Assessment of genetic diversity in *Pisum* germplasm for field pea improvement

Sajjad Ahmad

Master's of Science

Department of Plant Science,

Macdonald Campus, McGill University,

Ste-Anne-de-Bellevue, Québec, H9X 3V9

December, 2012

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

Abstract

Field pea (*Pisum sativum* L.) is an important cool-season crop cultivated globally for its protein-rich seed and soil fertility benefits. A segregating population was developed by Single Seed Descent (SSD) methods following controlled crosses between parents that can establish a source for future studies including genetic mapping. True hybrids were identified at seedling stage using polymorphic Simple Sequence Repeats (SSRs). SSRs or microsatellites are also valuable tools for assessing genetic diversity in plants as knowledge of genetic diversity is essential for the development of new desirable germplasm and elite breeding lines. Fifty microsatellites and four transposon-based markers (2 DNA-transposons & 2 RNA-transposons) were successfully employed in this study to assess genetic diversity in 35 diverse *Pisum* accessions. Pairwise genetic similarity ranged from 0.045 to 0.838. Polymorphic Information Content (PIC) ranged from 0.055 to 0.887 with a mean of 0.668. The Molecular markers explored in this investigation have potential to identify new resources for Marker-Assisted Selection (MAS) and future development of elite pea breeding lines in response to climate change and declining land, water and energy resources.

Résumé

Cultivé dans le monde entier dans les zones tempérées, le pois sec (*Pisum sativum* L.) est cultivé pour sa richesse en protéines et sa capacité d'améliorer les conditions du sol. Enfin de faciliter de futurs travaux de recherche, une population en ségrégation a été développée par « Single Seed Descent » après des croisements entre des lignes parentales. Les véritables hybrides ont été identifiés au stade plantule en identifiant les polymorphes ADN hautement répétitif (AHR). Les AHRs ou microsatellites sont des utiles précieux pour les évaluations de la diversité génétique car la connaissance de la diversité génétique est essentielle pour le développement de nouveau matériaux génétiques et des lignées d'élites. Cinquante microsatellites et quatre marqueurs de transposons (deux transposons d'ADN et deux transposons d'ARN) ont été employé avec succès afin d'évaluer la diversité génétique de 35 accessions de *Pisum*. Les statistiques par paires de la similarité génétique ont varié entre 0.045 et 0.838. L'information du contenu polymorphique (PIC) a varié entre 0.055 et 0.887 (moyen de 0.668). Les marqueurs moléculaires qui ont été exploré dans cet étude ont le potentiel d'identifier des nouvelles ressources pour la sélection assisté par marqueurs et pour le

développement de lignées élites en réponse aux changements climatiques, à la diminution des terres et de l'eau disponible et des ressources énergétiques.

ACKNOWLEDGEMENTS

In the name of God (Allah), The Most Gracious, The Most Merciful, All the praises and thanks be to God (Allah).

I gratefully acknowledge my parents, wife, three children, brother, sisters, mother-in-law, sister-in-law and brother-in-law for their sincere prayers, love, encouragement, and outstanding support throughout my progress as a student at McGill University, Canada. I am so blessed to have you as my loving family. I admire and love you all!

A deep and sincere gratitude is extended to my supervisor Dr. Jaswinder Singh for giving me a wonderful opportunity to work in his laboratory at McGill University, Canada. I am also very thankful to Dr. Singh for introducing me to the amazing world of molecular breeding. I thank him for all his guidance, patience, help and support that I received during the difficult times of my research project. Above all, Dr. Singh gave me confidence and friendship, which I appreciate from the core of my heart.

I would like to acknowledge with gratitude Dr. Mark Lefsrud, a member of my supervisory committee, for his direction, encouragement and kind assistance throughout my Master's program. My sincere gratitude is also extended to Dr. Philippe Seguin for his valuable and critical reviews,

comments and suggestions. I also thank Dr. Suha Jabaji for access and assistance to lab materials and equipment.

I hereby express my sincere appreciation to Dr. Manjit Singh for his brilliant support, technical guidance, continuous encouragement, friendship and help in this project as well as with manuscript writing. I am also extremely thankful to Dr. Kashmir Singh for his valuable suggestions, comments and corrections. Furthermore, I have been very privileged to work with my lab mates, Mr. Surinder Singh, Neil Dylan Lamb-Palmer, Ravneet Kaur, Prabhjot Nandha, and Marie-Josée Cardinal. I have learned so much about research, molecular techniques and also about life. You have been a precious support day in, day out, during last couple of years. It has been a pleasure to work with you guys. I am also indebted to Dr. Fazli Mabood for his consistent motivation, support and help when I needed it the most.

I would also like to thank Guy Rimmer for greenhouse assistance and Carolyn Bowes and Lynn Bachand for administrative support.

Finally, I wish to acknowledge the funding agencies: Natural Sciences and Engineering Research Council of Canada (NSERC), and Lefsrud Seed & Processor.

Contribution of Authors

Chapter 3: Hybridization of diverse pea (*Pisum sativum* L.) genotypes to generate a source population for trait inheritance studies.

The manuscript was authored by the candidate. The research idea was conceived and designed by Dr. Jawinder Singh, Department of Plant Science, and Dr. Mark Lefsrud, Department of Bioresource Engineering, Macdonald Campus of McGill University. The entire crossing/hybridization, experiments and data analyses were performed by the author.

Chapter 4.0: Assessment of genetic diversity in *Pisum sativum* L. accessions using microsatellite markers.

The findings of this chapter are published recently in the Canadian Journal of Plant Science (2012). This following manuscript was co-authored by the candidate, Neil Dylan Lamb-Palmer, Dr. Manjit Singh, Dr. Jaswinder Singh, Department of Plant Science, and Dr. Mark Lefsrud, Department of Bioresource Engineering, McGill University. Sajjad Ahmad, the primary author of manuscript, performed the experiment and data analyses. The research idea was conceived and designed by Dr. Jaswinder Singh. Dr. Mark Lefsrud had collected and provided diverse pea lines. Neil Dylan Lamb-Palmer and Dr. Manjit Singh also assisted in data analysis

especially in the categorization of diverse Pea lines. All the authors were involved in editing of the manuscript.

Chapter 5.0: Investigation of IRAP transposon-based molecular markers in pea for future breeding purposes.

The manuscript was authored by the candidate. The research idea was conceived and designed by Dr. Jawinder Singh, Department of Plant Science, Macdonald Campus of McGill University. The author performed the experiments and data analyses.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

af	Afila
AFLP	Amplified fragment length polymorphism
bp	base pair
CAPS	Cleaved amplified polymorphic sequence
cM	Centi morgan
D	Discriminating power
DHs	Doubled haploids
DNA	Deoxyribonucleic acid
EST	Expressed sequence tag
EtBr	Ethidium bromide
GMOs	Genetically modified organisms
IMP	Inter-*MITE* Polymorphism
IRAP	Inter-retrotransposon amplified polymorphism
Kb	Kilo base pair
KOAc	Potassium acetate
LINE	Long interspersed nuclear elements
LTRs	Long terminal repeats
MAS	Marker assisted selection

Mbp	Mega base pair
MITE	Miniature inverted-repeat transposable elements
PCR	Polymerase chain reaction
PIC	Polymorphic information content
QTL	Quantitative trait loci
RAPD	Random amplified polymorphic DNA
RBIP	Retrotransposon-based insertion polymorphism
REMAP	Retrotransposon microsatellite amplified polymorphism
RFLP	Restriction fragment length polymorphism
RILs	Recombinant inbred lines
RNA	Ribonucleic acid
SCAR	Sequence-characterized amplified region
SDS	Sodium dodecyl Sulphate
SINE	Short interspersed nuclear elements
SNP	Single nucleotide polymorphism
S-SAP	Sequence-specific amplification polymorphism
SSD	Single seed descent
SSLPs	Simple sequence length polymorphisms

SSR	Simple sequence repeats
STRs	Short tandem repeats
STS	Sequence-tagged sites
TBE	Tris Borate EDTA
TE	Transposable elements
TILLING	Targeted induced local lesions in genomes
TSDs	Target site duplications
UK	United Kingdom
UPGMA	Unweighted pair group method with arithmetic averages
USA	United States of America
UTR	Untranslated region

CHAPTER 1

GENERAL INTRODUCTION

1.1 Introduction

Field pea (*Pisum sativum* L.) is an economically valuable pulse crop grown around the globe for its protein rich seed and other soil restorative purposes (McPhee, 2003). Developing a segregating population for desirable traits including lipid content after hybridization of parental materials are among the initial steps in crop improvement programs. This population can also serve as a base for future studies including genetic mapping. Commercial interest in peas and other pulses as a protein source (Santalla et al., 2001) has been growing in recent years. In order to meet these demands, the development of new high-yielding cultivars with resistance to various abiotic and biotic stresses on a sustainable basis is greatly desired. However, limited parental material and intensive breeding for desirable trait combinations has gradually decreased genetic diversity in pea (Baranger et al., 2004). Genetic diversity in crops ensures allelic variation that allows novel gene combinations, favorable phenotypes and facilitates development of future breeding lines (Hawkes, 1991). Further improvements in capturing genetic diversity is obligatory in order to develop crop varieties with desirable traits including resistance to stresses,

higher yields or improved nutritive value (Able et al., 2007). Several approaches including molecular markers have been employed for assessing genetic diversity in plant populations. Among molecular markers, microsatellites are commonly used for probing the genetic makeup of many plant genomes including pea (Gong et al., 2010) due to their accuracy, reliability, co-dominancy, reproducibility and high polymorphism (Powell et al. 1996; Becher et al. 2000). Recently, transposable elements (TE)-based fingerprinting has also emerged as a marker system for varietal identification because of their ubiquity and formation of distinct DNA sequences during the integration process (Smykal 2006). An attempt was made in this study to further explore transposon-based markers in pea accessions that will lead to identification of new tools for accession fingerprinting and breeding.

The first objective of this study was to generate a mapping population by crossing four parents using a conventional breeding approach and exploration of polymorphic SSR markers to test hybridity at molecular level in order to advance only true hybrids for next generation. This population can be used by pea breeding community as a source for various future genetic studies including inheritance of important traits and genetic mapping. The second objective was to assess genetic diversity in diverse

pea genotypes using microsatellites and transposon-based molecular markers. This study will assist in the selection process of elite breeding lines by adding more breeding tools for development of high performing varieties on a sustainable basis particularly in response to climate change and declining water, land, and energy-resources.

1.2 General Hypotheses

The following hypotheses were tested in this research project:

1. Microsatellites or SSR markers variation in pea can categorise diverse and unknown genotypes.
2. Microsatellites can play an important role to identify hybrids reliably
3. DNA and RNA-based transposable elements exist in pea genome which can act as molecular markers
4. TEs are polymorphic across diverse pea genotypes

1.3 General Objectives

Aim 1: To develop new hybrids by crossing pea genotypes with variable lipid content.

Aim 2: Identification of true hybrids using molecular tools.

Aim 3: To develop segregating populations that can establish a source population for future genetic studies.

Aim 4: To identify polymorphic microsatellite markers for assessing genetic diversity in diverse pea accessions.

Aim 5: To identify transposon like structures in pea genome

Aim 6: To study the diversity of RNA-based transposons and DNA-based transposons in diverse pea breeding lines i.e. *Cyclop* and *Ogre* (RNA-based transposons) and *Mutator* and *MITES* (DNA-based transposons).

CHAPTER 2.0

LITERATURE REVIEW

2.1 Field Pea and its Economic Importance

Field Pea (*Pisum sativum* L.), a commercially important crop for food and feed, belongs to family Fabaceae (formerly Leguminosae), subfamily Papilionoideae. Field pea, classified as *Pisum sativum* L. is a cool-season legume or pulse crop. Pea is one of the world's oldest crops, cultivated as early as 9,000 years ago (PulseCanada 2010). It is native to Southwest Asia and wild species of pea have been found in Afghanistan, Iran and Ethiopia (Rodino et al., 2009). Gradually it was spread to across the world and is now grown in all climate zones including the tropics where it is cultivated at high-altitudes. Pea seeds are highly nutritious and contain 23–25% protein, 50% slowly digestible starch, 5% soluble sugars by mass and are also a source of fiber and essential minerals (Bastianelli et al., 1998). Increased commercial interest in pea has been developed in recent years due to increased demand for protein-rich food and feed (Santalla et al. 2001; (Pesta et al., 2012)

Peas have been grown in Canada, originating in the prairies, for over 100 years. After the Second World War, pea acreage in Canada was about

20,000 hectares with production concentrated in Manitoba (PulseCanada 2010). Pea production gradually increased during the 1970s and was introduced to the other parts of western Canada. This expansion was mainly due to increased export of dry peas to the European feed-pea market (www.agr.gov.sk.ca). Today, Canada is among the major pea-producing countries of the world with production mainly concentrated in the provinces of Saskatchewan, Alberta and Manitoba (Smýkal et. al., 2012). Field peas rank 4th as the world's largest legume crop in production after soybean, peanuts and dry beans (Yoshida et al., 2007).

2.2 Genetics of Pea

Pea (*Pisum sativum* L.) genomics have been well-studied (Samatadze et al., 2008) ever since the pioneering work of Gregor Mendel in the nineteenth century. Certain features of pea such as self-pollination, ease of cultivation, and easily distinguishable phenotypic traits that inspired Mendel to choose pea for his experiments and maintain the pea as a major focus of modern genetic studies. Many morphological characteristics have a simple inheritance and have played a role in increase of pea yield and production. Several early pea cultivars were tall with long vines (>100 cm) making them susceptible to lodging and disease

due to high humidity maintained in collapsed crop stands (McPhee, 2003). Later dwarf pea varieties that resist lodging were developed after the discovery of two genes, *afila (af)* and le. *Afila* gene (*af*) is responsible for converting leaflets into tendrils that intertwine with adjacent plants, increasing mutual support and resulting in upright plant canopy. This upright plant stand when combined with dwarf plant type (le), improved overall productivity due to resistance to lodging, reduction in disease and ease in harvesting (McPhee, 2003).

Quantitative traits in pea such as winter hardiness, tolerance to fungal diseases and seed yield are controlled by multiple genes (Krajewski et al., 2012). Generally, quantitative trait improvement is a time consuming process requiring several hybridization cycles and subsequent selection. Moreover, improvements in quantitative traits have proved difficult (Tar'an et al., 2005) due to environmental effects. Epistatic gene interaction also makes inheritance complex. Both low heritability and differences in environmental conditions contribute negatively to the identification of superior genotypes when using conventional methods. Advanced methods are, therefore, indispensable to estimate genetic gains that eliminate environmental effects. Like other crops, genetic maps have also been constructed in pea (Katoch et al., 2010: Loridon et al., 2005). The

consequential molecular markers are associated to quantitative trait loci (QTL) that allow breeders to identify superior genotypes prior to field evaluations thus eliminating environmental factors. Pea has a composite linkage map: SSR (228), other markers (231, mainly RADP), with 18552 number of registered ESTs (Sato et al in 2010).

2.3 Pea Floral Biology

Pea (*Pisum sativum* L.) belongs to a sub-family called *Papillonaceae*, a group named for the butterfly-like appearance of their flowers. Pea flower is borne in the leaf axil and has both male and female organs in the same flower. The female organ of the flower, called the pistil, has a sack at its base known as the ovary, which contains ovules (egg cells) usually 5-12 in number. The style is somewhat flattened and cylindrical which comes out of the ovary and bends at right angle to the ovary. On the top of the style there is a crown-like emblem known as the stigma which is sticky and hairy making the stigma appropriate for receiving and retaining pollen grains. In the pea flower, the pistil is surrounded by stamens in 9+1 arrangement. The filaments of 9 stamens are joined together while the 10^{th} stamen is free. The stamens are shorter than the style in early stages of growth but elongate with maturity. There are 5 petals in a 2+2+1

arrangement having 1 standard, 2 wings, and 2 keels that are fused except at their base. They cover the pistil and stamens.

2.4 Field Pea Breeding

Pea is a self-pollinated, cool-season, annual and diploid crop having 14 chromosomes (2n=14, n=7) with a genome size of about 5000 Mbp (Sato et al., 2010). Several methods are available for population improvement in pea (Muehlbauer et al., 1988). Due to continual self-pollination in pea, undesirable linkage can hamper recombination and hence crop improvement. To break linkage and to release desirable traits, F_2 generations are subjected to biparental mating. Biparental mating and backcrossing methods have been successfully used to transfer disease resistance from landraces into cultivars (Muehlbauer et al., 1992). Mutation breeding is also an effective tool for creation of novel alleles and increasing diversity. Improvement in yield, resistance to lodging (afila leaf type) and adjustments in maturity have been achieved with mutation breeding in legumes (Micke, 1988). Conventional breeding has been instrumental in developing many high performing cultivars but some agronomical traits are difficult to improve using conventional approaches. Advanced biotechnologies such as gene transformation and DNA markers

can be effective in supplementing traditional breeding techniques (Gunasekare 2007). MAS can hasten the selection process in plant breeding and consequently can shorten the development time of a new cultivar (Collard et al., 2008). Molecular markers have been frequently used in the past decade to select plants containing genes of interest (Edward et al., 2007). Marker use depends on linkage with a gene of interest. More closely linked markers are more reliable for superior plant selection. These markers help plant breeders to incorporate complex traits and help establishing the chromosomal locations using linkage maps.

Most field pea breeding programs in Canada have focused on the development of high yielding cultivars. Recent developed cultivars have approximately 20% higher yields than those released in early 1990s (Saskatchewan Agriculture and Food, 2006). Similarly, new cultivars have shown more resistance to lodging than the older cultivars of 1990s. This facilitates harvesting and reduces the severity of Ascochyta blight (Banniza et al., 2005). However, pea seed quality parameters such as seed size, shape, uniformity and colour for yellow and green food markets remain the main focus of traditional breeding. Furthermore, crude protein, lysine and energy content are key traits for monogastric feed markets. Development of low phytate lines are also in progress. Low phytate seeds

store phosphorus in a form available to monogastric animals which can not only reduce phosphorus pollution but also decrease the requirement for feed supplementation with phosphorus (Warkentin et al. 2012). Also pea cultivars are being tested for basal branching and their competitiveness with weeds. The leafed forage pea cultivars with longer vines performed better in suppressing weeds than semi-leafless grain cultivars (Spies et al. 2011). These competitive cultivars could then be grown with reduced seed applications thereby reducing the input cost. Inheritance studies of key traits related to seed shape, dimpling, and chlorophyll bleaching are under investigation with the intent to develop molecular markers for each trait (Ubayasena et al., 2010). The development and selection of pulse crop varieties with improved nitrogen fixation capabilities (Drew et al. 2008) is under discussion. Efforts were being made to develop doubled haploid (DH) pea lines in a collaborative work by Canadian and Australian researchers (Croser et al. 2006). DH technology allows the development of homozygous lines in one generation whereas conventional approach requires several generations of selfing in order to achieve homozygosity (Germana 2011). DH plants are produced when the chromosomes of haploid cells are doubled using various techniques for example use of cholchicine chemical (Wedzony et al.

2009). In genetics, production of DH plants is analogous to RILs generated by SSD method (Basu et al., 2011). Despite all benefits, doubled Haploidy is still in embryonic stages in field pea partly due to problems such as poor regeneration of fertile plants and most protocols are genotype specific posing threats to their wide application (Croser et al. 2006).

2.5 Genetic Diversity in Pea

Genetic diversity can be defined as intra-species variation and polymorphism at the DNA level. Allelic diversity can enable certain species to thrive in new and challenging environments, ensuring their long term survival. Genetic variation in crops, therefore, is regarded as essential for global food security (Able et al., 2007). Certain factors nonetheless tend to narrow the genetic diversity in crops. These include long term selection of high-performing cultivars, crop monoculture for increased crop uniformity and productivity, and intensive breeding. In addition, self-pollinating crops such as *Pisum sativum* L. develop increased homozygosity due to continual self-pollination (Cieslarova et. al., 2011). Although these factors ensure higher yields and production, they lead to unwanted genetic erosion (Akhalkatsi et al., 2010). Even during the Green Revolution in

1960s that ushered agriculture into new era of increased yields some groups argued that this came with substantial genetic erosion (Able et al., 2007).

Generally, both geneticists and plant breeders have emphasized the need for further improvement in capturing and harnessing genetic diversity (Able et al., 2007). Improving diversity in crops is essential for creation of novel and desirable germplasm as well as development of future breeding lines. Several methods are available to assess genetic diversity in plant species.

2.6 Molecular Markers

Molecular markers have been used by plant breeders to study plant genomic organization, trace genes of interest and facilitate the plant breeding process (Collard and Mackill, 2008; Meksem and Kahl, 2005). The underlying principle of markers is that an easy-to-observe trait (marker) is tightly linked to a more difficult-to-observe trait. Consequently by selecting the already detected marker (trait), breeders indirectly locate or select for the desirable trait (Semagn et al., 2006a). Molecular or genetic marker serves as a chromosomal landmark for tracing a particular region of DNA (Semagn et al., 2006a). Molecular markers are essential for processes such as genetic mapping, which is one of the vital ways of

retrieving key genomic information and an important tool to locate genes of interest. Genetic maps for various plant species have been constructed by means of diverse molecular marker systems (Meksem and Kahl, 2005). Genetic mapping involves determination of genomic location/position of genes and the distance between them with the aid of molecular markers. As an example, a genetic map can be described as analogous to signs or landmarks along a highway where the genes are “houses” and the landmarks are molecular markers (Collard et al., 2005). Moreover, molecular markers are also largely used as a tool to study genetic diversity through DNA sequence variation. Noteworthy molecular markers used in different breeding processes are: restriction fragment length polymorphism (RFLP) (Botstein et al., 1980), random amplified polymorphic DNA (RAPD) (Tullu et al., 2008; Tar'an et al., 2003), simple sequence repeats (SSR) also known as microsatellites (Tullu et al., 2008; Ubayasena et al., 2010), sequence-tagged sites (STS) (von Stackelberg et al., 2003; Palazzolo et al., 1991), amplified fragment length polymorphism (AFLP) (Ubayasena et al., 2010; von Stackelberg et al., 2003; Vos et al., 1995), single-nucleotide polymorphism (SNP), transposons-based molecular markers (Kalendar et al., 2011), sequence-characterized amplified region (SCAR) (Naqvi and Chattoo 1996), and cleaved amplified

polymorphic sequence (CAPS) (Barth et al., 2002). Molecular markers can be classified into either DNA-based or PCR-based molecular markers. PCR-based markers have an advantage as they require low quantities of DNA and are quick to assay. Important genetic markers including RFLP, RAPD, and SSR, with special reference to pea, are briefly discussed below.

2.6.1 Restriction Fragment Length Polymorphism (RFLP)

RFLPs are among the earliest described molecular markers. They were used for genetic mapping of a temperature-sensitive adenovirus serotype. Later, they were used for human genome mapping (Botstein et al., 1980) as well as for plant genomes (Helentjaris et al., 1986). RFLPs are produced by changes in DNA sequence, for instance, single-nucleotide mutations, insertions or deletions of DNA sequences of one to several hundred base pairs length, or rearrangements of DNA of large chromosomal regions. Such changes are associated with the gain, loss, or movement of some restriction sites thus acting as a basis for generating RFLPs. The technique involves digesting genomic DNA with specific restriction enzymes followed by hybridizing with probes. Genomic DNA, cDNA or expressed sequence tags (ESTs) could be used as probes in

RFLP. Given that RFLP markers are bi-allelic co-dominant, hybridizing with specific probe gives an advantage of identification of a unique locus and chromosomal position. RFLP markers are highly reproducible, show co-dominant inheritance, and have good transferability between laboratories. Moreover, these markers are relatively easy to score owing large size difference between fragments. Nevertheless, RFLP analysis requires more labour and time and requires expensive enzymes and probes.

In pea, Timmerman-Vaughan et al. (2004) utilized different markers including RLFPs to identify quantitative trait loci (QTLs) for Ascochyta blight resistance using F3 population, derived from a cross between partial resistance and susceptible pea breeding lines. They observed the total coverage of the linkage map as 930 cM with an average distance between markers of 10.8 cM. In another study, DNA-based (RLFPs) and PCR-based molecular marker techniques (RAPDs, ALFPs and SSRs) were compared for their effectiveness and applicability in assessing genetic diversity in pea. PCR-based techniques were more informative in 10 pea genotypes than cDNA-RLFP (Lu et al., 1996). Moreover, a linkage map of pea, consisting of 209 markers including RFLP, RAPD and AFLP and covers 1330 cM was constructed by Gilpin et al. (1997)

2.6.2 Random Amplified Polymorphic DNA (RAPD)

RAPD is among the earliest PCR-based markers used for genetic mapping and DNA-fingerprinting. The development of PCR-based markers such as RAPD that amplify target DNA segments before detection was great achievement in the generation of data points particularly when the RFLP analysis is labour-intensive and time consuming. Random primers usually within the range of 10-20 nucleotides are used for the generation of RAPD molecular markers (Jones et al., 1997). RAPD has several disadvantages in the procedure and because of this it is currently less popular in genetic mapping and genomic fingerprinting projects. RAPD has low reproducibility not only inter-laboratories but also within a laboratory over time. Under less stringent conditions, almost every aspect of PCR can affect reproducibility. The presence of low-intensity bands and the position and intensity of high-intensity bands are most notably affected by changes in PCR parameters. RAPDs were used in several pea studies. Pereira et al. (2010) worked on the identification of molecular markers associated with resistance of powdery mildew in pea (*Pisum sativum* L.) with the help of three DNA-markers techniques including RAPDs. RAPDs were also employed for varietal and genome identification. Four pea varieties and two genetic lines were analysed for genome and

chromosome polymorphism using RAPD-PCR analysis. The results showed high genomic polymorphism between-variety whereas within-variety and within-line polymorphism was low (Samatadze et al., 2008). Similarly, genetic relatedness in twenty-four pea varieties was studied using RAPD markers. RAPDs in this study were found an efficient marker system for polymorphic studies as 60 out of 80 primers gave clear bands with 74.8 % polymorphism (Choudhury et al., 2007). Likewise, genetic diversity in 148 *Pisum* lines was assessed using 121 different protein and PCR-based markers including RAPD. The molecular analysis classified pea genotypes mostly into groups consistent with their pedigree data and also clearly separated food, feed and fodder peas into distinct clusters (Baranger et al., 2004).

2.6.3 Simple Sequence Repeats (SSR)

Repetitive DNA sequences are commonly found in eukaryotic genomes. These short sequences are known as microsatellites (Litt and Luty, 1989), also known as variable nucleotide tandem repeats, simple sequence repeats (SSRs), and short tandem repeats (STRs) or simple sequence length polymorphisms (SSLPs).The number of these repeats can differ between individuals. Moreover, SSRs are highly polymorphic, plentiful, co-

dominant, PCR based and are widely used for genetic mapping, DNA fingerprinting, marker assisted selection (Hearne et al., 1992). Ubayasena et al. (2010) studied the genetics of cotyledon bleaching resistance and QTLs linked to this trait in green peas. The heritability estimates of bleaching for whole seed was 0.72 and for cotyledon greenness was 0.69, depicting as moderate. However, these estimates were increased when whole seed and cotyledons were subjected to accelerated bleaching conditions. The total coverage of map was over 890 cM and major QTLs were identified for resistance of cotyledon bleaching. Recently, Sarikamis et al. (2010) characterized 30 pea breeding lines and 10 commercial pea cultivars with 10 highly polymorphic SSR markers. Similarly, Cupic et al. (2009), Nasiri et al. (2009), and Zong et al. (2009) analyzed genetic diversity in pea (*Pisum sativum* L.) using SSR markers. Horacek et al (2009) studied genetic homogeneity as well as variation among varieties at DNA level of pea using RAPD and SSR markers. Earlier; Loridon et al. (2005) used 349 polymorphic SSR markers for genetic mapping studies in peas. Simple Sequence Repeats are now widely used and method of choice for assessing genetic diversity in crops including pea (Burstin et al. 2001; Loridon et al. 2005; Nasiri et. al. 2009; Sarikamis et al. 2010; Gong et al. 2010).

2.7 Transposon-Based Molecular Markers

Transposons were first discovered by Barbara McClintock in 1940s when working on corn (maize) genetic experiments (Jones, 2005). Transposons fall into two main classes. Class I transposons, also known as retrotransposons, transpose to new locations in the genome through RNA intermediate using replicative mechanisms (Kumar and Bennetzen, 1999; Sabot et al., 2004). This mechanism is analogous to infectious life cycle and structure of retroviruses (reviewed in detail by Beauregard et al., 2008). Class I transposons are further classified by the presence or absence of *Long Terminal Repeats* (LTR) that flank the main body of transposons. This kind of transposition causes an increase in the genome size due to increase element copy number (Sanmiguel and Bennetzen, 1998). Class II transposons are those that move directly as DNA, known as DNA transposons (Sabot et al., 2004). Unlike retrotransposons, DNA transposons were not found to greatly increase the genome size (Kunze et al., 1997). DNA transposons move by a "cut and paste" process: the transposon is cut out of its location and inserted into a new location (reviewed in Kalendar et al 2011). This process requires an enzyme known as transposase. Transposons are further divided into subclasses, superfamilies, families, and subfamilies based on certain criteria such as

(i) the presence, length, orientation, and sequence of their terminal repeats; (ii) encoding for proteins necessary to move DNA, for example, reverse transcriptase and integrase in retrotransposons (Class I) while transposase in DNA transposons (Class II); and (iii) the length and sequence of Target Site Duplications (TSDs) formed due to insertion (Grzebelus, 2006). Commonly known retrotransposon superfamilies are, *LINE* (*Long Interspersed Nuclear Elements*), *SINE* (*Short Interspersed Nuclear Elements*), *Ty3/gypsy-like,* and *Ty1/copia-like*. The well characterized superfamilies of DNA transposons in plants include *Ac/Ds, En/Spm, PIF,* and *Mutator*. Unlike these transposons, there are some transposons that can neither be assigned to Class I nor to Class II. Among these unclassifiable transposons are *MITEs* (*Miniature Inverted-Repeat Transposable Elements*) that are only several hundred base pairs in length.

Transposable elements based fingerprinting has recently emerged as a system for varietal identification (Smykal 2006). Transposition activity resulted in both DNA polymorphism at insertion sites and small scale reshuffling of genome, thus creating additional diversity (Bennetzen 2000). A role in generating diversity together with their ubiquitous nature is

making TEs important tools for use as molecular markers (Kumar and Hirochika 2001).

2.7.1 Identification of Transposon Insertion Sites

Several systems such as Inter-Retrotransposon Amplified Polymorphism (IRAP) and Retrotransposon Microsatellite Amplified Polymorphism (REMAP) have been used for the identification of polymorphic transposon insertion sites. Both IRAP and REMAP are based on the PCR amplification of DNA fragments (Kalendar et al., 1999), which lie between two retrotransposon insertion sites (IRAP) or between retrotransposon insertion site and a microsatellite site (REMAP). The presence or absence of PCR product detects polymorphism in both these systems. Lack of amplification, however, reveal that retrotransposons are absent at the particular locus. Numerous copies of retrotransposon *BARE-1,* in barley genome have made this system valuable for genetic mapping and development of molecular markers. Similar to IRAP, Inter-*MITE* Polymorphism (IMP) is another system which uses *MITE*-like transposons rather than retrotransposons. Other transposon based markers include Retrotransposon-Based Insertion Polymorphism (RBIP), Transposon

Display, and Sequence-Specific Amplification Polymorphism (S-SAP) (Jing et al., 2012).

Like other molecular markers, transposon-based markers in pea (*Pisum sativum* L.) are available in literature. For instance, Jing et al., (2010) genotyped 3020 samples of *Pisum* germplasm using 45 retrotransposons based insertion polymorphism (RIBP) markers with a Tagged Array Marker method. In another study, the utility of morphological, SSR, and retrotransposons insertion based polymorphism (RIBP) markers were evaluated in assessing genetic diversity of 164 pea accessions. Both RIBP (31 markers) and SSR (10 markers) produced Polymorphic Information Content (PIC) with values at 0.73 and 0.89 respectively (Symkal et al. 2008).

Characterization of *Stowaway MITEs* in pea identified 1500 copies of *Stowaway* elements in the haploid genome of pea (Macas et al., 2005). In addition, Macas et al., (2003) reported *Zaba* as a novel miniature transposable element, but are only moderately repetitive in pea genome when compared to *Medicago* and *Vicia* species.

2.8 Connecting Text

The following chapter examines the potential of hybridization of diverse pea (*Pisum sativum* L.) genotypes for lipid trait improvement in the subsequent generations as a long term goal. In order to improve desirable traits, a segregating population is developed by controlled crosses among genetically divergent parents. However, the progeny obtained from crosses among parents is usually a mixture of selfed and hybrid plants. Advancing true hybrids to next generation will ensure the development of successful segregating population. Hybridity was tested at the molecular level to identify true hybrids using SSR markers (Aim 1: To develop new hybrids by crossing pea genotypes with variable lipid content). Following hybridization among four parents, polymorphic SSR markers were screened for the selection of true hybrids from mix population of selfed plants and hybrids (Aim 2: Identification of true hybrids using molecular tools). True hybrids were then advanced for the next generation (Aim 3: To develop segregating populations that can establish a source population for future genetic studies). This segregating population can also be used for other purposes such as inheritance studies of important traits in subsequent generations.

CHAPTER 3.0

Hybridization of diverse pea (*Pisum sativum* L.) genotypes to generate a source population for trait inheritance studies

3.1 Abstract

Generation of mapping populations is an initial step in improving crop traits and is achieved by controlled crosses among selected parents. Among various mapping populations, the Single Seed Descent (SSD) method is a procedure for developing Recombinant Inbred Lines (RILs). Only selection of true hybrids, particularly in self-pollinated crops such as pea, can guarantee the successful development of a mapping population. Molecular markers such as SSRs can serve as powerful tools to analyze hybrids with enhanced accuracy even at early seedling stages. Seven crosses were made between four parents i.e. G611764, ILCA 5094, Dakota, and Wando. Hybridity was analyzed using SSR markers following screening for polymorphism among four parents. Among 14 SSRs, 3 highly polymorphic SSR markers were selected to test hybridity in progenies obtained from all crosses. Hybridity analysis revealed that 86.57% of progeny were true hybrids whereas only 13.43% were non-hybrid. True hybrids were then advanced for next generations that can establish a base

for future studies like genetic mapping. These advanced generations can also be used to analyse important agronomic traits including lipid content.

3.2 Introduction

Pea *(Pisum sativum* L) is an important legume crop in Canada with production concentrated in the provinces of Alberta, Manitoba, and Saskatchewan (Pulse Canada, 2010). Pea is a good source of protein, but studies on lipid content in pea is scarcely available in literature (Khodapanahi et al., 2012). Lipid content is genetically inherited trait and can be improved through breeding. Developing a segregating population is among the initial steps for trait improvement (Semagn et al., 2006b) and is achieved by controlled crosses among genetically divergent parents. There are various types of mapping populations such as F_2 population, F_2 derived F_3 (F_2:F_3) populations, Doubled Haploids (DHs), Backcross Inbred Lines (BILs), and Recombinant Inbred Lines (RILs). The aim of this study is to produce RILs by continuous selfing of each member of an F_2 population in order to achieve complete homozygosity. Single Seed Descent (SSD) method (Brim 1966) is widely used procedure for developing RILs. The most important advantage of RILs is that completely homozygous lines do not segregate further and can be propagated for an

indefinite period of time. Producing RILs is a time consuming process that requires many seasons and generations to develop. Before advancing the F_2 population, it is of immense importance to ensure the selection of true-hybrids, especially in case of self-pollinated crops with complete flower such as pea (*Pisum sativum* L.). Thus, proper emasculation at appropriate stage of female is required to prevent self-pollination. Identification of true-hybrids was traditionally based on morphological traits. However, morphological traits are largely affected by environment and in many crops these traits are not distinctive and non-observable. Therefore, the selection of true hybrids originated from morphologically similar population is difficult. The possibility of selecting selfed plants rather than true hybrids can adversely affect all stages of breeding program (Cordeiro et al. 2000). With the advent of molecular marker technology, new tools for hybridity evaluation have been developed that allow rapid, reliable, and inexpensive screening (Yashitola et al., 2002). Molecular marker-based analysis is independent of plant growth stage and environment. Due to greater polymorphism and co-dominance, microsatellite markers can be used as an effective tool in plant breeding programs to identify true hybrids (Hashemi et al., 2010; Selvakumar et al., 2012). This early identification of

true-hybrids can substantially reduce time and resources required for successful breeding programs (Sundaram et al., 2008).

The objectives of this research were; Aim1) to develop a segregating population for lipid content, Aim 2) to identify true hybrids using polymorphic SSR markers and, Aim 3) to advance only true hybrids for further generations in order to establish a source population for future genetic studies including lipid content.

3.3 Materials and Methods

The overall overview of the experimental strategy is shown in Figure 3.1.

3.3.1 Plant material

Seeds of four different pea breeding lines i.e. G611764, ILCA 5094 (round shaped), Dakota, and Wando (wrinkle shaped) were obtained from Dr. Mark Lefsrud, Bio-resource engineering Department, McGill University, Ste-Anne-de-Bellevue, Quebec, Canada. These seeds were planted during summer, 2010 at Horticultural Center, MacDonald campus, McGill University.

3.3.2 Emasculation and Pollination/hybridization

After the crop reached to flowering stage, emasculation at a proper time was accomplished. Emasculation is the removal of stamens/anthers with the help of forceps before shedding the pollens. It involves the opening of the pea flower with the forceps followed by holding anthers with forceps and pulling out stamens from the keels without damaging stigma. Pollination with desired pollens was completed by dehiscing fresh and mature pollens/anthers upon the stigma soon after emasculation. Maturity stage of both the male (when pollens are viable and mature) and female flowers (when the stigma is receptive) to be used for hybridization is critical for the success of any breeding program. Growth stages of both male and female flowers used for hybridization in our study can be seen in Figure 3.2. After the pollination, the flower was tagged as shown in Figure 3.3. Emasculation and pollination was conducted from 10.00 am to 12.30 pm. Seven crosses were made between these four parents (G611764 x ILCA 5094, G611764 x Wando, Wando x ILCA 5094, ILCA 5094 x Wando, G611764 x Dakota, and Dakota x G611764). The F_1 progenies obtained from all crosses were subjected to hybridity test using polymorphic SSR markers.

3.3.3 Genomic DNA Isolation

In order to confirm hybridity, DNA from young leaves of parents and F_1 plants was extracted using a standard phenol:chloroform extraction protocol (Singh et al. 2006) with minor alterations. Leaves from young plants at 8-10 leaf stage were collected in labelled micro-centrifuge tubes using snap freezing method (in liquid Nitrogen). The tissues were homogenized using TissueLyser II (Qiagen, Toronto, ON). Natural extraction buffer and 10% Sodium dodecyl Sulphate (SDS) was added to sample followed by incubation at 65°C for 15 min. Subsequently, 200 µl of 5M Potassium acetate (KOAc) was added, mixed and centrifuged at 14,000 rpm for 5 minutes to precipitate protein. The supernatant (950 µl) was transferred to new tubes followed by the addition of 450 µl of 1:1 Phenol:Chloroform into each tube. Sample was mixed and centrifuged for 5 minutes. Isopropanol (700 µl) was added and mixed to precipitate DNA and then centrifuged for 5 minutes. Supernatant was carefully removed and the pellet was washed twice with 70% ethanol. Pellet was dried by vacuum centrifugation. Pellet was resuspended in 65-70 µl Tris EDTA buffer with RNaseA (TER) and incubated at 37°C for 10 min. DNA sample was stored at 20 ^{o}C for future use. DNA quality and quantity was analyzed using Nanodrop spectrophotometer.

3.3.4 SSR Primers and Polymerase Chain Reaction (PCR)

Fourteen SSR markers, selected from primer sets used by Loridon et al. (2005), were screened (Table 3.1) for polymorphism among parental lines. PCR amplification reaction was conducted in a total volume of 25 µl using 1 µl of template DNA (25-35 ng), 0.4 µM each of forward and reverse primers and 5 µl of 5xC Taq-& LOAD™ Mastermix (MP Biomedicals; 1.5 µM $MgCl_2$, 200 µM dNTPs). Amplifications were performed on a C1000 Thermocycler (BioRad, Missisauga, ON) with the following profile: 95°C initial denaturation for 2 minutes, followed by 36 cycles of 30 seconds at 95°C, annealing at 50°C for 45 seconds and 72°C for 1 minute. PCR products were analyzed under UV light on a 3% agarose gel stained with ethidium bromide (EtBr) in Tris Borate EDTA (TBE) buffer. To determine the size of each amplified product a 1 Kb DNA Ladder (Invitrogen, USA) was used.

3.3.5 Identification of Polymorphic SSR Markers and Hybridity Analysis

Polymorphic SSR markers were screened until distinct and polymorphic bands between both the parents were obtained. Only polymorphic SSR markers were then selected for hybridity analysis. To confirm hybridity, all the F_1 progenies from each cross were analyzed alongside their respective

parents. The SSR banding profile of parents and their progenies were used to determine true hybrids. In case of true hybrids bands from both the donor and receptor were noted adjacent to their respective parents (as shown in Figure 3.5) whereas only band of receptor, and not a band of donor was noted in case of self-pollinated plants. Self-pollinated individual were excluded from further evaluation.

3.3.6 Generation of Segregating Population

The confirmed true hybrids of F_1 generation obtained from the crosses were self-pollinated in order to advance for F_2 generation in greenhouse. Leaf samples from F_2 individual plants were collected and DNA was isolated to test the hybridity of F_1 potential hybrids. The F_2 population was again self-pollinated to obtain F_3 generation in open field.

3.4 Results and Discussion

3.4.1 Crosses and Hybridization

Seven crosses (Table 3.2) were made between four parents, which were initially shown variation in oil content (Lefsrud and Singh, Personnel communication). In order to identify hybrid seeds for generation of

segregating populations, putative hybrid plants were analysed for the presence of co-dominant microsatellite markers. For confirmation of hybrids, polymorphic markers were initially identified as described in the following section.

3.4.2 Selection of Polymorphic SSR Markers

A total of fourteen SSR markers (Table 3.1) were selected from a set of primers used by Loridon et al., (2005) to identify polymorphism between each pair of parents used for hybridization. PCR product obtained from parents revealed various degrees of polymorphism as shown in Figures 3.4a and 3.4b. SSR markers showing polymorphism between parents of each cross were identified and selected for hybridity analysis of progenies (Table 3.2). Results also revealed that all parents used in this study were highly homozygous as is expected from self-pollinated crops.

A total of 54 bands were amplified in parents using 14 SSR markers which amplified products in case of each marker except SSR marker AA163.2 (that amplified bands only in Wando and Dakota). Among 14 SSR markers, four markers were found monomorphic or non-polymorphic (AB72, A9, AA90, and AA206). Five markers (AA103, AA67, AA205, AD147, and AC58) revealed unclear polymorphism whereas five (AA278,

AB91, AA285, AA163.2, and AA175) showed clear and unambiguous polymorphism for some parental genotypes. Among polymorphic SSR markers only three markers with clear-cut polymorphic bands were selected for final hybridity analysis (Table 3.2). To test the progeny for true hybrids between crosses G611764 x ILCA 5094 and Wando x ILCA 5094 (or their reciprocal, if any) SSR marker AA 175 was used. Similarly, to identify true hybrids in population obtained from crosses between G611764 x Wando and G611764 x Dakota, SSR marker AA 278 was used. SSR marker AA 285 was employed to confirm true hybrids in the progenies obtained from crosses between ILCA 5094 x Dakota. All the confirmed hybrids with their parents are given in Table 3.3.

3.4.3 Hybridity Analysis Using Polymorphic SSR Markers

Selected SSR markers were employed to confirm true hybrids (Table 3.2). DNA from parents and their progenies were PCR amplified using a particular SSR marker primer pair and were compared for hybridity confirmation. F_1 individuals obtained from crosses were considered true hybrids when it possessed alleles from both parents as shown in Figure 3.5. Self-pollinated plants will inherit a band from only one parent. Molecular analysis revealed that among seven crosses (Table 3.3), the

progeny from four crosses (G611764 x Dakota, Dakota x G611764, G611764 x Wando, ILCA 5094 x Dakota) were all confirmed as true hybrids with 100% frequency (Figure 3.6). However, some non-hybrids or selfed individuals were also observed during molecular testing of hybridity. For example, only 66% F_1 individuals were true hybrids in two crosses i.e. G611764 x ILCA 5094 and Wando x ILCA5094. Similarly, 80% F_1 plants obtained from cross between ILCA5094 x Wando were true hybrids. Total frequency of true hybrids in all the seven crosses were 86.57% despite care taken while selecting mature and viable pollens and receptive stigma. The growth stages of both male and female flowers used in hybridization are shown in Figure 3.2. The frequency of true F_1 hybrids in lentil (*Lens culinaris* Medik.) was low (20.8%) using RAPD and SSR markers (Solanki et al. 2010). Obtaining true hybrids in lentil is tedious due to its small flower size that encourages selfing and also making emasculation and hybridization difficult (Solanki et al. 2010). Results in this study showed high percentage of true hybrids (86.57%) as compared to low frequency of true hybrids (20.8%) in lentil (Solanki et al., 2010). But still identification of true hybrid is essential for developing a segregating population because even a low number of selfed plants can mislead subsequent genetic studies including inheritance or genetic mapping.

3.4.4 Advancing F1 Generation

Following the hybridity analysis, true F_1 generation hybrids obtained from various crosses were self-pollinated to produce F_2 population in the greenhouse and further advanced to F_3 by single seed decent procedure (Table 3.3). A total of 128 F_1 seeds were obtained from all the seven crosses. However, only healthy seeds from each cross were advanced for F_2 generation. For example, the F_1 seeds obtained from G611764xILCA 5094 cross were excluded from breeding program due to low number of healthy seeds (03 healthy seeds out of a total of 05 seeds). Following confirmation of hybrids, F_1 were selfed and a total of 627 seed were obtained from different crosses. F_2 generation were again selfed to produce F_3 generation by selecting only one seed from each plant of F_2 generation using Single Seed Descent method. A total of 364 F_3 seed were obtained and safely stored for advancing to next generation that will establish a source population for future genetic and inheritance studies such as freezing tolerance, lipid content and other agronomic traits. Analysis of lipid content in subsequent generation is due the fact that lipid content is mainly controlled by additive gene action (Delourme et al., 2006). The conversion of oilseed such as sunflower or oilseed rape from open-pollinated or pure line to hybrids did not greatly affect lipid.

Substantial heterosis or hybrid vigour in grain yield has been reported, whereas heterosis was low for lipid/oil content in oilseed rape (Qian et al., 2007). Similar results have been reported that heterosis in soybean was high for maturity, grain yield and plant height but low for oil and protein content (Lewers et al., 1998).

Table 3.1: SSR markers used to identify polymorphism among parents

Sr. no	SSR Marker	Polymorphism
1.	AB72	No polymorphism
2.	A9	No polymorphism
3.	AA90	No polymorphism
4.	AA206	No polymorphism
5.	AA103	Low level polymorphism
6.	AA67	Low level polymorphism
7.	AA205	Low level polymorphism
8.	AD147	Low level polymorphism
9.	AC58	Low level polymorphism
10.	AA278	High level polymorphism
11.	AB91	High level polymorphism
12.	AA285	High level polymorphism
13.	AA163.2	High level polymorphism
14.	AA175	High level polymorphism

Table 3.2: Selected SSR markers for hybridity testing in seven crosses

Sr. No	Crosses	SSRs
1.	G611764 x ILCA 5094	AA 175
2.	G611764 x Wando	AA 278
3.	ILCA 5094 x Dakota	AA 285
4.	Wando x ILCA 5094	AA 175
5.	ILCA 5094 x Wando	AA 175
6.	G611764 x Dakota	AA 278
7.	Dakota x G611764	AA 278

Table 3.3: Confirmed hybrids using various polymorphic SSR markers

Crosses	Hybrid Name	Crosses	Hybrid Name
ILCA 5094 x Dakota	ID 2	G611764 x Wando	GW 1
ILCA 5094 x Dakota	ID 4	G611764 x Wando	GW 2
ILCA 5094 x Dakota	ID 7	G611764 x Wando	GW 3
ILCA 5094 x Dakota	ID 9	G611764 x Wando	GW 4
G611764 x Dakota	GD 1	G611764 x Wando	GW 5
G611764 x Dakota	GD 2	G611764 x Wando	GW 6
G611764 x Dakota	GD 3	G611764 x Wando	GW 7
G611764 x Dakota	GD 4	G611764 x Wando	GW 8
Dakota x G611764	DG 1	Wando x ILCA 5094	WI 1
Dakota x G611764	DG 2	Wando x ILCA 5094	WI 3
Dakota x G611764	DG 3	Wando x ILCA 5094	WI 4
Dakota x G611764	DG 4	Wando x ILCA 5094	WI 8
G611764 x ILCA 5094	GI 1	Wando x ILCA 5094	WI 9
G611764 x ILCA 5094	GI 2	Wando x ILCA 5094	WI 10
G611764 x ILCA 5094	GI 3	ILCA 5094 x Wando	IW 1
ILCA 5094 x Wando	IW 3	ILCA 5094 x Wando	IW 2
ILCA 5094 x Wando	IW 4	ILCA 5094 x Wando	IW 5

Table 3.4: Advancing segregating population up to F_3 generations

Crosses	F_1	F_2	F_3
G611764xILCA 5094	5	-	-
G611764 x Wando	27	129	113
ILCA 5094 x Dakota	24	95	92
Wando x ILCA 5094	24	109	-
ILCA 5094 x Wando	9	93	89
G611764 x Dakota	23	129	-
Dakota x G611764	16	72	70
Total	**128**	**627**	**364**

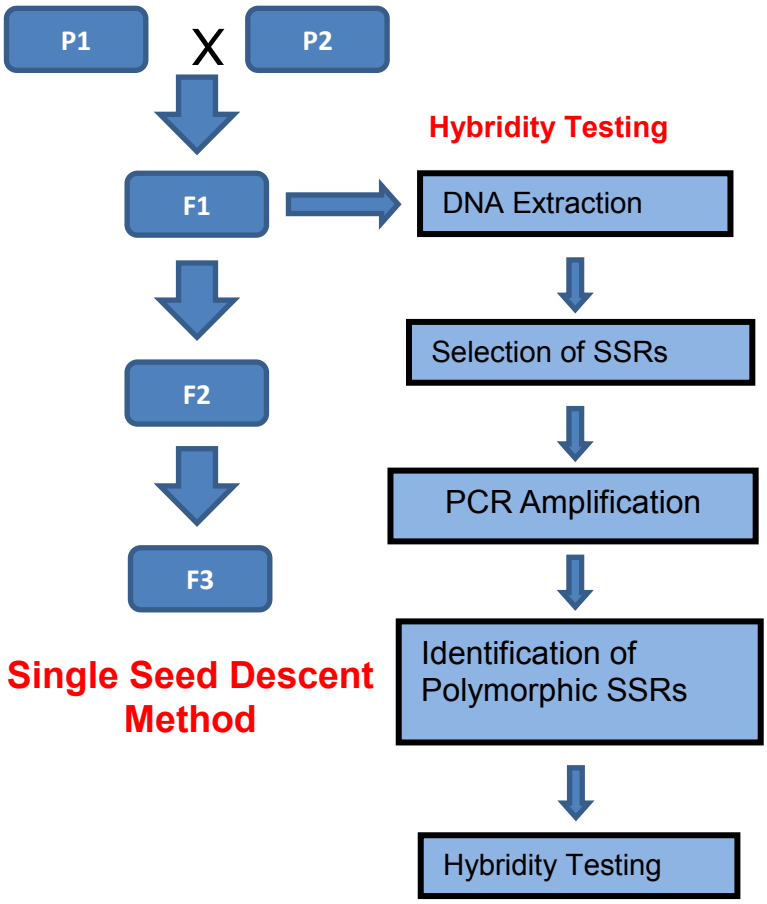


Figure 3.1: Flow chart of strategies used in the hybridity study

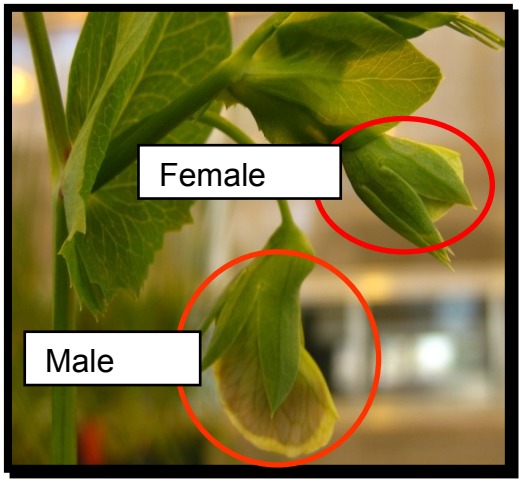


Figure 3.2: Developmental stages of female and male flower used for hybridization. The stigma of female flower should be receptive and pollens should be viable. Generally a male flower in pea breeding is more matured than female flower.

Figure 3.3: Proper tagging of crossed flower following pollination or hybridization. Tagging helps in identification of crosses.

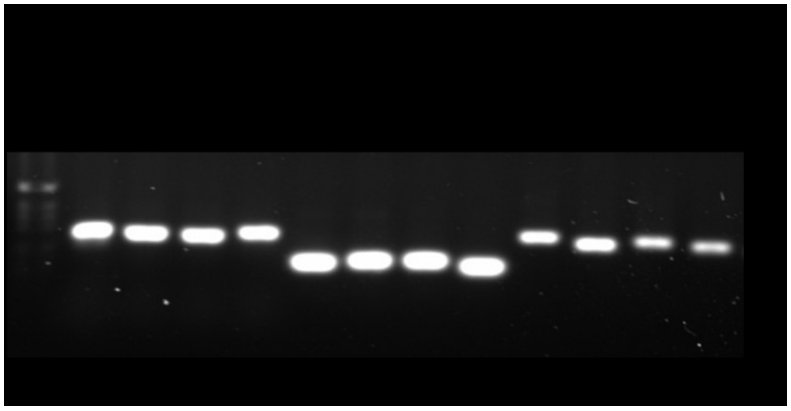


Figure 3.4a

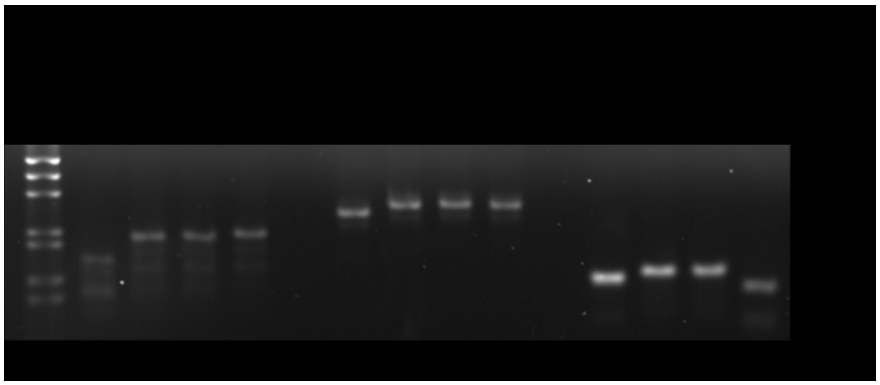


Figure 3.4b

Figure 3.4: Agarose gel (1%) showing DNA fingerprints of four parents (ILCA, Dakota, Wando and G611764). **Figure 3.4a:** SSR A9 is monomorphic, whereas marker AA 103 and AA 205 showed ambiguous polymorphism which is undesirable for hybridity testing. **Figure 3.4b:** SSR markers AB91 and AA175 showed unambiguous polymorphism between ILCA and the other three parents. AA278 showed unambiguous polymorphism between ILCA, G611764 and the other two parents.

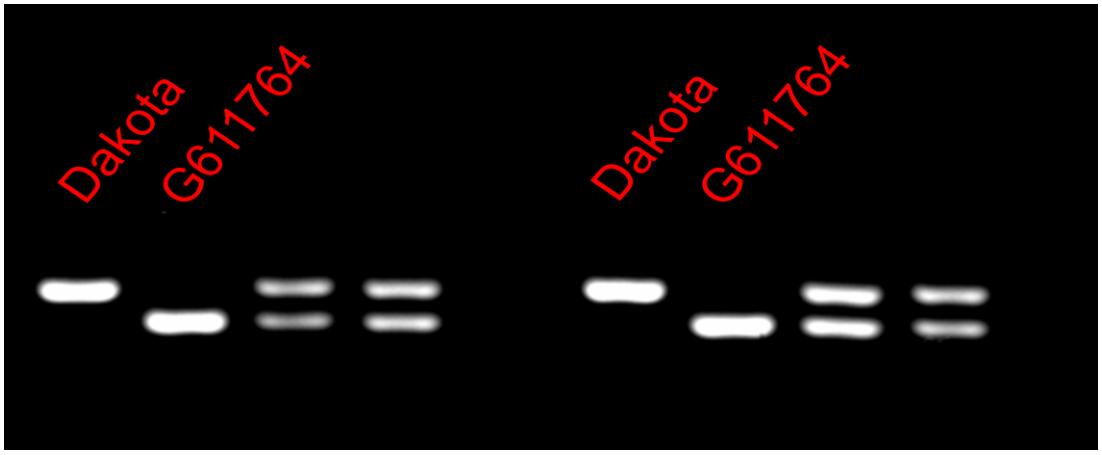


Figure 3.5: Agarose gel (3%) showing parents (Dakota and G611764) and F1 hybrid progeny. SSR markers are able to detect hybridity due to codominance. True hybrids are shown with a heterozygous banding pattern having alleles from both parents using SSR marker AA 278.

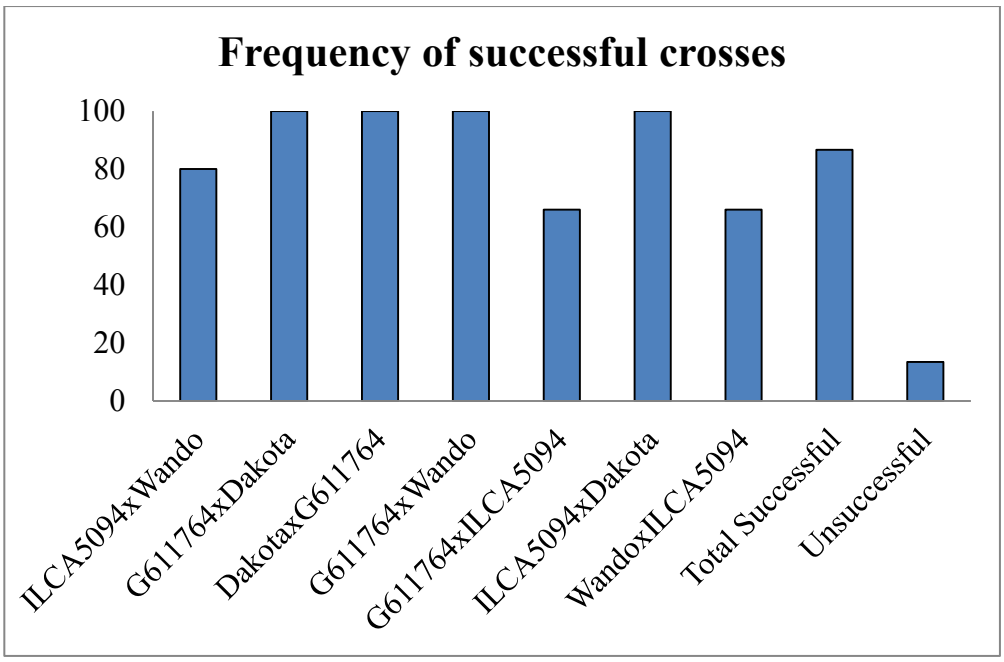


Figure 3.6: Frequency of successful and unsuccessful crosses. Total frequency of true hybrids in all the seven crosses is 86.57%. F1 individuals obtained from crosses were considered true hybrids if SSR alleles from both parents were amplified and. Self-pollinated plants amplify only a single parental band and were considered unsuccessful.

3.5 Connecting Text

The findings of the following chapter are published recently in the Canadian Journal of Plant Science (2012)*. This chapter focuses on assessment of genetic diversity using microsatellites or SSR markers in pea (Aim 1: To study the genetic diversity of diverse pea breeding lines using different microsatellites). For development of improved cultivars or inbred lines, divergent or contrasting parents are required. Therefore, knowledge of genetic diversity is essential for successful breeding programs. Genetic diversity studies at molecular level have other significant applications such as conservation and seed bank managements. Microsatellites are widely used to detect polymorphism in various crops including pea (Aim 2: To detect polymorphism at DNA level in pea genome). The data generated in this study (Aim 3: To generate molecular fingerprinting of pea accessions) can be used by pea community for breeding of elite progenies.

The following manuscript was co-authored by the candidate, Neil Dylan Lamb-Palmer, Dr. Manjit Singh, Dr. Jaswinder Singh, Department of Plant Science, and Dr. Mark Lefsrud, Department of Bioresource Engineering, McGill University. Sajjad Ahmad, the primary author of manuscript, performed the experiment and data analyses. The research idea was

conceived and designed by Dr. Jaswinder Singh. Dr. Mark Lefsrud has collected and provided diverse pea lines. Neil Dylan Lamb-Palmer and Dr. Manjit Singh also assisted in data analysis especially categorization of diverse Pea lines. All the authors were involved in editing of the manuscript.

This research was funded by Natural Sciences and Engineering Research Council of Canada (NSERC) and Lefsrud Seeds.

*Ahmad, S., Singh, M., Lamb-Palmer, N. D., Lefsrud, M., Singh, J. (2012) Assessment of genetic diversity in *Pisum spp* through microsatellite markers. *Can J Plant Sci.* 92:1075-1081).

CHAPTER 4.0

Assessment of genetic diversity in *Pisum sativum* L. accessions using microsatellite markers

4.1 Abstract

Field pea (*Pisum sativum* L.) is an important pulse crop and is well suited to the cool climatic conditions of Canada. Canadian growers would require improved cultivars periodically to keep pace with growing demand in this crop. However, limited parental material and low state of genetic diversity is available to pea breeders. On the other hand, understanding and kknowledge of genetic diversity is pre-requisite for the creation of new desirable germplasm and elite breeding lines of pea. Several molecular techniques are available that allow effective assessment of genetic architecture of closely related germplasm. In this study we intended to distinguish 35 pea accessions from different countries using 115 microsatellites located on different pea chromosomes. In total, 210 alleles were detected that ranged from 2 to 8 per locus with an average number of 4.2 alleles. Polymorphic Information Content (PIC) values varied from 0.055 to 0.887 with an average of 0.668. Discriminating power (D) values varied from 0.057 to 0.901 with an average of 0.686. Line AA38 (UK) and G611764 (Unknown) have closely similar set of marker profiles with

similarity values of 0.8367 whereas most genetically distant genotypes were Austrian Winter Pea (USA) and ILCA 5077 (Greece) with value of 0.0455. Unweighted Pair Group Method with Arithmetic Averages (UPGMA) cluster analysis grouped pea accessions into six major clusters, mostly consistent with their countries of origin. Majority of Canadian and European genotypes grouped separately suggesting both these groups are from genetically distinct gene pools. The genetically diverse groups identified in this study can be used to derive parental lines for pea breeding.

4.2 Introduction

Field pea (*Pisum sativum* L.) is an economically important cool season pulse crop cultivated around the globe for its protein rich seed and soil improving benefits (McPhee 2003) including its well recognized role in biological fixation of atmospheric nitrogen. Recent studies have shown the significance of pulses in human diet by reducing risk of type II diabetes and cardiovascular disease (Boye et al. 2010). Commercial interests in pea crop and other pulses have been greater than before following a demand increase in protein-rich food and feed (Santalla et al. 2001). Although many pea varieties have been developed with increased yield

potential, modified maturity, lodging resistant (afila type) and better nutritional value (Micke 1988) but pea growers would still require new high performing varieties periodically to cope with increased demand. Successful breeding programs, however, need diverse genetic resources to be used in various crosses in order to maximize genetic gain. Developing high yielding and uniform varieties though guarantee short term increase in yield, but generally these efforts resulted in unwanted loss of genetic diversity (Baranger et al. 2004) and genetic erosion (Akhalkatsi et al., 2010). Other factors such as self-pollination that increases homozygosity due to continual self-pollination may also contribute to loss of genetic variation (Cieslarova et. al., 2011).

The importance of genetic diversity in crops have been highlighted by both the geneticists and plant breeders and that further improvement in exploration of genetic diversity in crops is needed (Able et al., 2007). For example, insufficient genetic diversity leads to reduced breeding improvements as further breeding improvements have been stalled in case of *Phaseolus vulgaris* (McClean et al. 1993). There is scarcity of genetic diversity in *Pisum* and limited amount of parental material is available for pea breeding programs especially in Canada (Tar'an et al. 2005). Genetic diversity in crops is the variation of alleles that enable

species to withstand certain biotic and abiotic stresses and thus is crucial for long term survival of crops. Hence, knowledge of genetic variation is essential for creation of novel germplasm and future breeding lines. Data on genetic relatedness among parental lines facilitates selection of parents to be used in crosses for maximum genetic gain. Several methods are available to assess or capture diversity in diverse genotypes. Examples include allele mining or sequencing-based allele mining technique (Kumar et. al., 2010) and Targeted Induced Local Lesions IN Genomes (TILLING, McCallum et. al., 2000). These are promising methods for capturing genetic diversity but they are still costly and time consuming. Molecular markers on other hand can be efficiently employed for exploration of genetic diversity in crops that can provide authentic and reliable information about cultivar identification, independent of crop stage and environmental interactions. Molecular markers such as Random amplified polymorphic DNA (RAPD, Hoey et al. 1996) and Amplified fragment length polymorphism (AFLP, Simioniuc et al. 2002) were applied to discriminate between wild and cultivated pea species and to study genetic diversity in pea respectively. However, these markers were largely replaced by microsatellites or Simple Sequence Repeats (SSR) have due to their accuracy, reliability, co-dominancy, reproducibility and high polymorphism

and easily detectable by PCR (Powell et al. 1996; Becher et al. 2000). Numerous researchers have used SSR markers to study genetic diversity in peas (Burstin et al. 2001; Loridon et al. 2005; Nasiri et. al. 2009; Sarikamis et al. 2010; Gong et al. 2010).

In the present study, we assessed the genetic diversity of 35 pea genotypes using 115 SSR markers with the following main objectives:

Aim 1: To identify polymorphic microsatellite markers

Aim 2: To explore genetic diversity in diverse pea accessions.

Aim 3: To analyse microsatellites polymorphism and categorization of pea germplasm.

The majority of Canadian and European and genotypes from the USA were grouped into genetically distinct gene pools. These groups can be valuable for the selection of parental lines for future pea breeding.

4.3 Material and Methods

4.3.1 Plant Material

Seeds of 35 different field pea accessions (*Pisum sativum* L., Table 4.1) were acquired from the Plant Gene Resources of Canada (Saskatoon, SK) and the pea collection of the U.S. Department of Agriculture (Pullman,

WA). Seeds were grown in a research greenhouse located on the Macdonald Campus of McGill University, Ste-Anne-de-Bellevue, Quebec, Canada (Lat: 45 39' N). The experiment was conducted in a greenhouse (north south orientation) with raised benches (75cm from the floor). The seeds were placed in a peat-vermiculite soil-less mixture (1:1 sphagnum peat:vermiculite by volume contained in 0.75 L plastic pots. Each pot received a dose (at an equivalent rate of 3 kg m-3) of slow release fertilizer (10-10-10) and was hand irrigated daily with deionized water. The pot density on the benches was 64 m^{-2} for the experiment. The greenhouse section was heated by a hot-water distribution system consisting of overhead pipes, perimeter pipes and tubes embedded in the solid concrete floor. Air temperature at canopy height was maintained by set point at 20/18 (s.d. ±1) °C Day/Night. A 16-hour photoperiod was imposed to maintain temperature but no supplemental lighting was provided. The greenhouse was not controlled for humidity or carbon dioxide. The greenhouse section was outfitted with two horizontal airflow fans that operated continuously to improve mixing of the greenhouse air. The peas were grown until the 8-10 leaf stage for use in DNA extraction.

4.3.2 Genomic DNA Isolation

The reader is referred to section 3.3.3, chapter 3 for genomic DNA isolation protocol.

4.3.3 SSR Primers

A total of 115 SSR primer pairs were selected from a set of primers used by Loridon et al. (2005). These primer sets are located on various pea linkage groups. Initially all these markers were tested for clear banding patterns on a subset of 8 varieties. Only 50 out of 115 produced clear and highly reproducible banding patterns. For this reason, those 50 SSR markers with clear bands and polymorphism were selected and used for further analysis. The remaining SSR primers, which produced non-polymorphic bands or smear patterns, were dropped from the study. In case of smear banding pattern, efforts were made to adjust the annealing temperature and modified the DNA quantities, but the smear pattern was not significantly improved in these primers.

4.3.4 Polymerase Chain Reaction (PCR)

For PCR reaction protocol, the reader is referred to section 3.3.4, chapter 3. Agarose gel (2%) stained with ethidium bromide (EtBr) in Tris Borate

EDTA (TBE) buffer was used and the size of each amplified product was determined by matching with 1 Kb DNA Ladder (Invitrogen, USA).

4.3.5 Allele Scoring

The size of amplified band for each microsatellite marker was determined by matching it with DNA ladder. Amplified products from the microsatellite analyses were scored qualitatively for presence or absence of each marker allele. The SSR bands amplified by using the given primers were further treated as a unit character. Each SSR band was scored as present (1) or absent (0) for each genotype. An accession was assigned a null allele for a microsatellite locus whenever an amplification product could not be detected for a particular genotype marker combination.

4.3.5 Data Analysis

PIC value of each SSR marker was calculated according to the following formula.

Where, pi is the frequency of the ith allele of a given jth SSR marker across all 35 *Pisum* genotypes. D was calculated for each SSR marker

according to the following formula (Tessier et al.1999). Where, N is the total number of *Pisum* accessions and pi is the frequency of the ith allele of a given jth SSR marker.

Frequencies of null alleles were excluded while calculating discriminating power and PIC. For determination of genetic similarity and cluster analysis, the amplified fragments were compiled onto a binary data matrix with each band codified as presence (1) and absence (0) for each marker allele. The binary data matrix was further analysed by PAST software to determine genetic relationships among pea accessions using UPGMA and Jaccards's Index to develop a dendrogram.

4.4 Results and Discussion

4.4.1 Microsatellites Polymorphism

Molecular diversity studies can have significant applications in discriminating genotypes within and between populations using informative markers and/or combination of markers. Microsatellites are among the preferred markers technology recently been employed by various researchers for genetic diversity analysis in different crops

including barley (Russell et al. 1997), maize (Pejic et al. 1998), wheat (Bohn et al. 1999), rice (Temnykh et al. 2000), and soybean (Tantasawat et al. 2011). Microsatellites were employed in the present study to assay genetic diversity in 35 selected pea (*Pisum sativum*) accessions from various countries of origin (Table 4.1).

Pea accessions utilized in this study were differentiated by microsatellite markers, indicating their utility in assessing genetic diversity. A total of 115 SSR markers initially were used to analyze pea accessions. However, those markers that produced faint and unclear bands were removed from the study, leaving only 50 scorable SSR markers for further investigation. These 50 SSR markers generated polymorphic alleles revealing considerable variability and genetic diversity among the various pea genotypes. Detection of polymorphic allele suggest the existence of considerable genetic variability among the current group of various pea genotypes as shown in allelic profile of SSR marker in Figure (4.1) as an example of polymorphism. A total of 210 alleles were detected across 35 pea accessions. Number of alleles per locus ranged from 2 (AA206, AA163.2, AD147, AC75, AA103) to 8 (AA92) with a mean of 4.2 alleles (Table 4.2)

SSR marker AA92 amplified the maximum number of alleles (8) followed by markers AD 73 and D21 both with seven alleles each for a total of 14 alleles. Six alleles each were detected by 6 SSR markers (AA372.1, D23, AD270, AA5, AD59, and AA335) with 36 alleles in total. Eleven SSR markers revealed 5 alleles each (55 alleles altogether) whereas 12 markers amplified 4 alleles each making altogether 48 alleles. Maximum number of markers (13) amplified 3 alleles each whereas five markers amplified 2 alleles each for a total of ten alleles. We also observed high homozygosity in pea accessions as expected from self-pollinated crops such as pea (Cieslarova et. al., 2011). All pea accessions under our consideration were highly homozygous as expected suggesting these accessions were advanced breeding lines (Smith et al. 2000). No marker heterozygosity was found in present investigation contrary to other studies where heterogeneity in SSR marker analysis was observed while assessing historic and wild pea varieties (Loridon et al. 2005; Nasiri et al. 2009). Cross-pollinated species on the other hand tend to be heterozygous as much higher heterogeneity in SSR marker analysis was noted in case of rape (Pascher et al. 2010). Markers in this study were found to be multi-allelic detecting an average of 4.5 alleles per locus using 50 SSR loci. Cupic et al. (2009) reported exactly same results with an

average of 4.5 alleles per locus using 30 SSR markers in a population of 18 pea accessions. Zong et al. (2009) observed a slightly higher number of alleles per locus with an average of 4.9 alleles in 197 pea genotypes using 21 SSR markers. The average number of alleles per locus found in present study was almost at par with number of alleles reported by Cupic et al. (2009) and Zong et al. (2009) suggesting a similar genetic base of these pea accessions. However, a higher mean of 5.9 alleles per locus was detected in 20 pea varieties and 57 wild pea accessions using10 SSR markers (Nasiri et al., 2009). This higher mean of alleles per locus observed by Nasiri et al. (2009) is due to wild pea genotypes in their study thus making the average of total detected alleles higher. Generally, wild pea accessions display higher genetic diversity than cultivated cultivars. This also is suggestive of a broad genetic base of pea varieties and wild accessions used in Nasiri et al., (2009) study.

PIC and D values of each marker (Table 4.2) were used as decisive factor for the informativeness of each marker in resolving the diversity of *Pisum* accessions. The most informative loci were AA121 and AD148 with PIC value of 0.887 and 0.867 respectively. The mean PIC value for markers was 0.668. Among these SSR markers, the least informative locus was found to be AA206 with lowest PIC value of 0.055. Similarly the highest D

value of 0.901 was observed in marker AA121 followed by AD148 with D value of 0.8857. Marker AA206 showed lowest D value of 0.057. The average D value was 0.686. Both PIC and D scores of 35 microsatellites showed sufficient polymorphism to discriminate all 35 *Pisum* accessions in our study. AS mentioned earlier, marker AA121 was found highly informative with highest PIC value (0.887) and D value (0.901). However, Loridon et al. (2005) noted a slightly low level of polymorphism for the same AA121 marker (i.e. PIC=0.75). Lowest PIC (0.055) and D value was produced by SSR marker AA206 in our study making it the least informative locus across the 35 pea accessions. In contrast, Loridon et al. (2005) had a high level of polymorphism for this marker (AA206 PIC=0.73) suggesting that alleles of marker AA206 was uniformly distributed among pea accessions used by Loridon et al. (2005). Similarly, Nasiri et. al, (2009) observed higher level of polymorphism for marker AF004843 with 8 alleles while it was reported as low polymorphic in previous studies (4 alleles, Burstin et al. 2001). Interestingly, while analysing the same group of pea accessions using 15 SSR markers much smaller PIC and D values with a mean score of 0.460 and 0.475 respectively was reported (Ahmad et al 2012). In addition, the maximum number of alleles detected by Ahmad et al. (2012) again was low (i.e. 4 alleles). This indicates that SSR

markers used in present study have higher level of polymorphism among these loci which will facilitate the selection of informative markers for subsequent analysis of genetic variation in pea.

4.4.2 Genetic Diversity and Cluster Analysis

A pairwise genetic similarity (GS_j) matrix was calculated using Jaccard's coefficient. Genetic similarity values ranged from 0.0455 to 0.8367 with a mean similarity value of 0.2051. Line AA38 (UK) and G611764 (Unknown) were found to be genetically similar with highest pairwise genetic similarity value of 0.8367 followed by 0.5082 between Chinese Snow Pea (USA) and Red Small Pea (India). A pair wise similarity between Stella and Agaggiz and between Canstar and Stella (all from Canadian origin) was noted as 0.5 for these two pairs of genotypes. Genetically distant genotypes were found to be Austrian Winter Pea (USA) and ILCA 5077 (Greece) with lowest similarity value of 0.0455 followed by GSj value of 0.0674 between Line 22719 (Turkey) and Austrian Winter Pea (USA). The genetic similarity data may be valuable for designing breeding programs. An introduction of exotic breeding material into genetically similar genotypes, for example, Line AA38 (UK) and G611764 (Unknown) found in present study, may increase the genetic diversity that can be

instrumental in maximizing genetic gain by allowing favourable allele combinations. In order to broaden the genetic base of soybean (*Glysine max* L.) in the USA, two cultivars from China (previously not in ancestry of any U.S. cultivar or germplasm) were used as parental material. The resulting germplasm LG00-6313, recently released may provide new alleles for maximum genetic gain (Nelson and Johnson 2011).

Cluster analysis was performed using UPGMA to construct a dendrogram from the pairwise similarity matrix (Figure 4.2). Cluster analysis classified pea genotypes into six major groups that were mostly consistent with their geographical origins. The first cluster (I) consists of five pea genotypes, Agaggiz, Stella, Canstar, and Thunderbird (all from Canadian origin) except Maple pea NZ (USA). As Canada is sharing a long border with the USA, it is not uncommon to observe clustering of genotypes from these countries in same group. Cluster II is mainly dominated by genotypes with origin from the USA. Cluster II comprises seven accessions viz. Oregon Sugar II (USA), Galena (Unknown), Frosty (USA), Line 45760 (Unknown), Super Sugar Snap (USA), Dakota (Canada), and Wando (USA). Again, only one Canadian pea line in the USA dominated pea cluster is not surprising as both countries being closely neighbours. Cluster III consists of four accessions mostly from European origin i.e. ILCA 5077 (Greece),

ILCA 5075 (Syria), ILCA 5052 (Cyprus), and Line 22722 (Turkey). Cluster IV can be considered as inconsistent i.e. genotypes from various origins were intermixed in this group. This cluster is relatively a big group with ten genotypes namely, Big Pea (Costa Rica), Dull White Pea (India), ILCA 5094 (Albania), Line 5115 (Spain), and Line 22719 (Turkey), Green Small Pea (India), Line 31657 (Unknown), ILCA 5117 (Iran), Chinese Snow Pea (USA), and Red Small Pea (India). This inconsistent group probably may share a common parent resulted from germplasm exchange programs among different countries (Kuleung et al 2006). Semileafless (*afila*) leaf trait have been introduced in other varieties through germplasm exchange programs. Similarly, exchange of breeding material between China and Australia has been reported by Zong et al., (2009) for broadening their respective pea breeding gene pools. Six lines were grouped in Cluster V with Line 25579 (Unknown), G611764 (Unknown), Line AA38 (UK), Line 29559 (Unknown), ILCA 5089 (Albania), and ILCA 3005 (Greece). Only two accessions i.e. Austrian Winter Pea (USA), and ILCA 5032 (Yugoslavia) were grouped in cluster VI. Interestingly, Line 112340 with unknown country of origin is a solitary member and was assigned to a separate cluster (which was ignored).

Using unique genetic gene pools can lay the foundation of genetic improvements and can be useful in future breeding programs. Cluster analysis classified pea genotypes into various groups and inter-introgression of pea accessions, for example, from groups of Canadian, European and/or the USA origin can be instrumental in widening the genetic base and can increase genetic diversity in future varieties. Similarly, Islam et al. (2004) also suggested that in common beans, the narrow genetic base of Andean gene pool can be improved by introduction of genes from Meso-American gene pool. The data generated here can be useful in breeding programs, genetic diversity studies, and conservation of germplasm as well as better management of seed banks. This study also showed the importance and utility of SSR markers in genetic diversity studies.

Table 4.1: Pea accessions used in present study

Sr. No	Pea Genotypes	Accession No.	Country of Origin
1	Red Small Pea	PI 471294	India
2	Line 29559	Unknown	Unknown
3	ILCA 5077	PI505112	Greece
4	Big Pea	PI 262189	Costa Rica
5	ILCA 5094	PI 505127	Albania
6	Line 25579	Unknown	Unknown
7	ILCA 5117	PI 505146	Iran
8	Dull White Pea	PI 471312	India
9	Austrian Winter Pea	PI 517922	USA, Idaho
10	Agaggiz	Unknown	Canada
11	ILCA 5115	PI 505144	Spain
12	Chinese Snow Pea	PI 279933	USA, New York
13	ILCA 5032	PI 505074	Yogoslavia
14	Stella	Unknown	Canada
15	Thunder Bird	Unknown	Canada
16	Canstar	Unknown	Canada
17	Line 22722	PI 343990	Turkey
18	Maple pea NZ	PI 236494	USA, Iowa
19	ILCA 5052	PI 505092	Cyprus
20	ILCA 5089	PI 505122	Albania
21	ILCA 3005	Unknown	Greece
22	Oregon Sugar II	Unknown	USA
23	ILCA 5075	PI 505111	Syria
24	G 611764	Unknown	Unknown
25	Line AA38	PI 269762	UK
26	Galena	Unknown	Unknown
27	Frosty	Unknown	USA
28	Super Sugar Snap	Unknown	USA
29	Line 45760	Unknown	Unknown
30	Green Small Pea	PI 471211	India
31	Line 22719	PI 343988	Turkey
32	Line 31657	Unknown	Unknown
33	Dakota	Unknown	Canada
34	Wando	Unknown	USA
35	Line 112340	Unknown	Unknown

Table 4.2: SSR markers from different chromosomes and their Polymorphic Information Content (PIC) and Discriminating power (D) values over 35 pea genotypes

Marker	Sequence	Linkage group	PIC	D	Allele no
AA92	F aaggtctgaagctgaacctgaagg	III	0.8588	0.8807	8
	R gcagcccacagaagtgcttcaa				
AD73	F cagctggattcaatcattggtg	III	0.8392	0.8639	7
	R atgagtaatccgacgatgcctt				
D21	F tattctcctccaaaatttcctt	II	0.7935	0.8168	7
	R gtcaaaattagccaaattcctc				
AA372.1	F gagtgaccaaagttttgtgaa	II, VI	0.8065	0.8286	6
	R ccttgaacccatttttaagagt				
D23	F atggttgtcccaggatagataa	II	0.7967	0.8185	6
	R gaaaacattggagagtggagta				
AD270	F ctcatctgatgcgttggattag	III	0.7853	0.8067	6
	R aggttggatttgttgtttgttg				
AA5	F tgccaatcctgaggtattaacacc	III	0.7788	0.8000	6
	R catttttgcagttgcaatttcgt				
AD59	F ttggagaatgtcttctctttag	VI	0.7118	0.7294	6
	R gtatattttcactcagaggcac				
AA335	F acgcacacgcttagatagaaat	VI	0.6890	0.7092	6
	R atccaccataagttttggcata				

Table 4.2, continued

AD148	F gaaacatcattgtgtcttcttg	II	0.8669	0.8857	5
	R ttccatcacttgattgataaac				
AB146	F ggaaattggaaggagctatttgaag	V	0.8171	0.8387	5
	R gtgcataagcatttgattagatgacc				
AA480	F caattttatgctacacatactccct	II	0.7878	0.8084	5
	R tacagaagcatttgtgcagttgt				
A5	F gtaaagcataaggggattctcat	II	0.7869	0.8101	5
	R cagcttttaactcatctgacaca				
AB140	F ccagattcatgaagggcataca	III	0.7543	0.7748	5
	R gatgaaatttcgtttctctctgtctc				
AD56	F gaaacattggttgaagagcgag	VII	0.7527	0.7748	5
	R gttgtcgcgtgaacacaagtaa				
AD146	F tgctcaagtcaatatatgaaga	VII	0.7355	0.7546	5
	R caagcaaatagttgttttgtta				
AA491	F gaggtggtgttgaatttgtg	III	0.7331	0.7546	5
	R cctaattttacccctctctctct				
AD61	F ctcattcaatgatgataatccta	III	0.7331	0.7546	5
	R atgaggtacttgtgtgagataaa				
AB40	F aaatagacccttgtgtagaagc	II, VII	0.7053	0.7261	5
	R ggaaaagtgggttttgaa				
AA339	F gtgtagaagtattttacttgatg	II	0.5633	0.5798	5
	R catctattgaaggaaaattat				

Table 4.2, continued

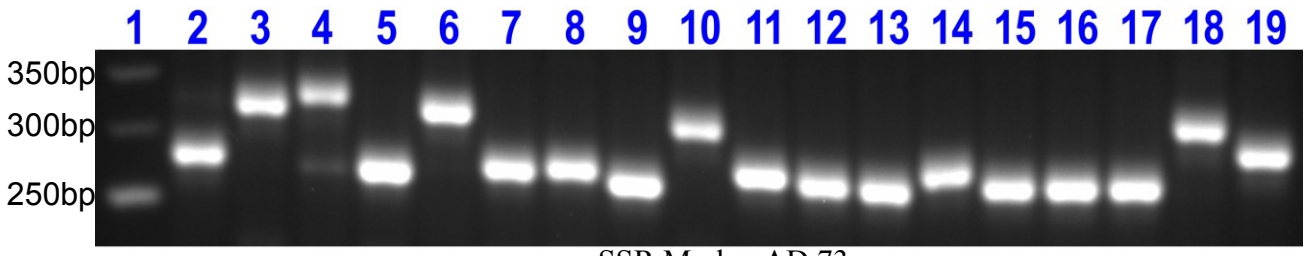
AA175	F ttgaaggaacacaatcagcgac	III	0.8465	0.8639	4
	R tgcgcaccaaactaccataatc				
AA473	F caatcgatcagacagtccccta	II	0.7943	0.8134	4
	R aagctcacctggttatgtccct				
AA122	F gggtctgcataagtagaagcca	IV	0.7151	0.7361	4
	R aaggtgtttcccctagacatca				
AA153	F tttgatagtccgacttttccat	II	0.7102	0.7294	4
	R gtgacaaaagaattcaaaacgc				
PSGAPA1	F gacattgttgccaataactgg	V	0.6971	0.7176	4
	R ggttctgttctcaatacaag				
AB53	F cgtcgttgttgccggtag	III	0.6922	0.7109	4
	R aaacacgtcatctcgacctgc				
AA238	F tatcatcaaggtccaatttagt	II	0.6180	0.6353	4
	R agctaaatcgtacctaatctgt				
AA446	F ttagcttgcagcccactc	VII	0.6008	0.6168	4
	R atccgacccatggattta				
AD83	F cacatgagcgtgtgtatggtaa	II	0.5861	0.6033	4
	R gggataagaagagggagcaaat				
AB141	F atcccaatactcccaccaatgtt	III	0.5706	0.5866	4
	R agacttaggcttcccttctacgactt				
AA135	F ccgttacacatcattaagatg	VII	0.5208	0.5361	4
	R tccatatccagattagtcaga				

Table 4.2, continued

A9	F: gtgcagaagcatttgttcagat	IV	0.651	0.669	4
	R: cccacatatatttggttggtca				
AA121	F tccataccttagtgttaaa	I	0.8873	0.9008	3
	R actaataaggtaaacatgtg				
AD60	F ctgaagcacttttgacaactac	VI	0.7608	0.7782	3
	R atcatatagcgacgaatacacc				
AB23	F tcagcctttatcctccgaacta	V	0.6922	0.7109	3
	R gaacccttgtgcagaagcatta				
A6	F cttaagagagattaaatggacaa	III	0.5796	0.5950	3
	R ccaactcataataaagattcaaa				
AA67	F: cccatgtgaaattctcttgaaga	I	0.513	0.528	3
	R: gcatttcacttgatgaaatttcg				
AA205	F: tacgcaatcatagagtttggaa	II	0.625	0.648	3
	R:aatcaagtcaatgaaacaagca				
AB72	F:atctcatgttcaacttgcaaccttta	II	0.660	0.679	3
	R: ttcaaaacacgcaagttttctga				
AA175	F: ttgaaggaacacaatcagcgac	III	0.312	0.322	3
	R: tgcgcaccaaactaccataatc				
AA285	F: tcgcctaatctagatgagaata	IV	0.526	0.541	3
	R: cttaacattttaggtcttggag				
AC58	F: tccgcaatttggtaacactg	V	0.614	0.632	3
	R: cgtccatttcttttatgctgag				

Table 4.2, continued

AD51	F: atgaagtaggcatagcgaagat R: gattaaataaagttcgatggcg	VI	0.336	0.350	3
AC76a	F: cccaatccaataaataaagaaa R: aatggttgttatgccatttt	VI	0.594	0.612	3
AA90	F: cccttaccatatttcgtttct R: tgcgactccattctagtattg	VII	0.486	0.501	3
AA206	F: ctgagaactcaacgctcagacg R:cgagggtcgagttctgagattt	VII	0.055	0.057	2
AA163.2	F: tagtttccaattcaatcgacca R: agtgtattgtaaatgcacaaggg	V	0.245	0.257	2
AD147	F: agcccaagtttcttctgaatcc R: aaattcgcagagcgtttgttac	I	0.472	0.487	2
AC75	F: cgctcaccaaatgtagatgataa R: tcatgcatcaatgaaagtgataaa	I	0.408	0.420	2
AA103	F: aagtgtgaaagtttgccaggtc R: cgggtacgggttatgttgtc	VI	0.408	0.420	2
		Mean	**0.668**	**0.686**	**4.2**



SSR Marker AD 73

Figure 4.1: Allelic profile of SSR marker AD 73 on 2 % agarose gel showing polymorphism across 18 pea accessions. **Pea genotypes:** 1.DNA ladder, 2. Red Small Pea (India), 3. Line 29559 (Unknown), 4. ILCA 5077 (Greece), 5. Big Pea (Costa Rica), 6. ILCA 5094 (Albania), 7. Line 25579 (Unknown), 8. ILCA 5117 (Iran), 9. Dull White Pea (India), 10. Austrian Winter Pea (USA), 11. Line 295115, 12. Agaggiz (Canada), 13. Chinese Snow Pea (USA), 14. ILCA 5032 (Yugoslavia), 15. Stella (Canada), 16. Thunder Bird (Canada), 17. Canstar (Canada), 18. Line 22722 (Turkey), 19. Maple pea NZ (USA).

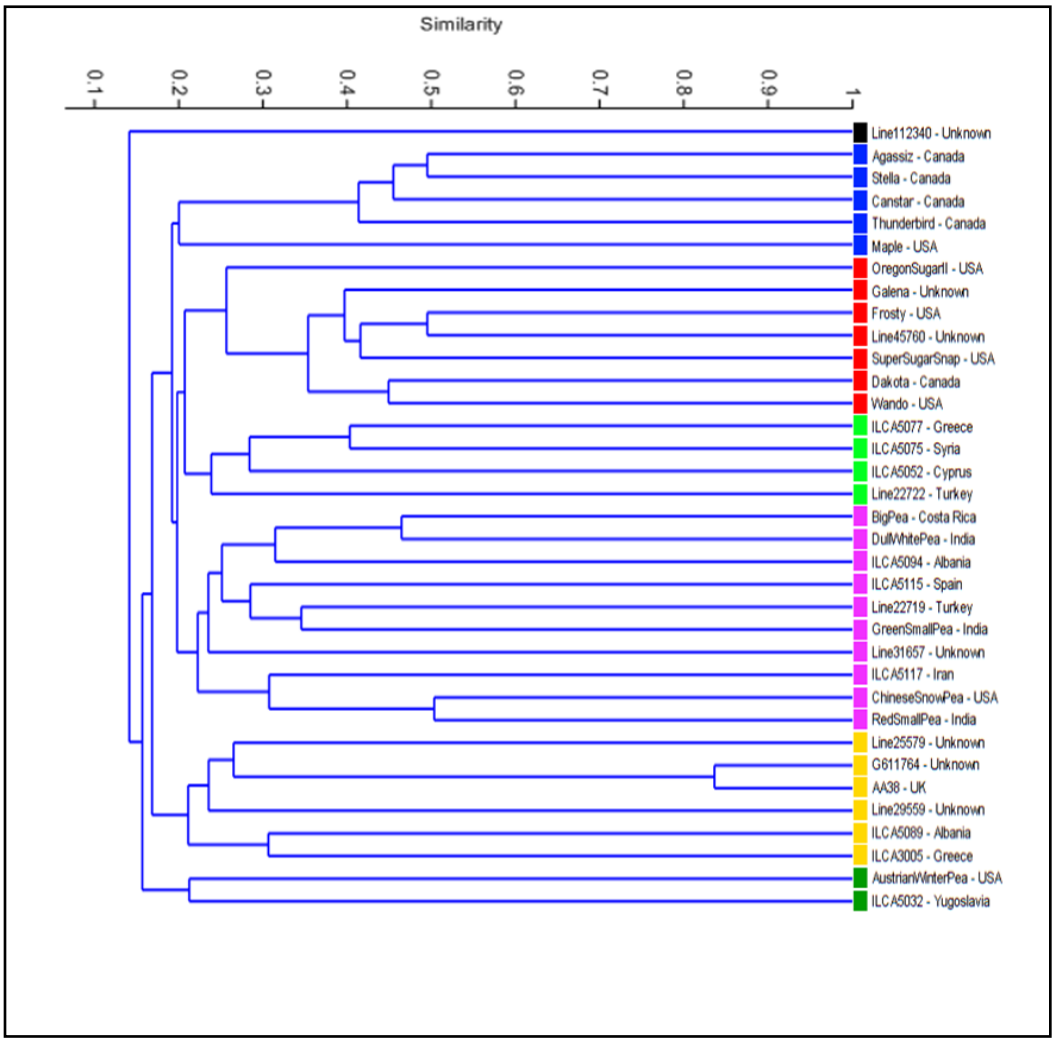


Figure 4.2: Dendrogram showing classification of 35 *Pisum* accessions based on molecular analysis obtained from 50 SSR markers. Genetic relationships among pea accessions were evaluated using an unweighted pairgroup method of arithmetic averages (UPGMA) and Jaccards's Index to develop a dendrogram. Coloured blocks indicate grouping of accessions into genetically similar clusters.

4.5 Connecting Text

Transposable or mobile genetic elements known as transposons constitute a major portion of many eukaryotic genomes. These elements are present in genomes of all higher plants with different origins and nucleotide sequences (Aim 1: To identify transposon like structures in pea genome)

There are two major classes of transposons viz. Class I (Retrotransposons or RNA-based transposons) and Class II (DNA transposons). Due to movement and abundant presence of transposons in genome, they can be effectively used in studies like fingerprinting and varietal identification as an alternative method to other available molecular marker technologies such as microsatellites (Aim 2: To study the diversity of RNA-based transposons and DNA-based transposons in diverse pea breeding lines i.e. *Cyclop* and *Ogre*, RNA-based transposons and *Mutator* and *MITES*, DNA-based transposons; Aim 3: Characterization of Transposon DNA sequences to explore their potential for use as molecular markers).

The following manuscript was authored by the candidate. The research idea was conceived and designed by Dr. Jawinder Singh, Department of Plant Science, Macdonald Campus of McGill University and Dr. Mark

Lefsrud, Department of Bio-Resource Engineering, Macdonald Campus of McGill University. S. Ahmad performed the experiments and data analyses. This research was funded by Natural Sciences and Engineering Research Council of Canada (NSERC), and Lefsrud Seeds

CHAPTER 5.0

Investigation of IRAP transposon based molecular markers in pea for future breeding purposes

5.1 Abstract

Transposition activity of transposons creates DNA polymorphism and their abundant presence in genomes are making transposons a promising marker system for varietal identification and fingerprinting. In this study, we employed four transposon-based markers (two DNA- and two RNA-transposons) to evaluate the effectiveness of Inter-Retrotransposon Amplified Polymorphism (IRAP) transposon system in assessing genetic diversity in pea accessions. A total of 28 alleles were detected across the 35 pea accessions with number of alleles per locus ranged from 5 (*Mutator*) to 9 (*Cyclops*). RNA transposons produced a higher number of polymorphic alleles (*Ogre*: 8, *Cyclops*: 9) than DNA transposon markers (*Mutator*: 5, *MITE*: 6). Overall mean PIC value and D value obtained from both these two types of transposon markers was 0.810 and 0.817 respectively. Genetic similarity values ranged from 0.143 to 0.823 with a mean similarity value of 0.403. Cluster analysis classified pea genotypes into six major groups that were somewhat consistent with their geographical origins. Despite low number of markers used in this study, all

35 accessions were differentiated and the overall molecular analyses and data generated higher PIC and D values that can be useful for MAS-based breeding programs in pea.

5.2 Introduction

Transposable or mobile elements known as transposons constitute the major portion of the genome (Neumann et al., 2001; Munoz-Lopez and Garcia-Perez 2010). The major two classes of transposons viz. Class I (Retrotransposons) and Class II (DNA transposons) move in the genome. The former transposes via RNA intermediate, often referred as copy-and-paste mechanism (Bennetzen, 2000) and the later move directly as DNA, using an excision-reintegration or cut-and-paste method. Due to replicative transposition mechanism of retrotransposons, they are found in high copy numbers while the DNA transposons are moderately repeated in genome because of their cut-and-paste system of transposition (Bennetzen, 2000). Numerous families of retrotransposons and DNA transposons have been identified since the early discovery of transposons by Barbara McClintock in 1940s. For example, some of retrotransposon super-families are, *LINE* (Long Interspersed Nuclear Elements), *SINE* (Short Interspersed Nuclear Elements), *Ty3/gypsy-like* and *Ty1/copia-like* (Macas et al. 2007). The

families of DNA transposons in plants are *Ac/Ds, En/Spm, PIF,* and *Mutator*. There are some unclassifiable transposons that are small, having several hundred base pairs (Jiang et al., 2003).

Several molecular marker technologies are available for fingerprinting and assessment of genetic diversity in various crops. An alternative marker system is always of interest even if a polymorphic and efficient marker system such as microsatellites is available. Transposable elements (TE)-based fingerprinting has recently emerged as a marker system for varietal identification following recognition of their abundance (Smykal 2006). Transposition activity resulted in both DNA polymorphism at insertion sites and small scale reshuffling of genome, thus creating diversity (Bennetzen 2000). A role in generating diversity together with their ubiquitous nature is making TEs important tools for use as molecular markers (Kumar and Hirochika 2001). For instance, Jing et al (2010) genotyped 3020 samples of *Pisum* germplasm for 45 retrotransposons based insertion polymorphism (RIBP) markers. Characterization of *Stowaway MITEs* in pea identified 1500 copies of *Stowaway* elements in the haploid genome of pea (Macas et al. 2005). In addition, Macas et al (2003) reported *Zaba* as the novel miniature transposable element, but are moderately repetitive in pea genome as compared to other *Medicago* and *Vicia* species.

Inter-Retrotransposon Amplified Polymorphism (IRAP) and Retrotransposon Microsatellite Amplified Polymorphism (REMAP) systems have been used for the identification of polymorphic transposon insertion sites (Kalendar et al., 1999). This project aimed at the assessment of genetic diversity using transposons from both Class I and Class II transposons in diverse pea genotypes. In this study, we have analysed transposons from diverse pea germplasm to investigate their utility as future molecular markers with the following main objectives:

Aim 1: To identify transposon like structures in pea genome

Aim 2: To study the diversity of RNA-based transposons and DNA-based transposons in diverse pea breeding lines i.e. *Cyclop* and *Ogre* (RNA-based transposons) and *Mutator* and *MITES* (DNA-based transposons).

5.2 Methods and Materials

The reader is referred to Chapter 4.0 for methods and materials section as these are exactly same as mentioned in Chapter 4 (section Methods and Material) except the section of primer designing.

5.2.1 Primer Selection

The transposon primers were selected from a set of primers used by Smykal (2006). The primers sequences are given in Table 5.1.

5.3 Results and Discussion

5.3.1 Transposon marker polymorphism

As described earlier, molecular diversity studies can have significant applications in discriminating genotypes within and between populations. While microsatellites are among the preferred marker types recently employed by various researchers for genetic diversity analysis (Russell et al. 1997, Tantasawat et al. 2011), here we assess the effectiveness of IRAP transposon markers. Two DNA (*MITE and Mutator*) and two RNA transposons (*Ogre and Cyclop*) were assessed to quantify genetic diversity in the same 35 selected pea (*Pisum sativum*) accessions from various countries of origin (Table 4.1). All pea accessions utilized in this study were differentiated using these IRAP markers, indicating their comparable utility with microsatellites in assessing genetic diversity. Previous studies have made similar suggestions (Grzebelus 2006). These 4 IRAP markers generated polymorphic alleles revealing considerable

variability and genetic diversity similar to the previous microsatellite study. Transposons markers typically produced a higher number of polymorphic alleles, as shown in the allelic profile of the transposons *MITE, Mutator, Ogre, and Cyclop* in Figures 5.1, 5.2, 5.3, and 5.4 respectively, indicating the use of transposons equally valuable for studies like varietal identification. Similar observations were also reported by Sant'Ana et al. (2012). A total of 28 alleles were detected across the 35 pea accessions. Number of alleles per locus ranged from 5 (*Mutator*) to 9 (*Cyclops*) with a mean of 7 alleles (Table 5.1). RNA transposons were observed to produce a higher number of polymorphic alleles (*Ogre*: 8, *Cyclops*: 9) than DNA transposon markers (*Mutator*: 5, *MITE*: 6). Scoring the transposon allelic profile of *Pisum* accessions proved significantly more difficult than with microsatellites due to some banding pattern ambiguity and smearing. The RNA transposons assessed in this study produced relatively ambiguous amplification than DNA transposons. This is a contrasting observation made by Smykal (2006) where the DNA transposon *MITE* was found to produce ambiguous banding pattern in comparison with RNA- based transposons like *Ogre* or *Cyclops*.

PIC and D values of each marker (Table 5.1) were used to assess the diversity of Pisum accessions as described previously in microsatellite

study (Symkal 2008; Ahmad et al., 2012). RNA transposons were found to be the most informative, with Cyclop having a PIC value of 0.870 and a D value of 0.875, followed by Ogre with a PIC value of 0.855 and a D value of 0.864. The DNA transposons had relatively lower scores with Mutator having a PIC value of 0.778 and a D value of 0.786, and MITE with a PIC value of 0.736 and a D value of 0.745. The mean PIC value for transposon markers was 0.810 and the mean D value was 0.817. These values are somewhat higher than average PIC and D scores found with microsatellite markers observed in Chapter 4. However, it is possible that these scores may be inflated due to ambiguity in allele scoring. Also the number of transposons markers used in this study is small (4 transposon species) as compared to the previous study of microsatellites (50) in Chapter 4. However, Smykal (2006) noted a substantially lower level of polymorphism for both Cyclop (PIC=0.299) and Ogre (PIC=0.236) over 33 Pisum accessions. This is likely due to the use of closely related accessions in their study. Converselly, Campbell (2011) reported also reported a similar low PIC scores for IRAP markers in barley despite a higher average number of alleles (15).

5.3.2 Genetic diversity and cluster analysis

A pairwise genetic similarity (GS_j) matrix was calculated using Jaccard's coefficient. Genetic similarity values ranged from 0.143 to 0.823 with a mean similarity value of 0.403. Line AA38 (UK) and G611764 (Unknown) were again found to share a high genetically similarity with the highest pairwise genetic similarity value of 0.823 followed by 0.8 between Chinese Snow Pea (USA) and Red Small Pea (India) and between DullWhitePea (India) and BigPea (Costa Rica). Genotype pairs with the lowest genetic similarity were found to be Super Sugar Snap II (USA) and Red Small Pea (India) with lowest similarity value of 0.146 followed by GSj value of 0.167 pairs Wando (USA) and Austrian Winter Pea (USA) and between ILCA 5094 and Red Small Pea (India). Genetic similarities were on average higher using transposon markers when compared with previous microsatellite data. However, this is likely due to the limited number or transposon markers assessed here. Overall, these four transposon markers were able to differentiate between all accessions. The genetic similarity data may be valuable for designing breeding programs.

Cluster analysis was performed using UPGMA to construct a dendrogram from the pairwise similarity matrix (Figure 5.5). Cluster analysis classified pea genotypes into six major groups that were somewhat consistent with

their geographical origins. The first cluster (I) consists of six pea genotypes, Chinese Snow Pea (China), Red Small Pea (India), Line 12340 (Unknown), ILCA 5077 (Greece), ILCA 3005 (Greece), and ILCA 5075 (Syria). This cluster did not show any geographical consistency, though all three ILCA lines originate from countries bordering the Mediterranean Sea. Cluster II contains two genotypes of South European origin: ILCA 5082 from the former Yugoslavia and Line 22722 from Turkey. Cluster III is another small cluster containing two accessions of unknown origin: Line 25579 and Line 31657. Cluster IV can be considered as inconsistent i.e. genotypes from various origins were intermixed in this group. This large cluster groups twelve genotypes: Big Pea (Costa Rica), Dull White Pea (India), Stella (Canada), Thunderbird (Canada), ILCA 5115 (Spain), ILCA 5094 (Albania), Maple (USA), ILCA 5052 (Cyprus), Agassiz (Canada), Canstar (Canada) Austrian Winter Pea (USA), and ILCA 5089 (Albania). This inconsistent group is likely due to the limited number or transposon markers assessed here, though germplasm exchange programs may have an effect as discussed in Chapter 4 (Kuleung et al., 2006). Four lines were grouped in Cluster V with ILCA 5117 (Iran), G611764 (Unknown), Line AA38 (UK), Line 22719 (Turkey). The remaining nine accessions were grouped into cluster VI: Oregon Sugar II

(USA), Galena (Unknown), Frosty (USA), Super Sugar Snap (USA), Line 45760 (Unknown), Wando (USA), Dakota (Canada), Line 29559 (Unknown), and Green Small Pea (India). Cluter VI is dominated by North American lines with the exception of Green Small Pea from India and three unknown lines.

Our findings in this study are in some agreement with previous work by Smykal (2006). While all four of the IRAP markers used in this study proved to be substantially more informative than observed by Smykal (2006) and Campbell (2011), we did discern that RNA transposon markers were more informative than DNA transposons in terms of PIC and D scores. Additionally, IRAP markers produced with both forward and reverse primer pairs were found to be more informative than markers produced with a single primer (*MITE*). These four IRAP markers were found to be on average more informative than the SSR markers described in Chapter 4 and were able to distinguish between all 35 *Pisum* accessions with only slightly higher pairwise similarity values. While we agree with Smykal (2006) and Campbell (2011) that IRAP markers provide many benefits over SSR markers in terms of cost and time required to obtain comparable resolution in fingerprinting and diversity studies, we observed a major drawback to IRAP in the prevalence of ambiguous

bands and smearing. This impedes the scoring process and can result in non-reproducible diversity findings. DNA transposon markers were found to have substantially less smearing and fewer ambiguous bands than RNA transposons, which is counter to the findings of Smykal (2006). While the overall molecular marker analyses and data generated in this research will be useful for pea breeding, marker assisted selection (MAS), genetic conservation and seed banks management, due to low number of transposons-based markers used in this study, proper conclusions and confident analyses cannot be drawn.

Table 5.1: Transposon markers and their Polymorphic Information Content (PIC) and Discriminating power (D) values for 35 pea genotypes

Marker	Sequence	PIC	D	Allele no	Base Pair range
MITE	P: CTGTGAATTTTTCCTTGCCTCCCTC	0.736	0.745	6	2000 – 590
Mutator	P: GGGAATTCGACGAAATGGAGGC	0.778	0.786	5	760 – 425
Ogre	F: TCGCGAGACCATGTCTTTTCCCAGGTTTAC	0.855	0.864	8	815 – 80
	R: GTGGGCTGGGCTTTAGTGAGATGCTTTCC				
Cyclops	F: CGATATCTCACAATCCCTGTGGAGAC	0.870	0.875	9	755 – 105
	R: GCAAGGAAACGGAGTGAAAGATGC				
	Mean	**0.810**	**0.817**	**7**	-

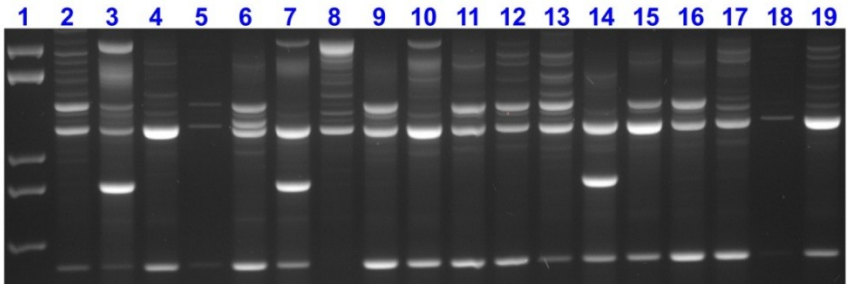


Figure 5.1: Agarose gel (2%) showing Allelic profile of DNA Transposon Marker MITE (See pea genotype legend at figure 5.4.)

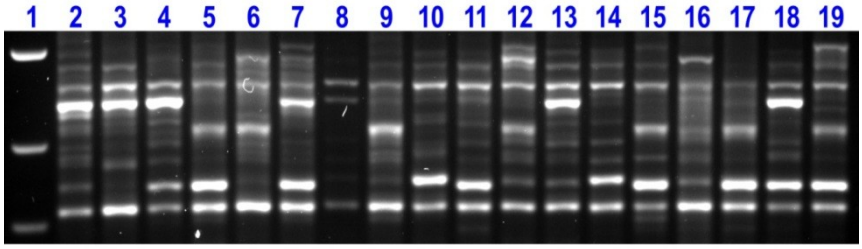


Figure 5.2: Agarose gel (2%) showing allelic profile of DNA Transposon Marker *Mutator* (See pea genotype legend at figure 5.4.)

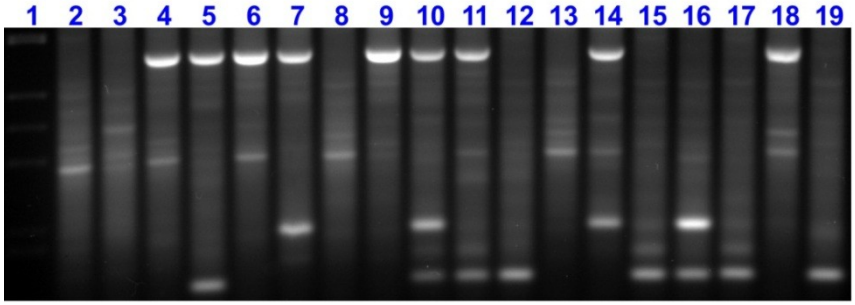


Figure 5.3: Agarose gel (2%) showing allelic profile of RNA Transposon Marker Ogre (See pea genotype legend at figure 5.4.)

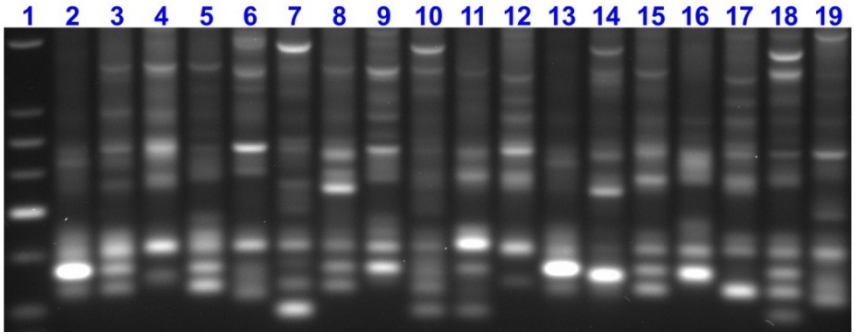


Figure 5.4: Agarose gel (2%) showing allelic profile of RNA Transposon Marker Cyclop

Pea genotypes: 1.DNA ladder, 2. Red Small Pea (India), 3. Line 29559 (Unknown), 4. ILCA 5077 (Greece), 5. Big Pea (Costa Rica), 6. ILCA 5094 (Albania), 7. Line 25579 (Unknown), 8. ILCA 5117 (Iran), 9. Dull White Pea (India), 10. Austrian Winter Pea (USA), 11. Line 295115, 12. Agaggiz (Canada), 13. Chinese Snow Pea (USA), 14. ILCA 5032 (Yugoslavia), 15. Stella (Canada), 16. Thunder Bird (Canada), 17. Canstar (Canada), 18. Line 22722 (Turkey), 19. Maple pea NZ (USA).

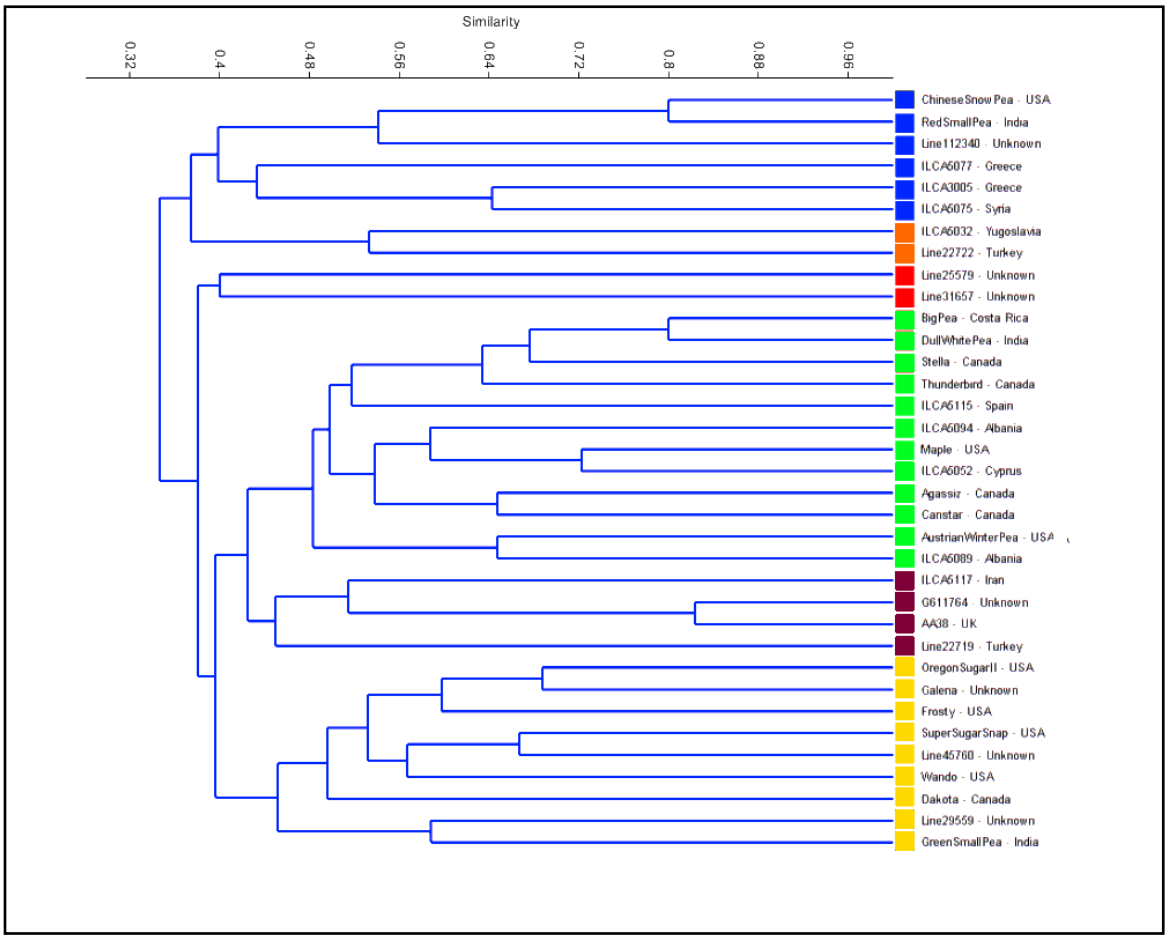


Figure 5.5: Dendrogram showing classification of 35 *Pisum* accessions based on IRAP analysis of 2 DNA and 2 RNA-transposon. Genetic relationships among pea accessions were evaluated using an unweighted pairgroup method of arithmetic averages (UPGMA) and Jaccards's Index to develop a dendrogram. Coloured blocks indicate grouping of accessions into genetically similar clusters.

CHAPTER 6.0

General Conclusions and Future Research

The main focus of this study was to assess genetic diversity in the diverse field pea germplasm (*Pisum sativum* L.) using molecular tools such as SSRs and transposons. This study also generated genetic populations using Single Seed Descent (SSD) from true hybrids of diverse genotypes variable in lipid content. These segregating populations can be used to identify transgresive segregents for lipid traits that may lead to increased commercial interest in field pea breeding favorable, for example, for the bio-energy market.

The generation of segregating populations following controlled crosses among desirable parents is essential for any trait improvement. The Single Seed Descent (SSD) method was successfully used to develop Recombinant Inbred Lines (RILs).

The advent of molecular markers has expedited plant breeding programs and played a pivotal role in a variety of studies including varietal identification, genetic diversity, and linkage map construction. Molecular markers detect sequence variation among parents and their progenies for identification of true hybrids. Microsatellite markers in this study were screened for polymorphism among parents. Polymorphic SSR markers

were found to be an effective tool for selection of true hybrids within segregating populations that usually contain a mixture of selfed and hybrid plants. This technology has complimented an error prone morphological-based selection of hybrids dependent on plant growth stage as well as environmental conditions. Here, we successfully developed 34 hybrids following their confirmation at a molecular-level using polymorphic SSR markers. New populations, each containing 70-113 recombinants, were successfully maintained up to F_3 using SSD method. These populations can be utilized by the pea community for further investigations such as genetic mapping and inheritance studies.

Microsatellites can also be used for probing the genetic architecture of closely related germplasm due to their reliability, co-dominancy, and high polymorphism. Microsatellites were employed to assess the genetic diversity in 35 diverse accessions of pea. Information about genetic diversity is important for successful breeding programs. PIC and D values of each marker (Table 4.2) were used as decisive factor for polymorphism. The 50 SSR markers assessed were able to discriminate all pea accessions under the present investigation. This indicated that SSR markers employed in this study had a high level of polymorphism. Moreover, cluster analysis grouped pea accessions into genetically distinct

clusters, mostly consistent with their countries of origin (Figure 4.2). Grouping of Canadian and European genotypes into separate clusters was suggestive of genetically distinct gene pools. Genetic diversity and variability among the breeding material can be enhanced by inter-group crossing of these lines or by introduction of an exotic breeding material within each cluster of genetically similar genotypes. The genetically diverse groups identified in this study can be used to derive parental lines for pea breeding.

Transposable elements (TEs) comprise a major portion of most eukaryotic genomes. Therefore, the emergence of TEs-based fingerprinting marker system for varietal identification is not surprising. In this study, we explored transposon-based molecular markers in pea genome by selecting two transposons from each class of transposons i.e. Class I (Retrotransposons) and Class II (DNA transposons). We observed almost similar PIC and D values for both classes of transposons (Table 5.1). Transposon-based molecular markers allow fast and efficient fingerprinting method as higher allele number and PIC values were observed than SSR markers. Moreover, fewer transposon-based molecular markers can efficiently assess genetic diversity and can easily distinguish between closely related cultivars. Overall molecular marker analyses and data

generated in this research will be useful for pea breeding, marker assisted selection (MAS), genetic conservation and seed banks management.

The present research opened avenues for further investigations and research projects. For example, the new source segregating populations developed in this study provide a base for genetic mapping studies. Economically important traits, such as lipid content, can also be analysed in subsequent generations.

New microsatellite markers are being developed every year. Considering the genome size of field pea (about 5000 Mbp), newly developed SSR markers could be employed to examine the relatedness of accessions used in this study. Moreover, using more markers will cover larger portions of pea genome and will strengthen the data generated in this research. These markers can also be associated with economically important traits which can be helpful for their use in Marker Assisted Selection.

Further investigation of Class I (Retrotransposons) and Class II (DNA transposons) can be helpful for the development of new transposon-based molecular markers. Future identification of native active transposon system can help understand complex and large pea genome for genetic and evolutionary studies.

References

Able, J. A., Langridge, P. and Milligan, A. S. (2007). Capturing diversity in the cereals: many options but little promiscuity. Trends in Plant Science **12**(2):71-79.

Ahmad, S., Singh, M., Dylan Lamb-Palmer, N., Lefsrud, M., & Singh, J. (2012). Assessment of genetic diversity in 35 Pisum sativum accessions using microsatellite markers. Canadian Journal of Plant Science, 92:1075-1081.

Akhalkatsi, M., Ekhvaia, J., Mosulishvili, M., Nakhutsrishvili, G., Abdaladze, O. and Batsatsashvili, K. (2010). Reasons and Processes Leading to the Erosion of Crop Genetic Diversity in Mountainous Regions of Georgia. *Mountain Research and* Development 30(3):304-310.

Agarwal, M., Shrivastava, N. and Padh, H. (2008). Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Reports* **27**(4):617-631.

Banniza S., Hashemi P., Warkentin T.D., Vandenberg A., Davis A.R. (2005). The relationships among lodging, stem anatomy, degree of lignification, and resistance to mycosphaerella blight in field pea (Pisum sativum) (vol 83, pg 954, 2005). *Canadian Journal of Botany-Revue Canadienne De Botanique* **83**:1365-1365.

Baranger A., Aubert G., Arnau G., Laine A.L., Deniot G., Potier J., Weinachter C., Lejeune-Henaut I., Lallemand J., Burstin J. (2004). Genetic diversity within Pisum sativum using protein- and PCR-based markers. *Theoretical and Applied Genetics* **108**:1309-1321.

Barth, S., Melchinger, A. E., & Lübberstedt, T. H. (2002). Genetic diversity in Arabidopsis thaliana L. Heynh. investigated by cleaved amplified polymorphic sequence (CAPS) and inter-simple sequence repeat (ISSR) markers. *Molecular Ecology*, **11**(3):495-505.

Bastianelli, D., Grosjean, F., Peyronnet, C., Duparque, M., Regnier, J.M. (1998). Feeding value of pea (*Pisum sativum*, L.)-1 Chemical composition of different categories of pea. *Anim. Sci.*, **67**(3):609–620.

Basu, S. K., Datta, M., Sharma, M. and Kumar, A. (2011). Invited Review Article Haploid production technology in wheat and some selected higher plants. *Australian Journal of Crop Science*, AJCS **5**(9):1087-1093.

Beauregard A., Curcio M.J., Belfort M. (2008). The Take and Give Between Retrotransposable Elements and their Hosts. *Annual Review of Genetics* **42**:587-617.

Becher, S.A., Steinmetz, K., Weising, K., Boury, S., Peltier, D., Renou, J.P., Kahl, G., and Wolff, K. (2000). Microsatellites for cultivar identification in Pelargonium. *Theoretical and Applied Genetics* **101:** 643-651.

Bennetzen J.L. (2000) Transposable element contributions to plant gene and genome evolution. *Plant Molecular Biology* **42**:251-269.

Bohn, M., Utz, H. F., and Melchinger, A. E. (1999). Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs, and SSRs and their us for predicting progeny variance. *Crop Science* **39:** 228-237.

Bohra, A., Dubey, A., Saxena, R. K., Varma Penmetsa, R., Poornima, K. N., Kumar, N., Farmer, A. D., Srivani, G., Upadhyaya, H. D., Gothalwal, R. and others (2011). Analysis of BAC-end sequences (BESs) and development of BES-SSR markers for genetic mapping and hybrid purity assessment in pigeonpea (Cajanus spp.). Bmc Plant Biology **11**(1):56.

Botstein D., White R.L., Skolnick M., Davis R.W. (1980). Construction of a genetic-linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**:314-331.

Boye J., Zare F., Pletch A. (2010). Pulse proteins: Processing, characterization, functional properties and applications in food and feed. *Food Research International* **43**:414-431.

Brim, C. A. (1966). A Modified Pedigree Method of Selection in Soybeans1. *Crop Science* **6**(2):220-220.

Burstin, J., Deniot, G., Potier, J., Weinachter, C., Aubert, G., and Baranger, A. (2001). Microsatellite polymorphism in Pisum sativum. *Plant Breeding* **120:**311-317.

Campbell, B.C., LeMare, S., Piperidis, G., & Godwin, I. D. (2011). IRAP, a retrotransposon-based marker system for the detection of somaclonal variation in barley. Molecular Breeding, **72**(2):193-206.

Chambers G.K., MacAvoy E.S. (2000). Microsatellites: consensus and controversy. Comparative Biochemistry and Physiology. *B-Biochemistry & Molecular Biology* **126**:455-476.

Choudhury P.R., Tanveer H., Dixit G.P. (2007). Identification and detection of genetic relatedness among important varieties of pea (Pisum sativum L.) grown in India. *Genetica* **130**:183-191.

Cieslarova J., Smykal, P., Dockalova, Z., Hanacek, P., Prochazka, S., Hybl, M., and Griga, M. (2011). Molecular evidence of genetic diversity changes in pea (Pisum sativum L.) germplasm after long-term maintenance. *Genetic Resources and Crop Evolution* **58:** 439-451.

Collard B.C.Y., Jahufer M.Z.Z., Brouwer J.B., Pang E.C.K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica* **142**:169-196.

Collard, B. C., & Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**(1491): 557-572.

Cordeiro, G. M., Taylor, G. O., & Henry, R. J. (2000). Characterisation of microsatellite markers from sugarcane (< i> Saccharum</i> sp.), a highly polyploid species. *Plant science* **155**(2):161-168.

Croser, J. S., Lulsdorf, M. M., Davies, P. A., Clarke, H. J., Bayliss, K. L., Mallikarjuna, N. and Siddique, K. H. M. (2006). Toward doubled haploid production in the fabaceae: Progress, constraints, and opportunities. *Critical Reviews in Plant Sciences* **25**(2):139-157.

Cupic T., Tucak M., Popovic S., Bolaric S., Grljusic S., Kozumplik V. (2009). Genetic diversity of pea (Pisum sativum L.) genotypes assessed by pedigree, morphological and molecular data. *Journal of Food Agriculture & Environment* **7**:343-348.

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

Drew, E., Gupta, V. and Roget, D. (2008). Herbicide use, productivity, and nitrogen fixation in field pea (Pisum sativum). *Crop and Pasture Science* **58**(12):1204-1214.

Edwards J, McCouch S, Guimarães E, Ruane J, Scherf B, Sonnino A, Dargie J (2007). Molecular markers for use in plant molecular breeding and germplasm evaluation. Marker-assisted Selection– *Current Status and Future Perspectives in Crops, Livestock, Forestry and Fish* 29-49

Friedt, W., Snowdon, R., Ordon, F., & Ahlemeyer, J. (2007). Plant breeding: assessment of genetic diversity in crop plants and its exploitation in breeding. *Progress in Botany* 151-178.

Germana, M. A. (2011). Anther culture for haploid and doubled haploid production. *Plant Cell Tissue and Organ Culture* **104**(3):283-300.

Gilpin B.J., McCallum J.A., Frew T.J., Timmerman-Vaughan G.M. (1997). A linkage map of the pea (*Pisum sativum* L) genome containing cloned sequences of known function and expressed sequence tags (ESTs). *Theoretical and Applied Genetics* **95**:1289-1299.

Gong, Y.-m., Xu, S.-c., Mao, W.-h., Hu, Q.-z., Zhang, G.-w., Ding, J., and Li, Y.-d. (2010). Developing new SSR markers from ESTs of pea (Pisum sativum L.). *Journal of Zhejiang University-Science B* **11:** 702-707.

Grzebelus, D. (2006). Transposon insertion polymorphism as a new source of molecular markers. *Journal of fruit and ornamental plant research*, **14**: 21.

Gunasekare, M. T. K. (2007). Applications of molecular markers to the genetic improvement of Camellia sinensis L. (tea) - A review. *Journal of Horticultural Science & Biotechnology* **82**(2):161-169.

Hashemi, S. H., Mirmohammadi-Maibody, S. A. M., Nematzadeh, G. A., & Arzani, A. (2010). Identification of rice hybrids using microsatellite and RAPD markers. *African Journal of Biotechnology* **8**(10).

Hawkes, J. G. (1991). The importance of genetic-resources in plant-breeding. *Biological Journal of the Linnean Society* **43**(1):3-10.

Hearne C.M., Ghosh S., Todd J.A. (1992) Microsatellites for linkage analysis of genetic-traits. *Trends in Genetics* **8**:288-294.

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

Hoey B.K., Crowe K.R., Jones V.M., Polans N.O. (1996). A phylogenetic analysis of Pisum based on morphological characters, and allozyme and RAPD markers. *Theoretical and Applied Genetics* **92**:92-100.

Horacek J., Griga M., Smykal P., Hybl M. (2009). Effect of Environmental and Genetic Factors on the Stability of Pea (Pisum sativum L.) Isozyme and DNA Markers. *Czech Journal of Genetics and Plant Breeding* **45**:57-71.

Islam, F. M. A., Beebe, S., Munoz, M., Tohme, J., Redden, R. J. and Basford, K. E. (2004). Using molecular markers to assess the effect of introgression on quantitative attributes of common bean in the Andean gene pool. *Theoretical and Applied Genetics* **108**(2):243-252.

Jiang, N., Bao, Z., Zhang, X., Hirochika, H., Eddy, S. R., McCouch, S. R., & Wessler, S. R. (2003). An active DNA transposon family in rice. *Nature* **421**(6919):163-167.

Jing R.C., Vershinin A., Grzebyta J., Shaw P., Smykal P., Marshall D., Ambrose M.J., Ellis T.H.N., Flavell A.J. (2010). The genetic diversity and evolution of field pea (Pisum) studied by high throughput retrotransposon based insertion polymorphism (RBIP) marker analysis. *Bmc Evolutionary Biology* **10**(1):44.

Jing, R., Ambrose, M. A., Knox, M. R., Smykal, P., Hybl, M., Ramos, A., Caminero, C., Burstin, J., Duc, G., van Soest, L. J. M. and others (2012). Genetic diversity in European Pisum germplasm collections. *Theoretical and Applied Genetics* **125**(2):367-380.

Jones C.J., Edwards K.J., Castaglione S., Winfield M.O., Sala F., vandeWiel C., Bredemeijer G., Vosman B., Matthes M., Daly A., Brettschneider R., Bettini P., Buiatti M., Maestri E., Malcevschi A., Marmiroli N., Aert R., Volckaert G., Rueda J., Linacero R., Vazquez A., Karp A. (1997). Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding* **3**:381-390.

Jones R.N. (2005). McClintock's controlling elements: the full story. Cytogenetic and Genome Research **109**:90-103.

Kalendar, R., Flavell, A. J., Ellis, T. H. N., Sjakste, T., Moisy, C. and Schulman, A. H. (2011). Analysis of plant diversity with retrotransposon-based molecular markers. *Heredity* **106**(4):520-530.

Kalendar, R., Grob, T., Regina, M., Suoniemi, A., & Schulman, A. (1999). IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. *TAG Theoretical and Applied Genetics*, **98**(5):704-711.

Katoch, V., Sharma, S., Pathania, S., Banayal, D. K., Sharma, S. K., & Rathour, R. (2010). Molecular mapping of pea powdery mildew resistance gene er2 to pea linkage group III. *Molecular Breeding*, **25**(2):229-237.

Khodapanahi, E., Lefsrud, M., Orsat, V., Singh, J., & Warkentin, T. D. (2012). Study of Pea Accessions for Development of an Oilseed Pea. *Energies* **5**(10):3788-3802.

Krajewski, P., Bocianowski, J., Gawlowska, M., Kaczmarek, Z., Pniewski, T., Swiecicki, W. and Wolko, B. (2012). QTL for yield components and protein content: a multienvironment study of two pea (*Pisum sativum* L.) populations. *Euphytica* **183**(3):323-336.

Kuleung, C., Baenziger, P. S., Kachman, S. D. and Dweikat, I. (2006). Evaluating the genetic diversity of triticale with wheat and rye SSR markers. *Crop Science* **46**(4):1692-1700.

Kumar A., Bennetzen J.L. (1999). Plant retrotransposons. *Annual Review of Genetics* **33**:479-532.

Kumar, A. and Hirochika, H. (2001). Applications of retrotransposons as genetic tools in plant biology. *Trends in Plant Science* **6**(3):127-134.

Kumar, G. R., Sakthivel, K., Sundaram, R. M., Neeraja, C. N., Balachandran, S. M., Rani, N. S., Viraktamath, B. C. and Madhav, M. S. (2010). Allele mining in crops: Prospects and potentials. *Biotechnology Advances* **28**(4):451-461.

Kunze R., Saedler H., Lonnig W.E. (1997). Plant transposable elements. *Advances in Botanical Research* **27**:331-470.

Laucou V., Haurogne K., Ellis N., Rameau C. (1998). Genetic mapping in pea. 1. RAPD-based genetic linkage map of Pisum sativum. *Theoretical and Applied Genetics* **97**:905-915.

Lewers K.S., St Martin S.K., Hedges B.R., Palmer R.G. (1998) Testcross evaluation of soybean germplasm. *Crop Science* **38**:1143-1149.

Litt M., Luty J.A. (1989). A hypervariable microsatellite revealed by invitro amplification of a dinucleotide repeat within the cardiac-muscle actin gene. *American Journal of Human Genetics* **44**:397-401.

Loridon K., McPhee K., Morin J., Dubreuil P., Pilet-Nayel M.L., Aubert G., Rameau C., Baranger A., Coyne C., Lejeune-Henaut I., Burstin J. (2005). Microsatellite marker polymorphism and mapping in pea (Pisum sativum L.). *Theoretical and Applied Genetics* **111**:1022-1031.

Lu J., Knox M.R., Ambrose M.J., Brown J.K.M., Ellis T.H.N. (1996). Comparative analysis of genetic diversity in pea assessed by RFLP- and PCR-based methods. *Theoretical and Applied Genetics* **93**:1103-1111.

Macas J., Koblizkova A., Neumann P. (2005). Characterization of Stowaway MITEs in pea (Pisum sativum L.) and identification of their potential master elements. *Genome* **48**:831-839.

Macas J., Neumann P., Navratilova A. (2007). Repetitive DNA in the pea (*Pisum sativum* L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and Medicago truncatula. *Bmc Genomics* **8**(1):427.

Macas J., Neumann P., Pozarkova D. (2003) Zaba: a novel miniature transposable element present in genomes of legume plants. *Molecular Genetics and Genomics* **269**:624-631.

Maksem K. and Kahl, G. (2005). The handbook of plant genome mapping : genetic and physical mapping / edited by Khalid Meksem and Günter Kahl. *Weinheim: Wiley-VCH*, ©2005.

Manigbas, N. L., & Villegas, L. C. (2004). Microsatellite markers in hybridity tests to identify true hybrids of sugarcane. *Philippine Journal of Crop Science* **29**(2): 23-32.

McCallum, C. M., Comai, L., Greene, E. A. and Henikoff, S. (2000). Targeted screening for induced mutations. *Nature Biotechnology* **18**(4):455-457.

McClean, P. E., Myers, J. R., & Hammond, J. J. (1993). Coefficient of parentage and cluster analysis of North American dry bean cultivars. *Crop science* **33**(1):190-197.

McPhee, K. (2003). Dry pea production and breeding: A minireview. *Journal Of Food Agriculture And Environment* **1**:64-69.

Micke, A. (1988). Genetic improvement of food legumes in developing countries by mutation induction. In: Summerfield, R.J. (ed.). World Crops: Cool season food legumes. *Dordrecht: Kluwer Academic Publishers*. p. 1030-1047.

Muehlbauer, F.J. (1992). Use of introduced germplasm in cool-season food legume cultivar development. In: Shands, H.L. and Weisner, L.E. (eds.). Use of Plant Introductions in Cultivar Development, Part 2. *Crop Science Society of America Special Publication Number* 2: 49-73.

Muehlbauer, F.J., Redden, R.J., Nassib, A.M., Robertson, L.D., and Smithson, J.B. (1988). Population improvement in pulse crops: an assessment of methods and techniques. In: Summerfield, R.J. (ed.). World Crops: Cool Season Food Legumes. *Kluwer Academic Publishers*: 943-966.

Munoz-Lopez M., Garcia-Perez J.L. (2010). DNA Transposons: Nature and Applications in Genomics. *Current Genomics* **11**:115-128.

Naqvi, N. I., & Chattoo, B. B. (1996). Development of a sequence characterized amplified region (SCAR) based indirect selection method for a dominant blast-resistance gene in rice. *Genome* **39**(1): 26-30.

Nasiri J., Haghnazari A., Saba J. (2009). Genetic diversity among varieties and wild species accessions of pea (Pisum sativum L.) based on SSR markers. *African Journal of Biotechnology* **8**:3405-3417.

Nelson, R. L. and Johnson, E. O. C. (2011). Registration of Soybean Germplasm Line LG00-6313. *Journal of Plant Registrations* **5**(3):406-409.

Neumann P., Nouzova M., Macas J. (2001). Molecular and cytogenetic analysis of repetitive DNA in pea (Pisum sativum L.). *Genome* **44**:716-728.

Orgel L.E., Crick F.H.C. (1980). Selfish DNA: the ultimate parasite. *NATURE* **284**:604-607.

Palazzolo M.J., Sawyer S.A., Martin C.H., Smoller D.A., Hartl D.L. (1991). Optimized strategies for sequence-tagged-site selection in genome mapping. *Proceedings of the National Academy of Sciences of the United States of America* **88**:8034-8038.

Pallavi, H. M., Gowda, R., Vishwanath, K., Shadakshari, Y. G. and Bhanuprakash, K. (2011). Identification of SSR markers for hybridity and seed genetic purity testing in sunflower (Helianthus annuus L.). *Seed Science and Technology* **39**(1):259-264.

Pascher, K., Macalka, S., Rau, D., Gollmann, G., Reiner, H., Glossl, J., and Grabherr, G. (2010). Molecular differentiation of commercial varieties and feral populations of oilseed rape (Brassica napus L.) *BMC Evolutionar Biology* **10:**63.

Pejic, I., Ajmone-Marsan, P., Morgante, M., Kozumplick, V., Castiglioni, P., Taramino, G., and Motto, M. (1998). Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. *Theoretical and Applied Genetics* **97:**1248-1255.

Pereira G., Marques C., Ribeiro R., Formiga S., Damaso M., Sousa M.T., Farinho M., Leitao J.M. (2010). Identification of DNA markers linked to an induced mutated gene conferring resistance to powdery mildew in pea (Pisum sativum L.). *Euphytica* **171**:327-335.

Pesta, A., Titlow, A., Hansen, J., Berger, A., Furman, S., Luebbe, M., Erickson, G. and Jenkins, K. (2012). Effects of feeding field peas in combination with distillers grains plus solubles in finishing and growing diets on cattle performance and carcass characteristics. *The Professional Animal Scientist* **28**(5):534-540.

Powell, W., Machray, G.C., and Provan, J. (1996). Polymorphism revealed by simple sequence repeats. *Trends in Plant Science* **1:** 215-222.

PulseCanada (2010) http://www.pulsecanada.com/

Qian W., Sass O., Meng J., Li M., Frauen M., Jung C. (2007). Heterotic patterns in rapeseed (Brassica napus L.): I. Crosses between spring and Chinese semi-winter lines. *Theoretical and Applied Genetics* **115**:27-34.

Rodino A.P., Hernandez-Nistal J., Hermida M., Santalla M., De Ron A.M. (2009). Sources of variation for sustainable field pea breeding. *Euphytica* **166**:95-107.

Russell, J. R., Fuller, J. D., Macaulay, M., Hatz, B. G., Jahoor, A., Powell, W., and Waugh, R. (1997). Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theoretical and Applied Genetics* **95:** 714-722.

Sabot F., Simon D., Bernard M. (2004). Plant transposable elements, with an emphasis on grass species. *Euphytica* **139**:227-247.

Sant'Ana, G., Ferreira, J., Rocha, H., Borém, A., Pasqual, M. and Cançado, G. (2012). Comparison of a retrotransposon-based marker with microsatellite markers for discriminating accessions of Vitis vinifera. *Genetics and Molecular Research* **11**(2):1507-1525.

Samatadze T.E., Zelenina D.A., Shostak N.G., Volkov A.A., Popov K.V., Rachinskaya O.V., Borisov A.Y., Tihonovich I.A., Zelenin A.V., Muravenko O.V. (2008). Comparative genome analysis in pea Pisum sativum L. Varieties and Lines with chromosomal and molecular markers. *Russian Journal of Genetics* **44**:1424-1430.

Sanmiguel P., & Bennetzen J.L. (1998) Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. *Annals of Botany* **82**:37-44.

Santalla M., Amurrio J.M., De Ron A.M. (2001). Food and feed potential breeding value of green, dry and vegetable pea germplasm. *Canadian Journal of Plant Science* **81**:601-610.

Sarikamis G., Yanmaz R., Ermis S., Bakir M., Yuksel C. (2010). Genetic characterization of pea (*Pisum sativum*) germplasm from Turkey using morphological and SSR markers. *Genetics and Molecular Research* **9**:591-600.

Saskatchewan Agriculture and Food website www.agr.gov.sk.ca 'Pulses'

Sato S., Isobe S., Tabata S. (2010). Structural analyses of the genomes in legumes. *Current Opinion in Plant Biology* **13**:146-152.

Selvakumar, P., Ravikesavan, R., Gopikrishnan, A., Thiyagu, K., Preetha, S., & Manikanda Boopathi, N. (2010). Genetic purity analysis of cotton (Gossypium spp.) hybrids using SSR markers. *Seed Science and Technology* **38**(2):358-366.

Semagn K., Bjornstad A., Ndjiondjop M.N. (2006a). An overview of molecular marker methods for plants. *African Journal of Biotechnology* **5**:2540-2568.

Semagn K., Bjornstad A., Ndjiondjop M.N. (2006b). Principles, requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology* **5**:2569-2587.

Simioniuc D., Uptmoor R., Friedt W., Ordon F. (2002). Genetic diversity and relationships among pea cultivars revealed by RAPDs and AFLPs. *Plant Breeding* **121**:429-435.

Singh J., Zhang S., Chen C., Cooper L., Bregitzer P., Sturbaum A., Hayes P.M., Lemaux P.G. (2006). High-frequency Ds remobilization over multiple generations in barley facilitates gene tagging in large genome cereals. *Plant Molecular Biology* **62**:937-950.

Skrypetz, S. (2004). Dry peas: situation and outlook. Agriculture and Agri-Food Canada, Market Analysis Division. *Bi-weekly Bulletin* **17**:1–10.

Smith, J. S. C., Kresovich, S., Hopkins, M. S., Mitchell, S. E., Dean, R. E., Woodman, W. L., Lee, M., and Porter, K. (2000). Genetic Diversity among Elite Sorghum Inbred Lines Assessed with Simple Sequence Repeats. *Crop Science* **40:** 226-232.

Smykal P. (2006). Development of an efficient retrotransposon-based fingerprinting method for rapid pea variety identification. *Journal of Applied Genetics* **47**:221-230.

Smykal P., Hybl M., Corander J., Jarkovsky J., Flavell A.J., Griga M. (2008). Genetic diversity and population structure of pea (*Pisum sativum* L.) varieties derived from combined retrotransposon, microsatellite and morphological marker analysis. *Theoretical and Applied Genetics* **117**:413-424.

Smykal, P., Aubert, G., Bustin, J., Coyne, C., Ellis, N., Flavell, A., ... & Warkentin, T. (2012). Pea (Pisum sativum L.) in the genomics era. *Agronomy* **2**(2): 74-115.

Solanki, R.K., Singh, S. and Kumar, J. (2010). Molecular marker assisted testing of hybridity of F_1 plants in lentil. *Food Legumes* **23**(1):21-24.

Sosulski, F.W. and McCurdy, A.R. (1987). Functionality of flours, protein fractions and isolates from field peas and faba bean, *J. Food Sci.* **52**:1010–1014.

Spies, J. M., Warkentin, T. D. and Shirtliffe, S. J. 2011. Variation in Field Pea (Pisum sativum) Cultivars for Basal Branching and Weed Competition. *Weed Science* **59**(2):218-223.

Sundaram, R., Naveenkumar, B., Biradar, S., Balachandran, S., Mishra, B., IlyasAhmed, M., Viraktamath, B., Ramesha, M. and Sarma, N. (2008). Identification of informative SSR markers capable of distinguishing hybrid rice parental lines and their utilization in seed purity assessment. *Euphytica* **163**(2):215-224.

Tantasawat, P., Trongchuen, J., Prajongjai, T., Jenweerawat, S., and Chaowiset, W. (2011). SSR analysis of soybean (Glycine max (L.) Merr.) genetic relationship and variety identification in Thailand. *Australian Journal of Crop Science* **5:**283-290.

Tar'an, B., Zhang, C., Warkentin, T., Tullu, A., and Vandenberg, A. (2005). Genetic diversity among varieties and wild species accessions of pea (Pisum sativum L. based on molecular markers, and morphological and physiological characters. *Genome* **48:**257-272

Tar'an, B., Warkentin T., Somers D.J., Miranda D., Vandenburg A., Blade S., Woods S., Bing D., Xue A., DeKoeyer D., Penner G. (2003). Quantitative trait loci for lodging resistance, plant height and partial resistance to mycosphaerella blight in field pea (*Pisum sativum* L.). *Theoretical and Applied Genetics* **107**:1482-1491.

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

Tessier, C., David, J., This, P., Boursiquot, J. M. and Charrier, A. (1999). Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. *Theoretical and Applied Genetics* **98**(1):171-177.

Timmerman-Vaughan G.M., Frew T.J., Butler R., Murray S., Gilpin M., Falloon K., Johnston P., Lakeman M.B., Russell A., Khan T. (2004). Validation of quantitative trait loci for Ascochyta blight resistance in pea (Pisum sativum L.), using populations from two crosses. *Theoretical and Applied Genetics* **109**:1620-1631.

Tullu A., Tar'an B., Warkentin T., Vandenberg A. (2008). Construction of an Intraspecific Linkage Map and QTL Analysis for Earliness and Plant Height in Lentil. *Crop Science* **48**:2254-2264.

Ubayasena L., Bett K., Tar'an B., Vijayan P., Warkentin T. (2010) Genetic control and QTL analysis of cotyledon bleaching resistance in green field pea (Pisum sativum L.). *Genome* **53**:346-359.

von Stackelberg M., Lindemann S., Menke M., Riesselmann S., Jacobsen H.J. (2003). Identification of AFLP and STS markers closely linked to the def locus in pea. *Theoretical and Applied Genetics* **106**:1293-1299.

Vos P., Hogers R., Bleeker M., Reijans M., Vandelee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., Zabeau M. (1995). AFLP - A new technique for dna-fingerprinting. *Nucleic Acids Research* **23**:4407-4414.

Warkentin, T., Delgerjav, O., Arganosa, G., Rehman, A., Bett, K., Anbessa, Y., Rossnagel, B. and Raboy, V. 2012. Development and characterization of low-phytate pea. *Crop Science* **52**(1):74-78.

Wedzony, M., Forster, B. P., Zur, I., Golemiec, E., Szechynska-Hebda, M., Dubas, E. and Gotebiowska, G. (2009). Progress in Doubled Haploid Technology in Higher Plants. *Advances in Haploid Production in Higher Plants*:1-33.

Wessler S.R., Bureau T.E., White S.E. (1995). LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Current opinion in genetics & development* **5**(6):814-821.

Yashitola, J., Thirumurugan, T., Sundaram, R. M., Naseerullah, M. K., Ramesha, M. S., Sarma, N. P., & Sonti, R. V. (2002). Assessment of purity of rice hybrids using microsatellite and STS markers. *Crop Science* **42**(4):1369-1373.

Yoshida H., Tomiyama Y., Tanaka M., Mizushina Y. (2007). Distribution of fatty acids in triacylglycerols and phospholipids from peas (Pisum sativum L.). *Journal of the Science of Food and Agriculture* **87**:2709-2714.

Zong X., Redden R.J., Liu Q., Wang S., Guan J., Liu J., Xu Y., Liu X., Gu J., Yan L., Ades P., Ford R. (2009). Analysis of a diverse global Pisum sp collection and comparison to a Chinese local P. sativum collection with microsatellite markers. *Theoretical and Applied Genetics* **118**:193-204.