSR markers for maize. Plant Molecular Biology. 48(5-6): 463-481.

Sharpe, A. G. 2012. New molecular marker platforms for analysis of phenotypic traits in Brassicas. Paper presented at the National Research Council. Saskatoon, SK, Canada.

Sikora, P., Chawade, A., Larsson, M., Olsson, J., and Olsson, O. 2012. Mutagenesis as a tool in plant genetics, functional genomics, and breeding. International Journal of Plant Genomics. Article ID: 314829.

Simons, G., Groenendijk, J., Wijbrandi, J., Reijans, M., Groenen, J., Diergaarde, P., Van der Lee, T., Bleeker, M., Onstenk, J., de Both, M., Haring, M., Mes, J., Cornelissen, B., Zabeau, M., and Vos, P. 1998. Dissection of the Fusarium I2 Gene Cluster in Tomato Reveals Six Homologs and One Active Gene Copy. The Plant Cell. 10: 1055-1068.

Sjödahl, S., Rödin, J., and Rask, L. 1991. Characterization of the 12S globulin complex of *Brassica napus*. European Journal of Biochemistry. 196(3): 617-621.

Slominski, B. 2015. Canola science cluster research report. Canola Council of Canada, Winnipeg, MB, Canada.

Snogerup, S. 1980. The wild forms of the *Brassica oleracea* group (2n=18) and their possible relations to the cultivated ones. In Tsunoda, S., Hinata, S., and Gómez-Campo, C. (Ed.), Brassica crops and wild allies: biology and breeding. p. 121-132. Japan Scientific Societies Press. Tokyo, Japan.

Stadler, L. 1928. Mutations in barley induced by x-rays and radium. Science. 68(1756): 186.

Statistics Canada. 2014. *Aquaculture Statistics*. Statistics Canada, Ottawa, ON, Canada. URL: http://www.statcan.gc.ca/pub/23-222-x/23-222-x2014000-eng.htm. (7 November 2017)

Staub, J. E., Serquen, F. C., and Gupta, M. 1996. Genetic markers, map construction, and their application in plant breeding. HortScience. 31(5): 729-741.

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.

Stefansson, B. R. and Kondra, Z. P. 1975. Tower summer rape. Canadian Journal of Plant Science. 55: 343-344.

Stein, H. H., Berger, L. L., Drackley, J. K., Fahey Jr, G., Hernot, D. C., and Parsons, C. M. 2008. Nutritional properties and feeding values of soybeans and their coproducts. In Soybeans chemistry, production, processing, and utilization. p. 613-660. AOCS Press. Urbana, IL, USA.

Stein, H. H., Seve, B., Fuller, M., Moughan, P., and De Lange, C. 2007. Invited review: Amino acid bioavailability and digestibility in pig feed ingredients: Terminology and application. Journal of Animal Science. 85(1): 172-180.

Tan, S., Evans, R. R., Dahmer, M. L., Singh, B. K., and Shaner, D. L. 2005. Imidazolinone-tolerant crops: history, current status and future. Pest Management Science. 61(3): 246-257.

Tan, S. H., Mailer, R. J., Blanchard, C. L., and Agboola, S. O. 2011. Canola proteins for human consumption: extraction, profile, and functional properties. Journal of Food Science. 76(1): R16-R28.

Tan, S. H., Mailer, R. J., Blanchard, C. L., and Agboola, S. O. 2014. Emulsifying properties of proteins extracted from Australian canola meal. LWT-Food Science and Technology. 57(1): 376-382.

Tautz, D., and Renz, M. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Research: 12(10): 4127-4138.

Teh, L., and Mollers, C. 2016. Genetic variation and inheritance of phytosterol and oil content in a doubled haploid population derived from the winter oilseed rape Sansibar x Oase cross. Theoretical and Applied Genetics. 129(1): 181-199.

Temnykh, S., Park, W. D., Ayres, N., Cartinhour, S., Hauck, N., Lipovich, L., Cho, Y. G., Ishii, T., and McCouch, S. R. 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). Theoretical and Applied Genetics. 100(5): 697-712.

Thomas, C. M., Vos, P., Zabeau, M., Jones, D. A., Norcott, K. A., Chadwick, B. P., and Jones, J. D. G. 1995. Identification of amplified restrction length polymorphism (AFLP) markers tightly linked to the tomato Xf-9 gene for resistance to *Cladosporium fulvum*. The Plant Journal. 8(5): 785-794.

Till, B. J., Cooper, J., Tai, T. H., Colowit, P., Greene, E. A., Henikoff, S., and Comai, L. 2007. Discovery of chemically induced mutations in rice by TILLING. BMC Plant Biology. 7(1): 1.

Till, B. J., Reynolds, S. H., Weil, C., Springer, N., Burtner, C., Young, K., Bowers, E., Comodo, C. A., Enns, L. C., Odden, A. R., Greene, E. A., Comai, L., and Henikoff, S. 2004. Discovery of induced point mutations in maize genes by TILLING. BMC Plant Biology. 4(1): 12.

Tsunoda, S., Hinata, K., and Gómez-Campo, C. 1980. Brassica crops and wild allies: biology and breeding: Japan Scientific Societies Press, Toyko.

Tylutki, T. P., Fox, D. G, Durbal, V. M., Tedeschi, L. O, Russell, J. B., Van Amburgh, M. E., Overton, T. R., Chase, L. E., and Pell, A. N. 2008. Cornell Net Carbohydrate and Protein System: A model for precision feeding of dairy cattle. Animal Feed Science and Technology. 143(1): 174-202.

Tzeng, Y., Diosady, L. L., and Rubin, L. J. 1988. Preparation of rapeseed protein isolate by sodium hexamteaphosphate extraction, ultrafiltration, diafiltration, and ion-exchange. Journal of Food Science. 53(5): 1537-1541.

United States Department of Agriculture (USDA). 2016a. *Soybeans & Oil Crops: Canola*. United States Department of Agriculture, Washington, DC, USA. URL: http://www.ers.usda.gov/topics/crops/soybeans-oil-crops/canola.aspx. (7 November 2017)

United States Department of Agriculture (USDA). 2016b. *World Agricultural Supply and Demand Estimates*. United States Department of Agriculture, Washington, DC, USA. URL: https://www.usda.gov/oce/commodity/wasde/ (7 November 2017)

Uppström, B. (1995). Seed Chemistry. In Kimber, D. S. and McGregor, D. I. (Ed.), *Brassica* oilseeds: production and utilization. p. 217-242. CAB International. Wallingford, UK.

Uzunova, M., and Ecke, W. 1999. Abundance, polymorphism and genetic mapping of microsatellites in oilseed rape (*Brassica napus* L.). Plant Breeding. 118(4): 323-326.

Vles, R. O., Bijster, G. M., and Timmer, W. G. 1978. Nutritional Evaluation of Low-Erucic-Acid Rapeseed Oils. In B. J. Leonard (Ed.), Toxicological Aspects of Food Safety: Proceedings of the European Society of Toxicology. p. 23-32. Springer-Verlag Berlin Heidelberg. New York City, NY, USA.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., Friters, A., Pot, J., Paleman, J., and Kuiper, M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research: 23(21): 4407-4414.

Wanasundara, J. P. 2011. Proteins of *Brassicaceae* oilseeds and their potential as a plant protein source. Critical Reviews in Food Science and Nutrition. 51(7): 635-677.

Wanasundara, J. P. D., McIntosh, T., Perera, S. P., Withana-Gamage, T. S., and Mitra, P. 2016. Canola/rapeseed protein-functionality and nutrition. Oilseeds and fats, Crops and Lipis. 23(4): D407.

Wang, N., Wang, Y., Tian, F., King, G. J., Zhang, C., Long, Y., Shi, L., and Meng, J. 2008. A functional genomics resource for *Brassica napus*: development of an EMS mutagenized population and discovery of FAE1 point mutations by TILLING. New Phytologist. 180(4): 751-765.

Wang, X., Wang, H., Long, Y., Li, D., Yin, Y., Tian, J., Chen, L., Liu, L., Zhao, W., Zhao, Y., Yu, L., and Li, M. 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.

Warrington, C., Abdel-Haleem, H., Hyten, D., Cregan, P., Orf, J., Killam, A., Bajjalieh, N., Li, Z., and Boerma, H. 2015. QTL for seed protein and amino acids in the Benning x Danbaekkong soybean population. Theoretical and Applied Genetics. 128(5): 839-850.

Warwick, S. I., Francis, A., and Gugel, R.K. 2009. Guide to Wild Germplasm of Brassica and Allied Crops (tribe Brassiceae, Brassicaceae) 3rd ed. Agriculture and Agri-Food Canada Research Branch Publication. Ottawa, ON, Canada.

Welsh, J., and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research. 18(24): 7213-7218.

Wen, J., Xu, J. F., Long, Y., Wu, J. G., Xu, H. M., Meng, J. L., and Shi, C. H. 2016. QTL mapping based on the embryo and maternal genetic systems for non-essential amino acids in rapeseed (*Brassica napus* L.) meal. Journal of the Science of Food and Agriculture. 96(2): 465-473.

Wilde, H. D. 2015. Induced Mutations in Plant Breeding. In Al-Khayri, J. M., Jain, S. M., and Johnson, D. V. (Ed.). Advances in Plant Breeding Strategies: Breeding, Biotechnology and Molecular Tools. p. 329-344. Springer-Verlag Berlin Heidelberg. New York City, NY, USA.

Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A., and Tingey, S. V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research. 18(22): 6531-6535.

World Health Organization (WHO). 2007. Protein and amino acid requirements in human nutrition. World Health Organization Technical Report Series. 935:1.

Wu, G. 2009. Amino acids: metabolism, functions, and nutrition. Amino Acids. 37(1): 1-17.

Wu, G. 2010. Functional amino acids in growth, reproduction, and health. Advances in Nutrition: An International Review Journal. 1(1): 31-37.

Wu, G. 2013. Amino acids: biochemistry and nutrition. CRC Press. Boca Raton, FL, USA.

Wu, G. 2013. Functional amino acids in nutrition and health. Amino Acids. 45(3): 407-411.

Wu, G., Fanzo, J., Miller, D. D., Pingali, P., Post, M., Steiner, J. L., and Thalacker-Mercer, A. E. 2014. Production and supply of high-quality food protein for human consumption: sustainability, challenges, and innovations. Annals of the New York Academy of Sciences. 1321(1): 1-19.

Wu, G., Wu, Z., Dai, Z., Yang, Y., Wang, W., Liu, C., Wang, B., Wang, J., and Yin, Y. 2013. Dietary requirements of "nutritionally non-essential amino acids" by animals and humans. Amino Acids. 44(4): 1107-1113.

Wu, J., and Muir, A. 2008. Comparative structural, emulsifying, and biological properties of 2 major canola proteins, cruciferin and napin. Journal of Food Science. 73(3): C210-C216.

Wu, Y., Bhat, P. R., Close, T. J., and Lonardi, S. 2008. Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. PLoS Genet. 4(10): e1000212.

Wurschum, T., Liu, W., Maurer, H. P., Abel, S., and Reif, J. C. 2012. Dissecting the genetic architecture of agronomic traits in multiple segregating populations in rapeseed (*Brassica napus* L.). Theoretical and Applied Genetics. 124: 153-161.

Xu, J. F., Long, Y., Wu, J. G., Li, J. R., Huang, Z. R., and He, Q. L. 2013. QTL mapping based on embryo and maternal genetic systems for oil and protein contents in rapeseed (*Brassica napus* L.). China Agriculture Science. 47(8): 1471-1480.

Yang, W., Xu, L., Li, C., and Beauchemin, K. 2013. Short communication: Effects of supplemental canola meal and various types of distillers' grains on growth performance of backgrounded steers. Canadian Journal of Animal Science. 93(2): 281-286.

Yen, J.T. 2000. Anatomy of the digestive system and nutritional physiology. In Lewis, A. J. and Southern, L. L. (Ed.). Swine Nutrition, Second Edition. CRC Press. Boca Raton, FL, USA.

Young, V. R., and Pellett, P. L. 1994. Plant proteins in relation to human protein and amino acid nutrition. The American Journal of Clinical Nutrition. 59(5): 1203S-1212S.