Studies on Transcription Factors Involved in Seed Oil Biosynthesis

BY

JESSICA LEANNE PRYSTENSKI

A Thesis
Submitted to the Faculty of Graduate Studies
in Partial Fulfillment of Requirements
for the Degree of

MASTER OF SCIENCE

Department of Plant Science
University of Manitoba
Winnipeg, Manitoba

TABLE OF CONTENTS

ACKOWLEDGEMENTS

I would like to thank both my co-supervisors Dr. Muhammad Tahir and Dr. Claudio Stasolla for providing direction, and support throughout my research. I am very grateful for their constant presence and initiative throughout my graduate program.

Many people working in both Dr. Stasolla's and Dr. Tahir's labs have helped complete this research. A special thanks to Kevin Baron for his contribution to this work. I was very fortunate to have Kevin in my lab, he was always willing to teach me a new technique, discuss any problems I was facing, and provide insight and suggestions to improve my research. Ralph Kowatsch was very helpful, answered numerous questions, and provided humor throughout the last two years. Other members of Dr. Tahir's and Dr. Stasolla's labs have also contributed to this research: Mohamed Elhiti, Susan Slater, Roger Watts, Nasir Javed, Doug Durnin, Lauren Erickson, Jaime Prystenski and Emily Smith.

Thank you to my committee members Dr. Peter McVetty and Dr. Sylvie Renault for taking the time to meet with me and providing suggestions to improve my research. Many people throughout the department have helped make this research possible, the Department of Plant Science Office, Greenhouse Staff, and professors who taught me classes throughout this program. Thank you to all my fellow graduate students for your friendship and support. Finally, I would like to thank my family and friends who have supported me throughout the last two years of graduate school.

This work was supported by the Natural Sciences and Engineering Research Council (NSERC) and industrial and government partners, through the Green Crop Networks (GCN) Research Network.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABI3, ABSCISIC ACID INSENSITIVE3
ACCase, acetyl-CoA carboxylase
ACP, acyl carrier protein
CBF, CCAAT box binding factor
cDNA, complementary deoxyribonucleic acid
CoA, Coenzyme A
CPT, cholinephosphotransferase
DAG, diacylglycerol
DGAT, diacylglycerol acyltransferases
DH, 3-hydroxyacyl-CoA-dehydrase
DHAP, dihydroxyacetonephosphate
DNA, deoxyribonucleic acid
ER, enoyl-CoA reductase
EST, expressed sequence tag
FAD, oleate desaturase
FAD2, oleate desaturase
FAD3, linoleate desaturase
FAS, fatty acid synthase
FAT, acyl-ACP thioesterase
FUS3, FUSCA3
G3P, glycerol-3-phosphates
G3PAT, glycerol-3-phosphate acyl transferase
G3PDH, glycerol-3-phosphate acyltransferase
KAS, ketoacyl-ACP synthase
KCS, 3-ketoacyl-CoA synthase
KR, 3-ketoacyl reductase
L1L, LEAFY COTYLEDON1-like
LB, Luria Bertani
LEC1, LEAFY COTYLEDON1
LEC2, LEAFY COTYLEDON2
LPA, lysophosphatidic acid
LPAAT, lysophosphatidic acid acyltransferase
LPCAT, lysophosphatidylcholine acyltransferase
NCBI, National Center for Biotechnology Information
OPPP, oxidative pentose phosphate pathway
PA, phosphatidic acid
PAP, phosphatidic acid phosphohydrolase
PC, phosphatidylcholine
PCR, polymerase chain reaction
PDAT, phospholipid: diacylglycerol acyltransferase
PDF PP2A, phosphatase 2A subunit
RNA, ribonucleic acid
TAG, triacylglycerols
VLCFA, very long chain fatty acids
WRI1, WRINKLED1

ABSTRACT

Prystenski, Jessica Leanne. M.Sc., The University of Manitoba, November, 2010.
Studies on Transcription Factors Involved in Seed Oil Biosynthesis. Supervisors: Muhammad Tahir, and Claudio Stasolla.

Canola *(Brassica napus)* oil has immense economic value due to its use as food. The production of biodiesel has led to the recent increase in the demand for canola oil, and further reinforced the need to improve the oil content and composition of canola seeds. The content and quality of canola oil is dependent on the process of fatty acid biosynthesis and accumulation of triacylglycerols (TAGs) during seed development. The biochemical pathways leading to the synthesis of fatty acids and TAGs are thoroughly understood, and many of the genes encoding the metabolic enzymes involved in this pathway have been isolated and functionally characterized. Despite comprehensive understanding of fatty acid and TAG biosynthesis, alternate factors and genes involved in the modification of seed oil content and composition are being investigated. This has initiated a new sector of research focusing on a series of transcription factors: LEAFY COTYLEDON1 (LEC1), LEAFY COTYLEDON2 (LEC2), FUSCA3 (FUS3) and WRINKLED1 (WRI1), which play a key role during embryo and seed development. Previous studies in *Arabidopsis* have demonstrated that the noted transcription factors influence the synthesis and accumulation of oil during seed development. These findings have prompted the design of two studies with the objective of further characterizing the role of these genes. In the first study, a homolog of *Arabidopsis LEC1* was isolated from *Brassica napus (Bn)* microspore derived embryos. *BnLEC1* contained the same distinct features identified in *LEC1* genes from other plant species. *BnLEC1* was inserted into an expression vector, which was subsequently used to transform *Arabidopsis* plants. The transgenic lines were characterized by differences in seed oil composition, and one line showed a reduction in the number of seeds counted per silique. Expression analysis revealed that the transgene was not expressed in the transgenic lines. To account for the discrepant findings, the possibility of endogenous gene silencing mechanisms suppressing transgene expression has been discussed. As an extension of this research, a second study assessed the differential expression of LEC1, LEC2, FUS3 and WRI1 in double haploid *Brassica napus* lines characterized by different levels of seed oil content. The expression level of the noted genes was determined at 7, 14, 21 and 28 days after pollination. Numerous changes in the transcript level were observed, but the trends were not consistent among high and low oil content lines.

1.0 INTRODUCTION

Canola *(Brassica napus)* is an economically important oilseed crop grown worldwide. In 2009, approximately 6 million hectares of canola were harvested in Canada, producing roughly 11 million tonnes of seed (Canola Council of Canada, 2009a). The Canadian canola industry produces oil mainly to be consumed as food. However, canola oil is also used for the production of biodiesel, and a number of industrial applications including the manufacturing of paints and lubricants (Scarth and Tang, 2006; Weselake et al., 2009). Biodiesel production is a major cause of the growing demand for canola oil. As a result, it is predicted that by 2015, a 30% increase in production will be required to meet the new requirements of the canola industry (Canola Council of Canada, 2007). Conventional plant breeding and genetic approaches have been used in the past to improve the fatty acid composition and content of canola oil to make it suitable for various end uses (Scarth and Tang, 2006; Weselake et al., 2009). Increased understanding of the genetic regulation of seed oil biosynthesis in the model plant *Arabidopsis,* combined with recent advances in plant genetics, genomics and molecular biology, have initiated new avenues of research which focus on improving canola oil by exploiting the genetic regulation of fatty acid biosynthesis

The metabolic pathways leading to fatty acid and triacylglycerols (TAG) synthesis require the participation of many proteins encoding enzymes and regulatory components (Li-Beisson et al., 2010). To date, more than 600 genes have been identified in acyl-lipid metabolism in *Arabidopsis thaliana* (Li-Beisson et al., 2010). Many of the genes participating in the pathways leading to fatty acid and TAG synthesis have been isolated and further characterized (Li-Beisson et al., 2010). While the fatty acid

biosynthesis and the enzymes involved in this pathway are thoroughly understood, research suggests that there is a threshold level regarding the influence these enzymes have upon seed oil biosynthesis and accumulation during seed development (Thelen and Ohlrogge, 2002; Mu et al., 2008; Baud and Lepiniec, 2009; Liu et al., 2009; Weselake et al., 2009). This has prompted numerous research efforts to understand the alternate factors influencing fatty acid and TAG synthesis (Mu et al., 2008; Liu et al., 2009). Fatty acids and TAGs accumulate during embryo maturation and seed development (Baud et al., 2008; Braybrook and Harada, 2008). As a result, the maturation phase of embryo development is of particular interest when trying to improve the seed oil content and composition. Recently, emphasis has been placed on manipulating fatty acid biosynthesis through a regulatory network of transcription factors: LEAFY COTYLEDON1 (LEC1), LEAFY COTYLEDON2 (LEC2), FUSCA3 (FUS3), WRINKLED1 (WRI1), and ABSCISIC ACID INSENSITIVE3 (ABI3), which control embryo development and seed maturation (Meinke, 1992; Meinke et al., 1994; Parcy et al., 1994; Stone et al., 2001; Cernac and Benning, 2004; Gazzarrini et al., 2004). These transcription factors interact to regulate different aspects of embryo growth and maturation (Meinke, 1992; Meinke et al., 1994; Parcy et al., 1994; Stone et al., 2001; Cernac and Benning, 2004; Gazzarrini et al., 2004; Braybrook and Harada, 2008).

The role of embryogenesis related transcription factors in oil biosynthesis and accumulation during seed development has been mainly studied in *Arabidopsis*. Transgenic *Arabidopsis* plants down-regulating or over-expressing these transcription factors have been developed to study the function of these genes (Meinke, 1992; Meinke et al., 1994; Parcy et al., 1994; Stone et al., 2001; Cernac and Benning, 2004; Gazzarrini

et al., 2004; Braybrook and Harada, 2008). Several transgenic plants generated from these studies show altered fatty acid composition in seed and during seedling development (Meinke, 1992; Meinke et al., 1994; Stone et al., 2001; Cernac and Benning, 2004; Mu et al., 2008; Liu et al., 2009). These studies have led to the conjecture that homologs of these transcription factors in *Brassica napus* may play the same role in fatty acid biosynthesis that has already been demonstrated in *Arabidopsis*. As a result, the hypothesis for this research is that the seed oil content and fatty acid composition of oil in *Brassica napus* can be improved through the manipulation of the expression of these transcription factors.

1.1 Research objectives

Genetic evidence has in fact revealed that the expression of *WRI1, LEC2, FUS3* and *ABI3* is induced by LEC1 (Baud et al., 2008; Mu et al., 2008; Santos-Mendoza et al., 2008). Based on the results from the studies described above, the purpose of the proposed research is to assess the effects of altered *Brassica napus LEC1* expression on the fatty acid composition of *Arabidopsis* seeds. This objective will be reached in several steps:

1. Isolation of *BnLEC1* from *Brassica napus* tissue
2. Ectopic expression of *BnLEC1* in *Arabidopsis*
3. Analysis of fatty acid composition in the transgenic plants

As an extension of these studies, the expression level of *LEC1, LEC2, FUS3* and *WRI1* will be measured in the developing seeds of *Brassica napus* plants characterized by different oil composition.

2.0 LITERATURE REVIEW

2.1 The history of canola oil

Canola *(Brassica napus)* is a Canadian-developed crop that has become increasingly prominent in markets worldwide. The first canola cultivar Tower, was registered in 1974 by Dr. Baldur Stefansson, a professor at the University of Manitoba (Przybylski et al., 2005; Scarth and Tang, 2006; Canola Council of Canada, 2010a). Canola was developed through the conventional breeding of rapeseed varieties (Przybylski et al., 2005). Canola differs from rapeseed because the meal contains less than 30 µmole/g of glucosinolates, and the oil has less than 2% erucic acid (Przybylski et al., 2005). According to the Canola Council of Canada (2010b), typical Canola oil is composed of 61% monounsaturated fat, 32% polyunsaturated fat, and 7% saturated fat. Compared to other vegetable oils, canola oil is considerably lower in saturated fatty acids, and contains increased levels of monounsaturated and polyunsaturated fatty acids (Casséus, 2009). The consumption of saturated fatty acids has been linked to elevated blood cholesterol levels and increased risk of cardiovascular diseases (De Lorgeril and Salen, 2004). Because canola oil is low in saturated fatty acids, it is considered a healthy option (Canola Council of Canada, 2010b). Canola oil not only has good nutritional value for human consumption but is also valuable for industrial applications including the production of biodiesel, cosmetics, lubricants and paints (Przybylski et al., 2005; Weselake et al, 2009).

2.2 The economic importance of canola oil

According to the Canola Council of Canada (2009b), canola production alone contributed 14.1 billion dollars to the Canadian economy in 2008. In 2006, a total of 3.6 million tonnes of canola seeds were crushed in different plants predominantly located throughout western Canada (Casséus, 2009). Canada was ranked the second largest producer of canola and rapeseed in the world in 2006, behind China (Casséus, 2009). Production has nearly doubled from 3,782,906 hectares in 2001 to 5,027,643 in 2006, resulting in 9 million tonnes of canola seed (Casséus, 2009). Because of the growing demand for canola oil, both in Canada and on a global basis, the Canola Council of Canada (2009b) has set a goal of increasing production to 15 million tonnes by 2015. The industry also aims at increasing the average of total seed oil content to 45% by 2015 from the current average of 42.5% (Canola Council of Canada, 2007).

Worldwide breeding programs have been successful in improving the total seed oil content, and modifying fatty acid profile of *Brassica* seeds. Because of the numerous applications and uses of canola oil, there is a coinciding high demand for oils of varying fatty acid composition. Conventional and transgenic approaches are currently being used to improve the seed oil content and composition of *Brassica napus*.

2.3 Conventional breeding of *Brassica* oils

Conventional breeding manipulates genetics through sexual recombination, and this is typically achieved by a combination of crossing, selfing, and selecting for desirable phenotypes to improve plant performance (Sleper and Poehlman, 2006). This approach has been used to modify the seed oil content and composition, and to develop

different *Brassica* oils which meet the diverse industry needs (Burton et al., 2004; Scarth and Tang, 2006). The changes in fatty acid composition have been the result of both natural and induced genetic mutations (Przybylski et al., 2005; Sleper and Poehlman, 2006).

There are numerous examples of different canola oils that have been developed by applying conventional breeding techniques. Low erucic acid (C22:1) oils were established for their nutritional properties suitable for human consumption (Przybylski et al., 2005). High erucic acid (C22:1) rapeseed oils were developed for the production of lubricants and solvents (McVetty and Scarth, 2002; Scarth and Tang, 2006). *Brassica* oils containing low levels of linolenic acid (C18:3), were established to prevent oxidation and improve oil stability (McVetty and Scarth, 2002; Scarth and Tang, 2006). High oleic acid (C18:1) and low linolenic acid (C18:3) oils were developed in order to retain fatty acid structure stability upon exposure to extreme temperatures (Warner and Mounts, 1993; Topfer et al., 1995). Oils containing low levels of saturated fatty acids (C16:0) were developed to produce a healthy vegetable oil with less than 7% saturated fat (Burton et al., 2004; Scarth and Tang, 2006). Conventional breeding has been a very effective method for improving and altering the seed oil composition of *Brassica napus.*

2.4 Transgenic approaches to improve *Brassica* oils

Transgenic approaches differ from conventional breeding because genetic manipulation is not restricted to sexual recombination (Sleper and Poehlman, 2006). The insertion of transgenes into the *Brassica* genome is a technique used to improve seed oil content and composition. Transgenes have been constructed with genes isolated from

different plant species (Scarth and Tang, 2006). These genes usually encode for enzymes directly involved in the biosynthesis of fatty acids and triacylglycerols (Scarth and Tang, 2006). *Agrobacterium* mediated transformation is typically used to transform *Brassica* plants (Scarth and Tang, 2006).

The utilization of transformation techniques in *Brassica* has led to the development of several new oils that might not have been generated using traditional breeding methods (Scarth and Tang, 2006). Low levels of saturated fatty acids (C16:0) were developed by over-expressing *β-ketoacyl synthase II* (*KASII*) (Ohlrogge, 1994; Scarth and Tang, 2006). High lauric acid (C12:0) oil was established using *lysophosphatidic acid acyltransferase* (*LPAAT*) and *FatB* transgenes (Knutzon et al., 1999; Tang et al., 2004). Transgenes containing *FatB* were used to develop an oil containing high levels of caprylic (C8:0) and capric (C10:0) acid (Dehesh et al., 1996). High palmitic acid (C16:0) oils were achieved by over-expressing *FatB* (Voelker et al., 1997). Oil containing high levels of stearic acid (C18:0) was developed by over-expressing *acyl ACP thioesterase* (*FatA*), and suppressing *stearoyl-ACP desaturase* (Knutzon et al., 1992; Topfer et al., 1995; Scarth and Tang, 2006). Very high oleic acid (C18:1) oils were obtained by silencing *oleate desaturase* (*FAD2*) (Stoutjesdijk et al., 2000; Scarth and Tang, 2006). Similarly, transgenic plants expressing *LPAAT* were generated in an attempt to obtain a super high erucic acid oil (C22:1); however, this approach was not successful (Weier et al., 1997).

"Novel" fatty acids containing unique carbon lengths, or double bonds in distinct positions have become popular for a number of industrial applications (Scarth and Tang, 2006). As a result of limited germplasm, it is often difficult to apply conventional

breeding to produce the "novel" or "unusual" fatty acid of interest (Scarth and Tang, 2006). Most of the oils containing novel fatty acids have been developed by introducing transgenes which modify the location or orientation of the double bonds (Knutzon et al., 1999; Scarth and Tang, 2006). Examples of the novel fatty acids produced include: monoenoic acid, gamma-linolenic acid, very-long-chain polyunsaturated fatty acids, conjugated fatty acids, epoxy and hydroxyl fatty acids (Scarth and Tang, 2006).

2.5 Current challenges in improving seed oil content and composition

Both conventional breeding and transgenic approaches have their own set of challenges (Scarth and Tang, 2006). One of the drawbacks of conventional breeding is the restriction of available genetic resources (Sleper and Poehlman, 2006). Even when the desired germplasm is available, agronomic characteristics are often another major limiting factor (Scarth and Tang, 2006; Sleper and Poehlman, 2006). Conventional breeding programs are also restricted to naturally occurring or chemically induced mutations (Sleper and Poehlman, 2006).

The success of the transgenic approach is influenced by several key factors, including the number of transgenes inserted into the genome, the location of the insertion, and the stability of the trait (Scarth and Tang, 2006). The efficiency and success of fatty acid modifications are also dependent upon the gene selected to construct the transgene. Previous transformations have been effective in modifying the types of fatty acids produced (Dehesh et al., 1996; Voelker et al., 1997; Tang et al., 2004; Sharma et al., 2008). However, research points to the existence of a threshold regarding the effects that

these transgenes have upon oil accumulation during seed development (Thelen and Ohlrogge, 2002; Weselake et al., 2009).

2.6 Seed development

Understanding the process of seed development is crucial for improving oil content and composition in canola. There are two phases involved in seed development, namely embryo growth and maturation (Braybrook and Harada, 2008). Zygotic embryogenesis is the process in which a single cell undergoes differentiation to become a mature embryo (Braybrook and Harada, 2008). The accumulation of seed storage proteins and lipids occur during the maturation phase of embryo development (Braybrook and Harada, 2008).

The first phase of embryo development is marked by the double fertilization resulting in the production of the diploid zygote and triploid endosperm tissue (Harada, 1999). The zygote divides asymmetrically and develops apical and basal cells of the embryo (Harada, 1999). Apical cells produce the embryo proper while the basal cells give rise to the hypophysis and suspensor. Cell division continues until the embryo forms cotyledons, a shoot apical meristem, a root apical meristem, and an axis (Harada, 1999).

The second phase of embryogenesis is demarked by the cessation of embryo morphogenesis and the initiation of storage product accumulation (Harada, 1999). During this phase the embryo develops tolerance to desiccation, and precocious germination is inhibited (Harada, 1999). Starch, proteins and lipids accumulate within the seed during the process of embryo maturation (Harada, 1999; Braybrook and Harada,

2008). These macromolecules play an essential role in ensuring that the embryo survives the dormant state while maintaining an energy supply. These conditions allow the embryo to germinate under the appropriate environment (Harada, 1999; Braybrook and Harada, 2008).

2.7 Fatty acid biosynthesis

The process of embryogenesis is accompanied by the synthesis and accumulation of seed storage macromolecules which include proteins, starch, and lipids (Harada, 1999; Braybrook and Harada, 2008). Fatty acids and lipids are the storage macromolecules of interest when trying to improve the seed oil content and composition of canola. The process of fatty acid biosynthesis during seed development relies upon a number of organelles and metabolic pathways within the cell (Baud et al., 2002; Li-Beisson et al., 2010). The fatty acids found within seeds are predominantly stored in the form of triacylgylcerols or TAGs (Ohlrogge and Browse, 1995). TAGs are produced from the esterification of three fatty acids to a glycerol backbone (Ohlrogge and Browse, 1995). TAG biosynthesis during seed development utilizes sucrose imported from maternal tissue (Schwender et al., 2003). Sucrose synthetases and invertases cleave sucrose to generate a pool of hexose phosphates (Schwender et al., 2003). Hexose phosphates serve as the main energy source for both glycolysis and oxidative pentose phosphate pathway (OPPP) (Baud et al., 2007). The glycolytic pathway converts sucrose into pyruvate, and the OPPP produces energy in the form of nicotinamide adenine dinucleotide phosphate (Schwender et al., 2003; Baud et al., 2007). Glycolysis and OPPP are the main pathways providing a carbon source from which fatty acids are (Ohlrogge and Browse, 1995; Baud et al., 2007; Baud and Lepiniec, 2010).

2.7.1 *De novo* fatty acid biosynthesis

De novo fatty acid biosynthesis is mediated by fatty acid synthase (FAS) which is a multi-enzyme complex catalyzing a number of successive reactions, summarized in Figure 2.1 (Baud et al., 2008; Mu et al., 2008; Li-Beisson et al., 2010). Fatty acid biosynthesis begins in the plastids where acetyl-CoA is used by the enzyme acetyl-CoA carboxylase (ACCase) to produce malonyl-CoA (Li-Beisson et al., 2010). ACCase is thought to be the major rate limiting step of the *de novo* fatty acid biosynthesis because it controls the carbon flux entering into the pathway (Baud et al., 2008). The malonyl-CoA provides the carbon source required for the elongation of acetyl-CoA (Ohlrogge and Browse, 1995; Baud et al., 2008). In order for malonyl-CoA to enter the fatty acid biosynthetic pathway, the malonyl moiety has to be transferred to an acyl carrier protein (ACP) or protein cofactor (Ohlrogge, 1995; Baud et al., 2008; Li-Beisson et al., 2010). The enzyme malonyl-CoA: acyl carrier protein malonyltransferase catalyzes the malonyl transfer (Baud et al., 2008). Acetyl-CoA is considered the initial substrate of the fatty acid synthase (FAS) and undergoes elongation as a result of the activity of malonyl-ACP (Somerville et al., 2000; Baud and Lepiniec, 2010). Multiple condensation reactions occur between the acetyl-CoA and malonyl-ACP moieties (Baud et al., 2008; Li-Beisson et al., 2010). The first reaction is catalyzed by 3-ketoacyl-ACP synthase (KAS III) and this leads to the production of 4:0-ACP (Ohlrogge, 1995; Baud et al., 2008). The second condensation reaction is catalyzed by KAS I which produces 16:0-ACP (Ohlrogge, 1995; Pidkowich et al., 2007; Baud et al., 2008). The final condensation reaction is catalyzed by KAS II and produces 18:0-ACP (Ohlrogge, 1995; Pidkowich et al., 2007; Baud et al., 2008). Following this reaction, the 18:0-ACP can be desaturated to produce 18:1-ACP.

This final desaturation step is catalyzed by stearoyl-ACP desaturase (Browse and Somerville, 1991).

As a result of these reactions, 16:0-ACP and 18:1-ACP are the predominant products derived from the fatty acid synthase complex (Li-Beisson et al., 2010). The long-chain acyl groups of 16:0-ACP and 18:1-ACP undergo hydrolysis catalyzed by the acyl-ACP thioesterase (Fat) enzyme, which releases free fatty acids into the cytoplasm (Somerville et al., 2000; Baud et al., 2008; Li-Beisson et al., 2010). FatA typically hydrolyzes 18:1-ACP whereas FatB hydrolyzes 16:0-ACP (Bonaventure et al., 2003; Li-Beisson et al., 2010). Acyl-CoA synthetase which converts the free fatty acids into coenzyme A esters (Baud et al., 2008; Li-Beisson et al., 2010). These esters are exported from the cytoplasm to the endoplasmic reticulum for further modifications, including elongation and desaturation (Baud et al., 2008; Li-Beisson et al., 2010).

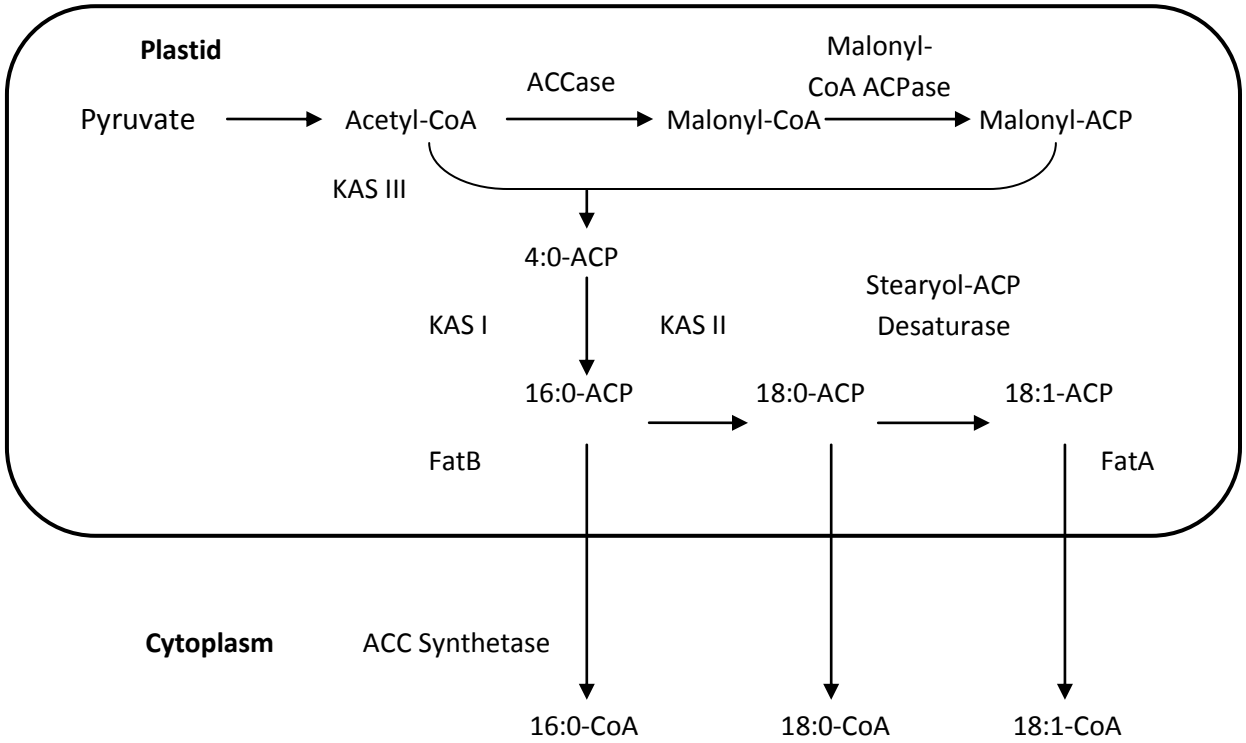

Figure 2.1. Biochemical pathway leading to the synthesis of fatty acids in the plastid. Glycolysis produces pyruvate which is used to generate acetyl-CoA. Acetyl-CoA is converted to malonyl-CoA by the enzyme ACCase (Acetyl-CoA carboxylase). Malonyl-CoA is modified to malonyl-ACP by the enzyme malony-CoA ACPase (Malonyl CoA acyl carrier protein malonyltransferase). The acetyl-CoA is elongated to 4:0-ACP (acyl carrier protein) by KAS III (3-ketoacyl-ACP synthase isoform III). KAS I converts 4:0-ACP to 16:0-ACP, and KAS II elongates 16:0-ACP to 18:0-ACP. Stearyol-ACP Desaturase produces 18:1-ACP from 18:0-ACP. FatA and FatB (acyl ACP thioesterase) free 18:1-ACP and 16:0-ACP respectively from the attached ACP group. The enzyme ACC synthetase (Acyl-CoA synthetase) converts the free fatty acids to 16:0-CoA, 18:0-CoA, and 18:1-CoA (Coenzyme A) esters, which are then exported to the cytoplasm. Adapted from Baud et al. (2008), Weselake et al. (2009) and Li-Beisson et al. (2010).

2.7.2 Further modifications of fatty acids: elongation and desaturation

16:0-CoA and 18:1-CoA can undergo fatty acid elongation and desaturation in the endoplasmic reticulum to produce fatty acid structures of varying composition (Somerville et al., 2000; Li-Beisson et al., 2010). Fatty acid elongation requires several steps to synthesize very long chain fatty acids (VLCFA) containing 20 to 24 carbons (Baud et al., 2008; Li-Beisson et al., 2010). The fatty acid elongase is a complex that is bound to the inner membrane of the endoplasmic reticulum and houses multiple enzymes required to carry out the modifications to the fatty acid groups (Somerville et al., 2000; Li-Beisson et al., 2010). Malonyl-CoA and 18:1-CoA are the initial substrates in fatty acid elongation, which is a multi-step process involving four reactions (Figure 2.2) (Ohlrogge and Browse, 1995; Somerville et al., 2000; Baud et al., 2008). The first reaction requires 3-ketoacyl-CoA synthase (KCS condensing enzyme) for the condensation of acyl-CoA with malonyl-CoA to produce 3-ketoacyl-CoA (von Wettstein-Knowles, 1982; Li-Beisson et al., 2010). This initial reaction is presumed to be the major control point in the elongation process of fatty acid biosynthesis (Baud et al., 2008). The second step is a reduction that requires 3-ketoacyl reductase (KR) and produces β-hydroxyacyl-CoA (von Wettstein-Knowles, 1982; Li-Beisson et al., 2010). In the third phase enyol-CoA is produced through a dehydration step requiring 3-hydroxyacyl-CoA-dehydrase (DH) (von Wettstein-Knowles, 1982; Li-Beisson et al., 2010). Finally, the enzyme enoyl-CoA reductase (ER) reduces enoyl-CoA to acyl-CoA (von Wettstein-Knowles, 1982; Li-Beisson et al., 2010). This series of reactions results in the production of enlogated fatty acids, which enter the acyl-CoA pool to be used in subsequent steps.

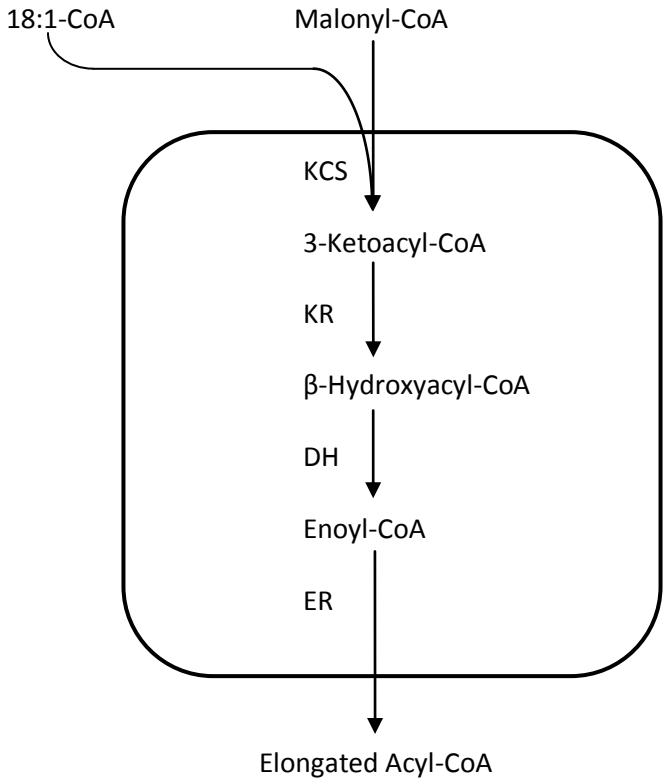

Figure 2.2. Biochemical pathway leading to the production of very long chain fatty acids in the endoplasmic reticulum. Malonyl-CoA and 18:1-CoA undergo a condensation reaction facilitated by KCS (2-ketoacyl CoA synthase) to produce 3-ketoacyl-CoA. KR (3-ketoacyl reductase) is required for the production of β-hydroxyacyl-CoA which is then converted to enoyl-CoA by DH (3-hydroxyacyl-CoA dehydrase). Enoyl-CoA produces an elongated acyl-CoA in a reaction requiring the enzyme ER (enoyl-CoA reductase). Adapted from Baud et al. (2008) and Li-Beisson et al. (2010).

Fatty acid desaturation is another modification occurring in the PC (phosphatidylcholine) pool located within the endoplasmic reticulum (Figure 2.3) (Li-Beisson et al., 2010). Linoleic (18:2) and linolenic (18:3) acids are the prominent unsaturated fatty acids in *Arabidopsis* seeds (Ohlrogge and Browse, 1995; Baud et al., 2008). Both linoleic (18:2) and linolenic (18:3) acids are formed as a result of the desaturation of oleic acid (18:1) (Baud et al., 2008). The two key enzymes responsible for the enlargement of the PC pool are cholinephosphotransferase (CPT), which uses

DAG as a substrate, and lysophosphatidylcholine acyltransferase (LPCAT) which uses fatty acids from acyl-CoA (Li-Beisson et al., 2010). The enzyme oleate desaturase (FAD2) catalyzes the reaction in which 18:1-PC is desaturated to 18:2-PC (Somerville et al., 2000; Baud et al., 2008). 18:2-PC is then desaturated to 18:3-PC by linoleate desaturase (FAD3) (Somerville et al., 2000; Baud et al., 2008).

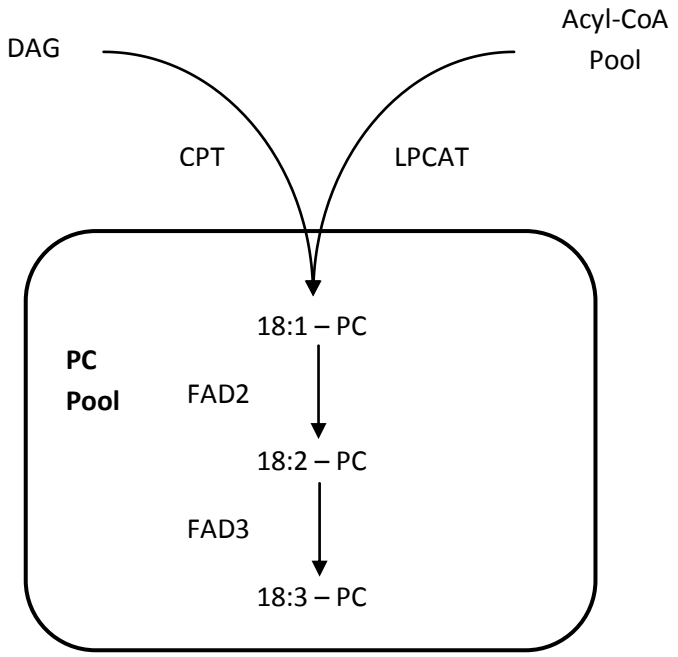

Figure 2.3. Fatty acid desaturation in the endoplasmic reticulum of *Arabidopsis* to produce linoleic and linolenic acids from oleic acid. The enzyme CPT (cholinephosphotransferase) is required for DAG (diacylglycerol) to enter into the PC (phosphatidylcholine) pool. Fatty acids from the acyl-CoA pool are converted by LPCAT (lysophosphatidylcholine) to be used in desaturation reactions. FAD2 (oleate desaturase) converts 18:1-PC to 18:2-PC which undergoes another desaturation step to produce 18:3-PC. This later reaction is catalyzed by FAD3 (linoleate desaturase). Adapted from Baud et al. (2008), and Li-Beisson et al. (2010).

2.7.3 TAG synthesis

TAG production also occurs within the endoplasmic reticulum through the Kennedy pathway (Figure 2.4) (Baud et al., 2008). In the Kennedy pathway, dihydroxyacetonephosphate (DHAP) is reduced to form glycerol-3-phosphate (G3P) in a reaction catalyzed by glycerol-3-phosphate acyltransferase (G3PDH) (Baud et al., 2008; Li-Beisson et al., 2010). G3P is converted to lysophosphatidic acid (LPA) by the enzyme glycerol-3-phosphate acyl transferase (G3PAT) (Ohlrogge and Browse, 1995; Baud et al., 2008). An acyl group is added to LPA by lysophosphatidic acid acyltransferase (LPAAT), resulting in the production of phosphatidic acid (PA) (Ohlrogge and Browse, 1995; Baud et al., 2008; Li Beisson et al., 2010). Phosphatidic acid phosphatase (PAP) dephosphorylates PA to produce diacylglycerol (DAG) (Baud et al., 2008; Weselake et al., 2009; Li-Beisson et al., 2010). An acyl group is added to DAG to produce TAG in a reaction catalyzed by the enzyme diacylglycerol acyltransferase (DGAT) (Ohlrogge and Browse, 1995; Baud et al., 2008; Li-Beisson et al., 2010).

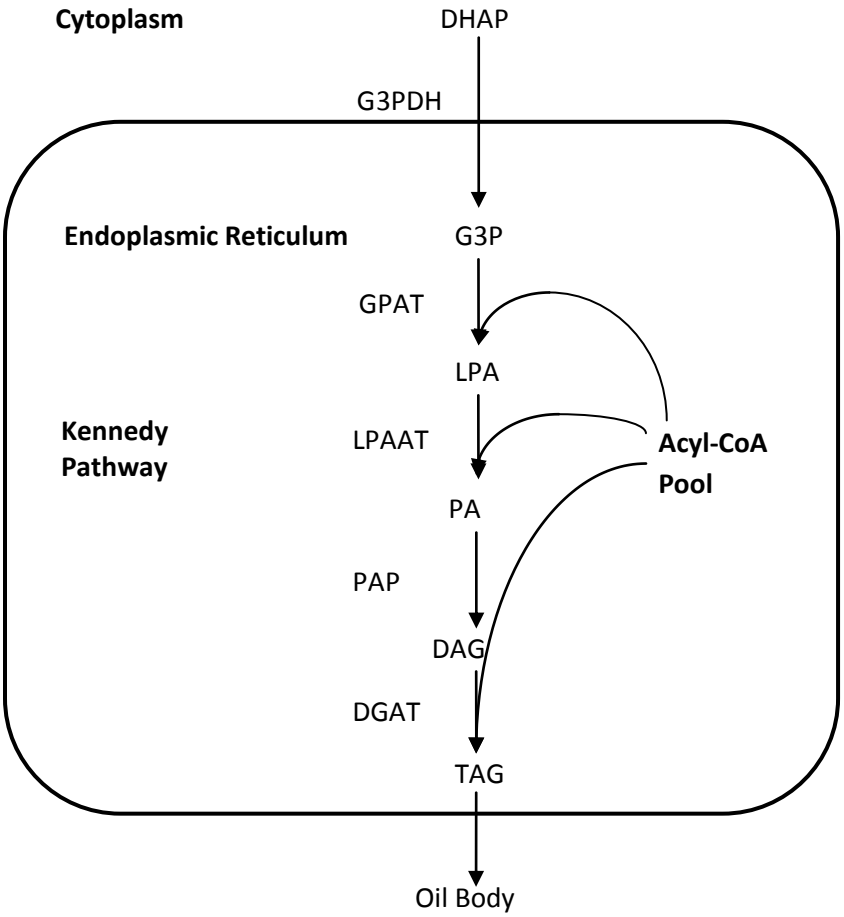

Figure 2.4. Biochemical pathway leading to the synthesis of TAG (triacylglycerol) in the endoplasmic reticulum. Glycolysis produces DHAP (dihydroxyacetone phosphate) which is converted to G3P (glycerol-3-phosphate) by G3PDH (glycerol-3-phosphate acyltransferase). The enzyme GPAT (glycerol-3-phosphate acyltransferase) is required to produce LPA (lysophosphatidic acid) from G3P. LPAAT (lysophosphatidic acid acyltransferase) catalyzes the reaction that produces PA (phospatidic acid) from LPA. PA is converted to DAG (diacylglycerol) by the enzyme PAP (phosphatidic acid phosphatase). TAGs (triacylglycerols) are produced from DAG in a reaction catalyzed by the enzyme DGAT (diacylglycerol acyltransferase). TAGs are stored as oil bodies which are exported from the endoplasmic reticulum. Adapted from Baud et al. (2008), Weselake et al. (2009) and Li-Beisson et al. (2010).

There are a number of possible routes allowing fatty acids to enter the pathway leading to TAG synthesis (Figure 2.5) (Ohlrogge and Browse, 1995; Weselake et al., 2009). Rather than acyl chains being directly available for the Kennedy pathway, some fatty acids undergo the process of desaturation (Browse and Somerville, 1991; Baud et al., 2008). The enzyme CPT is required for DAG to become incorporated into the PC pool where the remaining acyl groups are desaturated by the FAD enzymes (Li-Beisson et al., 2010). CPT catalyzes a reversible reaction where fatty acids from the PC pool can be used to produce DAG (Browse and Somerville, 1991; Baud et al., 2008). Fatty acids from the PC pool can also donate an acyl group to DAG to produce TAG, in a reaction catalyzed by phospholipid: diacylglycerol acyltransferase (PDAT) (Li-Beisson et al., 2010). The reaction catalyzed by PDAT also produces a second substrate, lysophosphatidylcholine (LPC) (Li-Beisson et al., 2010). LPC can also be produced through the hydrolysis of a fatty acid attached to PC in a reaction catalyzed by phospholipase (PLA_2) (Li-Beisson et al., 2010). LPC can re-enter the process of fatty acid desaturation in a reaction facilitated by lysophosphatidylcholine (LPCAT) (Baud et al., 2008; Weselake et al., 2009; Li-Beisson et al., 2010). LPCAT catalyzes the reaction that allows fatty acids from the acyl-CoA pool to enter into the PC pool (Weselake et al., 2009; Li-Beisson et al., 2010).

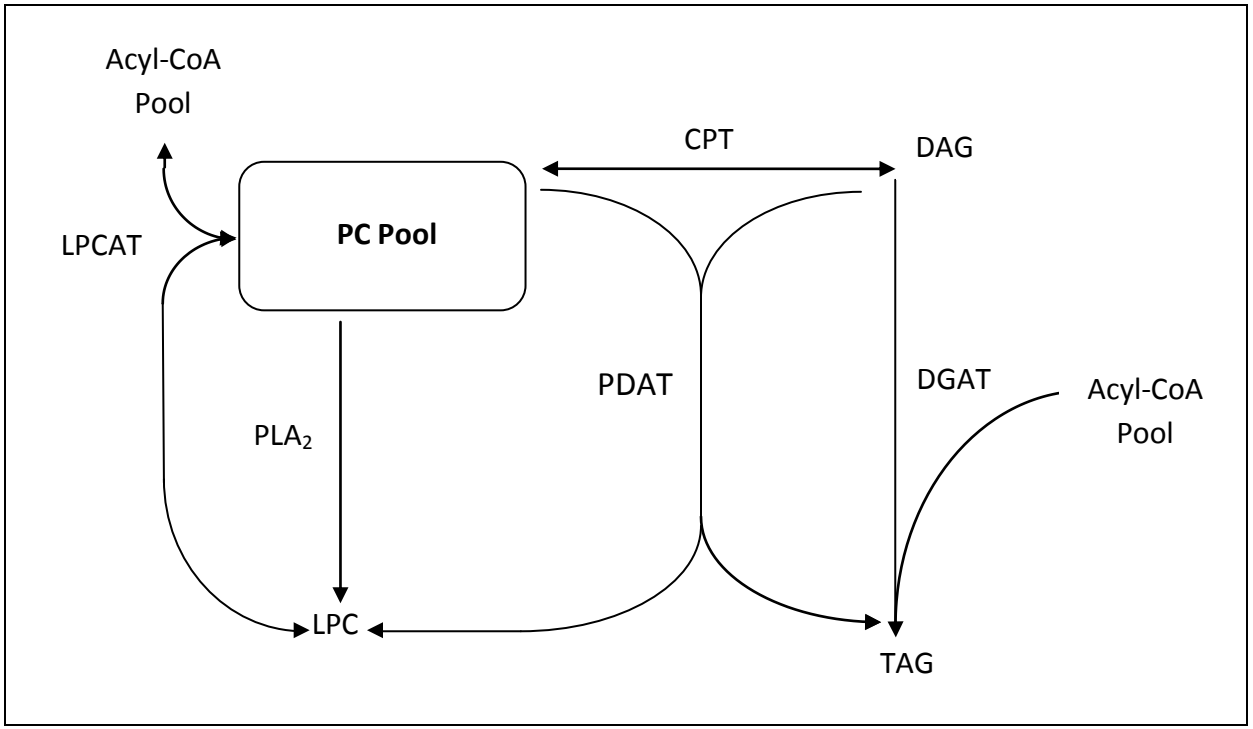

Figure 2.5. Summary of alternative pathways leading to TAG synthesis within the endoplasmic reticulum. DGAT (diacylglycerol acyltransferase) catalyzes the conversion of DAG (diacylglycerol) to produce TAG (triacylglycerols) in the Kennedy pathway. DAG can enter into the PC (phosphatidylcholine) pool through a reaction requiring the enzyme CPT (cholinephosphotransferase). Fatty acids from the PC pool can be converted to DAG by the enzyme CPT in a reversible reaction. Desaturated fatty acids can be directly utilized with DAG to produce TAG and LPC (lysophosphatidylcholine) in a reaction facilitated by PDAT (phospholipid: diacylglycerol acyltransferase). PLA_2 (phospholipase) also contributes to the production of LPC by hydrolyzing an acyl group from PC. LPC can undergo another desaturation step catalyzed by LPCAT (lysophosphatidylcholine acyltransferase). The reaction that allows fatty acids from the acyl-CoA pool to enter into the PC pool requires the enzyme LPCAT. Adapted from Baud et al. (2008), Weselake et al. (2009) and Li-Beisson et al. (2010).

2.7.4 TAGs are stored in oil bodies

Seed storage lipids are synthesized as TAGs and stored in oil bodies (Somerville et al., 2000; Li-Beisson et al., 2010). Oil bodies are specialized organelles consisting of a TAG core enclosed by a phospholipid monolayer of structural proteins predominantly composed of oleosins (Yatsu and Jacks, 1972; Ohlrogge and Browse, 1995; Baud et al., 2008; Baud and Lepiniec, 2010; Li-Beisson et al., 2010). Oil bodies are assembled with the phosphate groups of the membrane facing the cytosol and the polar tails oriented towards the center or the lumen where the TAGs accumulate (Yatsu and Jacks, 1972; Ohlrogge and Browse, 1995; Baud et al., 2008, Baud and Lepiniec, 2009).

The precise process of oil body production is unknown (Baud et al., 2008; Baud and Lepiniec, 2009). It is thought that TAGs congregate together within the lipid bilayer that is initially located within microdomains of the endoplasmic reticulum (Murphy and Vance, 1999; Baud et al., 2008). The accumulated TAGs are pinched off and surrounded by phospholipids containing proteins produced within the endoplasmic reticulum (Murphy and Vance, 1999; Baud et al., 2008; Li-Beisson et al., 2010). The proteins develop a surrounding coat around the TAGs, resulting in the production of an oil body (Murphy and Vance, 1999). Robenek et al. (2006) demonstrated an alternative method in which oil bodies develop from the endoplasmic reticulum. This study suggests that an adipophilin layer already present in the membrane of the endoplasmic reticulum transfers the produced phospholipids to the surface of the developing TAGs (Robenek et al., 2006; Baud et al., 2008).

2.7.5 TAG synthesis is a complex process

Figure 2.6 summarizes the numerous components and pathways involved in TAG synthesis, which vary depending on the fatty acid produced (Ohlrogge and Browse, 1995). The interaction among the many proteins and enzymes involved in the regulation of fatty acid synthesis is critical for improving and modifying the level and profile of oil in seeds. The pathway of fatty acid biosynthesis is well known and a lot of research has been conducted on many of the genes encoding for various enzymes and proteins directly participating in the different parts of fatty acid and TAG biosynthesis (Sharma et al., 2008; Baud and Lepiniec, 2009). However, despite this effort, the regulatory mechanisms of the pathway are still poorly understood (Thelen and Ohlrogge, 2002). Multiple studies have analyzed the differential expression of genes involved in the regulation of seed development (Mu et al., 2008; Liu et al., 2009; Le et al., 2010), and great interest has developed on a series of embryonic transcription factors, which may affect fatty acid biosynthesis (Mu et al., 2008; Liu et al., 2009).

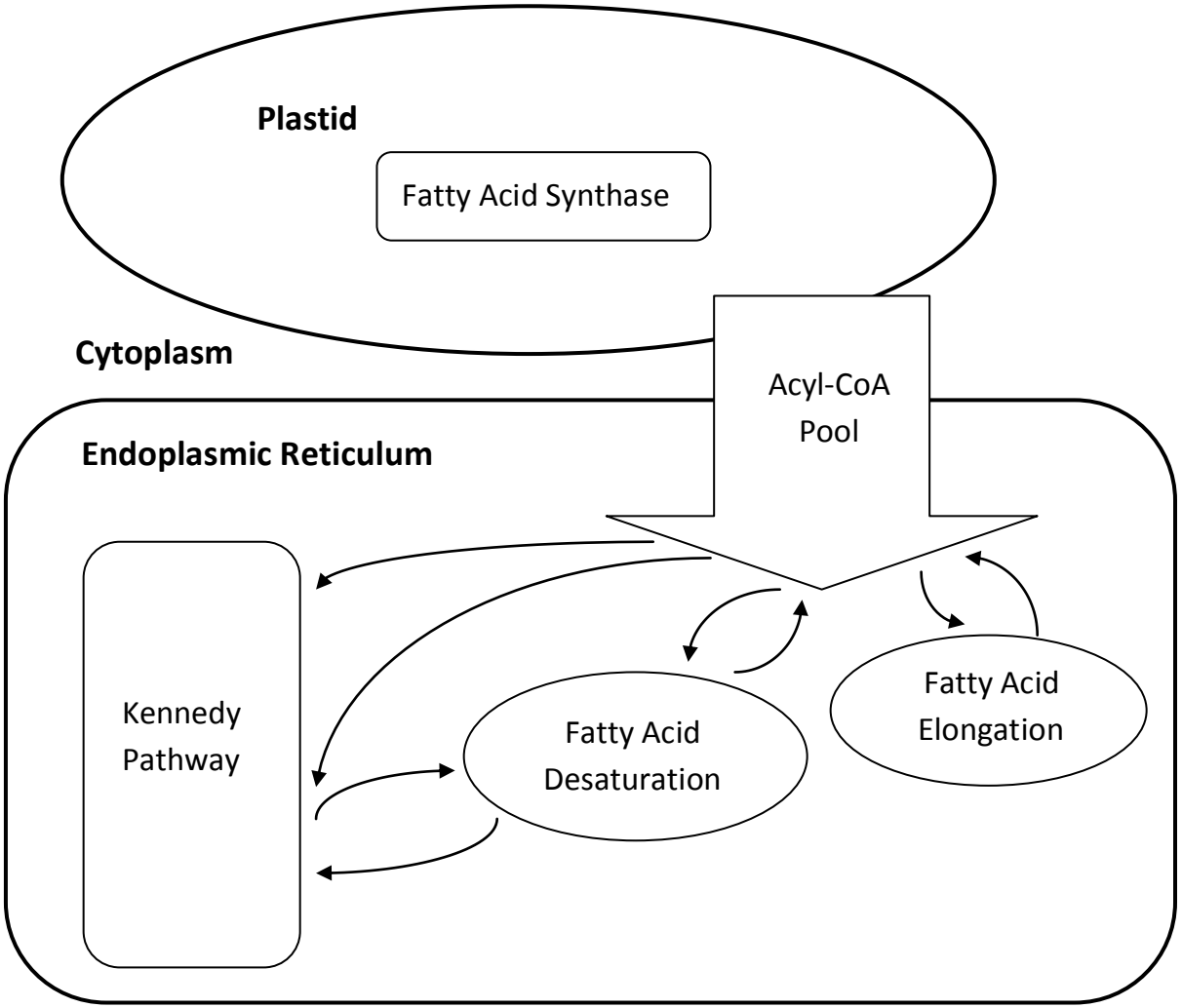

Figure 2.6. Summary of the major biochemical events leadings to fatty acid biosynthesis in cellular compartments. Fatty acid biosynthesis begins in the plastid with the production of an acyl-CoA pool utilized in subsequent reactions. Further reactions including fatty acid elongation and destauration occur in the endoplasmic reticulum. Through different routes fatty acids can enter the Kennedy pathway to produce TAGs. Adapted from Baud et al. (2008), Weselake et al. (2009), and Li-Beisson et al. (2010)

2.8 Regulatory network controlling embryogenesis

The synthesis and accumulation of fatty acids during seed development is associated with the physiological process of embryogenesis (Mu et al., 2008). A network of transcription factors regulates the process of embryo morphogenesis and maturation in *Arabidopsis*. These transcription factors include LEAFY COTYLEDON1 (LEC1), LEAFY COTYLEDON2 (LEC2), FUSCA3 (FUS3), WRINKLED1 (WRI1) and ABSCISIC ACID INSENSITIVE3 (ABI3) (Meinke, 1992; Meinke et al., 1994; Parcy et al., 1994; Stone et al., 2001; Cernac and Benning, 2004; Gazzarrini et al., 2004).

The interaction among LEC1, LEC2, WRI1, FUS3 and ABI3 has not been concretely determined, however numerous studies have generated and analyzed transgenic plants to determine the role of each gene (Meinke, 1992; Meinke et al., 1994; Parcy et al., 1994; Stone et al., 2001; Cernac and Benning, 2004; Gazzarrini et al., 2004). The transgenic plants typically over-express or down-regulate the gene of interest (Meinke, 1992; Meinke et al., 1994; Parcy et al., 1994; Stone et al., 2001; Cernac and Benning, 2004; Gazzarrini et al., 2004). Current genetic studies have established the proposed regulatory network based on changes in the expression level of the genes. However, the direct molecular interactions among the noted transcription factors still remain to be determined (Wang et al., 2007).

A proposed interaction among the embryonic transcription factors is summarized in Figure 2.7. LEC1 has been shown to positively regulate the transcription of *LEC2*, *FUS3*, *ABI3*, and *WRI1*. WRI1 regulates the expression of a number of genes involved in glycolysis (Santos-Mendoza et al., 2008). *LEC1* influences the expression of *LEC2*, an

upstream regulator of *ABI3* and *FUS3* (Kagaya et al., 2005b; Santos-Mendoza et al., 2008). ABI3 and FUS3 influence the production and accumulation of storage proteins during seed development (Kagaya et al., 2005b; Santos-Mendoza et al., 2008). FUS3 also influences the accumulation of fatty acids (Kagaya et al., 2005b; Santos-Mendoza et al., 2008). The interaction and regulatory functions fulfilled by LEC1, LEC2, WRI1, FUS3 and ABI3 are important for embryo and seed development (Meinke, 1992; Meinke et al., 1994; Parcy et al., 1994; Stone et al., 2001; Cernac and Benning, 2004; Gazzarrini et al., 2004). While some of the genes appear to be partially redundant in function they all play an essential role in this regulatory system.

Multiple aspects of the noted transcription factors (Figure 2.7) have been characterized during embryogenesis (Baud et al., 2008; Santos-Mendoza et al., 2008; Weselake et al., 2009). This has led to recent research focusing on the role of these factors in the biosynthesis and accumulation of fatty acids and proteins during seed development (Mu et al., 2008; Liu et al., 2009).

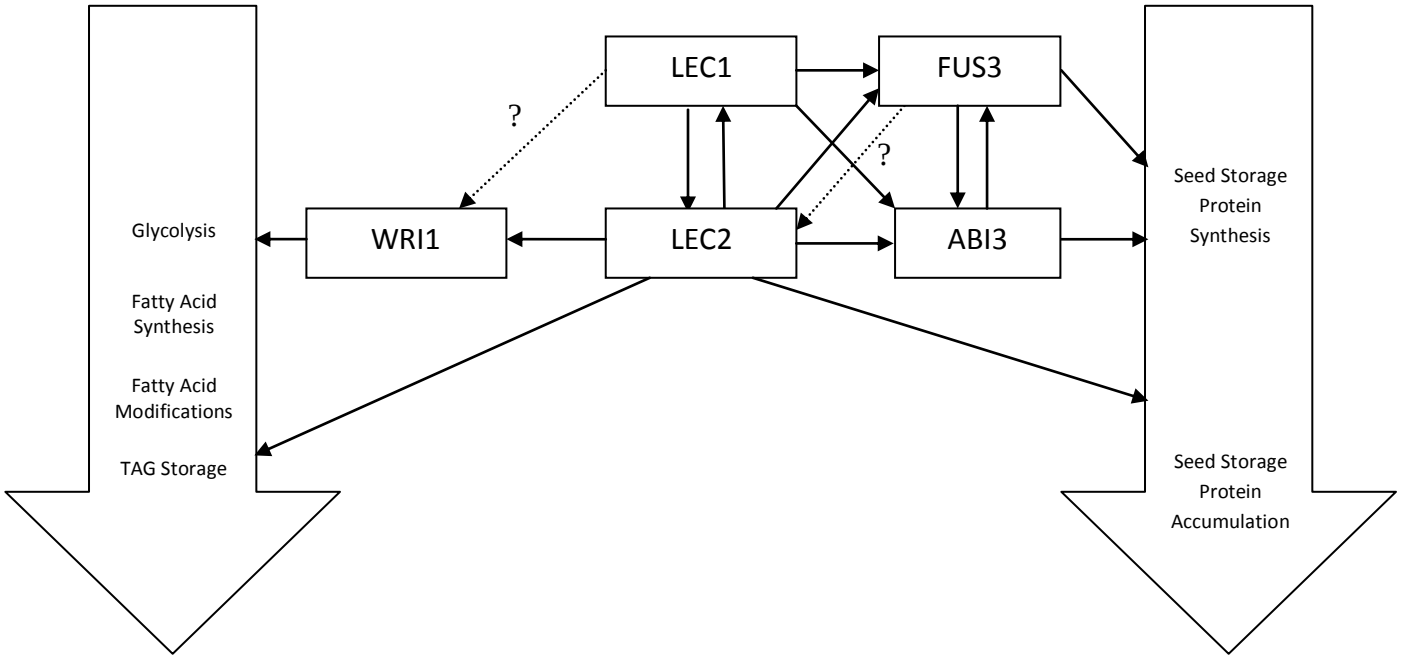

Figure 2.7. Proposed linkages and regulatory interactions occurring among the various transcription factors during embryo and seed development. LEC1 activates the expression of *LEC2*, which induces the transcription of *WRI1*. WRI1 is believed to alter the expression of genes involved in the late stages of glycolysis, which influence the synthesis of fatty acids and TAGs (triacylglycerols). LEC2 also influences the expression of *FUS3* and *ABI3*, which are involved in the synthesis, and accumulation of seed storage proteins. Adapted from Baud et al. (2008), Santos-Mendoza et al. (2008), and North et al. (2010).

2.8.1 LEAFY COTYLEDON1

LEAFY COTYLEDON1 (LEC1) is expressed throughout the entire process of embryogenesis from the very early to the later developmental stages (West et al., 1994). The LEC1 protein contains the HAP3 subunit of the CCAAT binding factor also known as the CBF (Lotan et al., 1998). The CBF allows LEC1 to be a specific transcriptional regulator that activates the expression of downstream genes containing the CCAAT recognition domain and involved in embryo development (Lotan et al., 1998).

The development and phenotypic analysis of transgenic plants containing mutations at the *LEC1* loci have been used to characterize the *LEC1* gene. *Arabidopsis* plants with a null *lec1* allele produced abnormal embryos characterized by a reduced hypocotyl and cotyledons that remained green on the plant until the later stages of maturation (Meinke et al., 1994). Other phenotypes of the *lec1* mutant seedlings included the development of trichomes on the adaxial surface of cotyledons and a decreased accumulation of proteins and lipids during seed development (Meinke et al., 1994). Compared to the wild type seeds, the seeds of the *lec1* plants are noticeably deformed and completely intolerant to desiccation (Meinke et al., 1994). Because of this intolerance, embryo rescue must be performed as a means of propagating the plants from one generation to the next (Meinke et al., 1994).

Available literature indicates that ectopic expression of *LEC1* is sufficient to initiate somatic embryogenesis even in an environment lacking auxin treatment (Lotan et al., 1998; Braybrook and Harada, 2008). Somatic embryogenesis is the process in which bipolar embryonic structures develop from either somatic or vegetative tissue without the fusion of gametes (Braybrook and Harada, 2008). The formation of somatic embryos

suggests that LEC1 regulates embryogenic competence within cells (Lotan et al., 1998; Stone et al., 2001). Stone et al. (2001) showed that LEC1 is responsible for promoting a cellular environment that induces embryonic growth by coordinating morphogenesis and maturation.

A recent study demonstrated that overexpression of *LEC1* in *Arabidopsis* not only induced the expression of oleosin and seed storage protein genes but also resulted in increased expression levels of other genes involved in fatty acid biosynthesis (Mu et al., 2008). In the same study, it was demonstrated that the over-expression of *BnLEC1* in *Arabidopsis* produced similar effects on oil biosynthesis as the over-expression of *AtLEC1*. Further microarray analysis showed that the overexpression of *LEC1* results in a concurrent over-expression of 58% of the genes encoding enzymes involved in the plastidal fatty acid synthetic pathway (Mu et al., 2008).

The phenotypes generated from over expressing or silencing *LEC1* have revealed the significant role that this transcription factor plays in the induction of numerous genes involved in embryo development (Meinke et al., 1994). Because of this regulatory property, it is important to understand the role of other genes participating in both embryo development and seed oil synthesis and accumulation.

2.8.2 LEAFY COTYLEDON2

LEAFY COTYLEDON2 (LEC2) is another major transcription factor participating in the genetic network controlling the process of embryo development. LEC2, FUS3, and ABI3 belong to the B3 transcription factor family, whose members contain a DNA binding region unique to plants (Stone et al., 2001; Baud et al., 2007).

The B3 domain is conserved in many genes such as *ABSCISIC ACID INSENSITIVE3 /VIVIPAROUS1,* and *AUXIN RESPONSE FACTOR1* (Reidt et al., 2001; Stone et al., 2001). The conserved structure of the B3 binding domain allows LEC2 to regulate the expression of downstream genes possessing the specific RY recognition site (Santos-Mendoza et al., 2008; Stone et al., 2008).

Transgenic and mutant analyses have revealed the function of *LEC2* during embryogenesis. *lec2* seedlings were characterized by an elongated hypocotyl and the accumulation of anthocyanin in the cotyledons (Meinke et al., 1994). Unlike *lec1* mutants, some *lec2* seeds are tolerant to desiccation and successfully germinate (Meinke et al., 1994). Interestingly, these seedlings develop trichomes on the adaxial side of the cotyledons but only when the immature embryos are germinated prematurely, whereas the seeds that are completely dried loose this phenotype (Meinke et al., 1994).

As observed for *LEC1*, the ectopic expression of *LEC2* has been found to induce genes required for the synthesis of auxin, which promotes somatic embryogenesis (Stone et al., 2008). LEC2 is thought to induce somatic embryogenesis in vegetative tissues by two mechanisms. The first involves an increase in the acquisition of embryonic competence through the activation of genes such as *ABI3* and *FUS3* (Stone et al., 2008). The second mechanism is exercised through the production of auxin which promotes the formation of competent cells (Stone et al., 2008). The over-expression of *LEC2* can induce unfertilized ovule integuments to accumulate lipid and protein reserves, a trait which is typically a characteristic of developing seeds (Stone et al., 2008). Stone et al. (2008) showed that the ectopic expression of *LEC2* activates *YUCCA* genes which are involved in the biosynthesis of auxin, the hormone commonly used to induce somatic

embryogenesis. The over expression of *LEC1* and *LEC2* are thought to establish an environment that increases the response of somatic cells to auxin and induces somatic embryogenesis (Braybrook and Harada, 2008).

Mendoza et al. (2005) over expressed *LEC2* under the control of an inducible promoter and observed an increase of very long chain fatty acids and TAGs in leaf tissue (Santos Mendoza et al., 2005). Seed specific transcripts coding for oleosins, a sucrose synthase and seed storage proteins, began to accumulate in the leaves of these transgenic plants. These results suggest that modifications of leaf metabolism lead to the accumulation of seed storage compounds (Santos Mendoza et al., 2005). LEC2 is also thought to influence fatty acid biosynthesis because it contributes to the regulation of WRI1, the transcription factor which activates several glycolitic genes (Baud et al., 2009).

2.8.3 WRINKLED1

WRINKLED1 (WRI1) belongs to a large family of transcription factors encoding for APETALA2/ethylene-responsive element binding (AP2/EREBP) proteins (Riechmann et al., 2000; Liu et al., 2009). Members of this family share the AP2 domain, which is important for binding to the promoter region of target genes to activate their expression. The direct targets of WRI1 are not entirely known, however, elevated levels of *WRI1* affect the expression of enzymes involved in glycolysis such as pyruvate kinase (Baud et al., 2008; Santos-Mendoza et al., 2008).

wri1 mutant plants were characterized by a wrinkled seed coat and a 80% reduction in total seed oil content (Focks and Benning, 1998). The authors found that

suppression of *WRI1* negatively influenced the expression of glycolytic enzymes (Focks and Benning, 1998). As a result, the developing embryos lacked the ability to efficiently produce and covert sucrose into the initial metabolites required for fatty acid biosynthesis (Cernac and Benning, 2004).

Liu et al. (2009) constitutively expressed *BnWRI1* in *Arabidopsis* and found a 10% to 40% increase in seed oil content in some of the transgenic lines (Liu et al., 2009). Cernac and Benning (2004) also over-expressed *AtLEC1* and observed a 10% to 20% increase in total seed oil content and TAG production in developing *Arabidopsis* seedlings.

Expression of *WRI1* not only affects embryo development but also germination (Cernac et al., 2006). Cernac et al. (2006) observed that WRI1 influences the ability of developing embryos to sense sugar in an ABI3 independent manner, possibly by modulating the sensitivity of the embryos to abscisic acid. By regulating glycolysis at a transcriptional level, WRI1 might be important in providing sugar molecules to be used in fatty acid biosynthesis (Cernac and Benning, 2004; Cernac et al., 2006; Liu et al., 2009).

2.8.4 FUSCA3

FUSCA3 (FUS3) is responsible for inducing the maturation phase in developing seeds of *Arabidopsis thaliana*. FUS3 encodes a B3 protein that accumulates primarily during seed development. The FUS3 protein recognizes and binds to the RY element CATGCA which is a sequence found in the promoters of many genes (Curaba et al., 2004; Santos-Mendoza et al., 2008).

fus3 mutants display several phenotypes that include desiccation intolerance, occasional viviparity, precocious germination, anthocyanin accumulation in developing seeds, and defective hormone synthesis and perception of light (Meinke et al., 1994; Harada, 2001). *Arabidopsis fus3* plants also show decreased protein and lipid accumulation (Meinke et al., 1994; Harada, 2001). Overall, *fus3* mutants display similar phenotypes to the *lec1* and *lec2* plants, however cotyledon identity is not as greatly affected as in the *lec* mutants (Vicient et al., 2000).

Increased expression of *FUS3* has been correlated with the expression of genes encoding the two main types of seed storage proteins (2S albumin and 12S cruciferin), and a number of enzymes involved in fatty acid biosynthesis (KAS I, KASII, KASIII, pyruvate dehydrogenase, acetyl-CoA carboxylase) (Wang et al., 2007). Transgenic studies have showed that FUS3 affects fatty acid biosynthesis, but the direct molecular interactions involved in this regulation are not fully understood (Vicient et al., 2000; Curaba et al., 2004). The activation of *FUS3* has been shown to be positively regulated by the expression of *LEC2* and *ABI3* (To et al., 2006).

2.8.5 ABSCISIC ACID INSENSITIVE3

ABSCISIC ACID INSENSITIVE3 (ABI3) proteins are essential regulators of seed development in *Arabidopsis* (Vicient et al., 2000; Santos-Mendoza et al., 2008). The ABI3 protein also contains a B3 binding domain that closely interacts with *FUS3* and *LEC2* (Vicient et al., 2000; Santos-Mendoza et al., 2008).

ABI3 has been shown to influence genes involved in the synthesis of storage proteins during seed development. In null *abi3* mutants chlorophyll fails to degrade

during the final stages of embryo development, and this results in seeds which remain green (Parcy et al., 1994). These seeds are also insensitive to ABA during germination (Parcy et al., 1994). None of the *abi3* mutants produced leafy cotyledons with trichomes, or viviparous embryos that germinated precociously (Meinke et al., 1994).

Parcy et al. (1994) found that overexpression of *ABI3* increased transcripts of genes encoding for seed storage proreins. These results are similar to those observed in over-expressors of *FUS3* (Wang et al., 2007), and likely reflect the fact that FUS3 activates the expression of *ABI3* (Parcy et al., 1994)

2.9 CCAAT binding factor family

The members of the CCAAT binding factor (CBF) family contain three main subunits, HAP2 (CBF-B, NF-YA), HAP3 (CBF-A, NF-YB), and HAP5 (CBF-C, NF-YC) (Sinha et al., 1995; Kim et al., 1996; Sinha et al., 1996; Romier et al., 2003). The three subunits interact to form a complex that binds to the CCAAT DNA motif in the promoter region of target genes (Sinha et al., 1995; Kim et al., 1996; Sinha et al., 1996; Romier et al., 2003). *LEC1* contains the HAP3 subunit of the CBF which consists of three regions referred to as the A, B, and C domains (Lotan et al., 1998). The B domain contains an amino acid residue that is required for the interaction of the HAP3 subunit with the promoters of target genes (Sinha et al., 1995; Kim et al., 1996; Sinha et al., 1996; Lotan et al., 1998; Romier et al., 2003). There are two subclasses of HAP3 subunits, the LEC1-type and the non-LEC1-type (Yamamoto et al., 2009). The LEC1-type consists of LEC1 and LEC1-like (L1L) (Kwong et al., 2003; Lee et al., 2003). The independent evolution of both the LEC1 and LEC1-like proteins from a common origin is

consistent with the notion that they carry out different functions during embryogenesis (Kwong et al., 2003; Lee et al., 2003; Xie et al., 2008; Yamamoto et al., 2009).

2.9.1 LEC1-like (L1L)

The LEC1 and LEC1-like proteins both contain the HAP3 subunit, and a specific amino acid residue that distinguishes their function from that of other proteins containing the HAP3 subunit (Kwong et al., 2003; Lee et al., 2003; Yamamoto et al., 2009). Even though LEC1 and L1L may be partially redundant in function, they are believed to have distinct roles during embryogenesis (Kwong et al., 2003; Lee et al., 2003; Yamamoto et al., 2009).

L1L RNAs accumulate during late embryogenesis whereas *LEC1* RNAs accumulate early in development (Kwong et al., 2003; Lee et al., 2003; Yamamoto et al., 2009). *L1L* RNAs have also been detected in vegetative organs at low levels, whereas *LEC1* is not present in vegetative tissue (Kwong et al., 2003). Finally, *lec1* mutants are characterized by an abnormal phenotype when endogenous *L1L* RNAs are still detectable, suggesting that the *L1L* gene is not sufficient to rescue the lack of *LEC1* expression (Kwong et al., 2003). Unlike *LEC1*, mutations at the *L1L* loci did not appear to have any effect on vegetative tissue (Kwong et al., 2003). While there appear to be functional redundancies between LEC1 and L1L, both seem to be influential in regulating embryogenesis (Kwong et al., 2003; Lee et al., 2003; Mu et al., 2008; Yamamoto et al., 2009).

3.0 CHAPTER 1: CLONING AND CHARACTERIZATION OF THE *BRASSICA NAPUS LEAFY COTYLEDON1*

3.1 ABSTRACT

LEAFY COTYLEDON1 (LEC1) is a transcription factor which plays a key role during embryo development and oil biosynthesis in *Arabidopsis.* Due to these functions, this gene has been isolated and characterized in different plant species. In order to further understand the role played by this transcription factor in canola *(Brassica napus)*, two full length *BnLEC1* genes were isolated from microspore derived embryos. Molecular characterization of *Bn*LEC1 revealed that the protein contains a HAP3 subunit of a CCAAT binding factor and a histone fold motif. The conserved domains are important for the role played by LEC1 during embryo and seed development. Phylogenetic analysis was also conducted to assess the evolution of *Bn*LEC1, and compare its sequence with that of other LEC1 homologs. This study demonstrated that the *Bn*LEC1 gene conserves the distinctive features and characteristics of other LEC1 proteins. It is therefore suggested that *BnLEC1* is an ortholog to the *Arabidopsis LEC1* and it fulfills the same function exercised by other LEC1 members during seed development and oil biosynthesis.

3.2 INTRODUCTION

LEC1 is a transcription factor that encodes a CCAAT binding factor (CBF) also known as NF-Y (Lotan et al., 1998). Three subunits have been identified within members of the CBF family, HAP2 (CBF-B, or NF-YA), HAP3 (CBF-A, NF-YB), and HAP5 (CBF-C, NF-YC) (Maity and de Crombrugghe, 1998; Mantovani, 1999). The interaction among the three subunits establishes a trimeric complex that binds to the

CCAAT motif located in the promoter region of downstream genes (Li et al., 1992; Lotan et al., 1998). HAP2, HAP3 and HAP5 contain a highly conserved core region that allows for the three subunits to interact (Xie et al., 2008). LEC1 conserves the HAP3 subunit of the CBF, which displays structural similarities with the histone fold motif (Li et al., 1992; Sinha et al., 1996; Lotan et al., 1998; Bolognese et al., 2000; Lee et al., 2003).

The HAP3 domain is composed of three regions, an amino terminal A domain, a highly conserved B domain, and a carboxyl terminal C domain (Li et al., 1992; Lotan et al., 1998; Xie et al., 2008). The B domain of the HAP3 subunit has been divided into two classes, the LEC1-type and the non-LEC1 type (Kwong et al., 2003; Lee et al., 2003). There are two main subclasses of the LEC1-type subunits, namely the LEC1 and the LEC1-like proteins (Kwong et al., 2003). Both LEC1 and LEC1-like proteins contain a B domain that is required and specific to the process of embryogenesis (Lee et al., 2003). The B domain contains the amino acids necessary for HAP3 to interact with the other HAP2 and HAP5 subunits (Lee et al., 2003). This allows the CBF complex to bind to the promoter region of genes containing the CCAAT motif as summarized in Figure 3.1 (Li et al., 1992; Xing et al., 1993; Sinha et al., 1996).

The HAP3 CBF is thought to be the feature of LEC1 which makes this protein function as a major transcription factor throughout embryogenesis (Lotan et al., 1998; Lee et al., 2003). The CBF complex of LEC1 is believed to play a role in regulating the expression of *LEC2, FUS3*, and *ABI3,* which are other transcription factors involved in embryo development (Li et al., 1992; Lotan et al., 1998; Kagaya et al., 2005a; Kagaya et al., 2005b; Yamamoto et al., 2009). These transcription factors (*LEC2, FUS3* and *ABI3)* contain a B3 domain, which interacts with the RY element found in the promoters of

many genes encoding seed storage proteins (Meinke, 1992; Meinke et al., 1994; West et al., 1994; Lotan et al., 1998; Reidt et al., 2000; Stone et al., 2001; Kroj et al., 2003; Mönke et al., 2004). Multiple studies suggest that these transcription factors (LEC1, LEC2, FUS3, and ABI3) regulate aspects of embryo and seed development, although their direct molecular interaction still remains to be determined (Santos-Mendoza et al., 2008; Yamamoto et al., 2009).

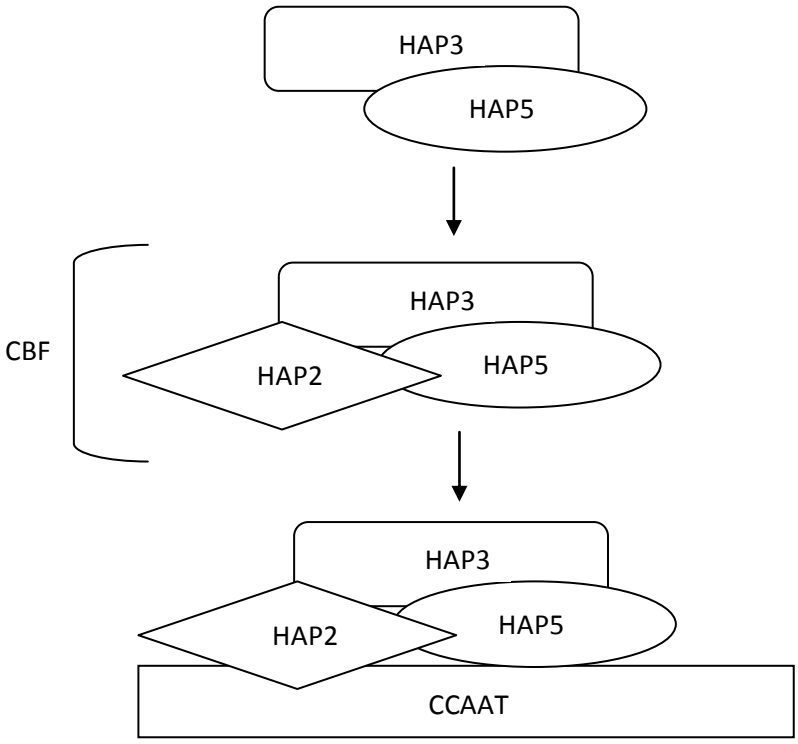

Figure 3.1. The interaction of the HAP3 and HAP5 domains form a dimer complex, which binds to HAP2 to establish a trimeric CCAAT binding factor (CBF). The CBF has a high affinity for CCAAT box motifs found in the promoter region of several genes. The binding of the trimer to the CCAAT motif activates the expression of the target gene. Adapted from Mantovani (1999), and Yazawa and Kamada (2007).

The function of *LEC1* appears to be conserved among plant species (Edwards et al., 1998; Xie et al., 2008). Based on phylogenetic analysis, the *LEC1*-type (*LEC1* and *LEC1*-like) *HAP3* genes of flowering plants are believed to have originated together throughout evolution (Xie et al., 2008). The divergence of the *LEC1*-type genes likely developed from the duplication and diversification of the function of the non *LEC1*-type *HAP3* genes (Kwong et al., 2003; Yang et al., 2005; Xie et al., 2008).

LEC1 proteins conserving the HAP3 subunit have been characterized in numerous plant species, including: *Arabidopsis thaliana* (NP_173616), *Glycine max* (ABW71515), *Glycine latifolia* (ABW71517), *Arachis hypogaea* (ADC33213), *Zea mays* (AAK955621), *Pistacia chinensis* (ADK91820), *Oryza sativa* (AAP22065), and *Brassica napus* (ACB12186). LEC1 has evidently become a transcription factor of great interest among many plant species. Because of the economic importance of canola seeds, exploring the genetic regulation of the *Brassica napus LEC1* gene appears to be a promising sector of research.

Previous molecular characterization of the HAP3 subunit of *LEC1* has provided insight on the evolution of this protein, which plays a key role seed development and oil biosynthesis (Lotan et al., 1998; Lee et al., 2003). Because the process of oil biosynthesis in canola is an area of great interest, the objective of this research is to isolate and characterize *Brassica napus LEC1* genes.

3.3 MATERIALS AND METHODS

3.3.1 Plant material

Plant material for the isolation of *LEC1* was generated using microspore cultures of the *Brassica napus* variety Topas. Five plants were grown in the greenhouse and at the onset of flowering, unopened buds (2 mm in length) were collected and kept on ice. The selected buds were utilized for microspore culture according to an established method (Belmonte et al., 2006). Microspore derived embryos were collected at day 36 of development and stored at -80°C. RNA was extracted from the collected tissue using the RNeasy® Plant Mini Kit (QIAGEN 27104) and the RNase-Free DNase set (QIAGEN 79254). The extracted RNA was then used as a template to synthesize cDNA using the SuperScript®II Reverse Transcriptase kit (Invitrogen 18064-014) according to the manufacturer's instructions.

3.3.2 Isolation of *BnLEC1*

Several steps were performed to obtain a full length clone of the *LEC1* gene. The National Center for Biotechnology Information (NCBI) database was searched for an *Arabidopsis thaliana LEC1* sequence (NM_102046.3) and an expressed sequence tags (ESTs) of *Brassica napus.* As of May 2008, an EST from *Brassica napus LEC1* (ES920786) was available on the NCBI database. These sequences served as a reference sequence for the design of the forward (5'-AAACGGCAGAGAAACAATGG-3') and reverse (5'-TCACTTATACTGACCATACTGGTC-3') primers that were utilized with the cDNA synthesized in the previous step in a PCR reaction to amplify *BnLEC1*.

The PCR reaction consisted of 10µl of Bioline™ BioMix (Bio-25012), 1 µl (10 pmole) of each primer, 7 µl of nuclease free water and 1 µl (700 ng) of cDNA. The PCR conditions included an initial denaturation cycle at 94°C for 1 minute, followed by 35 cycles of a denaturation period at 94°C for 30 seconds, annealing at 53°C for 30 seconds, elongation at 72°C for 2 minutes, and a final elongation cycle at 72°C for 10 minutes. The PCR product was separated by electrophoresis on a 1% agarose gel. The band was excised from the gel and collected in a 2 ml micro centrifuge tube. The QIAquick® Gel Extraction Kit (QIAGEN 28704) was used to extract the PCR product from the gel according to the instructions provided. In order to determine the approximate concentration of the extracted band, the product (1 µl) was run on a 1% agarose gel and compared to the standardized DNA concentrations of the Bioline Hyper Ladder™1 (Bio-22025).

The extracted PCR product was ligated into the pGEM®-T Easy Vector System II (Promega A1360) and transformed into JM109 Competent Cells (Promega L2001). Transformed colonies were selected on Luria Bertani (LB) medium (Sezonov et al., 2007) containing 100 µg/ml ampicillin, 0.5 mM isopropyl β-D-1-thiogalactopyranoside and 40 µg/ml bromo-chloro-indolyl-galactopyranoside. Colonies were screened by PCR using the M13 forward (5'-GTAAAACGACGGCCAGT-3') and reverse (5'-GCGGATAACAATTTCACACAGG-3') primers, and gene specific forward (5'-AAACGGCAGAGAAACAATGG-3') and reverse (5'-TCACTTATACTGACCATACTGGTC-3') primers to verify the presence of *BnLEC1* in the vector. Colonies containing the insert were grown in 5ml of LB medium containing 100 µg/ml of ampicillin as the selective antibiotic. Plasmid DNA was

extracted from the culture by following the protocol described in the QIAprep® Spin Miniprep Kit (QIAGEN 27104). The concentration of the plasmid DNA was adjusted to 100ng/µl and sent for sequencing to Macrogen USA (www.macrogenusa.com/). Sequencing revealed that two versions of *BnLEC1* had been cloned; they were denoted as *BnLEC1* version A and B.

3.3.3 Molecular characterization of *BnLEC1*

The cloned sequences were analyzed by multiple sequence alignment using the ClustalW (http://align.genome.jp/) and BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) programs. The sequences of the two versions of *BnLEC1* were initially compared to the original EST sequence (ES920786) used to design the primers for amplifying the gene. Shortly after, a second *Brassica napus* mRNA sequence (EU371726) became available (Mu et al., 2008). Nucleotide sequences were translated into amino acid sequences using the ExPASy Proteomics Server (http://expasy.org/tools/dna.html). The shared similarities among both the nucleotide and amino acid sequences of the different versions of the *BnLEC1* genes were determined. The conserved domains and structure of the gene were identified and visualized using the NCBI Search for Conserved Domains within in a Protein Sequence (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

3.3.4 Phylogenetic tree construction

Using the NCBI website, a BLAST search was run using the *Bn*LEC1 sequences to identify other LEC1-type genes sharing the HAP3 subunit. Sequences of *LEC1* and *LEC1*-like members containing the HAP3 subunit found in other species were collected and used to construct a phylogenetic tree. The phylogenetic tree was built using the MegAlign program with the ClustalW alignment method.

3.3.5 Construction of expression vectors

The two cloned versions of *BnLEC1* (version A and B) were both used to construct an expression vector in the sense orientation using the Gateway® Technology (Invitrogen 12535-029). This vector system requires the use of two independent reactions to obtain an expression construct. The first reaction is the BP reaction in which *BnLEC1* was incorporated into the pDONR^{TM}221 (Invitrogen 1235-029) Entry Vector by designing long primers that are specific to the binding regions *attB1* and *attB2* of the vector and *BnLEC1*. The forward primer

(5'–GGGGACAAGTTTGTACAAAAAAGCAGGCTACGGCAGAGAAACAATGGAACG-3')

and reverse primer

(5'-GGGGACCACTTTGTACAAGAAAGCTGGGTATCACTTATACTGACCATACT-3')

were used in a PCR reaction that included 10 µl of Bioline^{TM} BioMix (Bio-25012), 1µl (10 pmole) of each primer, 7 µl of nuclease free water and 1 µl (700ng) of cDNA. The PCR conditions consisted of an initial denaturation cycle at 94°C for 1 minute followed by 35 cycles of denaturation at 94°C for 30 seconds, annealing at 53°C for 30 seconds, elongation at 72°C for 2 minutes, and a final elongation cycle at 72°C for 10 minutes. The PCR product was then run on a 1% agrose gel containing 1 µg/ml of ethidium bromide to determine the correct band size (766 base pairs) which was excised from the gel and placed into a 2 ml micro centrifuge tube. QIAquick® Gel Extraction Kit (QIAGEN 28704) was used to purify the PCR product from the gel.

The PCR product was used in the initial BP reaction by following the instructions provided in the Gateway^{TM} Technology protocol manual. The BP reaction product was transformed into One Shot® OnmiMAX^{TM} 2-T1R Chemically Competent *E.Coli* cells

(Invitrogen C8540-03) which were grown on LB medium containing 50 µg/ml of kanamycin. Positive colonies were screened by PCR to verify the presence and orientation of *BnLEC1* within the pDONR entry vector. The M13 forward (5'-GTAAAACGACGGCCAG-3') primer and the *BnLEC1* reverse (5'-TCACTTATACTGACCATACTGGTC-3') primer were used to ensure the correct orientation of *BnLEC1* within the entry vector. A second PCR was conducted to screen for the presence of *BnLEC1* using full length gene specific forward (5'-AAACGGCAGAGAAACAATGG-3') and reverse (5'-TCACTTATACTGACCATACTGGTC-3') primers. The PCR reactions using both primer sets were executed under the same reaction parameters and conditions that were applied to construct the entry vector. The positive colonies were grown in 5 ml of LB medium containing 50 µg/ml of kanamycin. Plasmid DNA was extracted from the cultures using the QIAprep® Spin Miniprep Kit (QIAGEN 27104), and sent for sequencing at Macrogen USA to confirm that the full length *BnLEC1* was present in the entry vector.

The Entry Vector was then used in the LR reaction to insert *BnLEC1* into the Destination Vector pK2GW7 (Karimi et al., 2002) in the sense orientation. The LR reaction was performed according to the instructions provided in the Gateway® Technology protocol manual. The product of the LR reaction was electroporated into ElectroMAX™ DH10B™ *E. coli* cells (Invitrogen 18290-015) using a Bio-Rad Gene Pulser Xcell™ Electroporation system. Transformed cells were grown on LB medium containing 100 µg/ml of kanamycin and 50 µg/ml spectinomycin. Positive colonies were

screened for the presence and proper orientation of *BnLEC1* within the pK2GW7 vector by PCR using a 35S promoter forward (5'-TGGACCCCCACCCACGAG-3') primer (Elhiti et al., 2010) and a *BnLEC1* specific reverse (5'-TCACTTATACTGACCATACTGGTC-3') primer. Once the presence and orientation of the *BnLEC1* gene within the Destination Vector were confirmed, the positive colonies were grown in 5 ml of LB medium containing 100 µg/ml of kanamycin and 50 µg/ml of spectinomycin. Plasmid DNA was extracted and sequenced.

3.3.6 Electroporation of the pK2GW7 vector containing *BnLEC1* into *Agrobacterium tumefaciens* cells

The pK2GW7 Destination Vector containing *BnLEC1* was electroporated into *ElectroMAX™ Agrobacterium tumefaciens LBA4404* cells (Invitrogen 18313-015) according to the manufacturer's instructions, using the Bio-Rad Gene Pulser Xcell™ Electroporation System. The *Agrobacterium tumefaciens* cells were grown for 48 hours at 28°C on LB medium containing 100 µg/ml of kanamycin and 50 µg/ml of spectinomycin. The positive colonies were screened using the same primers, reaction, and PCR conditions previously described during the construction of the destination vector.

3.4 RESULTS

3.4.1 Comparison of the two *BnLEC1* clones

Two full length versions of the *BnLEC1* gene were cloned from microspore derived embryos of *Brassica napus*. The annotation of genes revealed that the coding region of both versions consisted of 693 nucleotides which translated to 230 amino acids. The nucleotide sequences of *BnLEC1* version A (GU945399) and version B (GU945398),

and amino acid sequences of version A (ADF81045) and version B (ADF81044) were submitted to the NCBI database. The nucleotide sequences of *BnLEC1* version A (GU945399) and version B (GU945398) were aligned and compared to other *BnLEC1* sequences available in the NCBI database (Figure 3.2). These sequences included the initial reference EST sequence (ES920786), the sequence released by Mu et al. (2008) (EU3717261) and another EST sequence (DY017663) available (Figure 3.2). An alignment was also performed using the amino acid sequences (Figure 3.3).

The two versions (A and B) of *BnLEC1* differed by 13 nucleotides (Figure 3.2) and 4 amino acids (Figure 3.3). *Bn*LEC1 version A was found to be 100% identical in nucleotide (Figure 3.2) and amino acid sequences (Figure 3.3) to the initial reference EST sequence (ES920786). *Bn*LEC1 version B was found to be 100% identical in terms of nucleotides (Figure 3.2) and amino acids (Figure 3.3) to the EST (DY017663). BnLEC1 (version B) was also very similar in nucleotide and amino acid sequence to the clone (EU3717261) published by Mu et al. (2008), differing by 5 nucleotides (Figure 3.2) translating into 2 amino acids (Figure 3.3).

Figure 3.2. Nucleotide alignment comparing five BnLEC1 sequences: GU945399 (version A), GU945398 (version B), EU3717261 and EST sequences ES902786, and DY017663**.** * indicate sequence similarities among all five clones.

Figure 3.3. Amino acid alignment of BnLEC1 ADF81045 (version A), ADF81044 (version B), ACB12186 and translated sequences of ES920786 and DY017663. The conserved domain of the HAP3 subunit is highlighted in grey. * indicate sequence similarities among all five clones.

3.4.2 Conserved domains and protein structure

The conserved domains and architecture of the two *Bn*LEC1 clones were found using the utility provided by the NCBI Search for Conserved Domains within a Protein Sequence. The BLAST search showed that both *Bn*LEC1 proteins contained the histone H2A superfamily domain in the central region of the protein. The superfamily includes the core histone (H2A, H2B, H3 and H4) subunits, which forms a nucleosome octamer (Figure 3.4). Due to these characteristics, *Bn*LEC1 has a histone fold motif (Li et al., 1992; Sinha et al., 1996; Lotan et al., 1998; Bolognese et al., 2000; Lee et al., 2003). *Bn*LEC1 also contains the HAP3 subunit of the CCAAT binding factor in the central region of the protein as shown in Figure 3.4. This trimeric complex consists of HAP2 (CBF-B and NF-YA), HAP3 (CBF-A and NF-YB), and HAP5 (CBF-C and NF-YC)

(Maity and de Crombrugghe, 1998; Mantovani, 1999). Histone fold motifs allow for HAP3 to interact with HAP5 and form a tight dimer which interacts with HAP2 (Figure 3.1) (Li et al., 1992; Sinha et al., 1996; Lotan et al., 1998; Mantovani, 1999; Bolognese et al., 2000; Lee et al., 2003; Xie et al., 2008). This demonstrated that both *Bn*LEC1 clones conserved the same subunits typical of other LEC1 proteins.

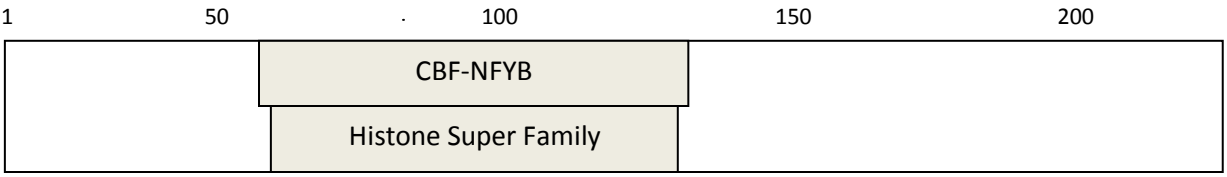

Figure 3.4. The major conserved domains of the BnLEC1 protein. Numbers indicate the amino acid position along the protein. The CBF (CCAAT binding factor) NFYB (HAP3) subunit occupies the central region of the protein. *Bn*LEC1 contains the histone super family (H2A, H2B, HA and HB).

3.4.3 LEC1 phylogenetic tree

A phylogenetic tree was constructed to demonstrate the likely evolution of the LEC1-type genes conserving the HAP3 subunit in different plant species (Figure 3.5). The tree includes both *LEC1* and *LEC1*-like genes. A list of proteins selected for the construction of the phylogenetic tree has been compiled in Table 3.1.

Sequence alignment by the ClustalW method showed that the three *Bn*LEC1 (1, 2, and 3) proteins share 98.3% (1 and 2), 99.1% (1 and 3) and 98.3% (2 and 3) identity (Figure 3.5 and 3.6). The three *Bn*LEC1 proteins displayed 73.6% (1 and 4), 73.6% (2 and 4), 73.6% (3 and 4), 77.4% (1 and 5), 77.4% (2 and 5), and 77.1% (3 and 5)

similarity with the two *At*LEC1 (4 and 5) amino acid sequences (Figure 3.5 and 3.6). The *Brassica napus* (12) and *Arabidopsis thaliana* (13) LEC1-like genes shared 82.8% (12 and 13) identity (Figure 3.5 and 3.6). The LEC1 and LEC1-like proteins appear to be highly conserved showing significant homology among the *Brassica napus* (12) and *Arabidopsis thaliana* (13) species. However the three *Brassica napus* LEC1 (1, 2, and 3) sequences and one LEC1-like (12) sequence are 47.8% (1 and 12), 46.9% (2 and 12), and 47.8% (3 and 12) similar. The two *Arabidopsis thaliana* LEC1 (4 and 5) sequences and one LEC1-like (13) share 41.8% (4 and 13) and 38.9% (5 and 13) identity (Figure 3.5 and 3.6).

Table 3.1. Selected LEC1 and LEC1-like protein sequences used for the construction of the phylogenetic tree.

Number	Accession Number	Protein	Species	Description
1	ADF81044	LEC1	*Brassica napus*	LEC1-1 transcription factor [Brassica napus] (Version B)
2	ADF81045	LEC1	*Brassica napus*	LEC1-2 transcription factor [Brassica napus] (Version A)
3	ACB12186	LEC1	*Brassica napus*	leafy cotyledon 1 transcription factor [Brassica napus]
4	AAC39488	LEC1	*Arabidopsis thaliana*	CCAAT-box binding factor HAP3 homolog [Arabidopsis thaliana]
5	NP_173616	LEC1	*Arabidopsis thaliana*	LEC1 (LEAFY COTYLEDON 1); transcription activator/ transcription factor [Arabidopsis thaliana]
6	AAP22065	LEC1	*Oryza sativa*	leafy cotyledon 1 [Oryza sativa Indica Group]
7	ABW71515	LEC1	*Glycine max*	transcription factor LEC1-B [Glycine max]
8	ABW71517	LEC1	*Glycine latifolia*	transcription factor LEC1-B [Glycine latifolia].
9	ADC33212	LEC1	*Arachis hypogaea*	leafy cotyledon 1-A [Arachis hypogaea]
10	ADK91820	LEC1	*Pistacia chinensis*	LEC1 Transcription Factor
11	AAK95562	LEC1	*Zea mays*	leafy cotyledon1 [Zea mays]
12	ACB12187	LEC1-like	*Brassica napus*	leafy cotyledon 1-like protein [Brassica napus]
13	AAN17924	LEC1-like	*Arabidopsis thaliana*	leafy cotyledon 1-like L1L protein [Arabidopsis thaliana]

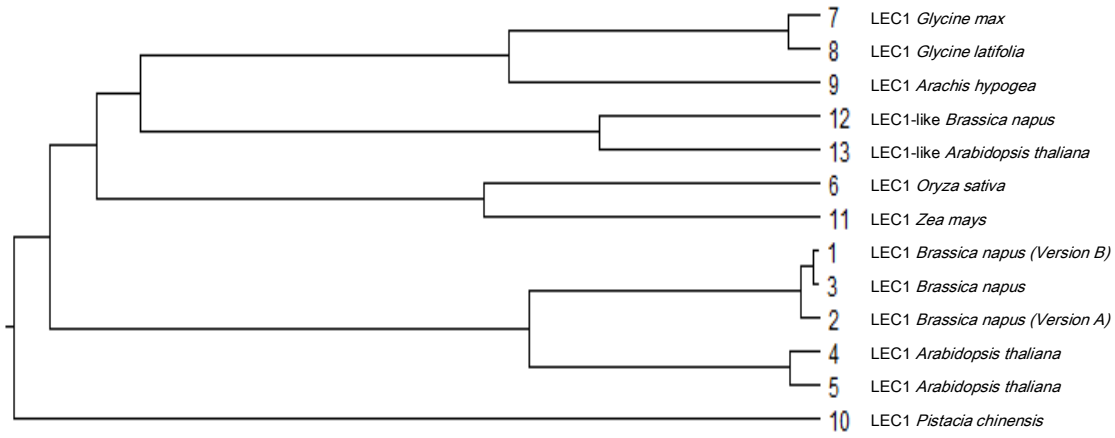

Figure 3.5. Phylogenetic tree constructed using LEC1 and LEC1-like proteins from different plant species**.** A list of the included proteins (with respective numbers) is shown in Table 3.1.

		Percent Identity													
Percent Divergence		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	
	1		98.3	99.1	73.6	77.4	43.9	46.9	44.8	44.2	45.7	44.3	47.8	40.0	**1**
	2	1.8		98.3	73.6	77.4	44.3	46.9	44.8	44.2	45.7	44.8	46.9	43.0	**2**
	3	0.4	1.3		73.6	77.1	45.0	46.9	44.6	45.1	45.0	44.2	47.8	39.8	**3**
	4	24.7	24.0	24.7		99.5	47.6	44.2	45.2	44.7	51.0	50.5	49.0	41.8	**4**
	5	25.7	25.1	25.7	0.5		41.6	43.8	44.6	44.2	49.2	45.4	48.8	38.9	**5**
	6	73.3	74.6	73.3	73.7	78.5		44.2	42.9	43.4	38.5	67.3	52.2	45.7	**6**
	7	77.9	77.9	77.9	76.3	83.1	66.9		97.3	66.4	42.9	44.2	57.4	53.1	**7**
	8	79.9	79.9	79.9	81.3	86.2	66.9	2.7		67.7	40.8	43.8	56.5	52.4	**8**
	9	69.4	70.6	69.4	71.0	77.4	65.8	26.2	26.4		51.3	45.6	52.2	51.8	**9**
	10	61.5	62.5	61.5	54.1	60.8	79.9	81.1	81.7	72.1		40.1	43.1	44.0	**10**
	11	77.3	78.6	77.3	76.5	83.7	28.6	73.1	71.9	73.0	77.3		51.7	45.3	**11**
	12	77.1	75.7	77.1	68.7	78.5	56.8	45.8	46.7	44.5	72.0	67.0		82.8	**12**
	13	87.9	86.5	87.9	80.6	90.5	70.4	57.0	57.7	57.1	79.9	78.6	19.0		**13**
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	

Figure 3.6. Identity and divergence matrix of the selected protein sequences included in the phylogenetic tree (Figure 3.5). Numbers in bold correspond to the proteins listed in Table 3.1.

3.4.4 Development of sense constructs

The Gateway Vector System was used to construct an expression vector containing the *BnLEC1* in the sense orientation. The pk2GW7 Destination Vector (Figure 3.7) contains a kanamycin resistant gene located upstream of the 35S promoter. The 35S promoter drives the expression of *BnLEC1* which is inserted into the construct between the *attB1* and *attB2* flanking regions. A 35S terminator is located at the end of the TDNA to stop the expression of the *BnLEC1* gene (Figure 3.7). Two sense constructs were developed using the two cloned versions of *BnLEC1*. The first construct was designed using *BnLEC1* version A (ADF81045) (Figure 3.8), whereas the second using *BnLEC1* version B (ADF81044) (Figure 3.9).

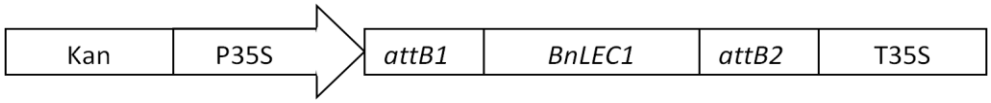

Figure 3.7. General layout of the TDNA insertion region of the expression vector pK2GW7. The kanamycin resistant gene is located upstream of the 35S promoter which drives the expression of *BnLEC1* flanked by the *attB1* and *attB2* sites. The T35S terminates the expression of the transgene.

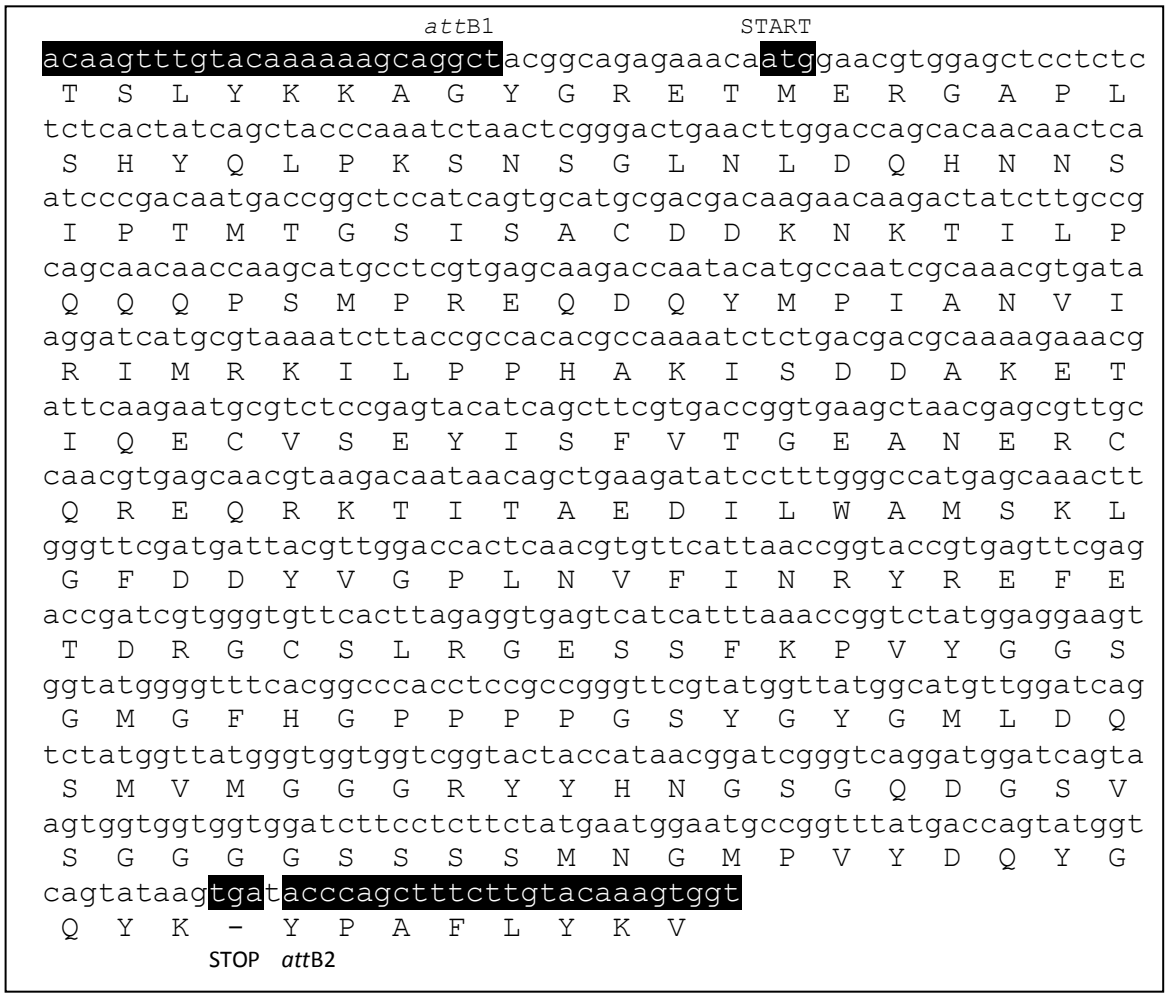

Figure 3.8. Sequence of *BnLEC1* version A (ADF81045) inserted into the pK2GW7 vector.

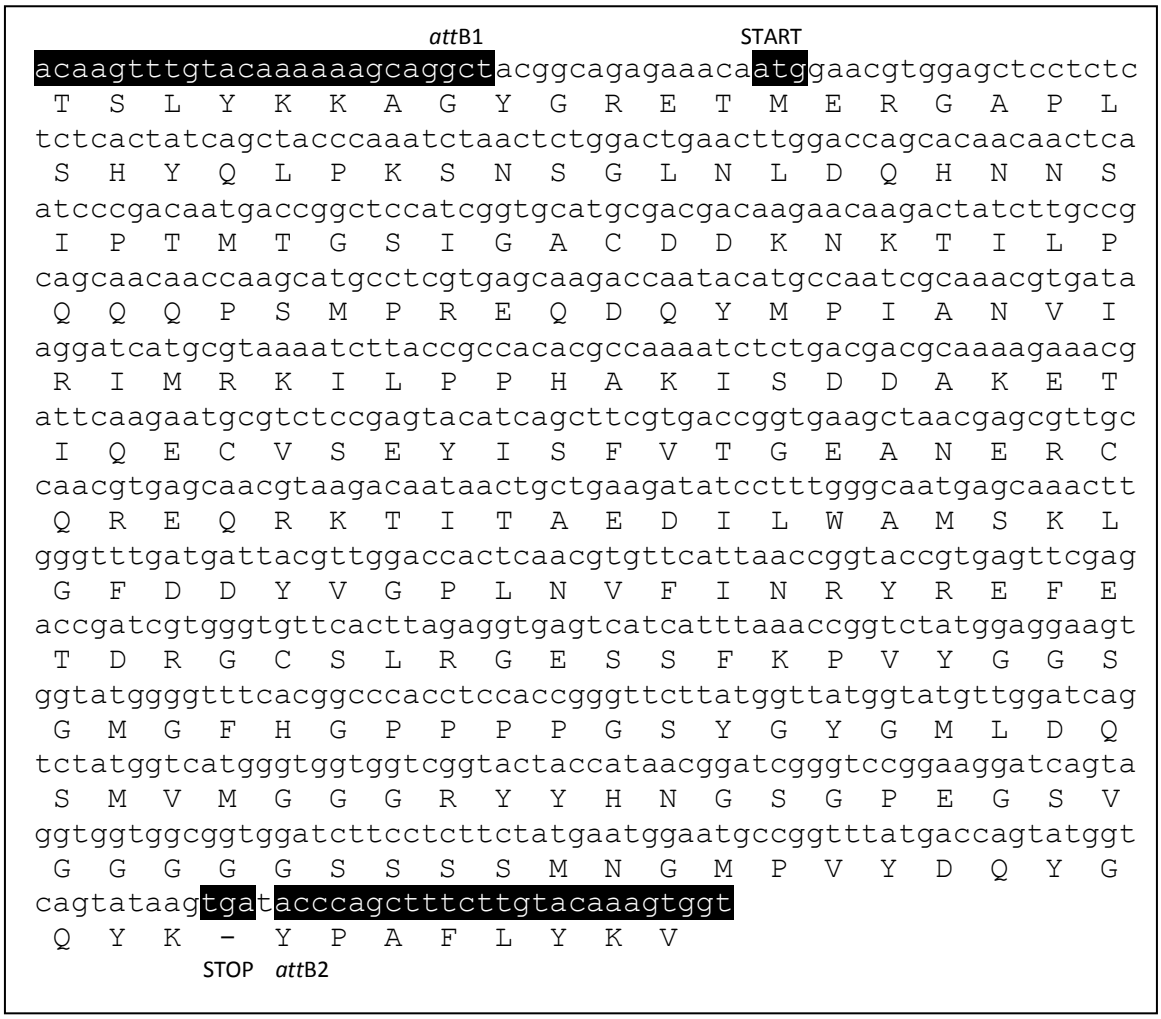

Figure 3.9. Sequence of *BnLEC1* version B (ADF81044) inserted into the pK2GW7 vector.

3.5 DISCUSSION

The *BnLEC1* protein contains a HAP3 subunit of the CCAAT box binding factor, which has previously been identified as a distinctive feature of other LEC1 proteins (Lotan et al., 1998). The HAP3 subunit is highly conserved, and is thought to be greatly influenced by the evolution of seed plants (Xie et al., 2008). A study by Xie et al. (2008) demonstrated that the LEC1-type (LEC1 and LEC1-like) HAP3 subunit may have initially arose from non seed plants (specifically, the lycophyte *Selaginella sinensis* and the fern *Adiantum capillus-veneris*). It is thought that LEC1 originally played an adaptive role in protecting the plant from vegetative desiccation (Xie et al., 2008). The function of LEC1 in non seed plants appears to have evolved into seed plants, as *LEC1*-type genes have been identified as major regulators of the early and later stages of seed development (Lotan et al., 1998; Xie et al., 2008).

Protein sequence comparison and conserved domain analysis further suggest that gene duplication has occured throughout the evolution of LEC1 (Xie et al., 2008). Gene duplication is an adaptive event that can occur from a single gene, a segment of a choromsome, or at the whole genome level (Yang et al., 2005; Xie et al., 2008). All levels of duplication can lead to the development of genes that may be redundant in function within the same species (Flagel and Wendel, 2009). The isolation of different copies of *BnLEC1* may be due to gene duplication, especially because *Brassica napus* is a polyploid species (Schranz and Osborn, 2004; Flagel and Wendel, 2009).

The evolutionary relationship between the *Bn*LEC1 and LEC1 proteins isolated from other species was studied through phylogenetic analysis (Figure 3.5). Members of the *Brassicacea* family, namely *Brassica napus* and *Arabidopsis thaliana* appear to share

significant sequence identity in LEC1 and LEC1-like proteins (Figure 3.4 and 3.5). The LEC1 and LEC1-like proteins both conserve the same HAP3 domain, and a specific amino acid residue that distinguishes their function from other genes containing the HAP3 subunit (Kwong et al., 2003; Lee et al., 2003; Yamamoto et al., 2009). The independent evolution of both LEC1 and LEC1-like proteins from a common origin, are consistent with the notion that the two genes are thought to play distinct roles in embryogenesis (Kwong et al., 2003; Lee et al., 2003; Xie et al., 2008; Yamamoto et al., 2009).

The presence of similar domains within the LEC1 proteins suggests a conserved function that is retained across species. The HAP3 domain interacts with HAP5 through histone fold motifs (H2A, H2B, H3 and H4) forming a dimer to which the HAP2 subunit binds (Sinha et al., 1995; Kim et al., 1996; Sinha et al., 1996; Mantovani, 1999; Lee et al., 2003; Romier et al., 2003). The trimer consisting of HAP3, HAP5 and HAP2 binds to specific DNA sequences by recognizing CCAAT-box motifs in the promoter region of target genes, allowing LEC1 to function as a transcription factor (Figure 3.1) (Lee et al., 2003). The CCAAT box motif is expected to be present in approximately 30% of promoters, and the specificity of the binding is determined by the DNA sequences flanking the activation site (Sinha et al., 1995; Mantovani, 1999; Romier et al., 2003). LEC1 is believed to activate the expression of downstream transcription factors such as LEC2, FUS3, WRI1 and ABI3 through this mechanism; however, the direct molecular interactions among these genes have not yet been demonstrated (Santos-Mendoza et al., 2008).

This study clearly shows that the two versions of *BnLEC1* contain the distinct features that are characteristics of previously identified *LEC1* genes from other species. This similarity suggests that *BnLEC1* might fulfill functions during embyo development similar to those observed in other *LEC1* genes. In order to examine this hypothesis, both *BnLEC1* versions (A and B) were used to prepare vectors to ectopically express the genes in *Arabidopsis* plants. The characterization of the transformed plants will be beneficial for understanding the role LEC1 plays during embryo and seed development.

4.0 CHAPTER 2: OVER-EXPRESSION OF *BRASSICA NAPUS LEAFY COTYLEDON1* IN *ARABIDOPSIS THALIANA*

4.1 ABSTRACT

LEAFY COTYLEDON1 (LEC1) is a major transcription factor regulating embryo morphogenesis and maturation. Fatty acids and triacylglycerols (TAGs) accumulate during the maturation phase of embryo and seed development. As a result, the genetic regulation of embryo maturation has become an area of great interest when trying to improve the seed oil content and composition in canola. A new sector of research focuses on determining the role of LEC1 in the regulation of seed oil biosynthesis. To further characterize the function of this gene, two versions of the *Brassica napus LEC1* were cloned. The *BnLEC1* genes were inserted into an expression vector, and seven *Arabidopsis* lines were successfully transformed with the construct using in planta transformation. The *Arabidopsis* lines were characterized by phenotypic differences in seed oil composition. One line also contained less seeds per silique compared to the wild type plants. Expression analysis demonstrated that the *BnLEC1* transgene was not expressed at significant levels. Speculations have been made to explain the inconsistencies between the lack of expression and phenotypic differences observed among lines. Unfortunately, no change in the expression level of *BnLEC1* in *Arabidopsis* plants prevented further functional characterizations of the gene during embryo and seed development.

4.2 INTRODUCTION

Understanding how LEC1 regulates embryogenesis has been the focus of many studies in plant development. Research on the *LEC1* gene began when Meinke (1992)

characterized the LEC1 phenotype in *Arabidopsis* plants. The transgenic plants with suppressed *LEC1* expression displayed phenotypes characterized by intolerance to seed desiccation and the formation of trichomes on cotyledons (Meinke et al., 1994; West et al., 1994). The *lec1* mutants had unique morphological characteristics pertaining to embryogenesis, seed dormancy and seedling development (Meinke, 1992; Meinke et al., 1994; West et al., 1994). These genetic studies suggest the essential role that LEC1 plays not only in the early and later stages of embryogenesis, but also in post-embryonic growth (West et al., 1994). This concept has been further supported by several over-expression studies. Lotan et al. (1998) ectopically expressed the *Arabidopsis LEC1* under the control of a 35S promoter into null *lec1 Arabidopsis* mutants. The transgenic plants displayed various phenotypes that differed among insertion lines (Lotan et al., 1998). The 35S::*LEC1* seedlings tended to be smaller with cotyledons that did not expand (Lotan et al., 1998). Numerous transgenic plants produced somatic embryos from the vegetative leaf tissue, and some seedlings displayed cotyledon-like organs in typical leaf locations, demonstrating that the plants retained embryonic characteristics (Lotan et al., 1998). Male sterility was observed in one insertion line, whereas others were capable of flowering and producing progeny (Lotan et al., 1998). These genetic analyses further revealed that the *LEC1* gene is a major regulator of embryo development (Meinke et al., 1994; West et al., 1994; Lotan et al., 1998).

The role of LEC1 during seed development was further determined when both *AtLEC1* and *BnLEC1* were over-expressed under the control of an estradiol inducible promoter in *Arabidopsis* (Mu et al., 2008). Interestingly, seedlings exposed to estradiol exhibited stunted growth, produced yellow cotyledons, and rarely grew true leaves (Mu et

al., 2008). In some instances, somatic embryos developed from the vegetative tissue of the transgenic lines (Mu et al., 2008). The fatty acid levels of the transgenic seedlings that over-expressed *AtLEC1* and *BnLEC1* were analyzed by gas chromatography-mass spectrometry. The results demonstrated the *LEC1* lines accumulated increased levels of fatty acids throughout various stages of germination (Mu et al., 2008). A similar fatty acid composition was observed in both *AtLEC1* and *BnLEC1* over-expressors, suggesting that the two genes were orthologs (Mu et al., 2008). To further understand the transcriptional changes induced by the elevated expression of *AtLEC1*, a microarray analysis was performed to identify differentially expressed genes within the transgenic lines (Mu et al., 2008). These studies revealed that some of the genes involved in carbohydrate metabolism, plastidial pathway and fatty acid synthesis were unregulated by *AtLEC1* (Mu et al., 2008). Alternatively, several genes involved in amino acid and protein metabolism were down-regulated (Mu et al., 2008). Overall, there studies suggest that LEC1 plays a role in the regulation of fatty acid biosynthesis during seed development.

Previous research has shown that a series of phenotypes have been obtained by manipulating the expression level of *LEC1* (Meinke et al., 1994; West et al., 1994; Lotan et al., 1998; Mu et al., 2008). These included the formation of somatic embryos from vegetative tissue, and increased fatty acid accumulation during seedling growth (Lotan et al., 1998; Mu et al., 2008). The variation in phenotype was possibly due to the different expression levels of the transgene. The existence of multiple *LEC1* orthologs in *Brassica napus* and *Arabidopsis thaliana*, adds complexity to further understanding the function of the gene.

Due to its involvement in embryogenesis and fatty acid biosynthesis, LEC1 is a good candidate for improving the oil content in oilseed plants such as canola (Meinke et al., 1994; West et al., 1994; Lotan et al., 1998; Mu et al., 2008). Therefore the objective of this chapter is to determine if the over expression of *BnLEC1* influences the fatty acid composition during seed development. Results from this study will further demonstrate if the *Brassica napus LEC1* genes (version A and B) fulfill the same function exercised by the *Arabidopsis thaliana LEC1* gene.

4.3 MATERIALS AND METHODS

4.3.1 *Arabidopsis thaliana* transformation

4.3.1.1 Growth conditions of *Arabidopsis* plants

Arabidopsis seeds (Columbia) were sterilized in 70% ethanol and 0.5% triton X-100 for 12 minutes, rinsed in 95% ethanol for 10 minutes, and dried on sterilized filter paper. The seeds were plated on *Arabidopsis* medium (Sauer and Friml, 2004) that contained half strength Murashige and Skoog Basal Medium with Gamborg's Vitamins (Sigma M0404), 0.05% MES, 1% sucrose and 0.8% agar. The seeds were stratified for 48 hours at 4°C and then transferred to room temperature with constant light. After producing two to four leaves, the seedlings were planted into 96 well flats containing LA4 Sunshine (Sun Gro) potting mix and covered with plastic wrap for seven days. The flats were kept in a growth cabinet and exposed to 16 hours of light, at an intensity of approximately 270 micro-einsteins per square meter per second. As the plants began to develop initial shoots, they were trimmed to encourage lateral shoot development.

4.3.1.2 Preparation of *Agrobacterium tumefaciens* containing the *BnLEC1* construct

Two liquid cultures of *Agrobacterium tumefaciens* containing version A or version B of the *BnLEC1* construct, were prepared according to a modified floral dip protocol (Clough and Bent, 1998). *Agrobacterium* was grown in 150 ml of LB medium containing 100 µg/ml of kanamycin and 50 µg/ml of spectinomycin, and vigorously shaken at 200 rpm for 48 hours at 28°C. The *Agrobacterium* cells were sedimented by centrifugation at 5000 rpm for 10 minutes, and the pellet was re-suspended in a solution containing 5% sucrose and 0.05% Silwet L-77.

4.3.1.3 In planta transformation of *Arabidopsis*

Once the *Arabidopsis* plants produced buds, the flats were sprayed with a solution containing *Agrobacterium*, sucrose and Silwet L-77, according to a transformation protocol (Clough and Bent, 1998). The flats were removed from the light, covered with plastic for 24 hours, and then returned to normal growing conditions for four days. The *Agrobacterium* solution was sprayed onto each flat every four days over a 20 day period. Once the plants began to display signs of maturity, the seeds were harvested and stored at 4°C.

4.3.1.4 Screening for transformants

In order to screen for transformants, the seeds were sterilized and plated on *Arabidopsis* medium (Sauer and Friml, 2004) supplemented with 80 µg/ml of kanamycin and 500 µg/ml of carbenicillin. The transformed seedlings were resistant to kanamycin, while the seedlings that did not contain the *BnLEC1* insertion did not possess the resistance to the antibiotic. To ensure that surrounding plants were not metabolizing the antibiotic, leading to false positives, the plants that remained green in the presence of

kanamycin were transferred onto fresh medium until two to four leaves developed. The plants that appeared to be transformed were transferred to 8.9 cm^2 pots containing LA4 Sunshine Mix (Sun Gro), covered with plastic, and grown under the same environmental conditions outlined in section 4.3.1.1.

4.3.1.5 Selfing of transformed *Arabidopsis* plants to obtain homozygous lines

The initial T0 lines were selfed to obtain T1 seeds. Ten T1 plants were grown in separate pots and covered with selfing cones to prevent cross-pollination. The plants were grown to maturity, at which point the T2 seeds from each of the ten plants derived from independent insertion lines were collected. Ten plants from each of the individually selfed T2 plants were grown to produce the T3 generation. Ten individually selfed sub-lines from each *Arabidopsis* line were developed. The selection and selfing process used to develop homozygous T3 lines from the initial T0 generation are summarized in Figure 4.1. The T3 generation of homozygous plants was used for phenotypic analysis. A wild type *Arabidopsis* line which was grown and selfed for three generations under the same environmental conditions served as the control.

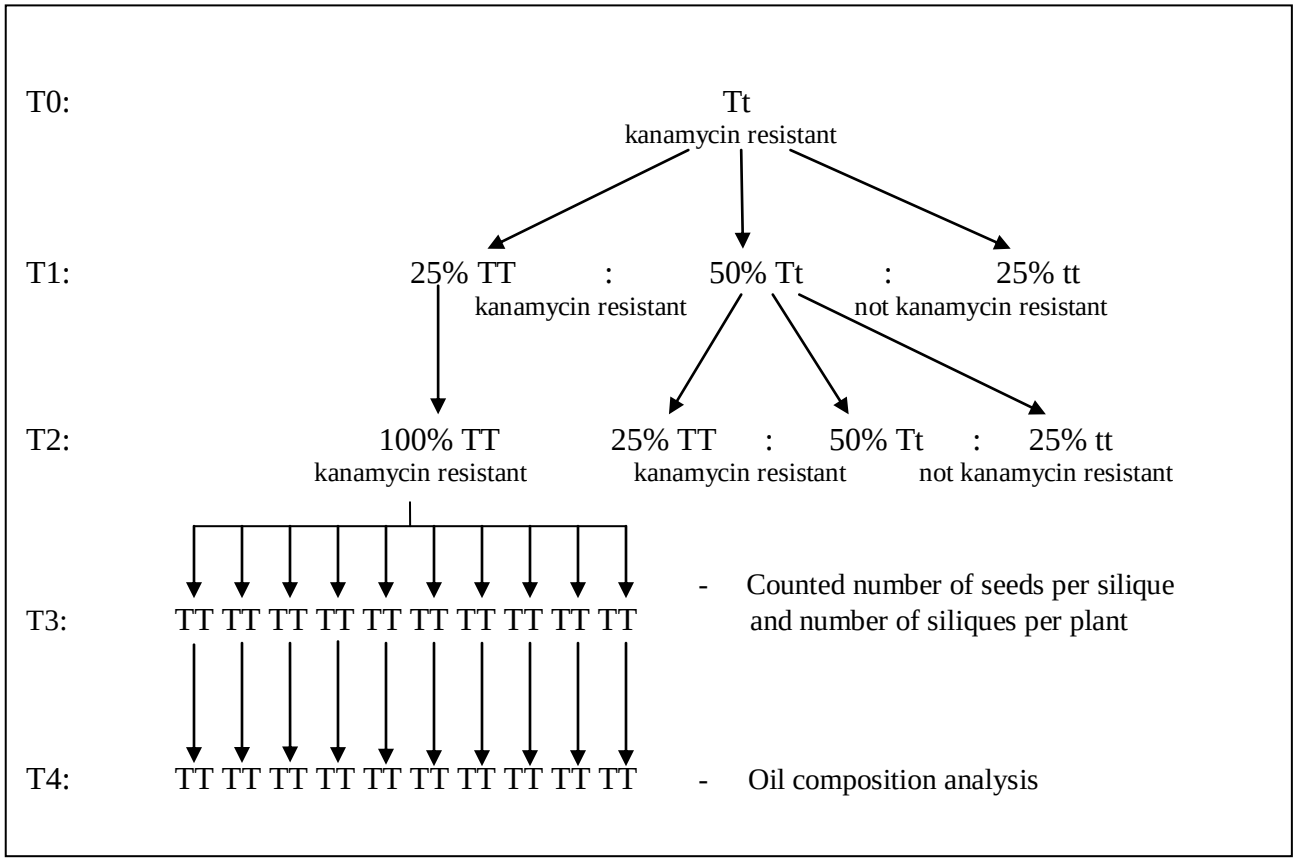

Figure 4.1. General procedure and segregational analysis of transformed plants used to develop independent homozygous lines (TT) which were then utilized for phenotypic analysis in the T3 and T4 generations. Kanamycin resistant plants (TT or Tt) were selected, and non resistant plants (tt) were eliminated from the selfing process.

4.3.1.6 Segregational analysis to determine the number of TDNA insertions

Segregational analysis was used to determine the number of TDNA insertions present in each of the transgenic *Arabidopsis* lines. During the T1 generation, the ratio of kanamycin resistant plants to non kanamycin resistant plants was recorded. A Chi Square test was performed to assess for the expected segregation ratio of 3:1 (resistant : non resistant plants) according to Mendelian genetics. The Chi square values were compared to the standardized critical value of 3.84 at 1 degree of freedom and 5% probability.

4.3.1.7 Immature seed rescue

The transgenic line A8 displayed a phenotype of inconsistent seed development and filling, only producing one or two seeds within each silique during the T0 generation. As a result, there was a concern that viable seeds would not be produced in this line. Immature seed rescue was performed to ensure that the plants of line A8 would progress to the T1 generation.

The procedure of immature seed rescue involved dipping a green *Arabidopsis* silique in 95% ethanol and placing it onto a sterilized slide located on the microscope stage. A surgical needle and forceps were then used to dissect immature seeds, which were grown on germination medium (Olsen et al., 1993) containing 80µg/ml of kanamycin. The plates containing seeds were incubated under a light bank that provided 16 hours exposure to light at room temperature. Once the seedlings developed two to four leaves, they were transferred to LA4 Sunshine Potting Mix (Sun Gro) and grown to maturity.

4.3.2 Genotyping of *Arabidopsis* plants

4.3.2.1 Genomic DNA extractions

One or two *Arabidopsis* leaves were placed in a 1.5 ml micro centrifuge tube, frozen in liquid nitrogen and ground using a micro pestle. A total of 500 µl of 2x cetyltrimethylammonium bromide buffer was added to the tissue which was incubated at 65°C for 90 minutes. After the addition of chloroform (500 µl) the slurry was centrifuged at 4500 rpm for 10 minutes at room temperature. Approximately 500 µl of supernatant was removed, and transferred into a new tube containing 500 µl of chloroform. The tube was centrifuged at 4500 rpm for 10 minutes at room temperature.

The supernatant was removed, 250 µl of isopropanol was added, and the tube was vortexed and left on ice for two hours. The tube was centrifuged at 4000 rpm for 30 seconds at room temperature to form a DNA pellet. The supernatant was removed, and the tube was left open to allow for the evaporation of excess isopropanol. A total of 50 µl of TE was added to each tube to dissolve the DNA pellet. The DNA was quantified using Ultraspec 2100 pro UV/visible spectrophotometer.

4.3.2.2 PCR conditions to genotype transgenic lines

The genomic DNA was used as a template for a series of PCRs to verify the presence of the transgene in the genome of the selected lines. The PCR reactions were performed using a 35S forward primer (5'-TGGACCCCCACCCACGAG-3') (Elhiti et al., 2010) and a *BnLEC1* reverse primer (5'-GATCCACCACCACCACTTACTGA-3'). The reverse primer was designed to hybridize only to the transgene (*BnLEC1*) and not to the native gene (*AtLEC1*). Each reaction consisted of 1 µl (200 ng) DNA, 1 µl (10 pmole) of both primers, 7 µl of nuclease-free water, and 10 µl of GoTaq® Hot Start Green Master Mix (Promega M512C). The PCR program involved an initial denaturation cycle of 94°C for 3 minutes, 37 cycles of denaturation at 94°C for 30 seconds, annealing at 53.5°C for 30 seconds, elongation at 72°C for 45 seconds, and a final 5 minute extension period at 72°C. The PCR reactions were conducted using templates from the 10 sub-lines developed from each transgenic *Arabidopsis* line. Water controls were also included.

4.3.3 Expression studies

4.3.3.1 RNA extraction

The expression level of the transgene was determined in *Arabidopsis* leaves and siliques. Two to three cauline leaves and four to five siliques were harvested from three sub-lines of each *Arabidopsis* line. The siliques were collected seven days after pollination. In the first trial, measurements were taken to ensure the siliques of lines A4, A7, A8, A9, B6, B7 and WT were longer than 1.0 cm, and line B3 was at least 0.7 cm in length. In the second trial the siliques of lines A4, A7, A8, A9, B6, B7 and WT were 1.0 to 1.5 cm long and those of line B3 were 0.7 to 1.0 cm in length. As the expression of *LEC1* varies throughout development (Braybrook and Harada, 2008), the measurements were taken to identify and eliminate any silique that may have been developmentally delayed as a result of being damaged from the pollination procedure.

The leaves and siliques were frozen in liquid nitrogen and RNA was extracted using the QIAGEN RNeasy® Plant Mini Kit (74904) and the RNase-Free DNase set (QIAGEN 79254). The concentration and quality of the RNA was checked using a NanoDrop Spectrophotometer (Thermo Scientific). The quality of the RNA was also checked by running 2 µl of RNA on a 1.3% agarose gel containing 1 µg/ml of ethidium bromide. The RNA was adjusted to 1µg/µl to make cDNA using the Applied Biosystems High Capacity RNA-to-cDNA Kit (4387406). The concentration of the cDNA was determined using the NanoDrop spectrophotometer (Thermo Scientific).

4.3.3.2 Expression level of the transgene

To quantify the expression level of the transgene, semi quantitative PCR reactions were performed. The reactions were conducted using 1 µl (700 ng) cDNA as a template,

1 µl (10 pmole) of each primer, 7 µl of nuclease-free water, and 10 µl of GoTaq® Hot Start Green Master Mix (Promega M512C). The forward (5'-CGGCAGAGAAACAATGGAAC-3') and reverse (5'-AGTACCGACCACCACCCATA-3') primers amplified both the native *AtLEC1* and the transgene *BnLEC1*. The expression level of *LEC1* was quantified relative to a stably expressed housekeeping gene encoding a protein phosphatase 2A subunit, PDF PP2A (AT1G13320) which was amplified by the forward (5'-TAACGTGGCCAAAATGATGC-3') and reverse (5'-GTTCTCCACAACCGCTTGGT-3') primer set (Czechowski et al., 2005). The linear range of the logarithmic PCR curve was determined by stopping the reaction at 20, 25, 30, 35, 40 cycles for both genes (*LEC1* and *PDF PP2A*). The PCR conditions for *LEC1* were performed with an initial denaturation cycle of 94°C for 1 minute, 32 cycles consisting of denaturation at 94°C for 30 seconds, annealing at 57°C for 30 seconds, elongation at 72°C for 30 seconds, and a final 10 minute extension at 72°C. The PCR program for the *PDF PP2A* involved an initial denaturation cycle of 94°C for one minute, 32 cycles consisting of denaturation at 94°C for 30 seconds, annealing at 57°C for 30 seconds, elongation at 72°C for 30 seconds, and a final 10 minute extension at 72°C.

The PCR product was run on a 1% agarose gel that contained 1 µg/ml of ethidium bromide. An image analyzer (Alpha Innotech Technology) was used to visualize and quantify the amplification products. The background subtraction setting was set to "autobase" to account for darkness or background variation within the gel. The *LEC1* was quantified as a percentage of the housekeeping gene *PDF PP2A*.

4.3.3.3 Statistical analysis of semi quantitative PCR results

The average for the first trial was obtained from three biological replicates (plants) and seven technical replicates (repeats of PCR), whereas the average for the second trial consisted of three biological replicates and three technical replicates. The analysis of means applied Least Significant Difference and Duncan's multiple range tests to determine if there were significant differences among the relative transcript levels between the transgenic and wild type lines. A comparison of means was performed using the SAS® 9.2 program.

4.3.4 Phenotypic analysis of homozygous plants

4.3.4.1 Oil composition analysis

Ten sub-lines developed from each of the *Arabidopsis* lines were used as replicates for the oil composition analysis. Each replicate consisted of seeds pooled from two plants in order to obtain a total of 100 ± 5 mg of T4 seeds. Oil was extracted using a modified protocol that was initially developed for rapeseed (Hougen and Bodo, 1973). Extraction of the oil breaks the ester bond between the glycerol and fatty acid, and facilitates the transesterification reaction of the triglyceride to the fatty acid methy esters (Hougen and Bodo, 1973). The solution was then inserted into autosampler vials containing a polypropylene insert. The fatty acid profile of each replicate was then determined using Gas Chromatography (Varian model 3900) which contained ultra high purity helium as the carrier gas. The fatty acids were separated based on chain length and degree of saturation. A flame ionization detector and the Varian Star Workstation software were used to measure the peak areas of each fatty acid. The fatty acid compositions of the transgenic seeds were compared to the wild type control.

4.3.4.2 Determination of the number of seeds per silique and the number of siliques per plant

Five siliques were randomly selected from three plants of each transgenic line and used to count the number of seeds per silique. The number of fully developed siliques per plant was counted from three plants of each *Arabidopsis* line.

4.3.4.3 Statistical analysis of phenotypic results

The data obtained from the oil composition analysis, number of seeds per silique and number of siliques per plant were analyzed using the SAS® 9.2 program. Duncan's multiple range and Least Significant Difference (LSD) tests were applied to determine the main effect of the mean for each transgenic line and the wild type control group. Both tests were run to ensure that similar results were obtained.

4.4 RESULTS

4.4.1 Segregational analysis of transgenic lines

Seven transgenic lines (Table 4.1) were obtained using the "in planta transformation method." Segregational analysis served as a basis for understanding the number of inserts present in the genome of the lines. Chi square tests were performed to determine if each of the transgenic lines segregated according to the expected 3:1 ratio in the T1 generation (Table 4.2). The segregation ratio of transgenic line A8 was actually tested using a T2 plant, because this line possessed an initial phenotype in the T0 where the seeds did not fill and did not develop uniformly within the silique. As a result, embryo rescue was performed to ensure that viable plants would be available for the next generation. Due to this problem, it was impossible to determine the segregation ratio of

the T1 seeds. Interestingly, the siliques of T1 plants developed seeds uniformly, thus allowing T2 seeds to be used for the segregational analysis.

Table 4.1. Summary of the seven transgenic lines that were obtained using the in planta transformation method. Lines A4, A7, A8, A9 were transformed with *BnLEC1* version A, and B3, B6, B7 with *BnLEC1* version B.

Transgenic Line	**Version of *BnLEC1* TDNA Insertion**
A4	*BnLEC1* Version A
A7	*BnLEC1* Version A
A8	*BnLEC1* Version A
A9	*BnLEC1* Version A
B3	*BnLEC1* Version B
B6	*BnLEC1* Version B
B7	*BnLEC1* Version B

Table 4.2. Chi Square tests performed to analyze the segregation ratio of the transformed plants. The analysis was performed on the T1 generation for lines A4, A7, A9, B3, B6, B7 and the T2 generation for line A8. Yate's correction factor was applied to calculate the Chi Square values at 1 df. N.S. indicates non-significant values.

Transgenic Line	Generation	Observed Ratio of Kanamycin Resistance : No Resistance	Expected Ratio	Chi Square Value
A4	T1	58 : 11	51.75 : 17.25	2.555 $^{N.S.}$
A7	T1	40 : 13	39.75 : 13.25	0.006 $^{N.S.}$
A8	T2	32 : 9	30.75 : 10.25	0.073 $^{N.S.}$
A9	T1	69 : 18	65.25 : 21.75	0.648 $^{N.S.}$
B3	T1	100 : 24	93 : 31	1.817 $^{N.S.}$
B6	T1	57 : 15	54 : 18	0.463 $^{N.S.}$
B7	T1	31 :13	33 : 11	0.273 $^{N.S.}$

The Chi Square tests show that all seven transgenic lines have a calculated value lower than the standardized value of 3.84 (Table 4.2). These results demonstrate that the transgenic lines segregated according to the expected 3:1 segregation ratio, thereby suggesting the presence of a single copy of *BnLEC1*.

4.4.2 Genotypic analysis of the transformed lines

A genotypic analysis was conducted to verify the presence of the *BnLEC1* TDNA insertion within the *Arabidopsis* genome by PCR. The results demonstrated that both the 35S promoter and *BnLEC1* were integrated into the *Arabidopsis* genome. This verified the presence of the *BnLEC1* transgene in plants displaying kanamycin resistance. Figure

4.2 shows the PCR results demonstrating the presence of *BnLEC1* within each of the transgenic lines, and the absence of the transgene in the wild type and water controls.

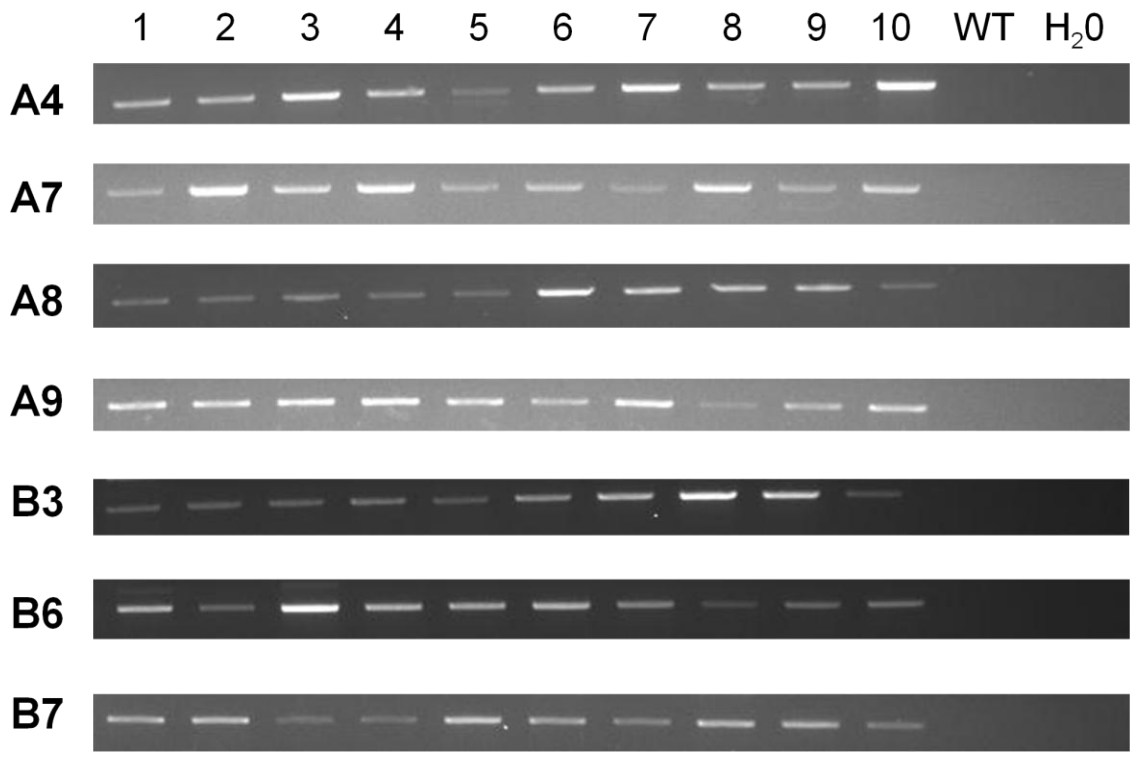

Figure 4.2. Genotypic analysis of the transgenic lines A4, A7, A8, A9, B3, B6, and B7. Lanes 1 to 10 show the 10 sub-lines, lane 11 is a wild type control and lane 12 is the water blank. DNA extracted from each line was used as a template in a PCR reaction using a 35S promoter and a *BnLEC1* specific primer. The amplified product indicates the presence of the transgene.

4.4.3 Expression analysis of the transgene

Two independent trials were conducted to determine the expression level of *BnLEC1* in the *Arabidopsis* lines. The first trial used RNA extracted from both siliques and leaves. In this initial trial, expression of *LEC1* was not observed in the leaves. As a result, only the RNA extracted from siliques was utilized in the second trial. No statistically significant differences were observed in the first trial (Figure 4.3). Large

variations between technical and biological replicates prompted the execution of a second trial. Within the second trial, the siliques were measured using a more stringent criterion to ensure consistency in development. The statistical analysis performed on the data obtained from the second trial showed that the expression level of *LEC1* in the transgenic line B7 was significantly higher compared to the wild type line (Figure 4.4). Because of the inconsistent results among the two trials, and the small increase in expression level exhibited by line B7, it cannot be concluded that the observations were valid. Therefore, it was assumed that the TDNA insertion does not significantly affect the overall expression level of *LEC1* in the transformed lines.

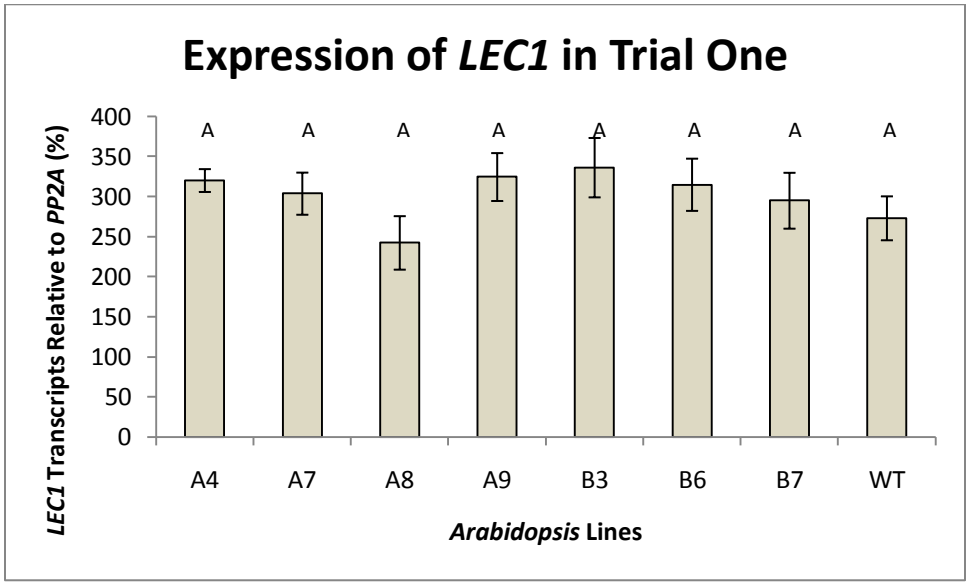

Figure 4.3. The expression level of *LEC1* in the transgenic lines A4, A7, A8, A9, B3, B6, B7 and wild type (WT) line in first trial. The analysis was conducted using siliques collected 7 days after pollination. Values ± SE are averages of three biological replicates and seven technical replicates. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

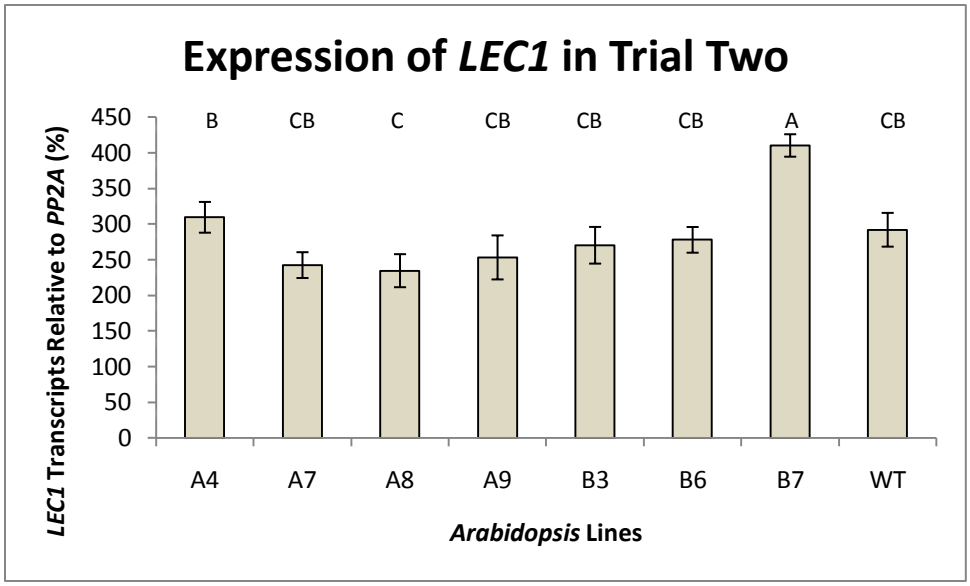

Figure 4.4. The expression level of *LEC1* in the transgenic lines A4, A7, A8, A9, B3, B6, B7 and wild type (WT) line in the second trial. The analysis was conducted using siliques collected 7 days after pollination. Values ± SE are averages of three biological replicates and three technical replicates. Uppercase letters indicate Duncan's Grouping used to determine significant differences

4.4.4 Phenotypic analysis of the transformed plants

In order to analyze the role of the *Brassica napus LEC1* on fatty acid biosynthesis, as well as embryo and seed development, several phenotypic analyses were conducted.

4.4.4.1 Oil composition analysis

Oil composition analysis revealed numerous significant differences among the fatty acid profiles of the transgenic lines when compared to the wild type line. These results are summarized in Figures 4.11 to 4.22. Transgenic line A4 showed a decrease in palmitoleic acid (0.262% ± 0.100), whereas line A7 displayed higher levels of linolenic acid (18.862% ± 0.146) and eicosadenoic acid (1.898% ± 0.017), and reduced levels of

oleic acid (13.734% ± 0.160) (Figures 4.5, 4.8, 4.10 and 4.13). Transgenic line A8 accumulated more linolenic acid (18.250% ± 0.103), and displayed a decrease in docosadenoic acid (0.431% ± 0.037), and oleic acid (14.490% ± 0.182) (Figures 4.8, 4.10 and 4.16). An increase in the amount of linolenic acid (18.949% ± 0.116), eicosenoic acid (20.967% ± 0.114), and eicosadienoic acid (1.894% ± 0.015), but a reduction in oleic acid (13.617% ± 0.115) was noted in line A9 (Figures 4.8, 4.10, 4.12 and 4.13). Transgenic line B3 displayed numerous changes which included higher levels of palmitic acid (7.856% ± 0.038), linoleic acid (27.884% ± 0.151), eicosadenoic acid (1.951% ± 0.018), and erucic acid (1.989% ± 0.195). A reduction in the amount of stearic acid (3.326% ± 0.038) and oleic acid (13.196% ± 0.138) was observed in this line (Figures 4.5, 4.7, 4.8, 4.9, 4.13 and 4.15). Line B6 accumulated more stearic acid (3.835% ± 0.050), oleic acid (17.550% ± 0.442), and linoleic acid (28.642 ± 0.181%) but less palmitic acid (7.342 ± 0.054%), linolenic acid (15.913% ± 0.235), arachidic acid (2.299% ± 0.043), eicosenoic acid (19.623% ± 0.215), eicosadenoic acid (1.423% ± 0.052), behenic acid (0.335% ± 0.018) and erucic acid (1.268% ± 0.042) (Figures 4.5, 4.7, 4.8, 4.9, 4.10, 4.11, 4.12, 4.13, 4.14 and 4.15). A reduction in stearic acid (3.405% ± 0.010) and oleic acid (14.478% ± 0.137) was observed in line B7 (Figures 4.7 and 4.8).

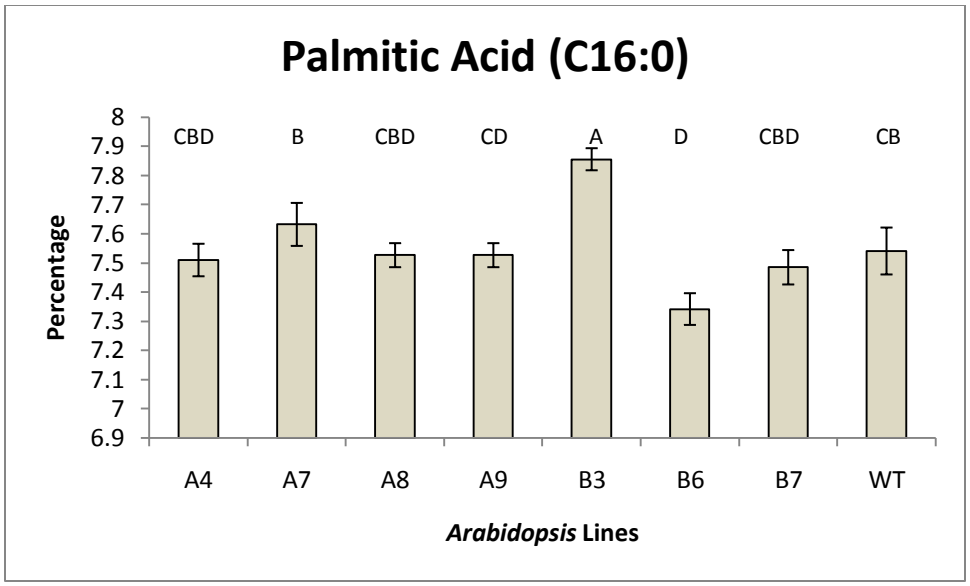

Figure 4.5. Palmitic acid (C16:0) level in the seven transgenic *Arabidopsis* lines (A4, A7, A8, A9, B3, B6, and B7) and the wild type (WT) line. Values ± SE are averages based on 10 replicates that contained seeds pooled from two plants. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

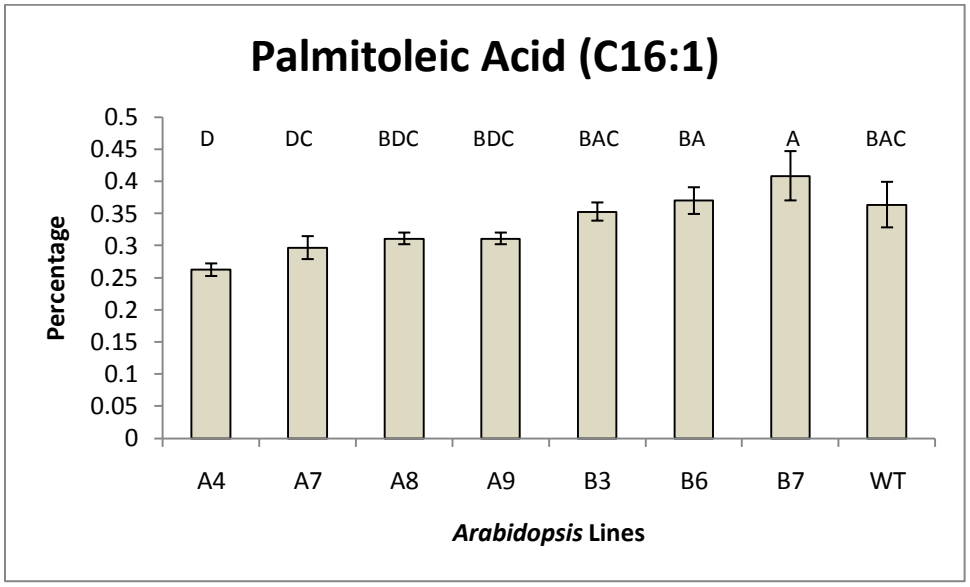

Figure 4.6. Palmitoleic acid (C16:1) level in the seven transgenic *Arabidopsis* lines (A4, A7, A8, A9, B3, B6, and B7) and the wild type (WT) line. Values ± SE are averages based on 10 replicates that contained seeds pooled from two plants. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

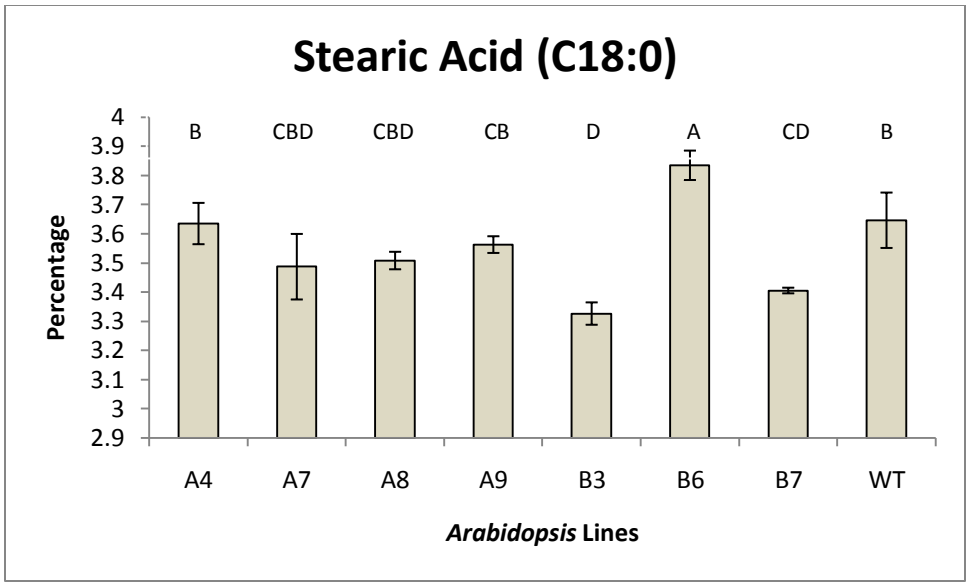

Figure 4.7. Stearic acid (C18:0) level in the seven transgenic *Arabidopsis* lines (A4, A7, A8, A9, B3, B6, and B7) and the wild type (WT) line. Values ± SE are averages based on 10 replicates that contained seeds pooled from two plants. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

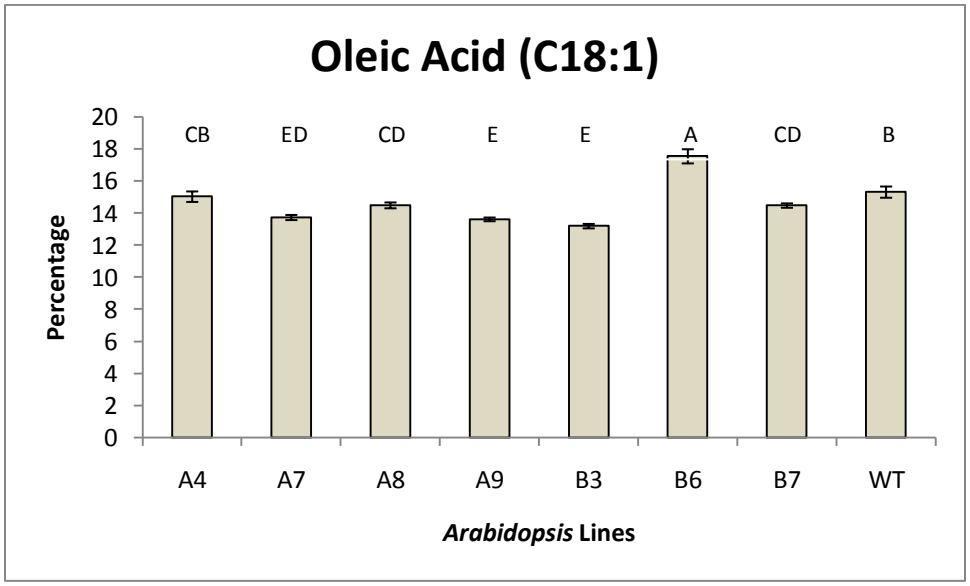

Figure 4.8. Oleic acid (C18:1) level in the seven transgenic *Arabidopsis* lines (A4, A7, A8, A9, B3, B6, and B7) and the wild type (WT) line. Values ± SE are averages based on 10 replicates that contained seeds pooled from two plants. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

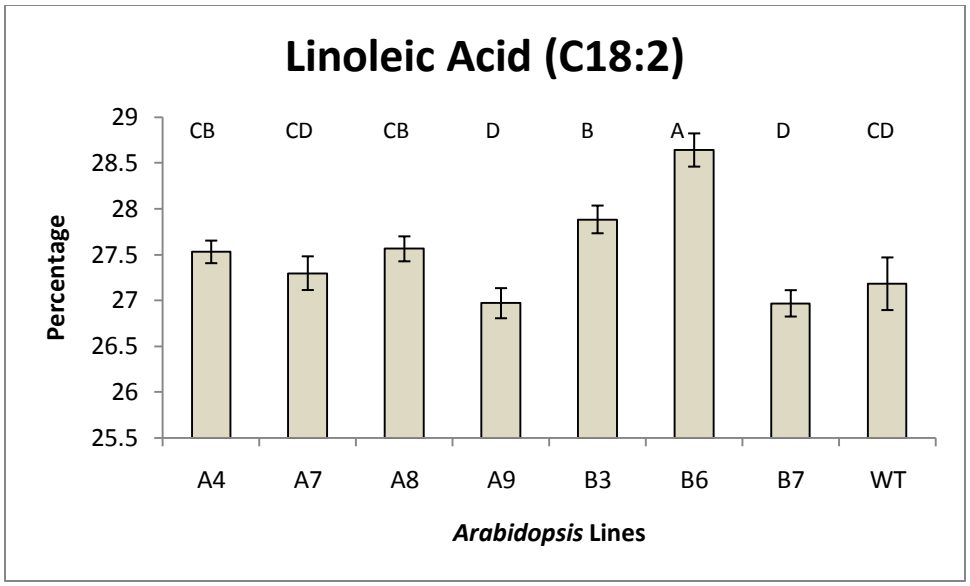

Figure 4.9. Linoleic acid (C18:2) level in the seven transgenic *Arabidopsis* lines (A4, A7, A8, A9, B3, B6, and B7) and the wild type (WT) line. Values ± SE are averages based on 10 replicates that contained pooled seeds from two plants. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

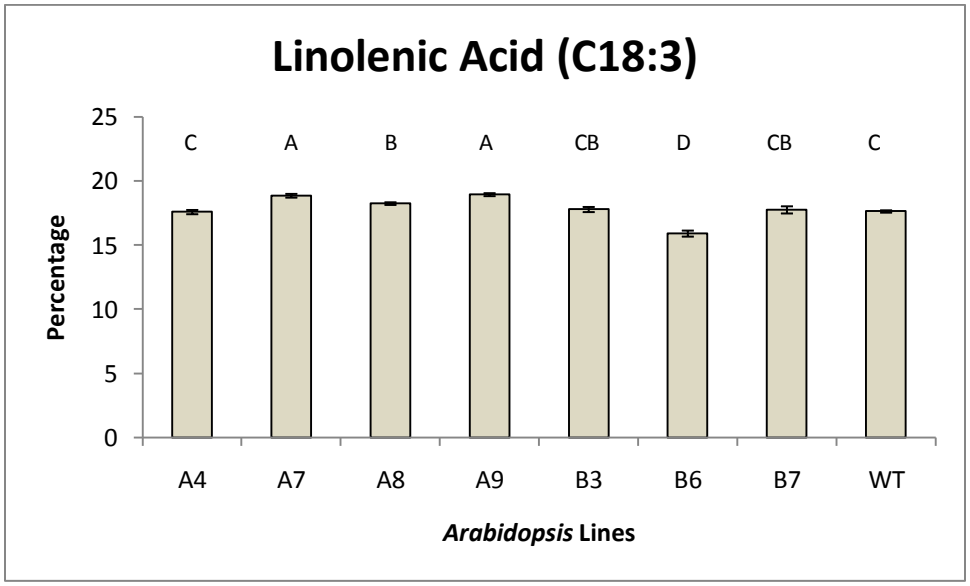

Figure 4.10. Linolenic acid (C18:3) level in the seven transgenic *Arabidopsis* lines (A4, A7, A8, A9, B3, B6, and B7) and the wild type (WT) line. Values ± SE are averages based on 10 replicates that contained pooled seeds from two plants. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

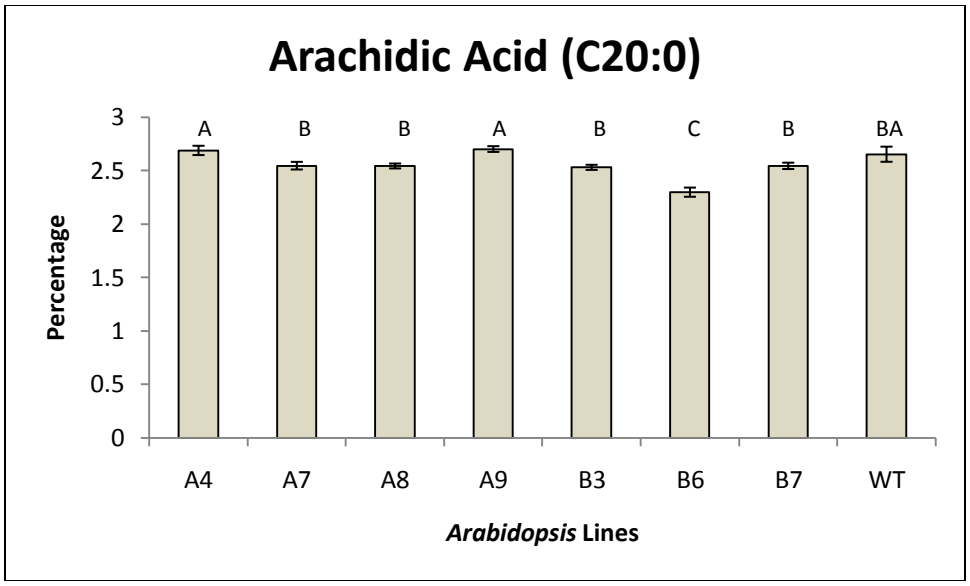

Figure 4.11. Arachidic acid (C20:0) level in the seven transgenic *Arabidopsis* lines (A4, A7, A8, A9, B3, B6, and B7) and the wild type (WT) line. Values ± SE are averages based on 10 replicates that contained seeds pooled from two plants. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

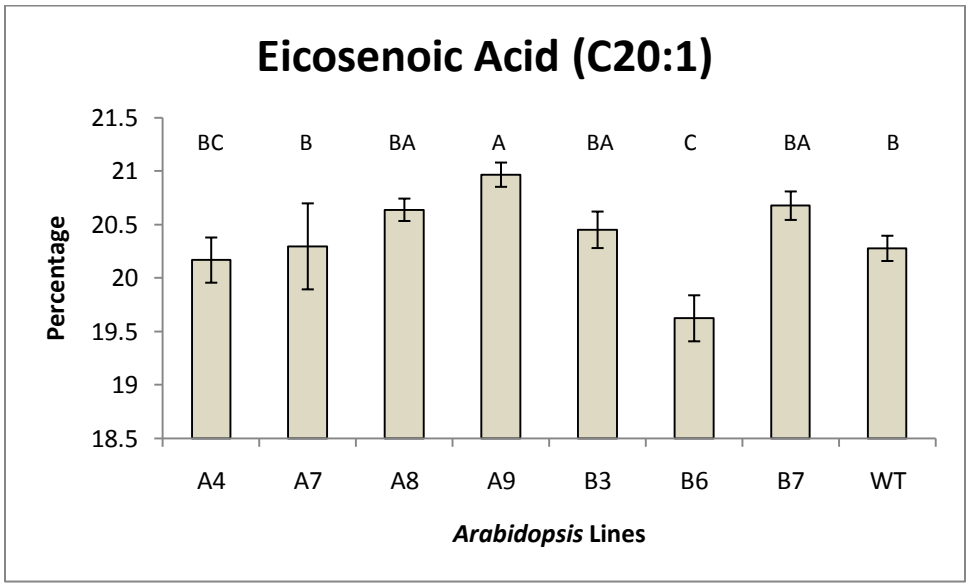

Figure 4.12. Eicosenoic acid (C20:1) level in the seven transgenic *Arabidopsis* lines (A4, A7, A8, A9, B3, B6, and B7) and the wild type (WT) line. Values ± SE are averages based on 10 replicates that contained pooled seeds from two plants. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

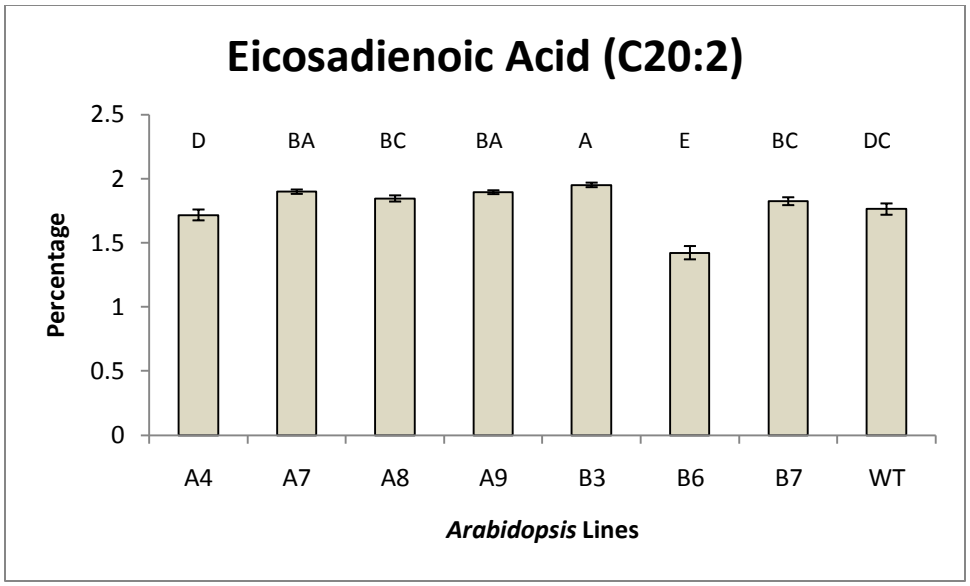

Figure 4.13. Eicosdienoic acid (C20:2) level in the seven transgenic *Arabidopsis* lines (A4, A7, A8, A9, B3, B6, and B7) and the wild type (WT) line. Values ± SE are averages based on 10 replicates that contained pooled seeds from two plants. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

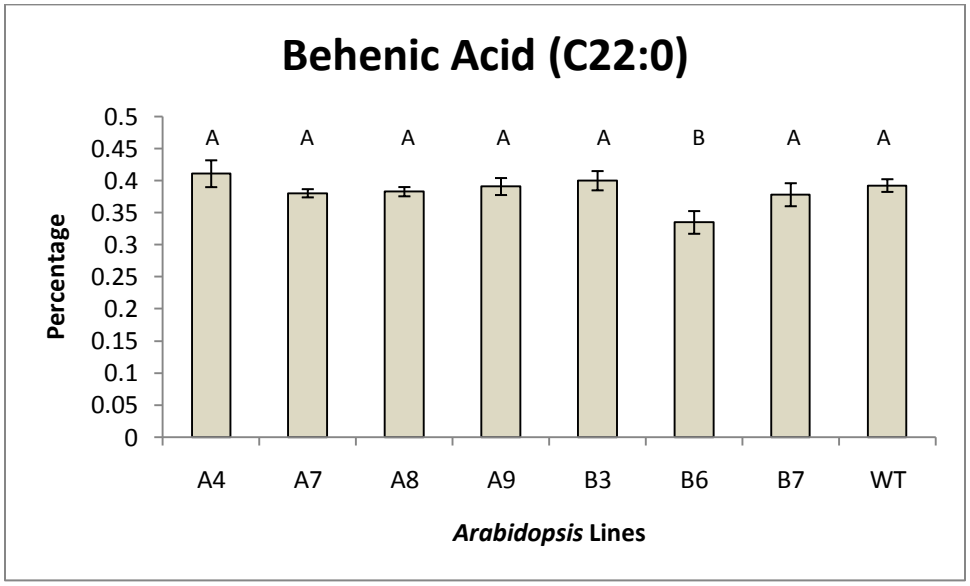

Figure 4.14. Behenic acid (C22:0) level in the seven transgenic *Arabidopsis* lines (A4, A7, A8, A9, B3, B6, and B7) and the wild type (WT) line. Values ± SE of each are averages based on 10 replicates that contained pooled seeds from two plants. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

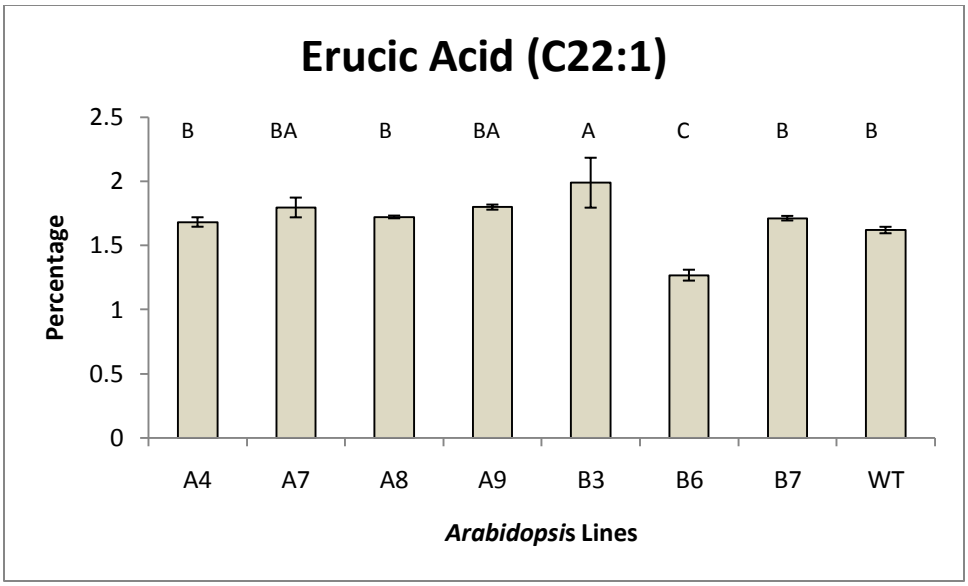

Figure 4.15. Erucic acid (22:1) level in the seven transgenic *Arabidopsis* lines (A4, A7, A8, A9, B3, B6, and B7) and the wild type (WT) control group. Values ± SE are averages based on 10 replicates that contained pooled seeds from two plants. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

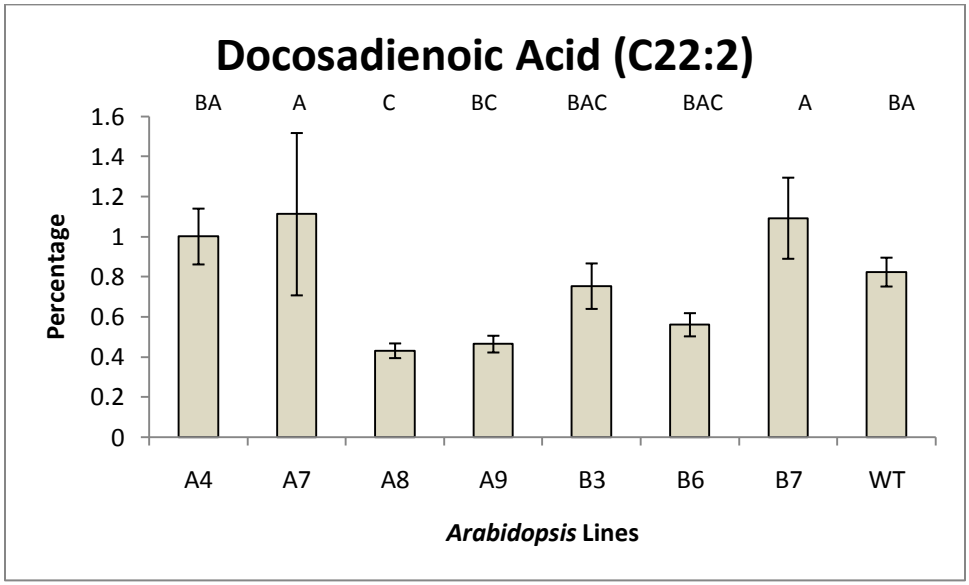

Figure 4.16. Docosadienoic acid (C22:2) level in the seven transgenic *Arabidopsis* lines (A4, A7, A8, A9, B3, B6, and B7) and the wild type (WT) line. Values ± SE are averages based on 10 replicates that contained seeds pooled from two plants. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

The results of the oil composition analysis indicate significant variation among the fatty acid profile of the transgenic lines. However, there is no correlation between the variation in seed oil composition and the expression level of the *BnLEC1* transgene.

4.4.4.2 Number of seeds per silique and number of siliques per plant

Duncan's analysis of means and the LSD tests demonstrated that the average number of seeds per silique in the transgenic line B3 was lower than the wild type line (Figure 4.17). This was interesting as the siliques of line B3 appeared to be consistently shorter in length relative to the wild type line. No statistically significant differences in the number of siliques per plant were observed among the *Arabidopsis* lines (Figure 4.18).

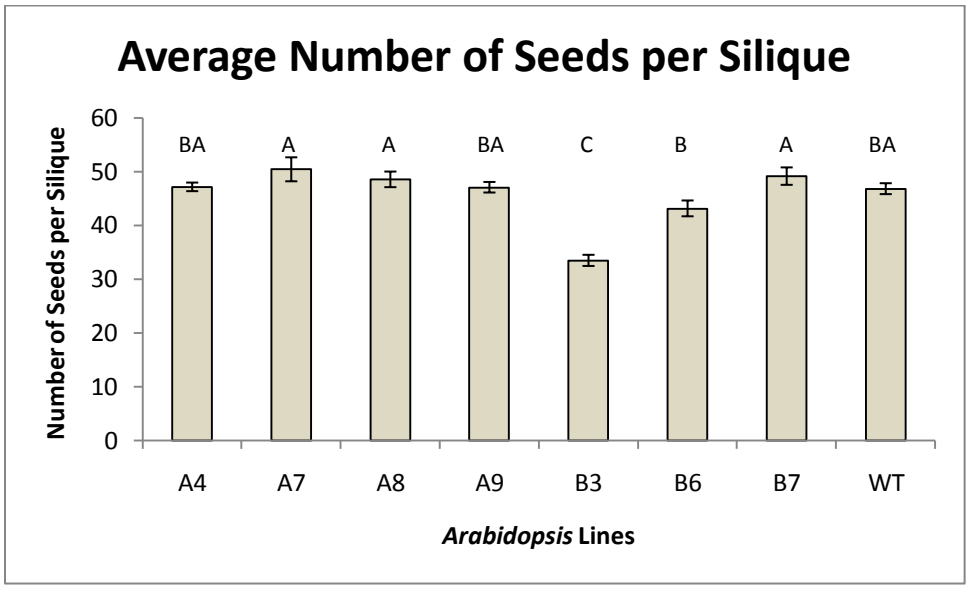

Figure 4.17. Number of seeds per silique in the seven transgenic lines (A4, A7, A8, A9, B3, B6, B7) and the wild type (WT) line. Values ± SE are averages obtained from five siliques of three plants within each line. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

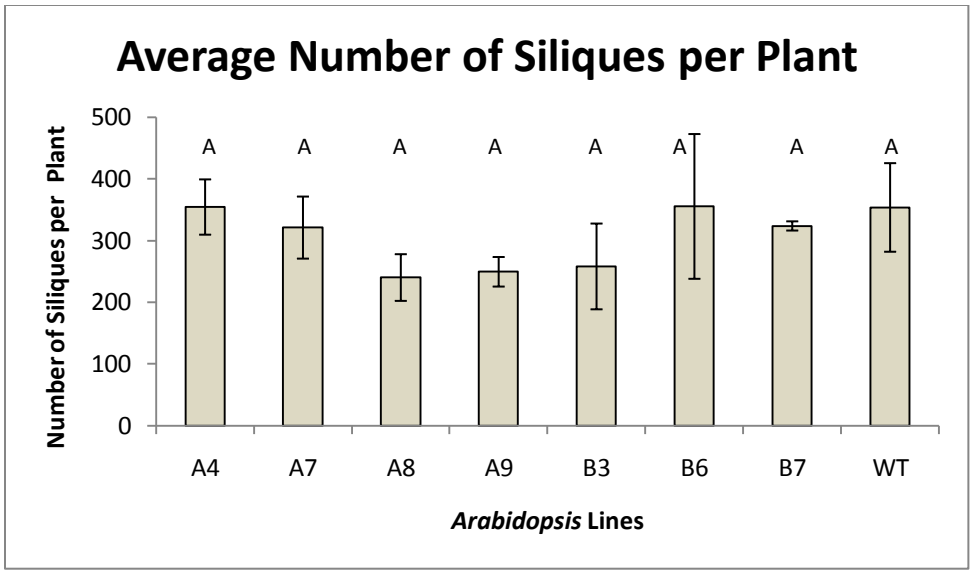

Figure 4.18. The average number of siliques per plant in the seven transgenic lines (A4, A7, A8, A9, B3, B6, B7) and the wild type (WT) line. Values ± SE are averages of the total number of fully developed siliques counted from three plants of each *Arabidopsis* line. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

4.5 DISCUSSION

To further understand the role of *BnLEC1* during seed development, in planta transformation was used to generate seven transgenic *Arabidopsis* lines. PCR analysis confirmed the presence of the 35S promoter and *BnLEC1* gene. Expression analysis revealed that the transgene was not expressed at significant levels within the transformed *Arabidopsis* lines. Interestingly, the transgenic lines still displayed differences in seed oil composition. As a result, a number of possible explanations have been explored to account for this discrepancy.

The technique of semi-quantitative PCR may not be sensitive enough to detect the expression level of *BnLEC1* within the *Arabidopsis* lines (Klein, 2002; Bustin et al.,

2005). The primer set used to identify the expression of *LEC1* did not distinguish between the native *AtLEC1* and the transgene *BnLEC1*. As a result, there is the possibility that a concomitant increase in *BnLEC1* expression and a decrease in *AtLEC1* expression would not be detected by the primers. Because LEC1 is a prominent transcription factor regulating embryogenesis (Meinke, 1992; West et al., 1994; Harada, 2001), even slight increases in expression may be significant enough to alter the seed oil composition of the *Arabidopsis* lines.

LEC1 is considered to be a major regulator of seed maturation and dormancy (Lotan et al., 1998; Braybrook and Harada, 2008). Because of this function, in planta transformation may not have been the best method of developing transgenic plants. The process of in planta transformation targets the floral organs of *Arabidopsis*, and the seeds of the transformed plants are screened on medium containing kanamycin. One particular concern is that the *BnLEC1* transgene may induce prolonged dormancy, thereby preventing germination and selection of transformed seeds. If this is true, it can be assumed that only seeds with very weak *BnLEC1* expression (not detectable by PCR) were able to germinate in the presence of kanamycin.

Another factor to consider is that plants possess gene silencing mechanisms, which are capable of repressing the expression of transgenes inserted into their genome (Dehio and Schell, 1994; Waterhouse et al., 2001). Gene silencing has challenged many scientists working with different plant species. Matzke et al. (1989) transformed tobacco plants with two different TDNA insertions, but one of the transgenes expression was suppressed (Matzke et al., 1989). The authors suspected that the significant homology among the two transgenes was the cause of gene silencing (Matzke et al., 1989).

Cosuppression is another form of transgene silencing, which was observed in transgenic petunia plants (Napoli et al., 1990; Van der Krol et al., 1990). *Petunia hybrida* plants were transformed with a gene influencing the intensity of flower pigmentation (Napoli et al., 1990; Van der Krol et al., 1990). Some of transgenic plants were characterized by an array of flower colors, whereas some plants lacked pigmentation (Napoli et al., 1990; Van der Krol et al., 1990). This led to the notion that the silencing of the transgene prevented the expression of the endogenous pigmentation gene (Napoli et al., 1990; Van der Krol et al., 1990). Both studies suggest that different degrees of gene silencing can occur among plants of different species. Currently there is no reliable method for preventing this event, but some of the potential factors leading to gene silencing have been identified (Angell and Baulcombe, 1997).

Gene silencing often occurs because of shared homology among a transgene and other genes within the plant genome. This typically happens when there is significant homology between the transgene and a related endogenous gene, or when multiple copies of TDNA insertions are present in the genome (Napoli et al., 1990; Van der Krol et al., 1990; Angell and Baulcombe, 1997; Stam et al., 1998). *BnLEC1* and *AtLEC1* share more than 73% amino acid similarity, which could have been a factor leading to silencing in the *Arabidopsis* lines.

Different mechanisms of gene silencing have been identified in plants. Transcriptional gene silencing occurs as a result of sequence specific DNA methylation (Stam et al., 1998; Cogoni and Macino, 1999). DNA methylation often occurs in the promoter region of genes, and interferes with the binding of transcription factors (Cogoni and Macino, 1999). Post-transcriptional gene silencing differs from transcriptional gene

silencing because the genes are normally transcribed into transcripts, which are then rapidly degraded (Cogoni and Macino, 1999). Both mechanisms are capable of silencing the expression of transgenes, and could have suppressed the expression of *BnLEC1* in *Arabidopsis.*

The expression level of a transgene is also greatly affected by the location of its insertion within the chromosome (Matzke and Matzke, 1998). Different chromosomal regions are characterized by varying levels of gene expression (Matzke and Matzke, 1998). There are highly active regions of the chromosome, such as the euchromatin, where numerous genes are continuously transcribed (Herman et al., 1990; Yan and Boyd, 2006). Conversely, repetitive sequences occurring within the heterochromatin, tend to suppress gene expression (Pröls and Meyer, 1992; Yan and Boyd, 2006). The sequence of the DNA flanking the TDNA insertion can also play a role in initiating gene silencing (Matzke and Matzke, 1998). Therefore, the location of transgene insertion can have a significant impact upon the expression level of the gene.

The fact that *LEC1* has only been identified in developing seed tissues, might have contributed to the suppression of the transgene (West et al., 1994; Lotan et al., 1998). The construct used to transform *Arabidopsis* contained a constitutive 35S promoter, which drives the expression of *BnLEC1* in tissues where endogenous *LEC1* transcripts are not present (Odell et al., 1985). This might have caused the suppression of *BnLEC1* in tissues and organs which do not normally express *LEC1*. Therefore, a 35S promoter may not be ideal for over-expressing *LEC1* in *Arabidopsis.* Mu et al. (2008) expressed both the *Arabidopsis* and *Brassica napus LEC1* genes under the control of an estradiol-inducible promoter. In the presence of estradiol, the transgenic seedlings grew

poorly and were characterized by yellow cotyledons which often did not develop into true leaves (Mu et al., 2008). Perhaps a seed specific promoter would be beneficial for over-expressing *BnLEC1* by increasing the accumulation of its transcript in tissues where the native gene is normally expressed.

The causes of transgene silencing have been explored thoroughly in this research. Despite the fact that the transgene was not expressed at significant levels, phenotypic differences were still observed among the *Arabidopsis* lines. It is possible that the transgene was incorporated into a region of the chromosome, where the flanking DNA regulated some aspect of embryo development or fatty acid biosynthesis. This "interference" could potentially influence the seed oil composition of the *Arabidopsis* lines. The significant differences in fatty acid profiles could also be a reflection of the natural variation in seed oil composition among *Arabidopsis* plants. The initial parent plants could have differed in seed oil composition, and these differences might have been inherited by the subsequent generations. The observed phenotypic differences may have occurred as a result of the groupings of the *Arabidopsis* lines.

Previous research has documented that the over-expression of *LEC1* in *Arabidopsis* induces changes in fatty acid composition and embryo development (Lotan et al., 1998; Mu et al., 2008). Perhaps the function of *LEC1* is too influential during seed maturation, making it an undesirable gene to over-express. It was initially thought that the over-expression of *BnLEC1* would trigger a series of changes among the downstream transcription factors (LEC2, FUS3, WRI1 and ABI3) leading to enhanced oil biosynthesis (Baud et al., 2008; Mu et al., 2008; Santos-Mendoza et al., 2008). Specifically, FUS3 and WRI1 are directly involved in the regulation of glycolytic

enzymes which establish a carbon pool that is utilized in fatty acid biosynthesis (Baud et al., 2007). Perhaps one of these transcription factors would be a stronger candidate for transformation studies.

Unfortunately, the functional role of *BnLEC1* could not be characterized due to the insignificant expression level of the transgene. Despite these unexpected results, this area of research still appears to be promising in providing tools for the manipulation of seed oil biosynthesis.

5.0 CHAPTER 3: EXPRESSION LEVELS OF SEED DEVELOPMENT TRANSCRIPTION FACTORS IN *BRASSICA NAPUS*

5.1 ABSTRACT

LEAFY COTYLEDON1 (LEC1), LEAFY COTYLEDON2 (LEC2), FUSCA3 (FUS3), and WRINKLED1 (WRI1) are transcription factors that regulate embryo and seed development in *Arabidopsis*. Fatty acids and triacylglycerols (TAGs) accumulate during the maturation phase of embryogenesis. As a result, recent studies suggest that these transcription factors influence the synthesis and accumulation of oil during seed development. In order to further characterize the role of LEC1, LEC2, FUS3 and WRI1, this research focuses on determining if there is a relationship between the expression of these genes and total seed oil content in canola. Semi-quantitative PCR analysis was used to determine the transcript level of *LEC1, LEC2, FUS3* and *WRI1* at 7, 14, 21, and 28 days after pollination in double haploid *Brassica napus* lines characterized by different levels of seed oil content (ranging from 42% to 50%). Results from this work indicate that the differences in oil content are not related to the differential expression of the genes analyzed. This implies the existence of other factors regulating fatty acid biosynthesis and TAG accumulation during seed development.

5.2 INTRODUCTION

Embryo growth and maturation are controlled by a network of transcription factors including, LEAFY COTYLEDON1 (LEC1), LEAFY COTYLEDON2 (LEC2), FUSCA3 (FUS3), and WRINKLED1 (WRI1) (Meinke, 1992; Meinke et al., 1994; Stone et al., 2001; Cernac and Benning, 2004; Gazzarrini et al., 2004). Previous work showed that the transcriptional changes induced by *LEC1, LEC2, FUS3* and *WRI1* during

different stages of embryo development, may affect the regulation of the metabolic pathways involved in seed oil biosynthesis (Mu et al., 2008; Santos-Mendoza et al., 2008; Sharma et al., 2008; Liu et al., 2009). This observation has established the baswork, whichent work which further addresses the role played by these transcription factors during the accumulation of seed storage products (Santos Mendoza et al., 2005; Mu et al., 2008; Sharma et al., 2008; Liu et al., 2009).

Multiple studies have demonstrated that LEC1 may be a factor involved in regulating fatty acid biosynthesis and TAG accumulation during seed development. Sharma et al. (2008) developed two independent transgenic lines over expressing diacylglycerol acyltransferase (DGAT) and observed a concomitant change in seed oil content and *LEC1* expression in the transformed plants. Another study by Mu et al. (2008) showed that the over-expression of *LEC1* in *Arabidopsis* increased the accumulation of fatty acids in both seeds and seedlings. These results were possibly due to the role exercised by *LEC1* on the expression of genes such as LEC2, FUS3 and WRI1 (Baud et al., 2008; Mu et al., 2008; Santos-Mendoza et al., 2008; Liu et al., 2009; Weselake et al., 2009).

Arabidopsis plants over-expressing *BnWRI1* exhibited a 10% to 40% increase in total seed oil content, and enlarged seed size (Liu et al., 2009). Numerous studies suggest that WRI1 affects fatty acid biosynthesis through the regulation of genes encoding enzymes involved in late glycolysis (such as pyruvate kinase) (Baud et al., 2008; Santos-Mendoza et al., 2008). These enzymes have been shown to influence TAG accumulation through the regulation of the carbon source available for *de novo* fatty acid biosynthesis (Baud et al., 2008; Baud and Lepiniec, 2009; Baud et al., 2009; Liu et al., 2009).

Correlations have also been made between elevated *FUS3* transcripts and the expression of seed storage protein genes (12S cruciferin and 2S albumin), as well as a number of genes encoding enzymes involved in fatty acid biosynthesis (KASI, KASII, KASIII, pyruvate dehydrogenase, acetyl-CoA carboxylase) (Wang et al., 2007; Mu et al., 2008). However, the direct mechanism by which FUS3 regulates the expression of these genes remains to be determined (Wang et al., 2007).

Another study conducted by Santos Mendoza et al. (2005) demonstrated that the over-expression of *LEC2* induced the accumulation of oil bodies within *Arabidopsis* leaves. This research suggests that LEC2 influences TAG accumulation during seed development (Santos Mendoza et al., 2005). Overall, these results indicate that changes in *LEC1, LEC2, FUS3* and *WRI1* expression may participate in the regulation of fatty acid synthesis and TAG accumulation (Santos Mendoza et al., 2005; Mu et al., 2008; Santos-Mendoza et al., 2008; Sharma et al., 2008; Liu et al., 2009).

To further explore this idea, the changes in relative transcript levels of *LEC1, LEC2, FUS3* and *WRI1* were measured in four double haploid *Brassica napus* lines differing in seed oil content. Expression analysis was conducted in developing seeds (7, 14, 21, and 28 days after pollination) by semi-quantitative PCR. The objective of this study is to determine if there is a relationship between the expression level of these transcription factors and the total seed oil content of the *Brassica napus* lines.

5.3 MATERIALS AND METHODS

5.3.1 Plant material

A double haploid population was developed from a cross between two canola (*Brassica napus*) varieties registered as Polo and Topas. According to the variety registration descriptions, Polo has a high oil content ranging from 46 to 48% (Agriculture Canada, 1987), whereas Topas has a low oil content of 42% (Mycogen Canada, 1994). A cross between Polo and Topas was used to establish a population that segregated for total seed oil content. Microspore-derived embryos from the F1 hybrid (Polo x Topas) were cultured to produce homozygous double haploid lines which had a total seed oil content ranging from 42% to 50%. Two low and two high oil lines were selected from the population, and used as the plant material in this experiment (Table 5.1). The seed oil content was measured by near infrared reflectance spectroscopy (NIRSystem model 6500). This model measured the reflectance energy readings from 6 g of intact seeds. WinISI II software version 1.04a was used to measure and analyze the spectral data in the wavelength range of 400 to 2500 nm. The total oil content of each line is expressed as a percentage of total dry seed mass. The establishment of the four lines was performed by other members of the lab.

Table 5.1. Oil content of the four double haploid *Brassica napus* lines used in this study. Values are expressed as a percentage of the total dry seed mass.

Double Haploid Line	Oil Content (%)
22	43.5
43	42.1
67	50.0
89	50.3

The selected seeds from two high oil lines (89 and 67), and two low oil lines (22 and 43) were grown in flats containing 48 cells filled with LA4 Sunshine potting mix (Sun Gro). The seeds were planted approximately 2 cm into the soil, placed in a growth room set at 22°C with a 16 hour light, 8 hour dark photoperiod. At the two to four leaf stage, the seedlings were transferred to pots (15 cm diameter x 14 cm deep) filled with black top soil, sand and peat moss (2:2:1). The seedlings were placed in a greenhouse where the average temperature was approximately 21°C and the photoperiod consisted of 17 hours light and 7 hours dark.

5.3.2 Collection of developing seeds

Once the plants began to flower, approximately 10 to 15 siliques were self pollinated and tagged with thread. The thread was color coded according to the day, which allowed for the siliques to be collected at the appropriate developmental stages of 7, 14, 21, and 28 days after pollination. To collect the tissue, the siliques were kept on ice and the immature seeds were excised out of the pod using a sterile blade and a surgical needle. The immature seeds were collected in 2 ml RNase free micro centrifuge tubes, placed on dry ice and stored at -80°C. Each pool of tissue included approximately 15 siliques containing about 10 to 15 seeds per silique, which were collected from five plants of the same double haploid line. Three pools of immature seeds served as biological replicates for each double haploid line during the defined developmental stages.

5.3.3 RNA extraction and cDNA synthesis

The immature seeds were ground to a powder using liquid nitrogen and a micro pestle in a 2 ml micro centrifuge tube. The QIAGEN RNeasy® Plant Mini Kit (QIAGEN

74904) was used to extract total RNA from each sample according to the protocol. The RNA was further purified using the QIAGEN RNase-Free DNase set (QIAGEN 79254). The concentration and quality of the RNA was checked using a NanoDrop Spectrophotometer (Thermo Scientific) and by gel electrophoresis (1.3% agarose gel containing 1 µg/ml of ethidium bromide).

The purified RNA was used to synthesize cDNA with a high capacity RNA-to-cDNA kit (Applied Biosystems 4387406). The concentration of the cDNA was determined using the NanoDrop Spectrophotometer (Thermo Scientific) and adjusted to a working concentration of 700 ng/µl. The cDNA was further utilized as the template for a series of semi quantitative PCR reactions.

5.3.4 Selection of transcription factors

Full length genes of *LEC1, LEC2, FUS3* and *WRI1* were isolated from *Brassica napus* based on the sequence similarity of these genes with their *Arabidopsis thaliana* homolgs and expressed sequence tags (ESTs) from *Brassica*. Other members of the lab performed this work, with the exception of *BnLEC1*. A list including the genes utilized in this study, primer sets used to amplify the genes of interest, and size of the amplified products is compiled in Table 5.2. *ACTIN* (AF111812) was utilized as the internal reference gene (Elhiti et al., 2010).

Table 5.2. Set of primers and size of amplification products used to determine the expression level of the selected genes.

Gene	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Fragment Size in Base Pairs
ACTIN	taaagtatccgattgagcatggtat	cgtaggcaagcttctctttaatgtc	455
LEC1	aaacggcagagaaaacaatgg	tcacttatactgaccatactggtc	710
LEC2	ccctttccctcttctaacgc	cagctccattttgcttcaca	729
WRI1	ccgactcaatcagagactcca	aagcaggacaacggagaaga	1100
FUS3	gaaggatgcctagacagaga	agaggagtatcgttggaggt	650

5.3.5 PCR conditions

The PCR reactions contained 1 µl (700 ng) cDNA, 1 µl (10 pmol) of each forward and reverse primer (Table 5.2), 10 µl of GoTaq® Hot Start Green Master Mix (Promega M5123) and 7 µl of nuclease-free PCR water. A PCR-100 Programmable Thermocycler (MJ Research Incorporated) was used to run the semi quantitative PCR reactions for each transcription factor.

Each PCR cycle consisted of an initial denaturation step at 95°C for 5 minutes followed by a cycle of 95°C for 30 seconds, annealing at 55°C for 30 seconds and extension at 72°C for one minute. The cycle was initially repeated 25, 30, 35, and 40 times to determine the linear phase of the logarithmic PCR curve. Based on optimized procedures previously established by other members of the lab, a total of 35 cycles was ideal for amplifying *LEC1* and *LEC2* and 30 cycles for *WRI1*, *FUS3*, and *ACTIN*. The PCR product was run on 1% agarose gel containing ethidium bromide at a final concentration of 1 µg/ml. An image analyzer (Alpha Innotech Technology) was used to expose and visualize the bands under the "auto expose setting". To quantify the area of each band, the background subtraction setting was set to "autobase" to account for

darkness or background variation among the gel. At the various developmental stages, the expression of each gene (*LEC1, LEC2, FUS3,* and *WRI1)* was quantified relative to that of the *ACTIN* gene. Three biological replicates (pools of immature seeds) and three technical replicates (repeats of PCR) were used to measure the expression of each gene at each developmental stage.

5.3.6 Statistical analysis of semi quantitative PCR results

Least Significant Difference and Duncan's multiple range tests were applied to determine significant differences among the relative transcript levels of the genes in the four double haploid lines at 7, 14, 21 and 28 days after pollination. The analysis of means was performed using the SAS® 9.2 program.

5.3.7 Analysis of seed development

The immature seeds were observed using a dissecting microscope to view the developmental stage of the embryos at 7, 14, 21, and 28 days after pollination. This was performed to ensure that the expression analysis was conducted on seeds of similar developmental stages.

5.4 RESULTS

5.4.1 Embryo morphology during silique collection

Globular embryos were first observed 14 days after pollination. At day 21, the immature seeds contained middle cotyledonary embryos which developed further to late cotyledonary embryos by day 28. These embryonic stages appeared to be consistent among the four double haploid lines.

5.4.2 Relative transcript levels of *LEC1, LEC2, FUS3* and *WRI1*

The results of the PCR studies (Figures 5.1, 5.2, 5.3, and 5.4) show the relative expression of each transcription factor at various developmental stages of seed development in the four *Brassica napus* lines characterized by different seed oil contents (Table 5.1)

At day 7 after pollination, the transcript levels of *LEC1* were highest in the lines with low oil content (43 and 22) (Figure 5.1). The transcript levels of this gene increased during the subsequent days, and by day 14 line 43 displayed the highest *LEC1* expression. The low oil lines (22 and 43) had elevated levels *LEC1* transcripts relative to the high oil line (89) at day 21. At day 28 after pollination, lines 22 and 89 exhibited the most pronounced expression level of *LEC1*. No significant differences were observed for lines 67 and 43 (Figure 5.1).

There were no significant differences in *LEC2* expression among the lines at day 7 after pollination (Figure 5.2). By day 14, lines 89 and 67 contained fewer *LEC2* transcripts than line 43. Upon further development, *LEC2* expression decreased in all lines, and at day 21 high levels of *LEC2* transcripts acummulated in line 43. Line 43 displayed the highest expression of *LEC2* 28 days after pollination (Figure 5.2).

The expression level of *FUS3* increased from day 7 to 14 in all lines, especially in line 43 (Figure 5.3). At day 14, *FUS3* expression in this line was significantly higher compared to other lines. Fluctuations in *FUS3* transcript levels were observed among all lines during the following days. A remarkable difference was observed at day 21 when line 22 exhibited the highest expression level (Figure 5.3).

At day 7, lines 67 and 22 displayed increased levels of *WRI1* expression relative to line 43 (Figure 5.4). Line 43 accumulated more transcripts than lines 89 and 22 after 14 days. At day 21, the expression level of *WRI1* was reduced in line 43 compared to lines 22 and 67. No differences were observed among the four lines 28 days after pollination (Figure 5.4).

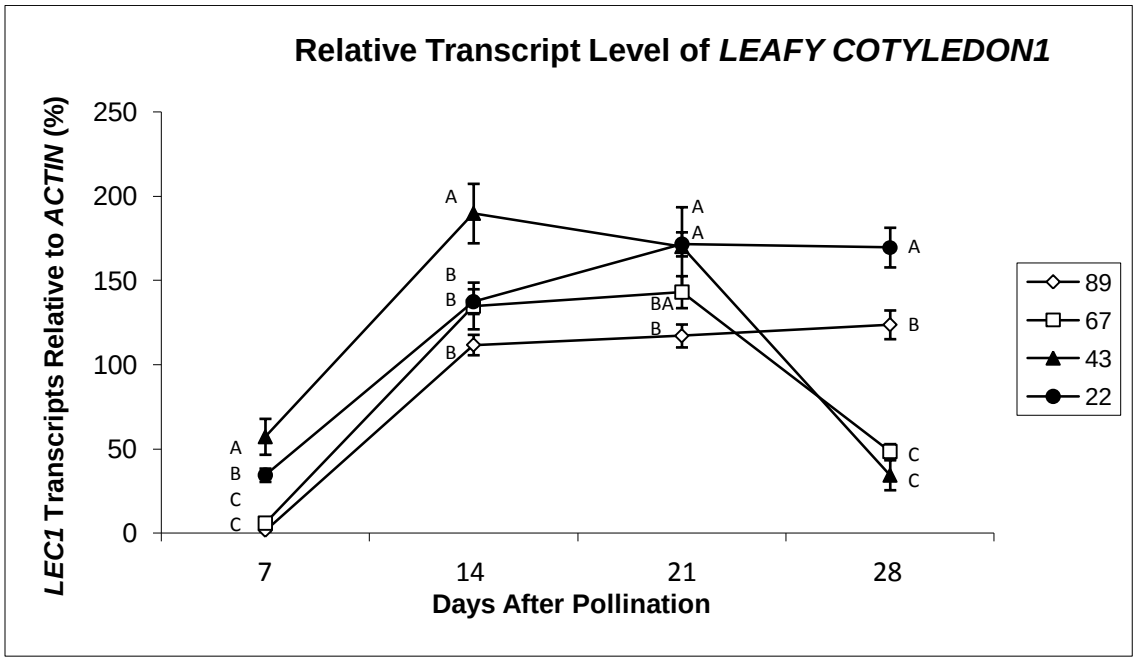

Figure 5.1. The transcript level of *LEC1* expressed as a percentage of the internal reference gene *ACTIN*. Values ± SE are averages based on three biological replicates and three technical replicates. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

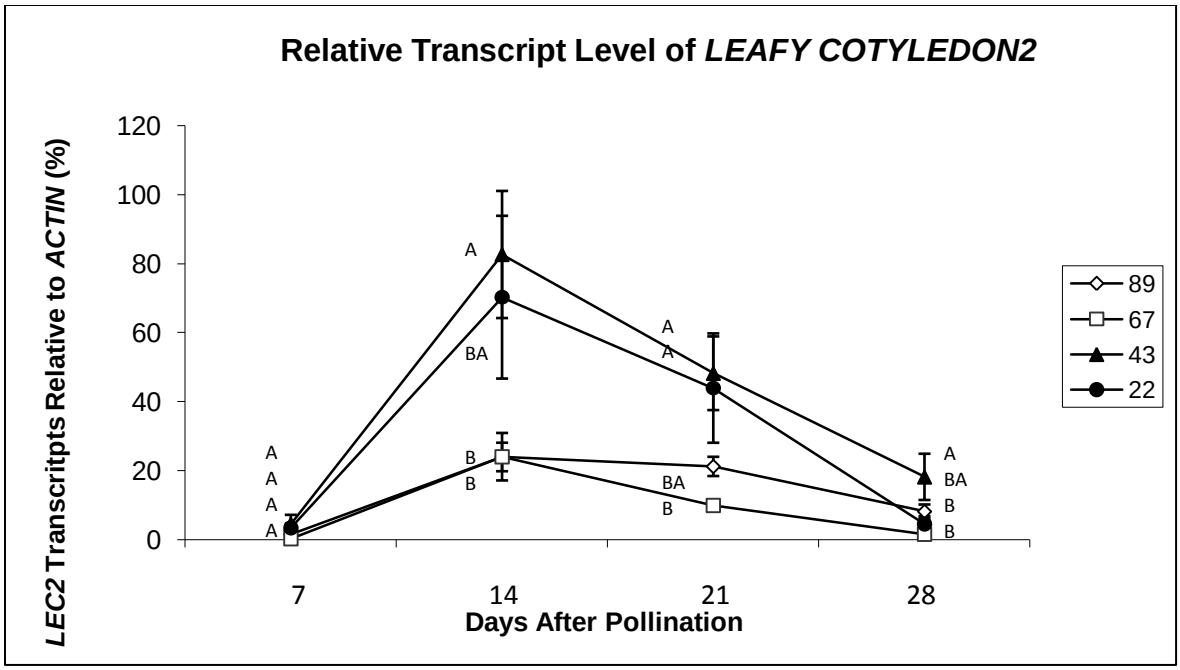

Figure 5.2. The transcript level of *LEC2* expressed as a percentage of the internal reference gene *ACTIN*. Values ± SE are averages based on three biological replicates and three technical replicates. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

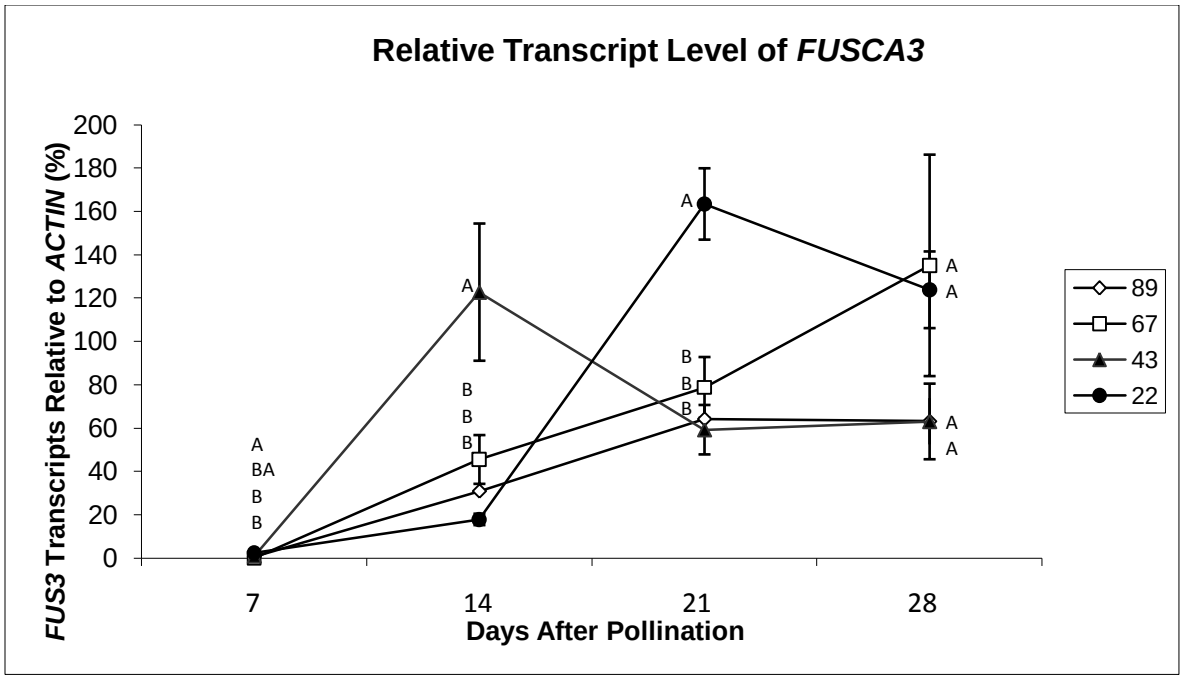

Figure 5.3. The transcript level of *FUS3* expressed as a percentage of the internal reference gene *ACTIN*. Values ± SE are averages based on three biological replicates and three technical replicates. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

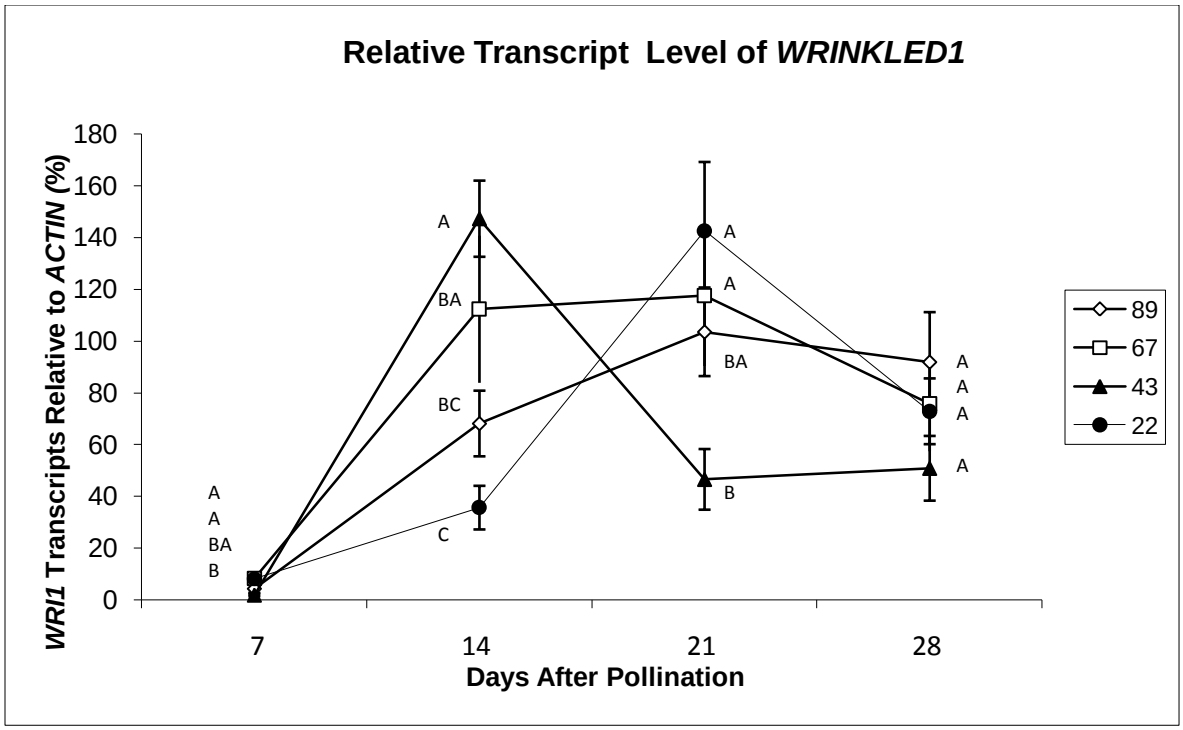

Figure 5.4. The transcript level of *WRI1* expressed as a percentage of the internal reference gene *ACTIN*. Values ± SE are averages based on three biological replicates and three technical replicates. Uppercase letters indicate Duncan's Grouping used to determine significant differences.

5.5 DISCUSSION

The results obtained from this study indicate that several changes in transcript levels of *LEC1, LEC2, FUS3* and *WRI1* occur in the seeds of the doubled haploid *Brassica napus* lines at 7, 14, 21, and 28 days after pollination. The general increase in expression exhibited by the four genes from day 7 to day 14 may be required to sustain the growth of the embryo during this time. In all four lines, embryo development is terminated at day 21, as indicated by the formation of the cotyledonary embryo. During the subsequent days (21 to 28) elongation occurs and storage products continue to accumulate within the seeds.

An interesting result was observed for *LEC1* 7 days after pollination, when the transcript level of this gene was higher in the two double haploid lines characterized by low oil content. During the subsequent days, the expression of *LEC1* did not correlate to the level of oil accumulated by the four lines. This was in contrast to other studies which demonstrated that elevated levels of *LEC1* and *LEC2* expression were associated to increased fatty acid biosynthesis (Santos Mendoza et al., 2005; Mu et al., 2008; Sharma et al., 2008). It is therefore suggested that the increased oil content observed in lines 67 and 89 is not a reflection of elevated *LEC1* or *LEC2* expression, and that other regulatory genes might be important for this process.

The changes in *FUS3* and *WRI* expression level do not correlate with the level of oil accumulated in the *Brassica napus* lines. Furthermore, the transcription profile of these genes appear to be different from that of *LEC1*, which is unexpected given the fact that *LEC1* induces the expression of *LEC2, FUS3* and *WRI1* (Baud et al., 2008; Santos-Mendoza et al., 2008). The results from this study show a simultaneous increase in the expression of *LEC1, LEC2, FUS3* and *WRI1* at day 14 in line 43, and at day 21 in line 22, but not during other developmental stages and in other lines. These findings are not consistent with the regulatory network shown in Figure 2.7 and suggest that the changes in the expression of *FUS3* and *WRI1* are not necessary for the accumulation of seed oil in *Brassica napus.*

Previous studies have shown that *WRI1* is involved in the regulation of genes participating in the glycolytic process, therefore influencing the availability of sucrose that can be utilized in fatty acid biosynthesis (Cernac and Benning, 2004; Liu et al., 2009). A recent publication documented an increase in the oil content of *Arabidopsis*

lines over-expressing *BnWRI1* (Liu et al., 2009). Increased *FUS3* expression has also been associated with increases in oil biosynthetic transcripts in *Arabidopsis* (Wang et al., 2007). Surprisingly, the results obtained from this experiment do not agree with the previous studies. This suggests the presence of additional components involved in various aspects of plant development, which influence the total seed oil content.

In conclusion, this study did not find any correlation between the expression levels of *LEC1, LEC2, FUS3* and *WRI1* and the amount of oil accumulated in *Brassica napus* seeds. These results, which do not agree with previous studies reporting that these transcription factors play a positive role on oil biosynthesis, suggest the presence of other regulatory mechanisms involved in seed oil accumulation.

6.0 GENERAL DISCUSSION AND CONCLUSIONS

The production of canola oil is dependent on the genetic potential of canola varieties or genotypes to produce high yields and seed oil content. The quality of canola oil is determined by its end use, and is dependent upon the fatty acid composition of the seed oil. The process of fatty acid biosynthesis and accumulation of triacylglycerols (TAGs) during seed maturation is genetically regulated, and requires the coordination of multiple biochemical pathways. Many of the genes encoding enzymes involved in these pathways have been thoroughly characterized and have demonstrated a threshold limiting the influence they have upon seed oil biosynthesis. Since the synthesis and storage of oil is correlated with various phases of embryo and seed development, it has been suggested that the genetic regulation of embryo morphogenesis and maturation may influence the synthesis and accumulation of oil during seed development. A new sector of research focuses on determining if a series of genetic regulators or transcription factors, which are known to control embryo maturation, including LEAFY COTYLEDON1 (LEC1), LEAFY COTYLEDON2 (LEC2), FUSCA3 (FUS3) and WRINKLED1 (WRI1), are also implicated in the regulation of fatty acid biosynthesis and accumulation during seed development.

Studies in *Arabidopsis* have shown that LEC1 is a major regulator of embryogenesis and plays a significant role in the synthesis and accumulation of fatty acids during seed development. As a result, two full length clones of *BnLEC1* (denoted as version A and B) were isolated and cloned from *Brassica napus* microspore derived embryos. The two versions of *BnLEC1* were characterized and compared to other

homologs of *LEC1* from *Brassica napus* and *Arabidopsis thaliana*. A phylogenetic tree was constructed to identify the conserved domains and protein similarity among LEC1 proteins from other plants species. This comparison demonstrated that the two versions of *BnLEC1* contained all the same molecular features as other LEC1 proteins.

Sense constructs were developed for *BnLEC1* (version A and B) using the pK2GW7 vector containing a 35S promoter. *Agrobacterium* mediated transformation was used to generate seven independent transgenic lines, four contained *BnLEC1* version A and three version B. Phenotypic differences in the seed oil composition were observed among the transgenic lines. In addition, one line displayed a decreased number of seeds per silique. The number of siliques per plant were also counted, and no significant phenotypic differences were observed. Expression analysis revealed that the transgene *BnLEC1* was not expressed at significant levels in *Arabidopsis.* As a result, no correlations could be made between the characterized phenotypes and expression level of the transgene. These unexpected results have been discussed in relation to endogenous gene silencing mechanisms, which often occur during transformation experiments.

Studies in *Arabidopsis* have shown that regulatory genes or transcription factors involved in embryogenesis and seed development can also influence the synthesis and storage of oil. This research assessed the change in relative expression of *LEC1, LEC2, FUS3* and *WRI1* in doubled haploid lines characterized by high and low oil contents. Significant changes in the transcript levels of *LEC1, LEC2, FUS3* and *WRI1* were observed during the different stages of development at 7, 14, 21 and 28 days after pollination, however there were no consistent trends in expression of the genes between the high and low oil content lines. Therefore, it is suggested that the differences in oil

content in the double haploid lines are not due to the altered expression of *LEC1*, *LEC2*, *FUS3* and *WRI1*. These results further imply that these transcription factors may not be the only regulators of fatty acid synthesis and oil accumulation during seed development.

The challenge of trying to improve seed oil content and modify the fatty acid composition of seed oil is very complex and multidimensional, due to the numerous physiological and genetic components involved in embryo and seed development. The experiments described above did not allow for the functional characterization of the embryonic transcription factors and their impact upon fatty acid biosynthesis and accumulation during seed development. Therefore, it is suggested that a different promoter is used to drive the over-expression of the *LEC1* in future transformation studies. Because *LEC1* transcripts are restricted to seed tissues, it may be beneficial to use a seed or tissue specific promoter rather than a constitutive promoter. This would increase *LEC1* expression in tissues where the native transcripts are found. This approach would reduce the likeliness of transgene silencing, which was a major challenge in this research. Since the goal of this research was to determine the role of *BnLEC1* in seed oil biosynthesis in *Brassica napus*, over-expressing *BnLEC1* in canola may be an alternative approach. Plant species can respond differently to the presence of the transgene, therefore over-expressing *BnLEC1* in canola may be successful.

LEC1 may also not be the best candidate to pursue in future work; since recent research in *Arabidopsis* suggests that other transcription factors such as WRI1 and FUS3 may have a more significant effect upon oil accumulation during seed development. WRI1 and FUS3 both influence fatty acid biosynthesis because of their regulation of glycolytic genes, which in turn influence production of sugars that can be utilized for

seed oil biosynthesis. Therefore, FUS3 and WRI1 may be stronger candidates for improving the seed oil content and composition of canola.

There are also numerous factors, besides the genes described above, which can influence embryo and seed development. For example, genes which affect sucrose availability for TAG synthesis and accumulation could be worth exploring to determine their role in the regulation of fatty acid biosynthesis and accumulation during seed development. Environmental factors such as nitrogen availability and plant photosynthetic efficiency, have been demonstrated to influence sucrose availability to initiate the metabolic pathways leading to seed oil biosynthesis. Therefore, genes involved in such metabolic processes are also worth investigating in future studies.

This research has focused on a network of transcription factors that regulate embryo development, specifically LEC1, LEC2, FUS3 and WRI1 with the goal to determine if they play a role in the biosynthesis and accumulation of oil during seed development. Embryo and seed development require the coordination of many metabolic pathways, and other physiological factors, which add complexity to understanding the genetic regulation involved in fatty acid biosynthesis and oil accumulation during seed development. Overall, the challenge of trying to improve the seed oil composition of *Brassica napus* is extremely complicated and may require a multidisciplinary approach.

LITERATURE CITED

Agriculture Canada – Food Production and Inspection Branch. 1987. Description of Variety (Topas).

Angell, S. and Baulcombe, D. 1997. Consistent gene silencing in transgenic plants expressing a replicating potato virus X RNA. EMBO J. 16:3675-3684.

Baud, S., Boutin, J., Miquel, M., Lepiniec, L. and Rochat, C. 2002. An integrated overview of seed development in *Arabidopsis thaliana* ecotype WS. Plant Physiol. Biochem. 40:151-160.

Baud, S., Dubreucq, B., Miquel, M., Rochat, C. and Lepiniec, L. 2008. Storage reserve accumulation in Arabidopsis: Metabolic and developmental control of seed filling. In: The Arabidopsis Book. American Society of Plant Biologists, Rockville, MD.

Baud, S. and Lepiniec, L. 2009. Regulation of de novo fatty acid synthesis in maturing oilseeds of *Arabidopsis*. Plant Physiol. Biochem. 47:448-455.

Baud, S. and Lepiniec, L. 2010. Physiological and developmental regulation of seed oil production. Prog. Lipid Research. 49:235-249.

Baud, S., Mendoza, M., To, A., Harscoet, E., Lepiniec, L. and Dubreucq, B. 2007. WRINKLED1 specifies the regulatory action of LEAFY COTYLEDON2 towards fatty acid metabolism during seed maturation in Arabidopsis. Plant J. 50:825-838.

Baud, S., Wuillème, S., To, A., Rochat, C. and Lepiniec, L. 2009. Role of WRINKLED1 in the transcriptional regulation of glycolytic and fatty acid biosynthetic genes in Arabidopsis. Plant J. 60:933-947.

Belmonte, M., Ambrose, S., Ross, A., Abrams, S. and Stasolla, C. 2006. Improved development of microspore derived embryo cultures of *Brassica napus* cv Topaz following changes in glutathione metabolism. Physiol. Plant.127:690-700.

Bolognese, F., Imbriano, C., Caretti, G. and Mantovani, R. 2000. Cloning and characterization of the histone-fold proteins YBL1 and YCL1. Nucleic Acid Res. 28: 3830-3838.

Bonaventure, G., Salas, J., Pollard, M. and Ohlrogge, J. 2003. Disruption of the *FATB* gene in Arabidopsis demonstrates an essential role of saturated fatty acids in plant growth. Plant Cell. 15:1-14.

Braybrook, S. and Harada, J. 2008. LECs go crazy in embryo development. Trends Plant Sci. 13:624-630.

Browse, J. and Somerville, C. 1991. Glycerolipid synthesis: Biochemistry and regulation. Annu. Rev. Plant Biol. 42:467-506.

Burton, J., Miller, J., Vick, B., Scarth, R. and Holbrook, C. 2004. Altering fatty acid composition in oil seed crops. Advan. Agron. 84:273-306.

Bustin, S., Benes, V., Nolan, T. and Pfaffl, M. 2005. Quantitative real-time RT-PCR -a perspective. J. Mol. Endocrin. 34:597-601.

Canola Council of Canada. 2007. Canola. Gowing Great 2015. http://www.canolacouncil.org/canola_growing_great_2015.aspx . Accessed 2010 October 21.

Canola Council of Canada. 2009a. Provincial Acreages and Yields. http://www.canolacouncil.org/acreageyields.aspx. Accessed 2010 October 21.

Canola Council of Canada. 2009b. Canada's canola industry – adding billions to the economy. http://www.canolacouncil.org/canola_council_of_canada.aspx Accessed 2010 October 21.

Canola Council of Canada. 2010a. Chapter 2 – canola varieties. http://www.canolacouncil.org/chapter2.aspx. Accessed 2010 November 2.

Canola Council of Canada. 2010b. Canola oil properties and uses. http://www.canolacouncil.org/canola_oil_properties_and_uses.aspx . Accessed 2010 October 21.

Casséus, L. 2009. Canola: A canadian success story. Statistics Canada. http://www.statcan.gc.ca/pub/96-325-x/2007000/article/10778-eng.pdf. Accessed 2010 October 31.

Cernac, A., Andre, C., Hoffmann-Benning, S. and Benning, C. 2006. WRI1 is required for seed germination and seedling establishment. Plant Physiol. 141:745-757.

Cernac, A. and Benning, C. 2004. *WRINKLED1* encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis. Plant J. 40:575-585.

Clough, S. and Bent, A. 1998. Floral dip: A simplified method foragrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J. 16:735-743.

Cogoni, C. and Macino, G. 1999. Homology-dependent gene silencing in plants and fungi: A number of variations on the same theme. Curr. Opin. Micro. 2:657-662.

Curaba, J., Moritz, T., Blervaque, R., Parcy, F., Raz, V., Herzog, M. and Vachon, G. 2004. *AtGA3ox2*, a key gene responsible for bioactive gibberellin biosynthesis, is regulated during embryogenesis by *LEAFY COTYLEDON2* and *FUSCA3* in Arabidopsis. Plant Physiol. 136:3660-3669.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. and Scheible, W. 2005. Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol. 139:5-17.

De Lorgeril, M. and Salen, P. 2004. Alpha-linolenic acid and coronary heart disease. Nutr. Metab. Cardiovasc. Dis. 14:162-169.

Dehesh, K., Jones, A., Knutzon, D. and Voelker, T. 1996. Production of high levels of 8: 0 and 10: 0 fatty acids in transgenic canola by overexpression of *Ch FatB2*, a thioesterase cDNA from *Cuphea hookeriana*. Plant J. 9:167-172.

Dehio, C. and Schell, J. 1994. Identification of plant genetic loci involved in a posttranscriptional mechanism for meiotically reversible transgene silencing. Proc. Natl. Acad. Sci. USA 91:5538-5542.

Edwards, D., Murray, J. and Smith, A. 1998. Multiple genes encoding the conserved CCAAT-box transcription factor complex are expressed in Arabidopsis. Plant Physiol. 117:1015-1022.

Elhiti, M., Tahir, M., Gulden, R., Khamiss, K. and Stasolla, C. 2010. Modulation of embryo-forming capacity in culture through the expression of *Brassica* genes involved in the regulation of the shoot apical meristem. J. Exp. Bot. 61:4069-4085.

Flagel, L. and Wendel, J. 2009. Gene duplication and evolutionary novelty in plants. New Phytologist. 183:557-564.

Focks, N. & Benning, C. 1998. *wrinkled1*: A novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism. Plant Physiol. 118:91-101.

Gazzarrini, S., Tsuchiya, Y., Lumba, S., Okamoto, M. and Mccourt, P. 2004. The transcription factor *FUSCA3* controls developmental timing in *Arabidopsis* through the hormones gibberellin and abscisic acid. Dev. Cell 7:373-385.

Harada, J. 1999. Signaling in plant embryogenesis. Curr. Opin. Plant Biol. 2:23-27.

Harada, J. 2001. Role of *Arabidopsis LEAFY COTYLEDON* genes in seed development. J. Plant Physiol.158:405-409.

Herman, L., Jacobs, A., Montagu, M. and Depicker, A. 1990. Plant chromosome/marker gene fusion assay for study of normal and truncated T-DNA integration events. Mol. Gen. 224:248-256.

Hougen, F. and Bodo, V. 1973. Extraction and methanolysis of oil from whole or crushed rapeseed for fatty acid analysis. J. Amer. Oil Chem. Soc. 50:230-234.

Kagaya, Y., Okuda, R., Ban, A., Toyoshima, R., Tsutsumida, K., Usui, H., Yamamoto, A. and Hattori, T. 2005a. Indirect ABA-dependent regulation of seed storage protein genes by FUSCA3 transcription factor in *Arabidopsis*. Plant Cell Physiol. 46:300-311.

Kagaya, Y., Toyoshima, R., Okuda, R., Usui, H., Yamamoto, A. and Hattori, T. 2005b. LEAFY COTYLEDON1 controls seed storage protein genes through its regulation of *FUSCA3* and *ABSCISIC ACID INSENSITIVE3*. Plant Cell Physiol. 46:399-406.

Karimi, M., Inzé, D., and Depicker, A. 2002. Gateway vectors for Agrobacterium-mediated plant transformation. Trends Plant Sci. 5:193-195.

Kim, I., Sinha, S., De Crombrugghe, B. and Maity, S. 1996. Determination of functional domains in the c subunit of the CCAAT-binding factor (CBF) necessary for formation of a CBF-DNA complex: CBF-B interacts simultaneously with both the CBF-A and CBF-C subunits to form a heterotrimeric CBF molecule. Mol. Cell. Biol. 16:4003-4013.

Klein, D. 2002. Quantification using real-time PCR technology: Applications and limitations. Trends Mol. Med. 8:257-260.

Knutzon, D., Hayes, T., Wyrick, A., Xiong, H., Maelor Davies, H. and Voelker, T. 1999. Lysophosphatidic acid acyltransferase from coconut endosperm mediates the insertion of laurate at the *sn-2* position of triacylglycerols in lauric rapeseed oil and can increase total laurate levels. Plant Physiol. 120:739-746.

Knutzon, D., Thompson, G., Radke, S., Johnson, W., Knauf, V. and Kridl, J. 1992. Modification of *Brassica* seed oil by antisense expression of a stearoyl-acyl carrier protein desaturase gene. Proc. Natl. Acad. Sci. USA. 89:2624-2628.

Kroj, T., Savino, G., Valon, C., Giraudat, J. and Parcy, F. 2003. Regulation of storage protein gene expression in *Arabidopsis*. Development 130:6065-6073.

Kwong, R., Bui, A., Lee, H., Kwong, L., Fischer, R., Goldberg, R. and Harada, J. 2003. LEAFY COTYLEDON1-LIKE defines a class of regulators essential for embryo development. Plant Cell 15:5-18.

Le, B., Cheng, C., Bui, A., Wagmaister, J., Henry, K., Pelletier, J., Kwong, L., Belmonte, M., Kirkbride, R. and Horvath, S. 2010. Global analysis of gene activity during

Arabidopsis seed development and identification of seed-specific transcription factors. Proc. Natl. Acad. Sci. USA. 107:8063-8070.

Lee, H., Fischer, R., Goldberg, R. and Harada, J. 2003. *Arabidopsis* LEAFY COTYLEDON1 represents a functionally specialized subunit of the CCAAT binding transcription factor. Proc. Natl. Acad. Sci. USA. 100:2152-2156.

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M., Arondel, V., Bates, P., Baud, S., Bird, D., Debono, A. and Durrett, T. 2010. Acyl lipid metabolism. In: The Arabidopsis book. American Society of Plant Biologists, Rockville, MD.

Li, X., Mantovani, R., Hooft Van Huijsduijnen, R., Andre, I., Benoist, C. and Mathis, D. 1992. Evolutionary variation of the CCAAT-binding transcription factor NF-Y. Nucleic Acids Res. 20:1087-1091.

Liu, J., Hua, W., Zhan, G., Wei, F., Wang, X., Liu, G. and Wang, H. 2009. Increasing seed mass and oil content in transgenic *Arabidopsis* by the overexpression of *wri1*-like gene from *Brassica napus*. Plant Physiol. Biochem. 48:9-15.

Lotan, T., Ohto, M., Yee, K., West, M., Lo, R., Kwong, R., Yamagishi, K., Fischer, R., Goldberg, R. and Harada, J. 1998. *Arabidopsis* LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. Cell 93:1195-1206.

Maity, S. and De Crombrugghe, B. 1998. Role of the CCAAT-binding protein CBF/NF-Y in transcription. Trend Biochem Sci. 23:174-178.

Mantovani, R. 1999. The molecular biology of the CCAAT-binding factor NF-Y. Gene 239:5-27.

Matzke, A. and Matzke, M. 1998. Position effects and epigenetic silencing of plant transgenes. Curr. Opin, Plant Biol. 1:142-148.

Matzke, M., Primig, M., Trnovsky, J. and Matzke, A. 1989. Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. EMBO J. 8:643-649.

McVetty, P. and Scarth, R. 2002. Breeding for improved oil quality in *Brassica* oilseed species. J. Crop Prod. 5:345-369.

Meinke, D. 1992. A homoeotic mutant of *Arabidopsis thaliana* with leafy cotyledons. Science 258:1647-1650.

Meinke, D., Franzmann, L., Nickle, T. and Yeung, E. 1994. *Leafy cotyledon* mutants of Arabidopsis. Plant Cell 6:1049-1064.

Mönke, G., Altschmied, L., Tewes, A., Reidt, W., Mock, H., Bäumlein, H. and Conrad, U. 2004. Seed-specific transcription factors ABI3 and FUS3: Molecular interaction with DNA. Planta 219:158-166.

Mu, J., Tan, H., Zheng, Q., Fu, F., Liang, Y., Zhang, J., Yang, X., Wang, T., Chong, K. and Wang, X. 2008. *LEAFY COTYLEDON1* is a key regulator of fatty acid biosynthesis in *Arabidopsis*. Plant Physiol. 148:1042-1054.

Murphy, D. and Vance, J. 1999. Mechanisms of lipid-body formation. Trends Biochem. Sci. 24:109-115.

Mycogen Canada, Inc. 1994. Request to the WCC / RCC for recommendation of candidate canola cultivar AG012 for registration in western Canada, Chatham, On.

Napoli, C., Lemieux, C. and Jorgensen, R. 1990. Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. Plant Cell 2:279-289.

North, H., Baud, S., Debeaujon, I., Dubos, C., Dubreucq, B., Grappin, P., Jullien, M., Lepiniec, L., Marion-Poll, A. and Miquel, M. 2010. Arabidopsis seed secrets unravelled after a decade of genetic and omics-driven research. Plant J. 61**:**971-981.

Odell, J., Nagy, F. and Chua, N. 1985. Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature 313:810-812.

Ohlrogge, J. 1994. Design of new plant products: Engineering of fatty acid metabolism. Plant Physiol. 104**:**821-826.

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

Olsen, L., Ettinger, W., Damsz, B., Matsudaira, K., Webb, A., and Harada, J. 1993. Targeting of glyoxysomal proteins in Ieaves and roots of a higher plant. Plant Cell 5: 941-952

Parcy, F., Valon, C., Raynal, M., Gaubier-Comella, P., Delseny, M. and Giraudat, J. 1994. Regulation of gene expression programs during Arabidopsis seed development: roles of the *ABI3* locus and of endogenous abscisic acid. Plant Cell 6:1567-1582.

Pidkowich, M., Nguyen, H., Heilmann, I., Ischebeck, T. and Shanklin, J. 2007. Modulating seed β-ketoacyl-acyl carrier protein synthase II level converts the composition of a temperate seed oil to that of a palm-like tropical oil. Proc. Natl. Acad. Sci. USA. 104:4742-4747.

Pröls, F. and Meyer, P. 1992. The methylation patterns of chromosomal integration regions influence gene activity of transferred DNA in *Petunia hybrida*. Plant J. 2:465-475.

Przybylski, R., Mag, T., Eskin, N.A.M. and McDonald, B.E. 2005. Canola Oil. p. 61-121. In: Bailey's Industrial Oil and Fat Produds. Volume 2. Sixth Edition. Shahidi F. ed. John Wiley and Sons Inc. Hoboken, NJ.

Rawsthorne, S. 2002. Carbon flux and fatty acid synthesis in plants. Prog. Lipid Res.. 41:182-196.

Reidt, W., Wohlfarth, T. and Ellerström, M. 2000. Gene regulation during late embryogenesis: The RY motif of maturation-specific gene promoters is a direct target of the FUS3 gene product. Plant J. 21:401-408.

Riechmann, J., Heard, J., Martin, G. and Reuber, L. 2000. *Arabidopsis* transcription factors: Genome-wide comparative analysis among eukaryotes. Science 290:2105-2110.

Robenek, H., Hofnagel, O., Buers, I., Robenek, M., Troyer, D. and Severs, N. 2006. Adipophilin-enriched domains in the ER membrane are sites of lipid droplet biogenesis. J. Cell Sci. 119:4215-4224.

Romier, C., Cocchiarella, F., Mantovani, R. and Moras, D. 2003. The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y. J. Biol. Chem. 278:1336-1345.

Santos-Mendoza, M., Dubreucq, B., Baud, S., Parcy, F., Caboche, M. and Lepiniec, L. 2008. Deciphering gene regulatory networks that control seed development and maturation in *Arabidopsis.* Plant J. 54:608-620.

Santos Mendoza, M., Dubreucq, B., Miquel, M., Caboche, M. and Lepiniec, L. 2005. LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in *Arabidopsis* leaves. FEBS Lett. 579:4666-4670.

Sauer, M. and Friml, J. 2004. In vitro culture of *Arabidopsis* embryos within their ovules. Plant J. 40:835-843.

Scarth, R. and Tang, J. 2006. Modification of *Brassica* oil using conventional and transgenic approaches. Crop Sci. 46:1225-1236.

Schranz, M. and Osborn, T. 2004. De novo variation in life-history traits and responses to growth conditions of resynthesized polyploid *Brassica napus* (Brassicaceae). Amer. J. Bot. 91:174-183.

Schwender, J., Ohlrogge, J. and Shachar-Hill, Y. 2003. A flux model of glycolysis and the oxidative pentosephosphate pathway in developing *Brassica napus* embryos. J. Biol. Chem. 278:29442-29453.

Sezonov, G., Joseleau-Petit, D. & D'ari, R. 2007. *Escherichia coli* physiology in Luria Bertani Broth. J. Bacteriol. 189:8746-8749.

Sharma, N., Anderson, M., Kumar, A., Zhang, Y., Giblin, E., Abrams, S., Zaharia, L., Taylor, D. and Fobert, P. 2008. Transgenic increases in seed oil content are associated with the differential expression of novel *Brassica*-specific transcripts. BMC Genomic 9:619-637.

Sinha, S., Kim, I., Sohn, K., De Crombrugghe, B. and Maity, S. 1996. Three classes of mutations in the A subunit of the CCAAT-binding factor CBF delineate functional domains involved in the three-step assembly of the CBF-DNA complex. Mol. Cell. Biol. 16:328-337.

Sinha, S., Maity, S., Lu, J. and De Crombrugghe, B. 1995. Recombinant rat CBF-C, the third subunit of CBF/NFY, allows formation of a protein-DNA complex with CBF-A and CBF-B and with yeast HAP2 and HAP3. Proc. Natl. Acad. Sci. USA. 92:1624-1628.

Sleper, D. and Poehlman, J. 2006. Biotechnology and Plant Breeding. In: Breeding Field Crops, 5th edition. Wiley-Blackwell, Ames, IA.

Somerville C., Browes J., Jaworski J.G., and Ohlrogge J.B. 2000 Lipids. p. 456-527 In: Biochemistry and Molecular Biology of Plants. Buchanan, B., Gruissem, W. and Jones, R. eds. American Society of Plant Physiologists, Rockville, MD.

Stam, M., Viterbo, A., Mol, J. and Kooter, J. 1998. Position-dependent methylation and transcriptional silencing of transgenes in inverted T-DNA repeats: Implications for posttranscriptional silencing of homologous host genes in plants. Mol. Cell. Biol. 18:6165-6177.

Stone, S., Braybrook, S., Paula, S., Kwong, L., Meuser, J., Pelletier, J., Hsieh, T., Fischer, R., Goldberg, R. and Harada, J. 2008. *Arabidopsis* LEAFY COTYLEDON2 induces maturation traits and auxin activity: Implications for somatic embryogenesis. Proc. Natl. Acad. Sci. USA.105:3151-3156.

Stone, S., Kwong, L., Yee, K., Pelletier, J., Lepiniec, L., Fischer, R., Goldberg, R. and Harada, J. 2001. *LEAFY COTYLEDON2* encodes a b3 domain transcription factor that induces embryo development. Proc. Natl. Acad. Sci. USA. 98:11806-11811.

Stoutjesdijk, P., Hurlestone, C., Singh, S. and Green, A. 2000. High-oleic acid Australian *Brassica napus* and *B. juncea* varieties produced by co-suppression of endogenous Δ12-desaturases. Biochem. Soc. Trans. 28:938-940.

Tang, J., Scarth, R. and McVetty, P. 2004. Stability of the expression of Acyl-ACP thioesterase transgenes in oilseed rape doubled haploid lines. Crop Sci. 44:732-740.

Thelen J., and Ohlrogge J. 2002. Metabolic engineering of fatty acid biosynthesis in plants. Metabol. Engineer. 4:12–21.

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

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

Van Der Krol, A., Mur, L., Beld, M., Mol, J. and Stuitje, A. 1990. Flavonoid genes in petunia: Addition of a limited number of gene copies may lead to a suppression of gene expression. Plant Cell 2:291-299.

Vicient, C., Bies-Etheve, N. and Delseny, M. 2000. Changes in gene expression in the *leafy cotyledon1 (lec1)* and *fusca3 (fus3)* mutants of *Arabidopsis thaliana*. J. Exp. Bot. 51: 995-1003.

Voelker, T., Jones, A., Cranmer, A., Davies, H. and Knutzon, D. 1997. Broad-range and binary-range acyl-acyl-carrier-protein thioesterases suggest an alternative mechanism for medium-chain production in seeds. Plant Physiol. 114:669-677.

von Wettstein-Knowles, P. 1982. Elongase and epicuticular wax biosynthesis. Physiol. Veg. 20:797–809.

Wang, H., Guo, J., Lambert, K. and Lin, Y. 2007. Developmental control of *Arabidopsis* seed oil biosynthesis. Planta 226:773-783.

Warner, K. and Mounts, T. 1993. Frying stability of soybean and canola oils with modified fatty acid compositions. J. Amer. Oil Chem. Soc. 70:983-988.

Waterhouse, P., Wang, M. and Lough, T. 2001. Gene silencing as an adaptive defence against viruses. Nature 411:834-842.

Weier, D., Hanke, C., Eickelkamp, A., Lühs, W., Dettendorfer, J., Schaffert, E., Möllers, C., Friedt, W., Wolter, F. and Frentzen, M. 1997. Trierucoylglycerol biosynthesis in transgenic plants of rapeseed *(Brassica napus l.).* Fett. Lipid 99:160-165.

Weselake, R., Taylor, D., Rahman, M., Shah, S., Laroche, A., McVetty, P. and Harwood, J. 2009. Increasing the flow of carbon into seed oil. Biotech. Advan. 27:866-878.

West, M., Yee, K., Danao, J., Zimmerman, J., Fischer, R., Goldberg, R. and Harada, J. 1994. *LEAFY COTYLEDON1* is an essential regulator of late embryogenesis and cotyledon identity in Arabidopsis. Plant Cell 6:1731-1745.

Xie, Z., Li, X., Glover, B., Bai, S., Rao, G., Luo, J. and Yang, J. 2008. Duplication and functional diversification of HAP3 genes leading to the origin of the seed-developmental regulatory gene, *LEAFY COTYLEDON1 (LEC1)*, in nonseed plant genomes. Mol. Biol. and Evol. 25:1581-1592.

Xing, Y., Fikes, J. and Guarente, L. 1993. Mutations in yeast HAP2/HAP3 define a hybrid CCAAT box binding domain. EMBO J.12:4647-4655.

Yamamoto, A., Kagaya, Y., Toyoshima, R., Kagaya, M., Takeda, S. and Hattori, T. 2009. Arabidopsis NF-YB subunits LEC1 and LEC1-LIKE activate transcription by interacting with seed-specific ABRE-binding factors. Plant J. 58:843-856.

Yan, C. and Boyd, D. 2006. Histone H3 acetylation and H3 K4 methylation define distinct chromatin regions permissive for transgene expression. Mol. Cell. Biol. 26**:**6357-6371.

Yang, J., Xie, Z. and Glover, B. 2005. Asymmetric evolution of duplicate genes encoding the CCAAT-binding factor NF-Y in plant genomes. New Phytol. 165:623-632.

Yatsu, L. and Jacks, T. 1972. Spherosome membranes: Half unit-membranes. Plant Physiol. 49:937-943.

Yazawa, K. and Kamada, H. 2007. Identification and characterization of carrot HAP factors that form a complex with the embryo-specific transcription factor C-LEC1. J. Exper. Bot. 58**:**3819-3828.