GENOME WIDE ASSOCIATION STUDY TO IDENTIFY SINGLE NUCLEOTIDE POLYMORPHISM MARKERS FOR FE, ZN, AND SE CONCENTRATION IN FIELD PEA SEEDS

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Master of Sciences
In the Department of Plant Sciences
University of Saskatchewan
Saskatoon

By

Dilanganie Dissanayaka

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other uses of materials in this thesis in whole or part should be addressed to:

Head
Department of Plant Sciences
University of Saskatchewan
Agriculture building, 51 Campus Drive
Saskatoon, Saskatchewan S7N5A8
Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9
Canada

Abstract

Micronutrient concentration is an important component of crop quality. Iron (Fe), zinc (Zn) and selenium (Se) are essential micronutrients for human health. Previous studies indicated substantial variation in micronutrient concentration in field pea (*Pisum sativum L.*). The objective of the current research was to evaluate a diverse pea panel for Fe, Zn, and Se concentration, and to identify associated single nucleotide polymorphism (SNP) markers using genome wide association study (GWAS). One hundred and seventy seven diverse pea accessions were assessed for Fe, Zn, and Se concentration from five location-years (2013 Saskatoon, 2014 Fargo, 2015 Saskatoon, 2016 Saskatoon and 2016 Rosthern) using atomic absorption spectroscopy. Accessions differed significantly for Fe and Zn in all location-years, but not for Se ($P \leq 0.05$). Year and location effects were significant for all three micronutrients. Mean concentration in the five location-year combined analysis were 50.46, 29.49, and 0.95, and the concentration (µg/g) ranged from 29.22 to 90.53, 12.83 to 51.47 and 0.06 to 8.75 in the ground whole pea seeds for Fe, Zn, and Se, respectively. Concentration of Fe, Zn, and Se were significantly correlated ($P \leq 0.05$) between each location with the exception of Se concentration between Fargo and Rosthern. Significant correlations were observed in all the location-years between Zn and Fe concentration ($P \leq 0.05$). Se significantly correlated with Fe in all the location years, except 2016 Rosthern ($P \leq 0.05$). The correlation between Se and Zn was only significant in 2014 Fargo and 2016 Saskatoon ($P \leq 0.05$). Fe and Zn concentration showed significant positive correlations ($P \leq 0.05$) with lodging and day to maturity (DTM), but not Se. In general, yield negatively correlated with Fe, Zn, and Se. The pea panel accessions clustered into nine major groups which could be considered consistent with the geographic origin of the accessions. Genotypic data generated from genotyping-by-sequencing were used in genome wide association study (GWAS) to identify significant SNP markers associated with Fe, Zn, and Se concentration. After Bonferroni correction, three significant markers for Fe concentration and seven significant markers for Zn concentration ($P \leq 0.05$, $-\log_{10} P$ value ≥ 5.46) were identified. Five markers for Fe, seven markers for Zn and four markers for Se (including few less significant markers) were validated on a pea recombinant inbred line population; PR-07 (Carrera × CDC Striker). The marker Sc1512_36017 co-localized with Sc11336_48840 on LGIIIb which was a flanking marker of quantitative trait loci (QTL) for seed Zn concentration. The markers identified from the present study can be used in marker assisted selection (MAS) in pea breeding to develop new varieties with high Fe, Zn, and Se concentration.

Acknowledgements

It is with immense pleasure I express my sincere gratitude to my supervisor Dr. Tom Warkentin for his kind guidance and inspiring supervision given throughout to achieve success in this masters research. Also, I would like to thank my committee members, Dr. Tim Sharbel, Dr. Rosalind Bueckert and Dr. Bunyamin Tar'an for their valuable advices on planning the research, data analysis, interpretation and thesis writing. Special thanks go to my external examiner Dr. Bifang Cheng.

I sincerely acknowledge my advisor Dr. Ambuj Jha for his patient guidance and friendly motivations throughout the research. A special acknowledgement is extended to Dr. Kishore Gali for sharing his knowledge and experience with me, especially related to genotype data. I extend my gratitude to all the other members in the 'pea group' for their valuable inputs and encouragements.

I am grateful to the technician Barry Goetz for his helpful assistance in using the atomic absorption spectrophotometer. Also, many thanks to the field lab crew for assisting me with seed sampling and agronomic data.

I would like to thank my funding agent Ministry of Agriculture, Government of Saskatchewan for the excellent financial support.

Finally, I am indebted to my family for their devotion, understanding, and care. Also, I sincerely acknowledge my well-wishers for their kind support throughout this research.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AAFC - Agriculture and Agri-food Canada

AAS - Atomic absorption spectrometry

AFLP - Amplified Fragment Length Polymorphism

ANOVA - Analysis of Variance

CAP - Cleaved Amplified Polymorphism

CDC - Crop Development Centre

CGIAR - Consultative Group for International Agricultural Research

CV - Coefficient of variation

DNA - Deoxyribonucleic acid

DTF - Days to flowering

DTM - Days to maturity

FAO - Food and Agriculture Organization

FAOSTAT - Food and Agriculture Organization Statistics

FDR - False discovery rate

FRET - Fluorescence resonance energy transfer

GBS – Genotyping by sequencing

GLM - General linear model

GWAS - Genome-wide association studies

ICP-MS - Inductively coupled plasma mass spectrometry

ISSR - Inter simple sequence repeats

KASP - Kompetitive Allele Specific PCR

LD - Linkage disequilibrium

LG - Linkage group

MALDI-TOF - Matrix-assisted laser desorption/ionization – time of flight

MAS - Marker assisted selection

MLM - Mixed linear model

NIH - National Institutes of Health

PCR - Polymerase chain reaction

PR - Pea recombinant

QTCAT - Quantitative trait cluster association test

QTL - Quantitative trait locus

RAPD - Random amplification of polymorphic DNA

RDA - Recommended daily allowance

RFLP - Restriction fragment length polymorphism

RIL - Recombinant inbred lines

SAS - Statistical Analysis System

SD - Standard deviation

SNP - Single nucleotide polymorphisms

SPG - Saskatchewan Pulse Growers

SSR - Simple sequence repeats

STS - Sequence-tagged sites

TASSEL - Trait analysis by association, evolution, and linkage

WFP - World Food Program

WHO - World Health Organization

1. Introduction

According to the World Health Organization (WHO), the micronutrients iron (Fe), zinc (Zn) and selenium (Se) are among the 20 essential nutrients which comprise the basis of all human nutrition. Even though micronutrients are required in trace concentration, they play a major role in many vital metabolic functions that are important for human growth and development (Bourre, 2006; Wakeel et al., 2018). A large proportion of the world's population suffers from mineral malnutrition, especially from Fe and Zn, as they are dependent on plant-based foods which often have a low mineral density (Waters & Grusak, 2008). Maternal and child undernutrition and micronutrient deficiencies have become a global burden and about 12% of deaths among under-5 children are related to deficiency of the four common micronutrients Fe, iodine, vitamin A, and Zn (Ahmed, 2012). Furthermore, micronutrient deficiencies, act as exacerbating factors in infectious and chronic diseases while greatly impacting morbidity, mortality, and quality of life (Tulchinsky, 2010). According to FAOSTAT, over 60% of the world's people are Fe deficient, 30% Zn deficient, and 15% are Se deficient (Thavarajah et al., 2011).

Ensuring sufficient intake of micronutrients is a general expectation for food security. Thus, micronutrient malnutrition has received considerable attention in policy discussions on food security (Meenakshi et al., 2010). Increased food production, supplementation, and food fortification are the main food-based strategies to combat micronutrient malnutrition. Supplementation is the most widely practiced intervention, while fortification can also be a potential strategy to target a large population (Bhutta et al., 2013). In supplementation, micronutrients are taken along with the major diet as pills or the powder. In fortification, cereals or the pulses are processed by blending with micronutrients (White, 2011). In addition to supplementation and fortification, biofortification is a balanced and economical approach and could aid in reducing micronutrient deficiencies globally (Bouis, 2002; 2011; 2017). Biofortification is an approach of improving the nutritional quality of food crops through agronomic practices, conventional plant breeding, or modern biotechnology (WHO, n.d.) According to Bouis (2011), there are a few major concerns to make the biofortication process a success. First, for the successful breeding of biofortified varieties, high nutrient density must be combined with other parameters such as high yield. Also, the efficacy of the biofortified crop must be demonstrated. Moreover, when the biofortified varieties are

developed by breeders, they have to be adopted by farmers and consumed by a significant proportion of the target population (Meenakshi et al., 2010; Bouis, 2011).

Field pea (*Pisum sativum* L.) is one of the major pulse crops in the world with production of approximately 14 million tonnes per year (FAOSTAT, 2016). Canada is the world's leading producer of field pea, followed by China and India (FAOSTAT, 2016). Previous studies at the University of Saskatchewan indicated substantial variation in micronutrient concentration in pulse crop varieties (Warkentin et al., 1996; Ray et al., 2014; Diapari et al., 2014; 2015). Pulses including field pea contain significant proportions of the recommended daily allowance (RDA) of micronutrients (Ray et al., 2014). Diapari et al. (2015) identified SNP markers associated with Fe and Zn concentration that can be applied in marker assisted selection. Considering the production of pea and the variation in micronutrient concentration in pea germplasm, pea has great potential for biofortification as an approach to address micronutrient malnutrition.

Evaluation of diverse pea accessions is required to select those with a high micronutrient concentration that could be utilized in a pea breeding program. Genome Wide Association Study (GWAS) is an experimental design used to identify the associations between genetic variants and traits in plants and animal. GWAS can identify single nucleotide polymorphism (SNP) markers for marker-assisted selection. GWAS is a powerful tool to discover the underlying genetics for a trait, however, biological factors including effect size, sample size, genetic heterogeneity, linkage disequilibrium and statistical tools are important factors to be considered (Korte & Farlow, 2013). The markers identified from GWAS can be incorporated in pea breeding programs to facilitate the selection for varieties with high micronutrient concentration that could assist in the biofortification of this crop.

The hypotheses tested in this research were that a substantial variation in Fe, Zn and Se concentration will be observed in a pea GWAS panel, and the SNPs associated with Fe, Zn and Se concentration in pea can be identified by GWAS. The objectives of this research were to evaluate the GWAS panel of 177 pea accessions grown in Saskatoon, SK, Rosthern, SK and Fargo, ND for Fe, Zn, and Se concentration, and to perform GWAS to identify SNPs associated with Fe, Zn, and Se concentration.

2. Literature Review

2.1. Field pea

Pea (*Pisum sativum* L.) is a cool season pulse crop which is classified under the Fabaceae (Leguminosae) family along with lentil, faba bean, grass pea, and chickpea. It is one of the oldest crops domesticated about 10,000 years ago. The center of pea genetic diversity is near the Tigris and Euphrates rivers which is a wide area of the Fertile Crescent through present day Turkey, Syria, Iraq, Israel and Lebanon. Pea has been cultivated in Europe for several thousand years and spread over all climatic zones including high elevated areas in tropical regions (Smartt, 1984; Gupta, 2004; Smýkal et al., 2010).

Pea as a legume has the ability to fix nitrogen (Phillips, 1980). This minimizes the external input of nitrogen to the cropping system which is an important ecological advantage (Smýkal et al., 2012). In Australia, North America, and Europe, field pea is usually grown in rotation with temperate cereals or oilseeds and in India it is rotated with rice (French, 2016).

Field pea is primarily used for human consumption. Pea provides a variety of nutrients such as proteins, complex carbohydrates, vitamin, and minerals such as Fe and Zn which are important for human nutrition (Roy et al., 2010; Ray et al., 2014; Diapari et al., 2014; Jha et al., 2015). Consumer acceptance for the pulses is influenced by the size, shape, colour and, chemical composition of the grain, by storage conditions and any pre-treatment before use. Moreover, the cooking time, texture, water absorption and dispersibility of solids are some other important qualities (Singh et al., 2000). Dry pea is typically consumed as dhal in Asian countries such as India and Bangladesh. In China, pea is fractionated into protein and starch fractions with the starch being used for production of vermicelli noodles, and the protein being sold in high value health food markets. Pea nutritional qualities such as moderate protein concentration, slowly digestible starch and high levels of soluble and insoluble fibre improve the potential for new food applications (Warkentin et al., 2015).

Field pea is an excellent livestock feed with high levels of carbohydrates, low in fiber and 86-87% total digestible nutrients. Field pea contains only 5-20% of the concentration of trypsin inhibitors compared to soybean which enables it to be directly fed to livestock without the requirement for the extrusion heating process (McKay et al., 2003).

2.2. Field pea production

Dry pea is the second most produced pulse crop in the world with a production of 14.4 million tonnes where the dry bean ranked as the first (FAOSTAT, 2016). By 2016, Canadian dry pea production was 4.6 million tonnes, followed by the Russian Federation, China, and India. Out of the 7.6 million ha of the pea harvesting area all over the world, Canada accounted for 1.7 million ha (FAOSTAT, 2016).

From the 1960s to 1980s, Eastern Europe had the greatest amount of the world pea production, then Western Europe led in the 1990s. With the opening of the European feed pea market in 1985, demand for peas more significantly increased. Since 1990s, North America, basically Canada became the largest producer (Warkentin et al., 2015). In the past two decades, field pea production in Canada, USA, and the Russian Federation has increased while the reverse trend was observed in Europe due to economic, biological, social, physical and technical factors (Smýkal et al., 2012). Field pea production in Western Canada has been continuously increasing since 1977 both in food and feed markets. In Canada, field pea production area was 74,400 ha in 1985 and 1,715,400 ha by 2017 (Statistics Canada, 2017).

Field pea is currently grown in Saskatchewan, Alberta, and Manitoba, collectively referred as Western Canada. Due to soil, climate, and the development of innovation networks, such as the Crop Development Centre at the University of Saskatchewan, the prairie provinces have the greatest pulse cultivation in Canada (Statistics Canada, 2011). Canada's top three pea export markets are India, China, and Bangladesh. In 2016, Canada's exports of pulses, including dry peas, to India were worth over $1.1 billion and accounted for 27.5% of Canada's global pulse exports. However, at the end of 2017, India announced a 50% tariff on dry pea imports from all countries and efforts have been made with the Government of India to provide stable, sustainable access for Canadian pulse exports to India (AAFC, 2017). In contrast, Canadian pea exports to China have risen substantially in recent years.

2.3. Micronutrients

2.3.1. Role of micronutrients in the human body

Micronutrients including iron (Fe), cobalt (Co), chromium (Cr), copper (Cu), iodine (I), manganese (Mn), selenium (Se), and zinc (Zn) are the chemical elements required in minuscule amounts for the normal growth and development of living organisms. Biochemical studies on micronutrients over the past 50 years have identified their necessity and role in the human diet. Severe deficiency of micronutrients causes characteristic disease states where

supplementation of the appropriate amount of micronutrient is required to overcome the disease. Highly complex and coordinated systems of redistribution of micronutrients have been observed in many diseases which provide evidence of the importance of micronutrients to the human body (Shenkin, 2005).

Fe, Zn, and Se are among the twenty essential nutrients in human nutrition (WHO, n.d.). Fe is vital for several metabolic functions in the human body. Fe in hemoglobin serves as a carrier of oxygen from the lungs to the tissues. Fe is required for important enzyme systems such as iron catalase, peroxidase, and cytochrome. These enzymes act as electron carriers and their role in oxidative metabolism is to transfer energy within the cell (Abbaspour et al., 2014).

Zn contributes to membrane signaling systems in cell growth and proliferation, hence, enhances the wound healing process (Prasad, 1996; MacDonald et al., 2000). Zn has an antioxidant role in protecting cells against free radical-induced oxidative damage (Bray & Bettger, 1990; Rostan et al., 2002; Prasad et al., 2004). Furthermore, Zn decreases the relative risk of prostate, hepatocellular and pancreatic cancers (Costello & Franklin, 2017).

Se protects the body tissues against oxidative stress (Tinggi, 2008) and has been implicated in maintenance of defense against infection, and modulation of growth and development. Se also has a protective effect on progression of carcinogenesis (Rayman, 2005; Zeng et al., 2008).

2.3.2. Micronutrient malnutrition

More than half of the world's population is suffering from micronutrient malnutrition. It is also known as 'hidden hunger', especially in Asia, Africa, and Latin America, where most diets are cereal-based and do not meet the recommended dietary allowances for key minerals. Deficiencies of essential minerals can lead to serious implications for children and pregnant women unless the problem is addressed carefully. The recommended daily allowance (RDA) for each mineral differs according to age and gender. According to the National Institute of Health (NIH), on average, the RDA for Zn is 11 mg/day for adult males and 8 mg/day for adult females. Recommendation for Se is similar for both males and females at 55 μg/day. The RDA of Fe for a female with pre- and post-menstruation and for an adult male is 8 mg/day. It would be 15-18 mg/day for women with menstruation.

Fe, iodine, folate, vitamin A, and Zn deficiencies are the most widespread nutrient deficiencies in the world, and they contribute to poor growth, intellectual impairments,

perinatal complications, and increased risk of morbidity and mortality (Bailey et al., 2015). Fe deficiency is the most common and widespread nutritional disorder in the world. Fe deficiency at critical times of growth and development can result in premature births, low birth weight babies, delayed growth and development, delayed normal infant activity, and poor cognitive skills (Allen, 2000; Lozoff et al., 2008; Bailey et al., 2015). Se deficiency is an important factor in the etiology of Keshan disease, Kashin-Beck disease, and myxedematous cretinism (Coppinger et al., 2001). Zn deficiency affects pregnant women significantly as it performs a major role in cell division. Inadequate Zn intake may result in loss of appetite, poor sense of smell and taste, pale skin, frequent infections, low fertility, stunted growth, mental problems, poor wound healing and a poor immune system (Roohani et al., 2013).

2.4. Plant uptake and biofortification

2.4.1. Micronutrient uptake by plants

As there is a close relationship between human malnutrition and micronutrient concentration in the edible crop plant parts, thus understanding the factors that govern plant micronutrient uptake is crucial. Plants require 17 essential nutrients for optimal growth and development to complete their life cycle. These nutrients are carbon (C), hydrogen (H), oxygen (O), nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), sulfur (S), zinc (Zn), copper (Cu), iron (Fe), manganese (Mn), boron (B), molybdenum (Mo), chlorine (Cl), and nickel (Ni). The role of the macronutrients in plants is producing the bulk of the carbohydrates, proteins, and lipids of plant cells, whereas micronutrients mostly participate in the enzyme activation processes of the plant. The nutrient availability of soils is affected by climate, soil, and plant factors and their interactions (Fargaria, 2016).

Plant characteristics such as tissue stoichiometry, biomass cycling rates, above and below ground allocation, root distribution, and maximum rooting depth are important to demonstrate the nutrient profiles of plants. Therefore, understanding the vertical distribution of soil nutrients is useful in agriculture (Jobbágy & Jackson, 2001). The root system in plants has evolved in order to access the maximum amount of nutrients from the soil. Root system architecture is the spatial arrangement of root tissue within the soil which is crucial for the uptake of water and nutrients. Water and nutrients are distributed heterogeneously throughout the soil profile that eventually affects plant fitness, crop performance, and grain yield (Gei & Powers, 2015; Rogers & Benfey, 2015).

Most mineral nutrients are absorbed by the plant through the rhizosphere. Rhizospheric bacteria participate in the geochemical cycling of nutrients, especially nitrogen, phosphorus, and micronutrients including Fe, Mn, Zn, and Cu, which eventually enhance the availability of nutrients in the soil (Dotaniya & Meena, 2015). The pulses have a mutualistic symbiotic association with Rhizobium bacteria (*Rhizobium leguminosarum*) which can fix free nitrogen from the air. Furthermore, soil organic matter is able to retain water and nutrients, provide a habitat for soil biota and improve soil structure; hence, it is important for sustainable yields in crops (Agegnehu et al., 2015).

Iron uptake by crop plants exceeds uptake of all other essential micronutrients except chlorine. Fe^{3+} (ferric ion) is the common form present in the soil and only the gramineous monocots such as wheat, oat, and barley have the ability to chelate, solubilize and uptake the soil Fe^{3+} by secreting mugineic acid family phytosiderophores into the rhizosphere (Kawai et al., 1988; Schenkeveld et al., 2014). However for other plants, Fe^{3+} has to be reduced to Fe^{2+} to facilitate uptake. Due to that necessity, iron uptake by these plants mainly depends on the soil redox potential and pH other than the plant's ability to secrete protons (H^{+}) to the rhizosphere and reduce soil Fe^{3+} to Fe^{2+}. When the pH decreases, Fe^{3+} tends to reduce to soluble Fe^{2+} and becomes more available for plant uptake (Morrissey & Guerinot, 2009; Fargaria, 2016; Selby-Pham, 2017).

Soil moisture conditions and soil chemical factors influence Zn uptake, however, the process of Zn uptake can be complex even among closely related genotypes because it is governed by plant-soil interactions (Ma et al., 2016). Zn is taken up mainly as Zn^{2+} by plant roots and in some cases, as organic ligand-Zn complexes. Based on the ligand secreted by plant roots, there are two possible physiological mechanisms, i.e., efflux of reductants, organic acids and H^{+} ions, or efflux of phytosiderophores (Gupta et al., 2016).

As reviewed by Gupta and Gupta (2017), selenium exists as two forms: inorganic forms, i.e., selenate (SeO^{2-4}), selenite (SeO^{2-3}), selenide (Se^{2-}), elemental Se, and the organic form. Depending on the plant species, phases of development and concentration of Se in the soil, salinity, soil pH, activity of membrane transporters and the translocation mechanisms of the plant, the process of Se uptake and translocation may vary (Gupta & Gupta, 2017).

2.4.2. Supplementation, fortification, and biofortification

According to WHO, "Fortification is the practice of deliberately increasing the content of an essential micronutrient, i.e., vitamins and minerals in a food, so as to improve the nutritional

quality of the food supply and provide a public health benefit with minimal risk to health". Supplementation, food fortification, and dietary diversification are current methods for combating micronutrient malnutrition.

When the diet alone fails to meet the nutrient requirements due to dietary restrictions or intolerances, micronutrient supplements, i.e. multivitamin with iron or other minerals, folate/folic acid, Fe and Zn are beneficial especially for the risk groups such as children and pregnant women (Wiltgren et al., 2015). Micronutrient supplements can be used along with the major diet, however, the feasibility and sustainability of supplementation as a mode of delivery in resource-poor settings are in question (FAO, 2010).

Partially pre-cooked and milled cereals, soybeans, common beans, or other pulses blended with micronutrients are called fortified foods which are used in food assistance programs by the World Food Program (WFP) to prevent and address nutritional deficiencies. Moreover, safe delivery systems, stable political policies, appropriate social infrastructure, and continued investment are required for dietary diversification (White, 2011).

According to WHO, "Biofortification is the process by which the nutritional quality of food crops is improved through agronomic practices, conventional plant breeding, or modern biotechnology". Thus, biofortification is a suitable approach and an upcoming strategy to improve micronutrients in diets (FAO, 2010). Through plant breeding, biofortification can improve the nutritional content of commonly eaten foods providing a comparatively inexpensive, cost-effective, sustainable, long-term means of delivering more micronutrients to the poor (Bouis, 2011).

The HarvestPlus Challenge Program is an initiative of the Consultative Group on International Agricultural Research (CGIAR), where over 100 scientists in approximately 50 research institutions and implementing agencies around the world are making collaborative efforts to breed and disseminate crops for better nutrition. The program was focused on six targeted staple foods: rice, wheat, maize, cassava, common beans and sweet potato (Pfeiffer & McClafferty, 2007). As reviewed by Bouis and Saltzman (2017), at this point, biofortified crops are grown and consumed by more than 20 million people in developing countries. To reach one billion people by 2030, increasing the consumer demand, implementing new policies and mainstreaming biofortified traits into public and private breeding programs are important.

2.4.3. Potential of field pea biofortification

Pulses are an important source of dietary proteins and slowly digestible starch, dietary fiber (Patterson et al., 2009; Roy, 2010), carotenoids (Ashokkumar et al., 2015), folates (Jha et al., 2015) and minerals (Bueckert et al., 2011; Thavarajah 2012). According to Ray et al. (2014), Saskatchewan-grown pulses are excellent sources of micronutrients, and in a single serving provide 50 to 100% of the RDA for potassium, magnesium, iron, manganese, selenium, and 25 to 35% for zinc. Agronomic interventions, breeding approaches, genetic modification and microbiological approaches are potential ways to enrich pulse crops (Singh et al., 2015). In addition, significant genetic variation was observed for the concentration of most minerals in Saskatchewan grown field pea (Ray et al., 2014), and therefore, field pea can be considered as a good candidate for crop biofortification.

2.5. Field pea breeding

2.5.1. Conventional breeding

Pea breeding has achieved many successes since Mendel's experiments in the mid-1800s. Varieties with high yield, lodging resistance, improved resistance to root rot diseases and pathogens have been introduced by pea breeding programs worldwide (Tayeh et al., 2015; Warkentin et al., 2015). For decades, controlled hybridization and selection has been performed on field pea in order to improve agronomic traits. In general, bulk population, pedigree selection and a combination of these two procedures are widely applicable in pea breeding around the world. Pea breeding programs in the USA, Australia, Europe, and India employ the pedigree breeding method. The F_2-derived family method is used at the University of Saskatchewan where approximately 300 new crosses are generated each year during the course of three crossing cycles (Warkentin et al., 2015).

Several limitations can be identified in conventional plant breeding. It is time-consuming and laborious, requiring many crosses and several generations of selfing. There is a possibility of transferring undesirable traits along with the traits of interest. The issues that are associated with conventional breeding can be addressed to some extent by incorporating molecular marker technology (Zargar et al., 2015). Marker-assisted selection (MAS) makes selection independent of the phenotypic expression of the traits. Marker systems differ based on factors including the crop species, heritability of the trait, and availability of the germplasm to exploit the variability. MAS may not immediately gain the interest of conventional breeders due to the requirement of highly skilled persons, specific infrastructure facilities, high cost,

and in some cases lack of reproducibility. Therefore, user-friendly and cost-effective MAS systems are required to achieve success (Kumar et al., 2011).

2.5.2. Marker-assisted selection

The pea genome size is 4.3 GB which is more than ten times larger than that of the model legume species *Medicago truncatula*, and approximately 85% of the pea genome consists of repetitive sequences. These challenges have contributed to a delay in the development and availability of genomic tools in pea, however, taking advantage of rapidly improving genotyping technologies, pea genome sequencing is currently in progress (Tayeh et al., 2015, Boutet et al., 2016).

A DNA marker is a fragment of DNA, which can be used to detect polymorphism between different genotypes or alleles. DNA markers such as restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), simple sequence repeat (SSR) and amplified fragment length polymorphism (AFLP) have been applied in pea (Collard et al., 2005; Gupta et al., 2008). In recent studies, single nucleotide polymorphisms (SNPs) are widely applied due to their abundant and uniformly distributed nature in the genome (Gupta et al., 2008; Leonforte et al., 2013; Boutet et al., 2016).

The use of DNA markers in plant breeding is called marker-assisted selection (MAS) which is a component of the new discipline of "molecular breeding". There is great potential to increase the efficiency and precision of conventional plant breeding by using molecular markers (Collard et al., 2005). MAS is based on pre-existing DNA sequences (DNA markers) located near a gene of interest, enabling the prediction of the presence of the particular gene in an individual progeny (McPhee, 2003). Due to the rapid improvement in genotyping procedures and statistical approaches, MAS strategy is now attracting many plant breeders towards its application in their breeding programs (Smýkal et al., 2016).

2.6. Genetic mapping

Many agricultural traits are quantitative in nature being influenced by many genes, the environment, and the genotype-by-environment interaction. To understand these quantitative traits, two broad classes of genetic mapping techniques are used, i.e., linkage mapping and association mapping. Both methods are based on the linkage disequilibrium (LD) between molecular markers and loci of interest. LD can be defined as the non-random association of alleles at different loci (Stich & Melchinger, 2010). The fundamental difference between the two mapping techniques is the mapping populations from where the LD is derived, which

directly determines the mapping resolution and power (Stich & Melchinger, 2010; Xu et al., 2017).

2.6.1. Linkage mapping

LD in linkage mapping is generated by the mating design and therefore, it can be defined as a family-based method. Linkage mapping is a useful tool to identify regions of the genome called quantitative trait loci (QTL) that co-segregate with a given trait either in F_2 populations or recombinant inbred line (RIL) families (Korte & Farlow, 2013). However, relatively low mapping resolution, low allele richness, and low speed are the major limitations in the linkage mapping (Xu et al., 2017). The limit on the mapping resolution is due to the amount of recombination that has occurred during the establishment of the RIL population. The allelic diversity that segregates between the parents of the particular F_2 cross or within the RIL population can only be evaluated in QTL mapping. Intercrossing multiple genetically diverse accessions before establishing the RILs and several generations of intercrossing when establishing the RIL population can address the above limitations to a certain extent, however, low speed could stand as a limitation (Korte & Farlow, 2013).

Although pea has lagged behind the major crops such as rice, wheat, and maize, with regard to the availability of genomic resources, early genetic maps were developed based on phenotypic and physiological characteristics (Blixt, 1974), isozymes, and early DNA-based markers such as RFLP, RAPD, ISSR, STS, CAP, and AFLP (Ellis et al., 1992; Weeden et al., 1996). Subsequently, SNP markers were used to construct pea linkage maps to identify QTLs.

To date, several pea linkage maps have been constructed to identify regions associated with agronomic traits, biotic and abiotic stress tolerances, and quality traits (Leonforte et al., 2013; Sudheesh et al., 2015; Ferrari et al., 2016; Timmerman-Vaughan et al., 2016). Tar'an et al. (2003) identified two QTLs for lodging resistance, three QTLs each for plant height and resistance to mycosphaerella blight, which accounted for 58%, 65% and 36% of the total phenotypic variation, respectively, from an evaluation of pea RILs from a cross between Carneval and MP1401 across western Canada. Moreover, from the same RIL population, four QTL each for grain yield and days to maturity, and three QTL for seed protein concentration were also identified and these genomic regions accounted for 39%, 45% and 35% of the total phenotypic variation, respectively (Tar'an et al., 2004). Jha et al. (2016; 2017) reported nine QTLs for ascochyta blight resistance, which individually explained 7.5 -

28% of the total phenotypic variation and two of them were fine mapped. Nine QTLs controlling yellow seed lightness, 3 for yellow seed greenness, 15 for seed shape, and 9 for seed dimpling were detected by Ubayasena et al. (2011) in two pea RIL population (derived from Alfetta × CDC Bronco and Orb × CDC Striker). Overall, 375 QTLs were identified for important traits such as days to flowering, days to maturity, lodging resistance, mycosphaerella blight resistance, seed weight, and grain yield based on three mapping populations, i.e., PR-02 (Orb x CDC Striker), PR-07 (Carerra x CDC Striker), and PR-15 (1–2347-144 x CDC Meadow) (Gali et al., 2018).

In a recent study, 21 QTLs (3 each on LG1, LGII, and LGVII; 6 QTLs on LGIV; 4 QTLs on LGV; and 2 QTLs on LGVI) were identified for Fe concentration in lentil by Aldemir et al. (2017). Ma et al. (2017) reported 46 seed mineral concentration (µg/g) QTLs and 37 seed mineral content (µg/ seed) QTLs using a linkage map constructed based on SNP markers for pea. The QTLs explained from 2.4% to 43.3% of the phenotypic variance (Ma et al., 2017).

2.6.2. Association mapping

In association mapping, ancestral recombination events and natural genetic diversity within a population are exploited for quantitative traits (Diapari et al., 2015). Advantages of association mapping over linkage mapping are higher resolution and greater allele numbers which have resulted from historic recombination events accumulated over hundreds of generations (Xu et al., 2017). Due to the advances in high-throughput genomic technologies, interest in identifying novel and superior alleles, and improvements in statistical methods, association mapping is indicating a continuous improvement (Mao et al., 2015). LD-based association mapping started with the model plant *Arabidopsis* and has now been extended to major crops (Kwon et al., 2012). Two major approaches are used in association mapping, i.e., genome-wide approach and candidate-gene approach. Polymorphisms in selected candidate genes that have identified roles in controlling phenotypic variation for specific traits are taken into consideration in candidate-gene association mapping (Abdurakhmonov & Abdukarimov, 2008). The candidate genes are selected based on the biochemical pathways, mutational analysis or linkage analysis. In contrast, genome-wide association mapping observes genetic variation in the whole genome to find signals of association for various complex traits where the prior knowledge of candidate genes is not required (Zhu et al., 2008).

The frequency of the variants in the sample and the difference of the effect size (phenotypic effect) are major dependant factors for the success of a GWAS. Therefore, both rare variants

and traits with small effect sizes hinder the ability of GWAS in providing insights into the genetic background of the trait. To mitigate each other's limitations, QTL mapping and GWAS could both be performed to address a particular segregating trait (Korte & Farlow, 2013).

2.7. Genotyping

2.7.1. SNP genotyping

SNP genotyping technology is rapidly advancing with the emergence of novel and cost-efficient methods. TaqMan, single-base extension-based assays, matrix-assisted laser desorption/ionization – time of flight (MALDI-TOF), mass spectrometry-based systems, the Invader assay, Pyrosequencing, capillary electrophoresis, microfluid-based chips, oligo nucleotide arrays are some of the previously used genotyping methods (He, 2014; Kim et al., 2016). SNP genotyping can be obtained either from a uniplex or multiplex genotyping platform.

Illumina GoldenGate is one of the multiplex genotyping assays that has been developed for application on a microarray platform which aims to minimizes the time, reagent volumes, and material requirements by allowing a high degree of loci multiplexing (1536-plex) during the extension and amplification steps (Illumina, Inc., 2006). In the last decade, Illumina GoldenGate assay has been used in many SNP genotyping projects on crop plants such as wheat (Akhunov et al., 2009; Trebbi et al., 2011), maize (Yan et al., 2010; Mammadov et al., 2010), pea (Deulvot et al., 2010; Diapari et al., 2015) and chickpea (Deokar et al., 2014; Diapari et al., 2014).

As sequencing costs have declined, sequence-based genotyping has gained prominence over array-based approaches. Genotyping by sequencing, or next-generation genotyping (GBS), is a genetic screening method for genotyping studies. GBS uses restriction enzymes (RE) for targeted complexity reduction and it generates high-quality polymorphism data by multiplex genotyping; major steps being DNA extraction, choosing RE, adaptor designing, library preparation, and sequencing, followed by sequence alignment and mapping (Elshire et al., 2011; Poland et al., 2012).GBS can be recognized for its important qualities of versatility, efficiency, reduction in per sample cost and high reproducibility of SNP discovery and genotyping (Sonah et al., 2013). GBS enables breeders to conduct genomic selection on a species without developing any prior molecular tools, and to determine population structure without prior knowledge of the genome or the diversity of the species. Hence, is an important

tool used in association studies even when the complete genome sequence is not available as the reference genome (Elshire et al., 2011). At present, GBS is widely being applied for SNP genotyping, generation of high-density linkage maps, and fine mapping of QTLs for various crops such as pea (Boutet et al., 2016; Annicchiarico et al., 2017; Gali et al., 2018), chickpea (Kujur et al., 2015; Singh et al., 2016) and lentil (Wong et al., 2015; Ates et al., 2016).

2.7.2. Kompetitive Allele Specific PCR (KASP) genotyping

The multiplex genotyping platforms such as GoldenGate assay and GBS are useful for large-scale genotyping studies. Despite the advantages of those methods, they are not cost effective and efficient for small-scale crop improvement applications. In such cases, KASP genotyping is a uniplex genotyping platform (Semagn et al., 2014). The method has been developed by KBioscience (Middlesex, UK) which is based on competitive allele-specific PCR. Bi-allelic scoring can be performed for SNP and insertions and deletions (Indels) at specific loci by using fluorescence detection (www.lgcgroup.com). KASP genotyping is widely applied in crop research for DNA polymorphism validation as an endpoint step for DNA marker discoveries (Graves et al., 2016; Chandra et al., 2017; Jha et al., 2017).

2.8. Phenotyping seed samples for micronutrient concentration

2.8.1. Phenotyping techniques

Use of high-quality phenotypic data is essential for successful association mapping for a particular trait. Transitioning from a candidate-gene to a genome-wide approach is also possible if robust phenotypic data were collected (Zhu et al., 2008). To detect the mineral concentrations in plant material, spectroscopic methods are widely applicable, such as atomic absorption spectroscopy (Thavarajah et al., 2007; Bueckert et al., 2011), near infrared reflectance spectroscopy (Stuth et al., 2003; Cozzolino et al., 2004), X-ray spectroscopy (MacLaren et al., 2012; Paltridge et al., 2012), and mass spectroscopic methods, such as inductively coupled plasma mass spectrometry (ICP-MS) (Retka et al., 2010; Enamorado et al., 2013). In the present study, atomic absorption spectroscopy was used based on previously developed protocols for pulse crops, and feasibility to access a suitable instrument.

2.8.2. Atomic absorption spectrophotometry

Atomic absorption spectrometry (AAS) is an analytical technique that measures the concentration of elements in a solution. Atomization converts the sample into free atoms irrespective of its initial state. Therefore, the chemical form of the element is not necessarily considered while preparing the samples. AAS is a spectro-analytical procedure for the

quantitative determination of chemical elements. According to the Royal Society of Chemistry, atoms of different elements absorb characteristic wavelengths of light (optical radiation) and the amount of light absorbed is proportional to the number of atoms in the sample. By evaluating several samples with known concentration of particular elements, calibration curves are constructed and the concentration of unknown samples is calculated based on the calibration curve.

Flame AAS is used for determination of comparatively high concentrations (tenths to tens μg/ml) of minerals such as Fe, Zn, Cu, Mg and K. The solution which contains the minerals is sucked up to the nebulizer and at the tip of the capillary, the solution is broken into small droplets and only about 1% of the sample is nebulized. Ethyne/air which produces a flame with a temperature of 2200–2400ºC, or ethyne/dinitrogen oxide, which produces a flame with a temperature of 2600-2800 ºC, are the commonly used flames to provide thermal energy to dry the sample droplets and to excite the atoms from the ground state. Hydride AAS is two-three times more sensitive than the flame AAS. Hydride AAS is applicable for minerals such as As and Se that are located in the region of wavelength < 200 nm with trace concentrations. The compound of the analyte is subjected to a reducing reaction with NaBH4 (sodium borohydride). Gaseous hydride of the analyte is formed in a hydride generator and gaseous products are delivered to the atomizer (Elwell et al., 2013).

2.9. Genome Wide Association Study (GWAS)

Over the last two decades, genome-wide association studies (GWAS) have evolved into a powerful tool starting from the first human sequence draft publication in 2001. As a global collaborative project, the high-density haplotype map of the human genome was constructed based on a large number of SNPs (Huang & Han, 2014). Since then, investigating the genetic architecture of human diseases such as cancer (Easton et al., 2007; Amos et al., 2008), Parkinson's disease (Nalls et al., 2014; Chang et al., 2017), and diabetes (Teumer et al., 2015; Cooper et al., 2017) has been conducted using GWAS.

Interestingly, GWAS has achieved greater success in plants than in humans. The cost for GWAS in crops is less than GWAS in human as it is not necessary to increase the population size and the number of markers continuously in order to overcome the missing heritability problem (Huang & Han, 2014). However, it is difficult to identify marker-trait associations with a small population, even in plants, if the functional locus has a very large effect and tested markers are in high LD with this locus (Zhu et al., 2008). The extent of LD determines

the required number of SNP markers and the mapping resolution in GWAS. In a homozygous genetic background, the recombination events are ineffective to cause LD decay. Therefore, self-fertilizing plants generally show less LD decay while the out-crossing crop species, such as maize show LD decay at short distance (Vos et al., 2017).

2.9.1. Population Structure

Geographical origin, local adaptation, and breeding history of assembled genotypes in an association mapping panel may cause false positive associations unless the population structure is taken into account (Zhu et al., 2008; Diapari et al., 2015). Some markers may appear to be significantly associated with the trait of interest when in reality, their frequency distribution is correlated with the population structure. However, if there is an average trait difference among the group of individuals, this frequency distribution may suggest the significance of markers (Sukumaran & Yu, 2013). To estimate the population structure and the familial relationship, SNPs are more appropriate than other molecular markers because of the low mutation rate and high-density availability of SNPs (Stich & Melchinger, 2010). When a quantitative trait locus (QTL), that is influencing a trait, is in linkage disequilibrium with the marker locus, the strength of marker-trait association can be inferred using statistical methods such as structured association, genomic control and mixed model approach (Zhu et al., 2008; George & Cavanagh, 2015).

The loci being correlated with the phenotype can be due to the true genetic variation or due to the false association patterns caused by the population structure of the germplasm. Population structure correction is applied in GWAS to avoid false genotype-to-phenotype associations using statistical models by distinguishing and isolating the true associations (Atwell et al., 2010; Rakitsch et al., 2012). However, in some cases, population structure correction is a stringent penalization, which can also mask the real associations. Therefore, it would be of great interest to develop novel, more powerful association methods that simultaneously correct for both population stratification and pedigree structure (Wu et al., 2011). Therefore, recently, statistical advancements have been initiated such as Quantitative Trait Cluster Association Test (QTCAT) enabling simultaneous multi-marker associations while considering correlations between markers (Klasen et al., 2016).

2.9.2. Statistical analysis in GWAS

Two statistical models are widely applied in associating molecular markers with a trait, i.e., General linear model (GLM) and Mixed linear model (MLM). In GLM, only the markers are

considered as fixed effects. In MLM in addition to markers, the principal component matrix (PCA) is also considered as a fixed factor. Furthermore, the random effects are taken into account by including the kinship matrix into the association (Diapari et al., 2014; 2015; Khazaei et al., 2017).

To reduce the chances of obtaining false-positive results (Type I errors) when multiple pairwise tests are performed on a single set of data, different correction methods are applied. The Bonferroni correction sets the significance cutoff at α/n where α is the significance level and n is the number of hypotheses, i.e., the number of markers in the case of association mapping. This is called the Bonferroni threshold level, which increases the probability of observing at least one strong significant result. Although the Bonferroni correction has become a popular correction method, its assumption that all tests are independent is not practical in many cases. In such situations, the Bonferroni correction could be extremely conservative and stringent, leading to a high rate of false negatives (Armstrong, 2014). Due to the controversy over Bonferroni correction, scientists are now testing other correction methods such as the false discovery rate (FDR).

2.9.3. GWAS application in crop plants

The first crop in which GWAS was used was maize (*Zea mays*). The maize genome experiences a rapid LD decay and in that case GWAS facilitated the genetic dissection of several complex traits, including kernel β-carotene and oil content (Cook et al., 2012; Li et al., 2013), flowering time (Brown et al., 2011; Li et al., 2016), drought tolerance (Mao et al., 2015; Wang et al., 2016), and disease resistance (Gowda et al., 2015; Olukolu et al., 2016).

GWAS is more effective in rice (*Oryza sativa*) than many other crops due to its evolution and domestication history. The properties of rice germplasm such as the unique combinations of locally adapted allele complexes, extended LD within some subpopulations, rapid LD decay within other subpopulations, and the abundance of well-partitioned genetic variation among subpopulations along with small genome size (380 Mb) provide ideal characteristics for GWAS. GWAS has been widely applied to the study of grain quality, flowering time, disease resistance and abiotic stress tolerance in rice (McCouch et al., 2016).

When considering the pulse crops, important traits have been studied using GWAS such as disease resistance in pea and common bean (Desgroux et al., 2016; Perseguini et al., 2016) and abiotic stress tolerance in chickpea (Thudi et al., 2014). Pulses are amenable to biofortification as they are among the major staple crops in the world. It can be recognized

that the attempts at identifying the underlying genetic structure of micronutrient concentration using GWAS have currently gained the interest of plant breeders and succeeded, especially in chickpea, lentil and field pea.

Chickpea is a pulse crop with a genome size of ~740 Mbp and the genomes of two chickpea cultivars (one desi and one kabuli) have been sequenced (Jain et al., 2013; Varshney et al., 2013). Substantial variability in chickpea germplasm has been indicated for seed Fe and Zn concentrations and 8 SNP loci associated with Fe and/or Zn concentration in chickpea seeds were identified by Diapari et al. (2014) using GWAS which was based on 1536 SNP markers. In a recent study, the association mapping approach (FDR cut-off ≤ 0.05) reported 16 genomic loci significantly associated ($P \leq 10^{-7}$) with seed Fe and Zn concentrations in chickpea (Upadhyaya et al., 2016).

DellaValle et al., (2013) demonstrated a significant biofortification potential and bioavailability of Fe in lentil. The lentil germplasm tested by Khazaei et al. (2017) exhibited a wide range of variation for seed Fe and Zn concentration with the detection of two SNP markers for seed Fe and one SNP for Zn concentration ($-\log_{10} P \geq 4.36$) in the marker-trait association study.

To date, there is only one GWAS study on field pea micronutrients which was conducted by Diapari et al. (2015), in which 94 diverse pea accessions were genotyped using 1,233 SNPs in an Illumina GoldenGate array. Nine SNPs ($-\log_{10}$ P value ≥ 4.2) were significantly associated with iron concentration in seeds, two SNPs with Zn concentration, but no SNPs were associated with Se concentration. SNP loci PsC22912p327 (LGI) and PsC9886p84 (LGIV) explained the highest phenotypic variation, 10.7 %, followed by PsC16473p224 (unmapped) that explained 9.8 %. PsC5316p234 (LGV) explained 9.6 %; PsC7893p98 (LGIII) and PsC8677p415 (LGIII) and PsC13009p652 (LGIV) explained 9.3 %; PsC25762p728 (unmapped) explained 8.5 % of phenotypic variation, suggesting that markers responsible for phenotypic variation of iron were well distributed over the genome. Two zinc-associated SNPs PsC7872p386 and PsC3195p368 (LGIII) explained 11.5 and 9.2 % of the genetic variation, respectively. To build on the research of Diapari et al. (2015) the current study evaluates marker-trait assocaitions in a larger pea germplasm collection, i.e., 177 accessions instead of 94 accessions, and uses 14,391 SNP markers instead of 1,233. Evaluating marker-trait assocaitions for micronutrient concentration in a large pea germplasm collection using next-generation sequencing techniques is beneficial and effective to more

precisely identify markers which can facilitate pea breeding. Furthermore, validation of markers in an unrelated population adds more reliability to the markers for future use in MAS.

3. Materials and Methods

3.1. Evaluation of diverse pea accessions for the concentration of Fe, Zn, and Se

3.1.1. Plant material

A genome wide association study (GWAS) panel of 177 pea accessions developed at the Crop Development Centre (CDC), University of Saskatchewan consisting of cultivars and landraces from North America, western and eastern Europe, and Australia were evaluated for Fe, Zn, and Se concentration (Appendix A). A total of 1770 seed samples derived from field trials conducted at Saskatoon and Rosthern locations in Saskatchewan and at Fargo, North Dakota were used (Table 3.1).

Table 3.1: Seed samples derived over five location-years

Year	Fargo, ND	Saskatoon,SK	Rosthern,SK	Sample number
2013	-	2 replicates	-	177 x 2 = 354
2014	2 replicates	-	-	177 x 2 = 354
2015	-	2 replicates	-	177 x 2 = 354
2016	-	2 replicates	2 replicates	177 x 4 = 708

3.1.2. Seed grinding

Based on a preliminary study, 60 seeds of each harvested sample were utilized (Appendix B). The 60 seeds were randomly sampled from a sub-sample of the mechanically harvested bulk seeds in each accession. The whole seed samples were ground into fine powder using a cyclone sample mill (UDY Corporation, Fort Collins, Colorado-USA) and the ground samples were stored at room temperature and ambient humidity until digestion.

3.1.3. Standard protocol for nitric acid digest for seeds

Finely ground seed samples were digested according to the methods of Thavarajah et al. (2007) in a Vulcan-84 digester and processed according to the methods described in Diapari et al. (2015). For each location, two replicates were digested separately and both replicates were analyzed at the same time with the atomic absorption spectrophotometer (nova 300, Analytic Jena AG, Germany).

A total of 0.3 g of the powdered sample was weighed into a digestion tube. Two digestion blocks (2 × 42 digestion tubes) were placed in the Vulcan-84 digester at a time. One set of 84 digestion tubes consisted of 4 blanks, 8 internal standard seed samples (ground CDC Meadow field pea) and 72 seed samples.

The digestion blocks with the digestion tubes were warmed to 86ºC. A volume of 6 ml HNO_3 (ACS grade) was added to each tube and they were kept in the block for 45 minutes. The tubes were swirled or shaken to allow the gas to escape and not to leave residue on the side of the tube. After 45 minutes, 3 ml of 30% H_2O_2 was added to the digest while they are still in the digestion block. After that, another 2 ml of 30% H_2O_2 was added. The reaction mixture was allowed to digest for another 65 minutes. Then, 3 ml of 6 M HCl was added and allowed to digest for another 5 minutes. The digestion tubes were taken off of the block digester and volume was topped up to 25 ml with distilled water. Samples were stored in plastic vials at room temperature for AAS analysis.

3.1.4. Obtaining atomic absorption spectroscopy (AAS) data

The "blank solution" was prepared by conducting the total nitric acid digestion for seeds in clean empty digestion tubes. The standard solutions were prepared by diluting 1000 mg/L stock solutions of Fe and Zn with the blank solution. The standards for Fe were 0 mg/L, 0.5 mg/L, 1.0 mg/L, 3.0 mg/L, and 4.0 mg/L and for Zn are 0 mg/L, 0.2 mg/L, 0.4 mg/L, 1.2 mg/L, and 1.6 mg/L. The Fe and Zn standards were combined when diluting so that the 4.0 mg/L Fe and the 1.6 mg/L Zn are in the same solution and so forth.

Five ml of each digested solution of the seed samples was transferred into 16 ml vials. For Fe analysis, internal standards were prepared in the same manner, additionally added 0.2 ml of 25 mg/L Fe standard solutions. The concentration of Fe was calculated based on the internal standard. For Zn analysis, internal standards were necessary. The digested solutions and the Fe internal standards were run through the AA air/acetylene flame for analysis.

The Se analysis was conducted using hydride analysis. For Se, the standards were made from 0-10 µg/L in the digested solution. The digested solutions were prepared by diluting it six times with 0.75 M HCl to a volume of 12 ml and then run on the hydride system. Internal standards were prepared for several samples by adding 0.1 ml of 1 mg/L Se standard solution to the digested solutions. The prepared solutions were run through the hydride system. For all Fe, Zn, and Se samples, two AAS readings were taken and the average was calculated in solid (µg/g).

3.1.5. Phenotypic data analysis

Analysis of variance (ANOVA) was performed for Fe, Zn, and Se concentration, the dependent variables, using the mixed model in SAS 9.4 (SAS Institute Inc., Cary, North Carolina, USA). The effect of genotype, genotype-by-location, genotype-by-year and

genotype-by-year-by-location interactions were analyzed for all the location-years. Genotype, location and year were considered as fixed effects, while replicates were taken into consideration as random effects. Five location-years were considered as five environments and the combined analysis of variance was also performed. Broad sense heritability and the percentage total sums of squares were calculated. The Pearson's correlation coefficients between each micronutrient, between each different location, between the agronomic traits (days to flowering, lodging, days to maturity and yield) and the micronutrient concentration were also calculated using the SAS 9.4 (SAS Institute Inc., Cary, North Carolina, USA).

3.2. GWAS to identify SNPs associated with Fe, Zn, and Se concentration

3.2.1. Genotyping

The GWAS panel accessions were genotyped using genotyping-by-sequencing (GBS; Elshire et al., 2011) to assess the level of genetic diversity based on SNP variation. The number of raw sequencing reads per accession was obtained for the GWAS panel. The tagged reads were aligned to the pea genome sequence (Source: International Pea Genome Sequencing Consortium) and the SNPs identified across the panel accessions were used for the diversity analysis (Gali et al, in prep).

3.2.2. Population structure and diversity analysis

Population structure for GWAS panel was assessed with genotyping data (SNP), the population structure matrix (Q) was obtained and kinship matrix (K) was generated using the fastSTRUCTURE (Stanford, CA, USA). The GWAS accessions were grouped into major clusters based on the distribution of SNPs (Gali et al, in prep).

3.2.3. Association study

Population structure was set as a fixed effect and the kinship among individuals was incorporated as the variance-covariance structure of the random effect for the individuals (Diapari et al., 2015). Trait Analysis by Association, Evolution, and Linkage (TASSEL, New York, USA) software was used to test the association between SNP markers against each of Fe, Zn, and Se concentration for location-years individually. Mixed linear model (MLM) was used for the association study (Bradbury et al., 2007). Manhattan plots were generated to show the distribution of the *P* values for the SNP markers. Bonferroni corrected threshold (*P* value/number of SNP markers) at $P \leq 0.05$ (P = 3.47E-06, $-\log_{10} P$ value ≥ 5.46) was used to declare the association of the markers with the trait. Significant SNPs, if any, associated with Fe, Zn, and Se concentration were identified. To observe the presence of the identified

significant markers across multiple location-years, the threshold was lowered (P = 6.95E-04, -$\log_{10}$ P value ≥ 3.16).

3.3. Validation of the identified SNP markers

3.3.1. Plant material

PR-07 recombinant inbred lines (RILs) population was used to validate the SNP markers identified in the pea GWAS panel of 177 accessions. The PR-07 population was developed by crossing the two parents CDC Striker and Carrera at the Crop Development Centre (CDC), University of Saskatchewan. Previously, 94 lines of PR-07 were evaluated for Fe, Zn, and Se concentration from the seeds obtained from field trials in six location-years (2010 Rosthern, 2010 SPG, 2011 Rosthern, 2011 SPG, 2012 Rosthern, and 2012 SPG with two biological replicates). The same RILs were genotyped using genotyping-by-sequencing (GBS) and used for identification of quantitative trait loci (QTLs) for Fe, Zn, and Se (Gali et al., 2018).

3.3.2. Kompetitive allele-specific PCR (KASP) genotyping

Out of the significant markers identified in the GWAS panel, sixteen potential markers were selected. When there were two or more significant SNP markers from the same scaffold, only one of them was selected. Also, the presence of the markers at multiple location-years was considered. Primers were designed for the selected markers for KASP genotyping using a primer picker software (KBiosciences, Middlesex, UK) based on the 50 nucleotide sequence flanking each side of the SNP. Two allele-specific oligonucleotides of about 40 nucleotides in length (forward primer for allele 1 and forward primer for allele 2) and 1 common oligonucleotide of about 20 nucleotides in length (common reverse primer) were designed.

The KASP assay mixture was prepared by mixing 12 µl each of 100 µM allele-specific primer 1 and allele-specific primer 2, 30 µl of the common primer, and 46 µl H_2O. The 384-well microlitre plates were used for the KASP assay. A total reaction volume was produced by adding 3 µl of template DNA (5 ng/ml), 3 µl of 2 × KASP reaction mix and 0.083 µl of KASP assay mixture. The plates were sealed with an optically clear seal. Thermal cycling consisted of: 21ºC of 2 min, 95ºC of 15 min, 9 cycles of 94ºC, 20 s and 65ºC, 30 s, 32 cycles of 94ºC, 30s and 57ºC, 1 min and then 12 cycles of 94ºC, 20s and 57ºC, 1 min (including 21ºC, 2 min followed by each 3 cycle) was perfomed in the C1000™ Thermal Cycler (CFX384™ Real-Time System, Bio-Rad Laboratories, Inc., USA) (Appendix C). Genotypic data were analyzed using fluorescence resonance energy transfer (FRET) readers and

fluorescence signals were viewed graphically as a scattered plot using the software SNPViewer (KBioscience, Middlesex, UK).

First, all the primers were genotyped with parental DNA to check whether the SNP locus is polymorphic for the parents of the population. Sc1512_36017 was the only marker polymorphic between CDC Striker and Carrera and used to genotype the 94 RILs of PR-07 according to the protocol mentioned earlier.

3.3.3. Validation of the SNP markers on PR-07 RILs

Based on the presence of allele 1 and allele 2, the KASP data were divided into two groups and "two groups mean t-test" was performed for individual location-years and the combined location-years of the PR-07 Zn concentration using SAS 9.4 (SAS Institute Inc., Cary, North Carolina, USA).

Genotypic data of Sc1512_36017 along with 3389 SNP markers used earlier for the linkage map of PR-07 (Gali et al., 2018) were used to construct a linkage map using MstMap. The parameters used for linkage map construction were, LOD of 10, no mapping distance threshold of 15 cM and no mapping missing threshold of 15%. The KASP marker data were added to the existing genotyping information and a linkage map was created to check whether each SNP marker is co-located with the QTL for a given micronutrient.

4. Results

4.1. Phenotypic data analysis

4.1.1. Descriptive statistics

Mean Fe concentration across the five location-years was 50.46 µg/g and the coefficient of variation (CV) was 15.1%. Highest standard deviation (SD) and highest coefficient of variation (CV) was observed in 2014 Fargo. The rest of the location-years had similar SD and CV % values. Lowest Fe concentration was reported in 2016 Saskatoon and highest Fe concentration was in 2014 Fargo; 29.22 µg/g and 90.53 µg/g, respectively (Table 4.1).

Table 4.1: Summary of the descriptive statistics of Fe concentration in seeds of pea GWAS panel over five location-years.

Year	Location	Mean ± SD (µg/g)	CV%	Range (µg/g)
2013	Saskatoon	45.95 ± 6.46	14.1	31.41 - 68.96
2014	Fargo	52.95 ± 8.97	17.0	36.72 - 90.53
2015	Saskatoon	48.80 ± 6.16	12.6	37.4 - 74.21
2016	Saskatoon	50.81 ± 6.68	13.2	29.22 - 74.76
2016	Rosthern	53.81 ± 6.66	12.4	39.73 - 74.52
2016	Ros /Sas	52.31 ± 6.83	13.1	29.22 - 74.76
2013/2015/2016	Saskatoon	48.51 ± 6.74	13.9	29.22 - 74.76
All	All	50.46 ± 7.61	15.1	29.22 - 90.53

CV – Coefficient of variation; SD – Standard deviation

Mean Zn concentration for the five location-years was 29.49 µg/g with a SD of 5.71 µg/g. The concentrations ranged from 12.83 µg/g to 51.47 µg/g: highest being 2014 Fargo and lowest being 2016 Rosthern. Same as Fe, Zn concentration across the genotypes in 2014 Fargo had the greatest CV% of 20.1% (Table 4.2).

Table 4.2: Summary of the descriptive statistics of Zn concentration in seeds of pea GWAS panel over five location-years.

Year	Location	Mean ± SD (µg/g)	CV%	Range (µg/g)
2013	Saskatoon	28.34 ± 3.98	14.0	16.78 - 39.89
2014	Fargo	23.22 ± 4.66	20.1	12.83 - 43.46
2015	Saskatoon	29.62 ± 3.66	12.3	16.64 - 44.08
2016	Saskatoon	32.79 ± 4.49	13.7	19.75 - 44.33
2016	Rosthern	33.55 ± 4.92	14.7	21.97 - 51.47
2016	Ros /Sas	33.17 ± 4.72	14.2	19.75 - 51.47
2013/2015/2016	Saskatoon	30.26 ± 4.46	14.8	16.64 - 44.33
All	All	29.49 ± 5.71	19.4	12.83 - 51.47

CV – Coefficient of variation; SD – Standard deviation

Se concentration had a substantially different pattern compared with Fe and Zn. Concentration across the five location-years ranged from 0.06 µg/g to 8.75 µg/g with a CV of 75.6%. The lowest CV% was observed at 2014 Fargo (Table 4.3).

Table 4.3: Summary of the descriptive statistics of Se concentration in seeds of pea GWAS panel over five location-years.

Year	Location	Mean ± SD (µg/g)	CV%	Range (µg/g)
2013	Saskatoon	1.09 ± 0.85	78.3	0.25 - 6.12
2014	Fargo	0.41 ± 0.09	22.6	0.22 - 0.88
2015	Saskatoon	1.73 ± 0.76	44.2	0.27 - 8.75
2016	Saskatoon	0.96 ± 0.37	39	0.33 - 3.28
2016	Rosthern	0.58 ± 0.3	51.8	0.06 - 1.9
2016	Ros/Sas	0.77 ± 0.39	50.5	0.06 - 3.28
2013/2015/2016	Saskatoon	1.26 ± 0.77	61.2	0.25 - 8.75
All	All	0.95 ± 0.72	75.6	0.06 - 8.75

CV – Coefficient of variation; SD – Standard deviation

4.1.2. Analysis of Variance (ANOVA)

The ANOVA for Fe and Zn concentration showed similar effects of year, location, genotype, and their interactions. The genotypes differed significantly in all the location-years. Combined analysis for Saskatoon and Rosthern in 2016 also showed a significant genotype effect and a significant location effect. Both the year and genotype effects were significant

for Saskatoon in 2013, 2015, and 2016. The interactions of the genotype-by-year or genotype-by-location were not significant (Table 4.4, Table 4.5).

Table 4.4: Analysis of variance (ANOVA) of Fe concentration in seeds of pea GWAS panel over five location-years.

Year	Location	F value				
		Y	L	G	G×Y	G×L
2013	Saskatoon	.	.	4.74***	.	.
2014	Fargo	.	.	5.03***	.	.
2015	Saskatoon	.	.	2.07***	.	.
2016	Saskatoon	.	.	1.88***	.	.
2016	Rosthern	.	.	3.4***	.	.
2016	Ros/Sas	.	67.69***	3.95***	.	0.93ns
2013/2015/2016	Saskatoon	87.24***	.	4.89***	1.04ns	.

Y –Year; L – Location; G – Genotype; *** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$; ns- not significant

Table 4.5: Analysis of variance (ANOVA) of Zn concentration in seeds of pea GWAS panel over five location-years.

Year	Location	F value				
		Y	L	G	G×Y	G×L
2013	Saskatoon	.	.	4.1***	.	.
2014	Fargo	.	.	3.24***	.	.
2015	Saskatoon	.	.	2.53***	.	.
2016	Saskatoon	.	.	4.28***	.	.
2016	Rosthern	.	.	5.16***	.	.
2016	Ros +Sas	.	10.58***	6.55***	.	1.03ns
2013/2015/2016	Saskatoon	209.27***	.	5.72***	1.28ns	.

Y –Year; L – Location; G – Genotype; *** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$; ns- not significant

For Se concentration, the genotype effect was not significant except at 2014 Fargo. Year and location significantly affected the Se concentration in both the combined analysis for 2016 and for Saskatoon. Genotype-by-year or genotype-by-location did not have a significant effect on Se concentration (Table 4.6).

Table 4.6: Analysis of variance (ANOVA) of Se concentration in seeds of pea GWAS panel over five location-years.

Year	Location	F value				
		Y	L	G	G×Y	G×L
2013	Saskatoon	.	.	1.06ns	.	.
2014	Fargo	.	.	1.99***	.	.
2015	Saskatoon	.	.	1.04ns	.	.
2016	Saskatoon	.	.	0.98ns	.	.
2016	Rosthern	.	.	0.86ns	.	.
2016	Ros/Sas	.	282.61***	0.99ns	.	0.85ns
2013/2015/2016	Saskatoon	185.41***	.	1.01ns	0.92ns	.

Y –Year; L – Location; G – Genotype; *** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$; ns- not significant

Combined analysis of variance for all the five location-years was performed considering the location-years as five different environments. Genotype, environment, and the genotype-by-environment interaction had significant effects on Fe and Zn concentration. Same as the individual location-year analysis, the genotype effect was not significant for Se concentration. The environment significantly affected the Se concentration (Table 4.7).

Table 4.7: Analysis of variance (ANOVA) for the five location-years (five environments) in seeds of pea GWAS panel.

Micronutrient	F values		
	E	G	G×E
Se	319.77***	1.1ns	0.91ns
Zn	677.56***	10.52***	1.44***
Fe	149.68***	9.78***	1.4***

E – Environment; G – Genotype; *** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$; ns- not significant

Broad sense heritability is considerably high for Zn and Fe concentration, 86.4% and 85.7% respectively, however, it is quite low for Se concentration (8.1%) (Table 4.8).

Table 4.8: Variance estimates and the broad sense heritability for Fe, Zn, and Se in seeds of pea GWAS panel over five location-years.

Source	Variance estimates (σ^2)		
	Se	Zn	Fe
E	0.2597	16.9822	10.0811
G	0.002592	8.1356	20.1224
G×E	0	1.8983	3.9481
Residual	0.2916	9.0168	25.6193
H^2	8.1%	86.4%	85.7%

E – Environment; G – Genotype; H^2 – Broad sense heritability

Based on the sum of squares calculated by the combined analysis, the percentage total sums of squares (Semipartial eta squares; η^2) were obtained. Out of the total variation in Fe, contribution of the genotype was 40.5%, followed by genotype-by-environment (23.1%), the residual error (22.2%) and environment (14.1%). For Zn, 28.6% of the total variation can be explained by genotype, 41.9 % by environment, 15.7% by genotype-by-environment, and 13.8% by residual error. Genotype explained 6.3% of the total variation in Se whereas environment contributed 41.6%, the genotype-by-environment 20.9%, and the residual error 20.9% of the total variation (Table 4.9).

Table 4.9: Percentage total sums of squares (Semipartial eta squares; η^2) for the five location-years (five environments) in seeds of the pea GWAS panel.

Source	Fe		Zn		Se	
	Type III SS	η^2 %	Type III SS	η^2 %	Type III SS	η^2 %
G	41178.95	40.5	16467.08	28.6	55.68	6.3
E	14316.15	14.1	24090.53	41.9	368.84	41.6
G×E	23537.79	23.1	9022.23	15.7	185.69	20.9
Error	22587.78	22.2	7935.86	13.8	276.8	31.2
Total	101719.4	99.9	57480.93	100.1	887.32	100

E – Environment; G - Genotype

4.1.3. Correlation analysis

4.1.3.1. Correlation between the micronutrient concentrations

Significant correlations were observed in all the location-years between Fe and Zn concentration. Fe concentration was significantly correlated with Se in all the location-years, except 2016 Rosthern. The correlation between Zn and Se was only significant in 2014 Fargo and 2016 Saskatoon. All the correlations were positive except the correlation between Zn and Se in 2016 Rosthern (Table 4.10).

Table 2: Pearson's correlation coefficient (r) between the Fe, Zn, and Se in seeds of field pea over five location-years.

Statistics	Pearson's correlation coefficient				
	2013S	2014F	2015S	2016R	2016S
Fe and Zn	0.63***	0.70***	0.48***	0.53***	0.41***
Fe and Se	0.21***	0.47***	0.12*	0.05ns	0.18***
Zn and Se	0.09ns	0.40***	0.10ns	-0.07ns	0.27***

2013S - 2013 Saskatoon, 2014F - 2014 Fargo, 2015S - 2015 Saskatoon, 2016R - 2016 Rosthern, 2016S - 2016 Saskatoon. *** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$; ns- not significant

4.1.3.2. Correlation between the locations

Three different locations were included in the five location-years of the study: Saskatoon, Rosthern, and Fargo. For Fe and Zn concentration, the correlation was significant between each of these locations. Se concentration in Saskatoon correlated significantly with both Rosthern and Fargo, however, Fargo and Rosthern did not correlate significantly (Table 4.11).

Table 4.11: Pearson's correlation coefficient (r) for Fe, Zn, and Se between the locations over five location-years.

Statistics	Pearson's correlation coefficient (r)		
	Fe	Zn	Se
Rosthern and Saskatoon	0.73***	0.62***	0.19*
Fargo and Rosthern	0.58***	0.53***	0.07ns
Fargo and Saskatoon	0.63***	0.77***	0.26***

*** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$; ns - not significant

4.1.3.3. Correlation between micronutrient concentrations and agronomic data

Lodging and DTM were significantly correlated with Fe concentration in all the location-years. The correlation between DTF and Fe concentration was positive in three out of five location-years. Yield showed a significant negative correlation with Fe in three out of five location-years: 2014 Fargo, 2015 Saskatoon, and 2016 Saskatoon (Table 4.12).

Table 4.12: Pearson's correlation coefficient (r) between the Fe concentration and agronomic traits over five location-years.

Location-Year	Pearson's correlation coefficient (r)			
	DTF	Lodging	DTM	Yield
2013S	0.18***	0.43***	0.58***	-0.07ns
2014F	0.42***	0.28***	0.55***	-0.48***
2015S	-0.11*	0.38***	0.12*	-0.29***
2016S	0.05ns	0.35***	0.26***	-0.11*
2016R	0.11*	0.37***	0.30***	0.07ns

DTF - Dates to flowering, DTM - Dates to maturity, 2013S - 2013 Saskatoon, 2014 F - 2014 Fargo, 2015S - 2015 Saskatoon, 2016R - 2016 Rosthern, 2016S - 2016 Saskatoon.*** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$; ns - not significant

Significant correlations were observed between DTF and Zn concentration in three out of five location-years: 2013 Saskatoon, 2014 Fargo, and 2015 Saskatoon. Lodging was significantly correlated with Zn in all the location-years. Correlation between DTM and Zn concentration was significant in all the location-years except in 2015 Saskatoon. In general, correlation between yield and Zn concentration was negative except in 2016 Rosthern and the negative correlations were significant in 2013 Saskatoon, 2014 Rosthern, and 2015 Saskatoon (Table 4.13)

Both DTF and lodging score were significantly correlated with Se concentration in 2014 Fargo but not in the other location-years. Correlation between DTM and Se concentration was significant in two location-years: 2013 Saskatoon and 2014 Fargo. Grain yield was significantly correlated with Se concentration in three out of five location-years, where the correlation in the two 2016 location-years were not significant. Three location-years showed a negative correlation with the Se concentration (Table 4.14).

Table 4.13: Pearson's correlation coefficient (r) between the Zn concentration and agronomic traits over five location-years.

Location-Year	Pearson's correlation coefficient (r)			
	DTF	Lodging	DTM	Yield
2013S	-0.12*	0.43***	0.23***	-0.19***
2014F	0.34***	0.33***	0.39***	-0.53***
2015S	-0.27***	0.19***	-0.10ns	-0.45***
2016S	0.03ns	0.27***	0.25***	-0.01ns
2016R	0.08ns	0.33***	0.18***	0.01ns

DTF - Dates to flowering, DTM - Dates to maturity, 2013S - 2013 Saskatoon, 2014 F - 2014 Fargo, 2015S - 2015 Saskatoon, 2016R - 2016 Rosthern, 2016S - 2016 Saskatoon. *** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$; ns - not significant

Table 4.14: Pearson's correlation coefficient (r) between the Se concentration and agronomic traits across Rosthern, SK, Saskatoon, SK and Fargo, ND in 2013, 2014, 2015, and 2016

Location-Year	Pearson's correlation coefficient (r)			
	DTF	Lodging	DTM	Yield
2013S	0.07ns	0.08ns	0.16**	0.15**
2014F	0.30***	0.16**	0.33***	-0.44***
2015S	-0.05ns	0.03ns	-0.02ns	-0.18***
2016S	-0.05ns	0.09ns	0.01ns	-0.05ns
2016R	0.10ns	0.04ns	0.10ns	0.05ns

DTF - Dates to flowering, DTM - Dates to maturity, 2013S - 2013 Saskatoon, 2014 F - 2014 Fargo, 2015S - 2015 Saskatoon, 2016R - 2016 Rosthern, 2016S - 2016 Saskatoon. *** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$; ns - not significant

4.2. Genotypic data analysis

The GWAS panel accessions were genotyped using genotyping-by-sequencing (Gali et al., in prep) to assess the level of genetic diversity based on SNP variation. The number of raw sequencing reads obtained was 3.6 to 12.9 million reads per accession with an average of 7.47 million tagged reads per accession. Of the tagged reads of individual accessions, 47-75% were aligned to the pea genome sequence (Source: International Pea Genome Sequencing Consortium). This resulted in the identification of 1632 to 5209 SNPs per accession, indicating the genetic diversity of the panel accessions. A total of 14,391 SNPs identified across the panel accessions were used for the diversity analysis. The GWAS accessions were grouped into nine major clusters based on the distribution of 14,391 SNPs (Gali et al., in prep) (Appendix D).

The grouping of accessions was generally consistent with geographic origin. Both the accessions in cluster 1 originated from Russia. The majority of the accessions originated from France were in cluster 3. Cluster 5 mostly consisted of Canadian accessions which were developed by the CDC, University of Saskatchewan. Cluster 4 and 7 mainly consisted of Australian origin accessions. All of the accessions originated from UK were in cluster 8 and 9. Almost all the accessions from Netherlands were in cluster 9. The majority of the accessions which had a high micronutrient concentration were in cluster 3, whereas the majority of the accessions which had a low micronutrient concentration were grouped in cluster 9 (Gali et al., in prep) (Appendix E).

4.3. Genome wide association study

In genome wide association study, several SNP markers were significantly associated with Fe, Zn, and Se concentration after Bonferroni correction ($-\log_{10} P$ value ≥ 5.46). None of these markers were significant in more than one location-year, therefore the threshold limit was reduced ($\log_{10} P$ value ≥ 3.16) to identify the presence of the significant markers at multiple location-years.

Two markers significantly associated with Fe concentration: Sc_219_42293 and Sc_219_42326 (at $-\log_{10} P$ value ≥ 5.46) in 2013 Saskatoon and those two markers appeared in 2014 Fargo and 2016 Rosthern at $-\log_{10} P$ value ≥ 3.16. Sc_6182_63782 was uniquely identified in 2015 Saskatoon (Table 4.15).

For Zn concentration, 7 SNP markers were identified at $-\log_{10} P$ value ≥ 5.46. All of these significant SNP markers were identified in the same location-year (2016 Rosthern). In addition, they were also significant in at least two other location-years, except Sc_1340_418746 which was only detected in 2016 Rosthern at $-\log_{10} P$ value ≥ 3.16 (Table 4.16).

For Se, 44 significant markers were detected at $-\log_{10} P$ value ≥ 5.46. Even though markers were significant in 2013 Saskatoon (2 markers) and 2015 Saskatoon (42 markers), Repeated occurrence across the location-years was not observed for any of those markers (Appendix F).

The locations of the markers in the pea genome by chromosomes are unknown at this time (source: International Pea Genome Sequencing Consortium), but this will be pursued in future research. The associated markers that were present in two or more location-years can be considered as having the potential to be developed and incorporated in marker assisted selection.

Table 4.15: SNP markers associated with Fe concentration in seeds of pea GWAS panel over five location-years.

Location-year	Marker	Marker R^2 %	*P* value	$-\log_{10}$	Other location-years that the marker is present ($-\log_{10} P$ value ≥ 3.16)
2013S	Sc_219_42293	15.5	5.57E-07	6.25**	2014F, 2016R
2013S	Sc_219_42326	15.5	5.57E-07	6.25**	2014F, 2016R
2015S	Sc_6182_63782	16.2	2.15E-06	5.67*	none

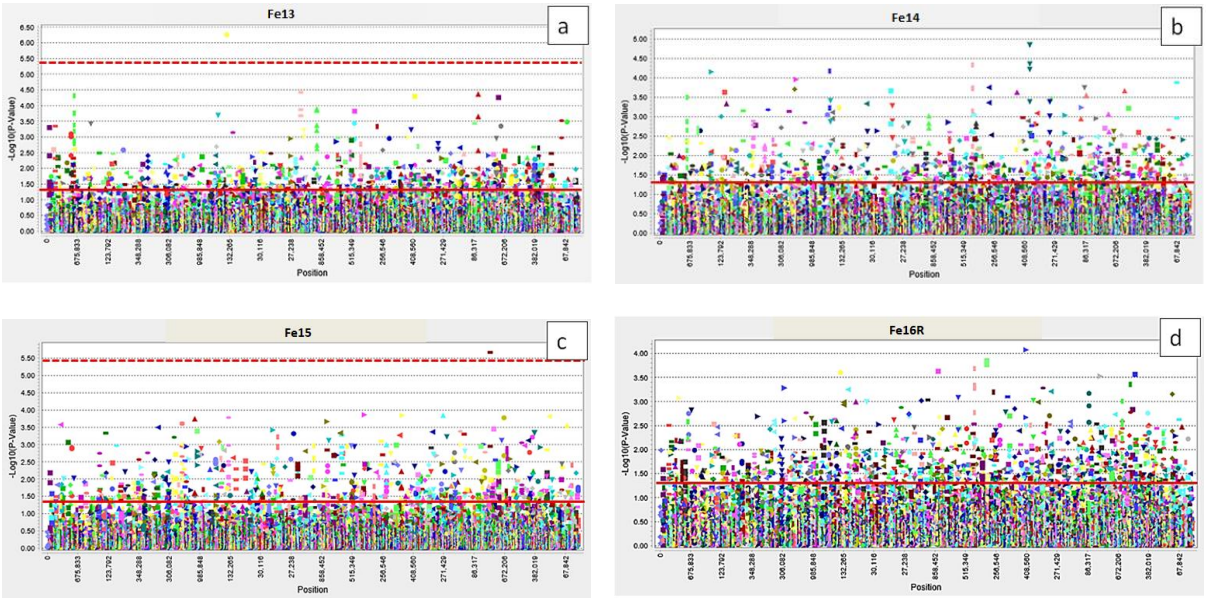
2013S - 2013 Saskatoon, 2014F - 2014 Fargo, 2015S - 2015 Saskatoon, 2016R - 2016 Rosthern, 2016S - 2016 Saskatoon. ** $P \leq 0.01$ ($-\log_{10} P$ value ≥ 6.16); * $P \leq 0.05$ ($-\log_{10} P$ value ≥ 5.46)

Table 4.16: SNP markers associated with Zn concentration in seeds of the pea GWAS panel over five location-years.

Location-year	Marker	Marker R^2 %	*P* value	-$\log_{10}$	Other location-years that the marker is present (-$\log_{10}$ *P* value ≥ 3.16)
2016R	Sc_8032_15394	17.3	1.44E-07	6.84**	2013S, 2014F
2016R	Sc_8032_15457	17.3	1.44E-07	6.84**	2013S, 2014F
2016R	Sc_8032_15361	17.4	8.58E-07	6.07*	2014F, 2016S
2016R	Sc_1512_36017	18.6	1.57E-06	5.80*	2014F
2016R	Sc_1512_35793	14.2	1.62E-06	5.79*	2014F, 2016S
2016R	Sc_1512_36043	17.8	2.75E-06	5.56*	2014F
2016R	Sc_1340_418746	19.3	3.28E-06	5.48*	none

2013S - 2013 Saskatoon, 2014F - 2014 Fargo, 2015S - 2015 Saskatoon, 2016R - 2016 Rosthern, 2016S - 2016 Saskatoon. ** $P \leq 0.01$ (-$\log_{10}$ *P* value ≥ 6.16), * $P \leq 0.05$ (-$\log_{10}$ *P* value ≥ 5.46)

Manhattan plots were developed for all five location-years for Fe, Zn, and Se (Figure 4.1, 4.2 and 4.3).



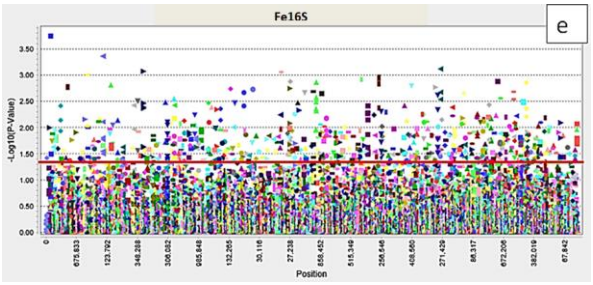


Figure 4.1: Manhattan plots of $-\log_{10} P$ values of the GWAS for Fe concentration using MLM; a) 2013 Saskatoon, b) 2014 Fargo, c) 2015 Saskatoon, d) 2016 Rosthern, and e) 2016 Saskatoon. The straight lines indicate the threshold of $-\log_{10} P$ value ≥ 1.30 and the dotted lines indicate Bonferroni corrected threshold of $-\log_{10} P$ value ≥ 5.46 at 0.05 level of significance, Y axis: $-\log_{10} P$ values, X axis: base pair numbers.

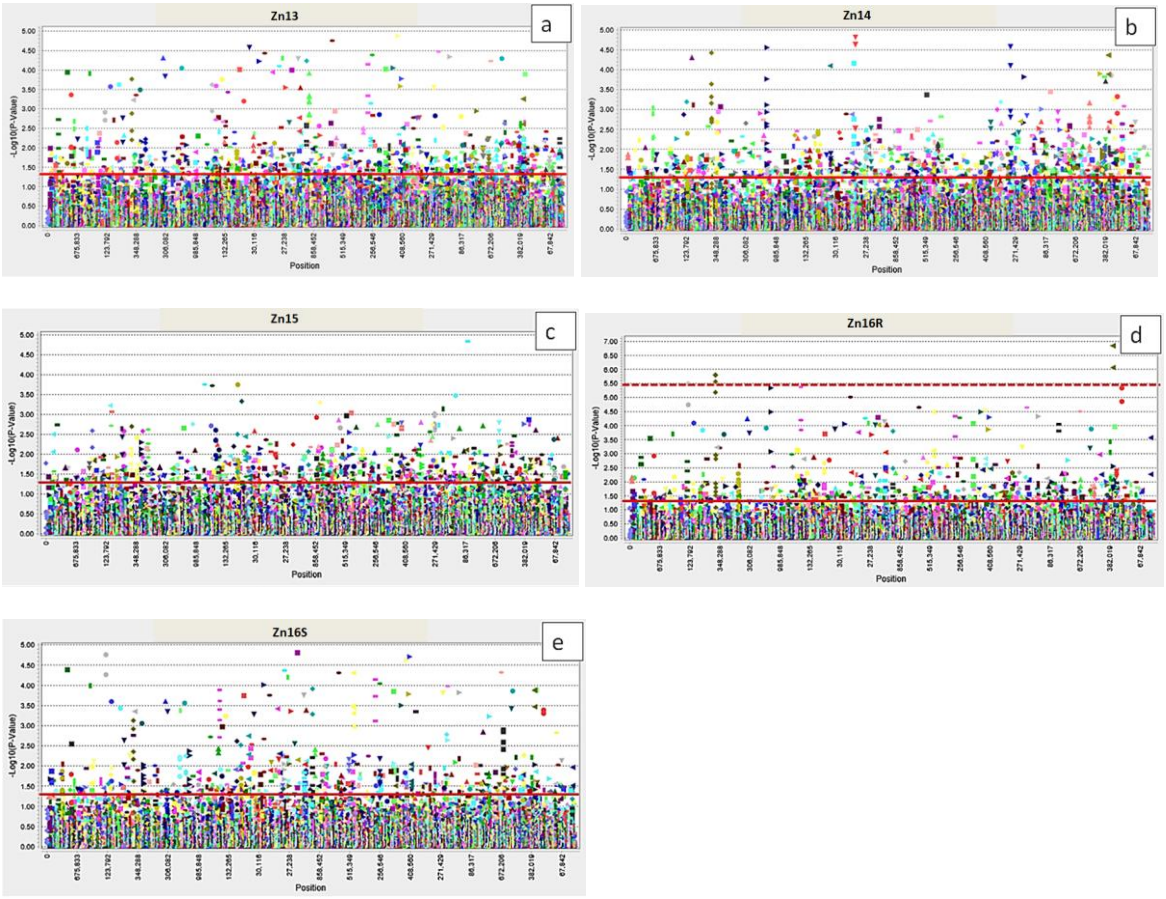


Figure 4.2: Manhattan plots of $-\log_{10} P$ values of the GWAS for Zn concentration using MLM; a) 2013 Saskatoon, b) 2014 Fargo, c) 2015 Saskatoon, d) 2016 Rosthern, and e) 2016 Saskatoon. The straight lines indicate the threshold of $-\log_{10} P$ value ≥ 1.30 and the dotted lines indicate Bonferroni corrected threshold of $-\log_{10} P$ value ≥ 5.46 at 0.05 level of significance. Y axis: $-\log_{10} P$ values, X axis: base pair numbers

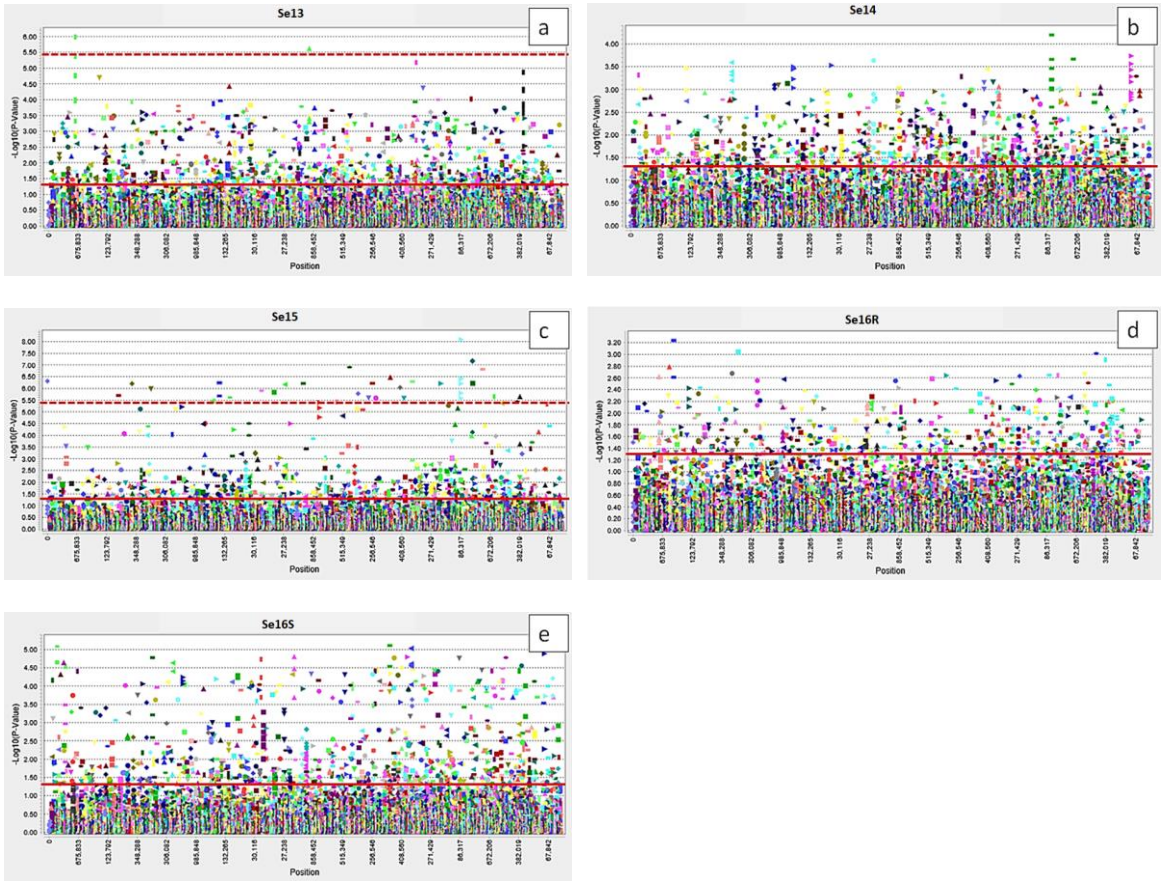


Figure 4.3: Manhattan plots of $-\log_{10} P$ values of the GWAS for Se concentration using MLM;a) 2013 Saskatoon, b) 2014 Fargo, c) 2015 Saskatoon, d) 2016 Rosthern, and e) 2016 Saskatoon. The straight lines indicate the threshold of $-\log_{10} P$ value ≥ 1.30 and the dotted lines indicate Bonferroni corrected threshold of $-\log_{10} P$ value ≥ 5.46 at 0.05 level of significance. Y axis: $-\log_{10} P$ values, X axis: base pair numbers

4.4. Validation of the identified SNP markers

Based on association studies in GWAS panel (Tables 15,16 and Appendix E), 16 SNP markers for Fe, Zn, and Se were selected for validation in an unrelated pea population, PR-07 (derived from a cross between Carrera and CDC Striker) using KASP assays. (Table 4.17).

Table 4.17: The list of promising markers resulted from the pea GWAS panel

No	Micronutrient	SNP marker
1	Fe	Sc_219_42326
2	Fe	Sc_6182_63782
3	Fe	Sc_890_604394
4	Fe	Sc_4420_108326
5	Fe	Sc_3584_82073
6	Zn	Sc_8032_15394
7	Zn	Sc_1512_36017
8	Zn	Sc_1340_418746
9	Zn	Sc_2113_215664
10	Zn	Sc_858_268917
11	Zn	Sc_433_404350
12	Zn	Sc_2859_396829
13	Se	Sc_1139_105036
14	Se	Sc_3049_27354
15	Se	Sc_5765_134167
16	Se	Sc_6064_83250

These 16 SNP markers were used to design primers to conduct validation of the markers. Thirteen markers succeeded from the quality control step where the other 3 which had extra degenerate nucleotides were not effective for primer designing (Table 4.18).

Table 4.18: List of allele specific primers designed for the 13 SNP markers (two forward primers and one common primer for each SNP locus).

No	Primer	Sequence (5' to 3')
1	Sc219_42326_ALT	GAAGGTGACCAAGTTCATGCTCCAGAGAACAACAACCCTGTCT
	Sc219_42326_ALC	GAAGGTCGGAGTCAACGGATTCCAGAGAACAACAACCCTGTCC
	Sc219_42326_C1	GACTATGTTGACATGAACTCAGGGAATTT
2	Sc6182_63782_ALT	GAAGGTGACCAAGTTCATGCTCATGAAGTGGTGCCCCTTTATCAA
	Sc6182_63782_ALA	GAAGGTCGGAGTCAACGGATTCATGAAGTGGTGCCCCTTTATCAT
	Sc6182_63782_C1	CMTGGGGTGCTGTTTGGAA
3	Sc890_604394_ALG	GAAGGTGACCAAGTTCATGCTTTWCTTTTGCTATTTTTTTAATCATAAATTTG
	Sc890_604394_ALT	GAAGGTCGGAGTCAACGGATTTWCTTTTGCTATTTTTTTAATCATAAATTTT
	Sc890_604394_C2	ACATTTGATCARATTTCGTGTATATA
4	Sc4420_108326_ALG	GAAGGTGACCAAGTTCATGCTCATCAAAAGACCAAGGGAACAATGTC
	Sc4420_108326_ALT	GAAGGTCGGAGTCAACGGATTACATCAAAAGACCAAGGGAACAATGTA
	Sc4420_108326_C1	GATAGAATAGTTTTTGACAATGAGATTCTT
5	Sc3584_82073_ALT	GAAGGTGACCAAGTTCATGCTATTCTGAAGCTTCAAAGAACAACACCT
	Sc3584_82073_ALA	GAAGGTCGGAGTCAACGGATTCTGAAGCTTCAAAGAACAACACCA
	Sc3584_82073_C1	GTAGAAACTGCAAATATGTGCCYGGTT
6	Sc8032_15394_ALA	GAAGGTGACCAAGTTCATGCTGAACTGGGATACAAYCACCA
	Sc8032_15394_ALG	GAAGGTCGGAGTCAACGGATTGAACTGGGATACAAYCACCG
	Sc8032_15394_C1	TACWAGTAGCAAGTCAAT
7	Sc1512_36017_ALC	GAAGGTGACCAAGTTCATGCTTCTTCACAAAACGAAGATGGTACTG
	Sc1512_36017_ALG	GAAGGTCGGAGTCAACGGATTCTTCACAAAACGAAGATGGTACTC
	Sc1512_36017_C1	TTCATCATTCTGAACACTGACATCATCYTT
8	Sc2113_215664_ALA	GAAGGTGACCAAGTTCATGCTAAATCTCAAGCTGATGTTGATCAGGTT
	Sc2113_215664_ALG	GAAGGTCGGAGTCAACGGATTATCTCAAGCTGATGTTGATCAGGTC
	Sc2113_215664_C1	TAATTGATTATCAGWAACATTAA
9	Sc858_268917_ALT	GAAGGTGACCAAGTTCATGCTGCTAGCTCTATTGCATGTGATYTA
	Sc858_268917_ALC	GAAGGTCGGAGTCAACGGATTGCTAGCTCTATTGCATGTGATYTG
	Sc858_268917_C2	AAAGATTCYAAATACATTAAATA
10	Sc433_404350_ALT	GAAGGTGACCAAGTTCATGCTAATGAATCAAAAGATTAAACARCAAACAATGAA
	Sc433_404350_ALC	GAAGGTCGGAGTCAACGGATTATGAATCAAAAGATTAAACARCAAACAATGAG
	Sc433_404350_C1	AGCGAGAAATCAACAACAATCAATACSATT
11	Sc2859_396829_ALT	GAAGGTGACCAAGTTCATGCTCCGACATGTGTCTGTGTCAATGT
	Sc2859_396829_ALC	GAAGGTCGGAGTCAACGGATTCCGACATGTGTCTGTGTCAATGC
	Sc2859_396829_C1	GATATAATACAACACTGACACGGACACTA
12	Sc1139_105036_ALC	GAAGGTGACCAAGTTCATGCTGTCTAGAAGGCCGGATTAGKTATG
	Sc1139_105036_ALA	GAAGGTCGGAGTCAACGGATTAGTCTAGAAGGCCGGATTAGKTATT
	Sc1139_105036_C1	CYGAGATAATTGCGAAATAAAAAATGAATA

13	Sc6064_83250_ALC	GAAGGTGACCAAGTTCATGCTAAGCCACACCAACATGTTRCTGG
	Sc6064_83250_ALT	GAAGGTCGGAGTCAACGGATTGAAGCCACACCAACATGTTRCTGA
	Sc6064_83250_C1	ATATCTAACTGCAGAATCTCCACTYAACTA

One SNP locus Sc_1512_36017 which was identified for Zn was polymorphic for the parents (CDC Striker and Carrera) of population PR07. For the rest of the twelve markers, the alleles of the particular SNP locus were either distributed across the plot or not closely grouped, thus were considered as not polymorphic for the parents. Therefore, those markers were not able to be cross validated with PR-07. The polymorphic marker Sc_1512_36017 was genotyped with the PR-07 RILs and significant allele discrimination could be observed (Figure 4.4). Out of 94 lines of PR-07 tested, 49 lines had allele C and 41 lines had allele G, while three were heterozygotes and one was undetermined. Five out of six location-years of the PR-07 RIL population showed a significant difference ($P \leq 0.001$) between the two allele groups in terms of Zn concentration. The highest group mean and standard deviation were observed in 2010 SPG, 4.5 µg/g and 3.1 µg/g, respectively, in the two allele groups. Combined analysis also showed a significant difference between the two allelic groups ($P \leq 0.001$) (Table 4.19).

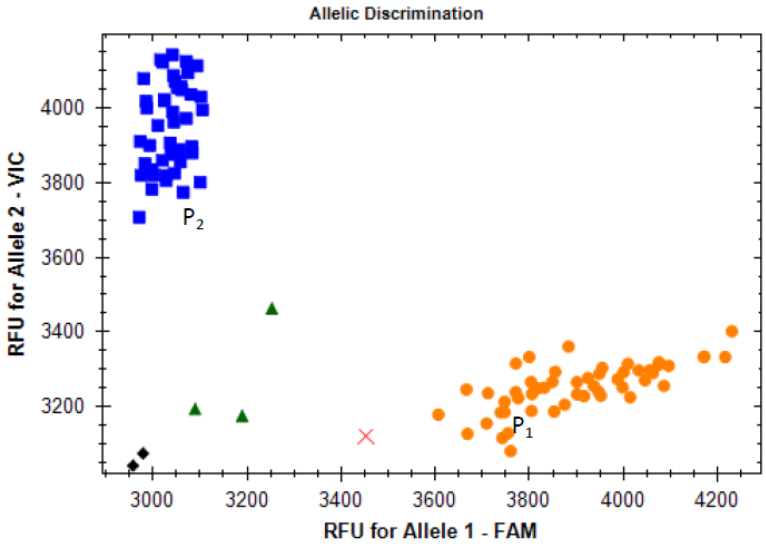


Figure 4.4: Allele discrimination plot of the KASP genotyping for the SNP locus Sc1512_36017 with the PR-07 RIL populationP_1 – Carrera; P_2 – CDC Striker; ● - Allele 1; ▲ - Heterozygotes; ■ - Allele 2; ◆ - Blanks; × - Undetermined, RFU; VIC; FAM).

Table 4.19: Descriptive statistics of the two allele groups for Zn concentration in PR-07 RIL population based on the presence of the alleles C and G in the SNP locus Sc1512_36017 over six location-years.

Location-year	Allele	N	Mean (µg/g)	SD	Minimum (µg/g)	Maximum (µg/g)	t-value
2010	C	49	32.8	1.8	28.0	36.6	
Rosthern	G	41	28.9	2.8	25.1	37.6	8.1***
2010	C	49	36.9	2.7	27.2	41.0	
SPG	G	41	32.5	3.5	26.6	40.6	6.8***
2011	C	49	31.4	2.1	26.9	37.5	
Rosthern	G	41	28.9	2.6	23.9	38.7	5.0***
2011	C	49	27.6	2.5	20.9	32.8	
SPG	G	41	25.5	2.7	17.4	33.1	3.8***
2012	C	49	33.2	2.8	26.6	39.3	
Rosthern	G	41	30.9	3.3	26.9	38.7	3.6***
2012	C	49	28.3	2.4	23.3	33.7	
SPG	G	41	29.1	3.0	24.2	36.7	1.4ns
Combined	C	49	31.7	1.4	27.6	35.0	
	G	41	29.3	2.0	26.3	35.2	6.7***

N – Number of RILs; SD – Standard deviation; *** $P \leq 0.001$; ** $P \leq 0.01$; * $P \leq 0.05$; ns - not significant

The Linkage map was re-constructed for PR-07 incorporating Sc1512_36017 and it provided the same order of markers, and clustering of sub-linkage groups as reported by Gali et al. (2018) (Appendix H). The marker Sc1512_36017 co-located with Sc11336_48840 on LGIIIb in PR-07 (Figure 4.5). Sc11336_48840 was the flanking marker of a QTL for seed Zn concentration in four of the six location-years in PR-07, and it explained 50.1% of total phenotypic variation (Gali et al. 2018).

LGIIIb

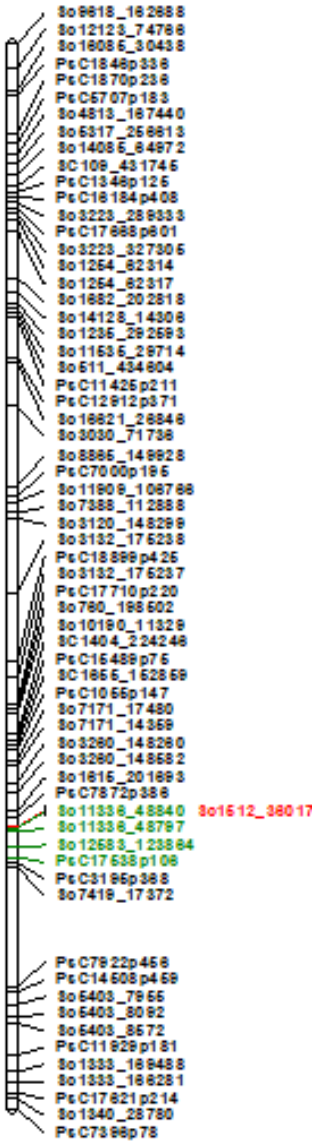


Figure 4.5: LGIIIb of PR-07.The markers in green represent the QTL interval (87.7-91.0cM) for the trait, Zn concentration, whereas the marker in red is co-localized marker Sc1512_36017 with Sc11336_48840

5. Discussion

Micronutrient concentration in pea seeds is a quantitative trait which is an important quality trait of the crop. Phenotyping of a diverse collection of pea varieties for micronutrient concentration using atomic absorption spectroscopy and their association with genotypic data using genome-wide association study (GWAS) was useful to identify single nucleotide polymorphism (SNP) markers that could assist in the biofortification of this crop.

5.1. Inferences on phenotypic data

In the present study, mean concentration for the combined analysis of the five location-years were 50.46, 29.49, and 0.95 and the concentration (µg/g) ranged from 29.22 to 90.53, 12.83 to 51.47 and 0.06 to 8.75 in the pea seed flour for Fe, Zn, and Se, respectively. In a similar study of 94 accessions of pea in Rosthern and Saskatoon in 2011 and 2012, mean concentrations (µg/g) were reported as 45.32, 33.74, and 0.88, and the concentration (µg/g) ranged from 25.71–93.68, 14.39–92.51, and 0.08–5.53, respectively for Fe, Zn, and Se (Diapari et al., 2015). Ray et al. (2014) reported the mean concentration (mg/ kg) as 53.6, 30.5, and 4.7 and the concentrations ranged from 47.7 to 58.1 for Fe, from 27.4 to 34.0 for Zn, and from 0.4 to 5.5 for Se in mg/ kg in a study using 17 pea cultivars grown at six locations in Saskatchewan in 2005 and 2006. The range of the Fe, Zn, and Se are noticeably higher in our study compared with Ray et al. (2014), which may be due to the diversity of the 177 GWAS accessions in our study.

In the present study, the mean Fe concentration in pea seeds harvested at Rosthern (53.81 ± 6.66 µg/g) was higher than those harvested at Saskatoon in the average of all three location-years (48.51 ± 6.74 µg/g). The same pattern occurred in Zn concentration: 33.55 ± 4.92 at Rosthern compared with 30.26 ± 4.46 at Saskatoon. A similar trend was reported for pea by Diapari et al. (2015) for Fe, 47.96 ± 5.95 and 42.65 ± 6.35, and for Zn, 35.77 ± 5.53 and 31.17 ± 5.95 for Rosthern and Saskatoon, respectively.

The CV% did not vary greatly among individual location-years for Fe and Zn, except for Fargo, which was somewhat greater than the others. However, for Se, the CV% differed considerably across the location-years. The lowest CV among the location-years was observed at Fargo. The CV% for the combined dataset were 15.1, 19.4, and 75.6, for Fe, Zn and Se, respectively. Similar CVs were reported by Diapari et al. (2015) for Fe, Zn, and Se (14.7%, 18.5%, and 91.5%).

The GWAS genotypes differed significantly ($P \leq 0.001$) for Fe and Zn concentration at all location-years. However, for Se, genotypes only differed significantly at Fargo. Year and location effects were significant for all three micronutrients, however, genotype-by-year and genotype-by-location effects were not significant ($P \leq 0.001$) for any of the location-years. Diapari et al. (2015) reported a similar result with significant genotype effect in pea ($P \leq 0.001$) for Fe and Zn, but not for Se, with significant year and location effects of all the three micronutrients. Thavarajah et al. (2010) also did not observe significant genotypic differences in total Se concentration among 17 field pea cultivars grown at six locations for two years in Saskatchewan. Significant genotype effects of Fe and Zn were observed in other pulse crops as well, such as lentil (Khazaei et al., 2017; Vandemark et al., 2018.), chickpea (Diapari et al., 2014; Vandemark et al., 2018) and soybean (King et al. 2013).

The variation in Fe and Zn concentration in seeds was mainly due to the genetic factor. From the broad sense heritability test, considerable genetic variation between individual accessions in the GWAS panel could be observed in Zn and Fe (86.4% and 85.7%, respectively) but not in Se (8.1%). In lentil, Khazaei et al. (2017) reported that the broad-sense heritability was higher for seed concentration of Fe (78% and 90%) compared with that of Zn (44% and 68%) in both years (2013 and 2014, respectively). When considering the cereal crops, sorghum showed 98% heritability for both Fe and Zn content (Badigannavar et al., 2016).

In terms of the percentage total sums of squares (Semipartial eta squares; η^2), the genotype explained 40.5%, 28.6%, and 6.3% for Fe, Zn, and Se respectively. Diapari et al. (2015) observed a similar result that Se concentration showed a very low genotypic contribution compared with Fe and Zn (60.3 %, 66.5%, and 2.7% for Fe, Zn, and Se respectively). In contrast, the genotypic contribution for Fe and Zn variation in the present study is less than the values reported by Diapari et al. (2015). Any difference between the η^2 compared with this previous study may be due to location and year effects. For example, in our study, Se concentration in Fargo, was differed substantially from the locations in Saskatchewan. Se concentration was less in North Dakota (0.33 µg/g) than Saskatchewan (0.73 µg/g) in chickpea as well (Vandemark et al., 2018).

Significant correlations were observed in all the location-years between Fe and Zn concentration ($P \leq 0.05$). Se concentration was significantly correlated with Fe in all the location-years, except 2016 Rosthern ($P \leq 0.05$). The correlation between Se and Zn was only significant in 2014 Fargo and 2016 Saskatoon ($P \leq 0.05$). When considering the 20

accessions with the greatest and the least concentration of Fe, Zn, and Se for the combined analysis across all the location-years, many of them were common between Fe and Zn. Within the 20 accessions with the greatest concentration, the accession 'Serpette D'auvergne' which was originated from France appeared in all three micronutrients (Appendix E). Similar correlations between Fe, Zn, and Se were observed by Diapari et al. (2015) and Ma et al. (2017). A positive correlation between Fe and Zn was also reported in lentil (Khazaei et al., 2017), chickpea (Diapari et al., 2014), maize (Akinwale & Adewopo, 2016) and durum wheat (Magallanes-López et al., 2017). The reason could be the same QTL controls Fe and Zn concentration, or linked QTLs control the two traits. Several studies have identified QTLs associated with grain Zn and Fe concentration that have been mapped in the same chromosomal region in wheat (Tiwari et al., 2009, Crespo-Herrera et al., 2017).

Concentration of Fe, Zn, and Se were significantly correlated ($P \leq 0.05$) between each location with the exception of Se concentration between Fargo and Rosthern. Diapari et al. (2015) also reported a strong correlation between Fe and Zn across locations, but weak to no correlation in the case of Se. In the present study, it is not a surprise to see the exception in Se when considering the range and CV% of the Se concentration in Fargo and Rosthern. Soil factors, precipitation, and temperature may have played a role for this exception. According to Environment Canada weather data, mean temperatures in Rosthern and Saskatoon for 2013, 2015 and 2016 were similar (~16 ºC). However, 2016 was a wetter year than other growing years of the study (184.1, 143.5, 220.1, and 240.3 in mm for 2013 Saskatoon, 2015 Saskatoon, 2016 Saskatoon and 2016 Rosthern, respectively). 2014 was considered as above average for precipitation and below average for temperature in North Dakota (National Climate Report, USA, 2014).

It is important to consider the effect of the environmental factors as well on the seed micronutrient concentration, especially the soil in the crop field. A concurrent study of soil in Rosthern and Saskatoon for the GWAS accessions showed that the CV was higher in Rosthern than Saskatoon for the Fe, Zn, and Se concentration in soil (Appendix I). The pea accessions grown in the same fields showed the same pattern for the Fe, Zn, and Se concentration in seeds. However, when considering the correlation between soil and seed concentration, noticeable relationship could not be demonstrated. Rosthern soil texture was mostly silt loam in 0-15 cm and 15-30 cm depths and silty clay loam in the 30-60 cm depth, while at Saskatoon the three depths are more similar with the texture, which is a mixture of silt, loam, and clay. Other than the concentration of the micronutrients in the soil, various

different factors such as nutrient interactions in soil, microbial activity, redox potential and aeration also regulate the availability of the micronutrients for the plants (Kumar et al., 2016).

Generally, Fe and Zn concentration showed a significant positive correlation with lodging and days to maturity in all the location-years but correlation was not significant for Se concentration. In the majority of the location-years for all three micronutrients, grain yield was negatively correlated with concentration (Appendix G). Accessions with high lodging scores, i.e., with weak stems, tended to be low yielding and with greater Fe and Zn concentration. Ma et al. (2017) demonstrated negative correlations between seed weight and all the mineral nutrient concentrations they studied including Fe and Zn in pea. Yield showed a significant positive correlation with Fe but significant negative correlation with Zn in lentil and chickpea (Vandemark et al., 2018). Fe and Zn concentration showed significant negative correlations with grain yield in wheat (Morgounov et al., 2007), and soybean (Oliveira et al., 2016). It has been hypothesized that when the yield increases, the concentrations of mineral elements in produce decreases because of a "dilution effect" caused by plant growth rates exceeding the ability of plants to acquire these elements; both environmental and genetic factors affect this phenomenon (Jarrell and Beverly, 1981; Davis et al. 2004; White et al., 2009).

In the present study, Fe showed significant positive correlation with days to flowering in many of the location-years, but the correlations were either negative or not significant for Zn and Se. A similar result was demonstrated in lentil and chickpea by Vandemark et al. (2018).

5.2. Potential SNP markers for micronutrient concentration

GWAS panel accessions were grouped into 9 clusters in the diversity analysis. The majority of the clusters included accessions with the same origin (Gali et al., in prep.). Diapari et al. (2015) reported that the grouping of 96 pea accessions was consistent with their geographic origin.

Association study resulted 3 SNP markers for Fe, 7 SNP markers for Zn, and 44 SNP markers for Se, which are significant (-log *P* value $\geq$ 5.46) for the concentration. All the three markers that were significant for Fe were identified from the location Saskatoon. Two out of the three markers identified for Fe were in the same scaffold Sc_219 and closely positioned in the sequence. When observing under the lower threshold (-log *P* value $\geq$ 3.16), Sc_219_42293 and Sc_219_42326 were appeared in three location-years out of five. Each of

these two markers accounted for 15.5% of the total phenotypic variation. Considering the facts, any of those two markers has a great potential to be used in MAS.

The seven markers identified significant for Zn concentration were distributed among three scaffolds, Sc_8032, Sc_1512, and Sc_1340. All the markers appeared common in at least two location-years, except Sc_1340_418746 which only appeared in 2016 Rosthern (at -log *P* value ≥ 3.16).

Even though, 44 SNP markers could be identified for the Se concentration, there are few facts that have to be considered. Se showed a high coefficient of variation in all the location-years, 78.3%, 22.6%, 44.2%, 39%, and 51.8% in 2013 Saskatoon. 2014 Fargo, 2015 Saskatoon, 2016 Saskatoon, and 2016 Rosthern, respectively. Also, the CVs are not in a similar range across the location-years compared with the other two micronutrients: Fe and Zn, which suggest that the Se concentration largely depends upon the location and the year. Further, Se shows a low broad sense heritability (8.1%). Moreover, ANOVA results provide evidence that genotypic effect is insignificant for the Se concentration ($P \leq 0.05$) in many of the cases. Therefore, the possibility of the markers identified for the Se concentration to be used in MAS is less compared with the markers identified for Fe and Zn concentration.

In a marker-trait association study conducted by Diapari et al. (2015) reports a total of nine SNPs (-log $_{10}$ *P* value ≥ 4.2) for Fe and two SNPs for Zn which were significantly associated with the concentration. However, none of the SNP markers were associated with the pea seed Se concentration after discarding the markers considering the high CV and the high location effect for Se (Diapari et al., 2015). In contrast, the current study succeeded to discover seven strongly associated markers for Zn at a higher threshold limit of -log *P* value ≥ 5.46 compared to a threshold limit of -log $_{10}$ *P* value ≥ 4.2 reported by Diapari et al. (2015) . In lentil, Khazaei et al. (2017) used a threshold of -log $_{10}$ *P* value ≥ 4.36 to detect strong associations and lower thresholds of -log $_{10}$ *P* value ≥ 3.36 and 3.06 to observe the common markers across the sites. Diapari et al. (2014) reported significant markers for chickpea at -log $_{10}$ *P* value ≥ 3.1. At a lower threshold of -log *P* value ≥ 3.16, we could observe more SNP markers associated with Fe and Zn for pea, however, the strongest markers were selected to be validated in an unrelated population as described in section 5.3.

5.3. QTLs for micronutrient concentration

For the validation, significant markers were identified based on the *P* value (-log *P* value ≥ 5.46). When a multiple marker appeared in the same scaffold closely positioned (100-150

base pair distance), one of the markers was chosen for the validation. Also, the markers which are non-significant for Bonferroni threshold, but the -log *P* value is greater than 3.16 were ranked in a descending order. Few noteworthy markers with higher -log *P* value at multiple location-years were also selected. A number of 16 markers were resulted to be validated. Due to the extra degenerate nucleotides beside the particular SNP, 3 out of 16 markers failed from the quality control and therefore, only 13 markers were validated with PR-07. Twelve out of the 13 markers were not polymorphic for the PR-07 parents: Carrera and CDC Striker at their particular SNP locus. The marker which was polymorphic (Sc1512_36017 with 18.6% contribution to the phynotypic variation) was genotyped with the PR-07 RILs and distinct allele discrimination could be observed.

QTL regions can be determined by linkage mapping the markers with the existing genotyping information to make inferences on strong applicability of the markers in MAS. The marker Sc1512_36017 was co-localized with Sc11336_48840 on LGIIIb in PR-07. Sc11336_48840 has identified by Gali et al. (2018) to be the flanking marker of QTL for seed Zn concentration in four of the six field trials. Diapari et al. (2015) mapped two Zn-associated SNPs on LGIII where PsC7872p386 and PsC3195p368 explained 11.5% and 9.2 % of the genetic variation, respectively. In Gali et al. (2018), the marker PsC3195p368 was also reported as flanking marker with Sc11336_48840 in the QTL interval of 87.9-91.1cM and the marker PsC7872p386 was closely positioned to the Sc11336_48840. The Sc11336_48840 that was strongly associated with seed Zn concentration had LOD 17.3 and it explained 50.1% of phenotypic variance. Ma et al. (2017) reported QTL, [Zn]-Ps3.1 for seed Zn concentration in a mapping population of pea (Aragorn × Kiflica) on LGIII, whereas Gali et al. (2018) reported another QTL for Zn concentration on LGIIIb in PR-02 RIL population of pea (Orb×CDC Striker). These literature support that the validated marker Sc1512_36017 has a great possibility to be used in marker assisted selection for Zn concentration in field pea.

To date, there is only one study which reported pea association mapping for micronutrients which used 94 pea accessions, genotyped by Illumina GoldenGate assay with 1233 SNP markers (Diapari et al., 2015). The current study is the first in which a large number of accessions were used (177), sequenced with next generation sequencing: GBS with 14,391 SNP markers. Also, we used data from three different locations covering Saskatchewan and North Dakota in four years: 2013, 2014, 2015, and 2016. In contrast, Diapari et al. (2015) used data from two locations in Saskatchewan in two years: 2011 and 2012. Moreover, the present study is the first pea association mapping study where the identified markers were

crossed checked and validated on an unrelated population. To date, the present study is the most robust marker-trait association study in pea for micronutrients.

5.4. Conclusion and future research

In this research, 3 significant markers were detected for Fe concentration and 7 significant markers for Zn concentration (-log10 *P* value ≥ 5.46). Five markers for Fe, 7 markers for Zn, and 4 markers for Se were validated in pea recombinant inbred line population PR-07 (derived from Carrera × CDC Striker). The marker Sc1512_36017 was identified to co-localize with Sc11336_48840 on LGIIIb which was earlier reported to be the flanking marker of a quantitative trait loci (QTL) for seed Zn concentration. To improve the efficiency and precision of conventional plant breeding, this marker has good potential to be used in marker assisted selection in pea breeding.

In future research, the identified markers can be tested in other parental population as well where the parents are polymorphic for the particular SNP locus. When the pea genome sequence is available, the SNP markers identified for micronutrients can be used for gene annotation which will allow identification of the coding and non-coding regions of the gene, gene locations and functions.

References

Agriculture and Agri-Food Canada (2008) available from http://www.agr.gc.ca/eng/?id=1299099172433 (Accessed on 15th April 2016).

Agriculture and Agri-Food Canada (2017) https://www.newswire.ca/news-releases/statement-by-the-government-of-canada-on-pulse-exports-to-india-660860313.html (Accessed on 14th September 2018).

Abbaspour, N., Hurrell, R., & Kelishadi, R. (2014). Review on iron and its importance for human health. *Journal of Research in Medical Sciences*, 19(2).

Abdurakhmonov, I. Y., & Abdukarimov, A. (2008). Application of association mapping to understanding the genetic diversity of plant germplasm resources. *International Journal of Plant Genomics*, 2008.

Agegnehu, G., Bass, A. M., Nelson, P. N., Muirhead, B., Wright, G., & Bird, M. I. (2015). Biochar and biochar-compost as soil amendments: effects on peanut yield, soil properties and greenhouse gas emissions in tropical North Queensland, Australia. *Agriculture, ecosystems & environment*, *213*, 72-85.

Ahmed, T., Hossain, M., & Sanin, K. I. (2012). Global burden of maternal and child undernutrition and micronutrient deficiencies. *Annals of Nutrition and Metabolism*, *61*(Suppl. 1), 8-17.

Akhunov, E., Nicolet, C., & Dvorak, J. (2009). Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. *Theoretical and Applied Genetics*, *119*(3), 507-517.

Akinwale, R. O., & Adewopo, O. A. (2016). Grain iron and zinc concentrations and their relationship with selected agronomic traits in early and extra-early maize. *Journal of Crop Improvement*, *30*(6), 641-656.

Aldemir, S., ATEŞ, D., TEMEL, H. Y., YAĞMUR, B., Alsaleh, A., Kahriman, A., ... & Tanyolac, M. B. (2017). QTLs for iron concentration in seeds of the cultivated lentil (Lens culinaris Medic.) via genotyping by sequencing. *Turkish Journal of Agriculture and Forestry*, *41*(4), 243-255.

Allen, L. H. (2000). Anemia and iron deficiency: effects on pregnancy outcome–. *The American journal of clinical nutrition*, *71*(5), 1280S-1284S.

Amos, C. I., Wu, X., Broderick, P., Gorlov, I. P., Gu, J., Eisen, T., ... & Sullivan, K. (2008). Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25. 1. *Nature genetics*, *40(5)*, 616-622.

Annicchiarico, P., Nazzicari, N., Pecetti, L., Romani, M., Ferrari, B., Wei, Y., & Brummer, E. C. (2017). GBS-based genomic selection for pea grain yield under severe terminal drought. *The plant genome*.

Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics*, *34(5)*, 502-508.

Ashokkumar, K., Tar'an, B., Diapari, M., Arganosa, G., & Warkentin, T. D. (2014). Effect of cultivar and environment on the carotenoid profile of pea and chickpea. *Crop Science*, *54(5)*, 2225-2235.

Ates, D., Sever, T., Aldemir, S., Yagmur, B., Temel, H. Y., Kaya, H. B., ... & Tanyolac, B. (2016). Identification QTLs controlling genes for se uptake in lentil seeds. *PLoS One*, *11*(3), e0149210.

Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., ... & Jiang, R. (2010). Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. *Nature*, *465*(7298), 627.

Badigannavar, A., Girish, G., Ramachandran, V., & Ganapathi, T. R. (2016). Genotypic variation for seed protein and mineral content among post-rainy season-grown sorghum genotypes. *The Crop Journal*, *4*(1), 61-67.

Bailey, R. L., West Jr, K. P., & Black, R. E. (2015). The epidemiology of global micronutrient deficiencies. *Annals of Nutrition and Metabolism*, *66*(Suppl. 2), 22-33.

Bhutta, Z. A., Das, J. K., Rizvi, A., Gaffey, M. F., Walker, N., Horton, S., ... & Maternal and Child Nutrition Study Group. (2013). Evidence-based interventions for improvement of maternal and child nutrition: what can be done and at what cost?. *The lancet*, *382*(9890), 452-477.

Blair, M. W., Astudillo, C., Grusak, M. A., Graham, R., & Beebe, S. E. (2009). Inheritance of seed iron and zinc concentrations in common bean (*Phaseolus vulgaris* L.). *Molecular Breeding*, *23(2)*, 197-207.

Blixt, S., (1974) The pea. In: King RC (ed) *Handbook of genetics, vol. 2. Plenum Press, New York*, pp 181–221

Bouis, H. E. (2002). Plant breeding: a new tool for fighting micronutrient malnutrition. *The Journal of nutrition*, *132*(3), 491S-494S.

Bouis, H. E., Hotz, C., McClafferty, B., Meenakshi, J. V., & Pfeiffer, W. H. (2011). Biofortification: a new tool to reduce micronutrient malnutrition. *Food and nutrition bulletin*, *32*(1_suppl1), S31-S40.

Bouis, H. E., & Saltzman, A. (2017). Improving nutrition through biofortification: A review of evidence from HarvestPlus, 2003 through 2016. *Global Food Security*, *12*, 49-58

Bourre, J. M. (2006). Effects of nutrients (in food) on the structure and function of the nervous system: update on dietary requirements for brain. Part 2: macronutrients. Journal of Nutrition Health and Aging, 10(5), 386.

Boutet, G., Carvalho, S. A., Falque, M., Peterlongo, P., Lhuillier, E., Bouchez, O., & Baranger, A. (2016). SNP discovery and genetic mapping using genotyping by sequencing of whole genome genomic DNA from a pea RIL population. *BMC genomics*, *17(1)*, 121.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633-2635.

Bray, T. M., & Bettger, W. J. (1990). The physiological role of zinc as an antioxidant. *Free Radical Biology and Medicine, 8(3),* 281-291.

Brown, P. J., Upadyayula, N., Mahone, G. S., Tian, F., Bradbury, P. J., Myles, S., ... & Rocheford, T. R. (2011). Distinct genetic architectures for male and female inflorescence traits of maize. *PLoS genetics, 7*(11), e1002383.

Bueckert, R. A., Thavarajah, P., & Pritchard, J. (2011). Phytic acid and mineral micronutrients in field-grown chickpea (Cicer arietinum L.) cultivars from western Canada. *European Food Research and Technology, 233(2),* 203-212.

Chandra, S., Singh, D., Pathak, J., Kumari, S., Kumar, M., Poddar, R., ... & Mukhopadhyay, K. (2017). SNP discovery from next-generation transcriptome sequencing data and their validation using KASP assay in wheat (Triticum aestivum L.). *Molecular Breeding*, *37*(7), 92.

Chang, D., Nalls, M. A., Hallgrímsdóttir, I. B., Hunkapiller, J., van der Brug, M., Cai, F., ... & Hinds, D. (2017). A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nature genetics, 49(10),* 1511.

Cichy, K. A., Caldas, G. V., Snapp, S. S., & Blair, M. W. (2009). QTL analysis of seed iron, zinc, and phosphorus levels in an Andean bean population. *Crop Science, 49(5),* 1742-1750. 19

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(1-2),* 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.

Cook, J. P., McMullen, M. D., Holland, J. B., Tian, F., Bradbury, P., Ross-Ibarra, J., ... & Flint-Garcia, S. A. (2012). Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiology, 158(2),* 824-834.

Cooper, N. J., Wallace, C., Burren, O. S., Cutler, A., Walker, N., & Todd, J. A. (2017). Type 1 diabetes genome-wide association analysis with imputation identifies five new risk regions. *bioRxiv, 120022.*

Coppinger, R. J., & Diamond, A. M. (2001). Selenium deficiency and human disease. In Selenium (pp. 219-233). *Springer* US.

Costello, L. C., & Franklin, R. B. (2017). Decreased zinc in the development and progression of malignancy: an important common relationship and potential for prevention and treatment of carcinomas. *Expert opinion on therapeutic targets*, *21*(1), 51-66.

Cozzolino, D., & Moron, A. (2004). Exploring the use of near infrared reflectance spectroscopy (NIRS) to predict trace minerals in legumes. *Animal Feed Science and Technology, 111(1),* 161-173.

Crespo-Herrera, L. A., Govindan, V., Stangoulis, J., Hao, Y., & Singh, R. P. (2017). Qtl mapping of grain zn and fe concentrations in two hexaploid wheat ril populations with ample transgressive segregation. *Frontiers in plant science, 8*, 1800.

Davis, D. R., Epp, M. D., & Riordan, H. D. (2004). Changes in USDA food composition data for 43 garden crops, 1950 to 1999. *Journal of the american College of nutrition, 23*(6), 669-682.

DellaValle, D. M., Thavarajah, P., Vandenberg, A., & Glahn, R. P. (2013). Lentil (Lens culinaris L.) as a candidate crop for iron biofortification: Is there genetic potential for iron bioavailability?. *Field Crops Research, 144*, 119-125.

Deokar, A. A., Ramsay, L., Sharpe, A. G., Diapari, M., Sindhu, A., Bett, K., ... & Tar'an, B. (2014). Genome wide SNP identification in chickpea for use in development of a high density genetic map and improvement of chickpea reference genome assembly. *BMC genomics, 15*(1), 708.

Desgroux, A., L'anthoëne, V., Roux-Duparque, M., Rivière, J. P., Aubert, G., Tayeh, N., ... & McGee, R. J. (2016). Genome-wide association mapping of partial resistance to Aphanomyces euteiches in pea. *BMC genomics, 17(1)*, 124.

Deulvot, C., Charrel, H., Marty, A., Jacquin, F., Donnadieu, C., Lejeune-Hénaut, I., ... & Aubert, G. (2010). Highly-multiplexed SNP genotyping for genetic mapping and germplasm diversity studies in pea. *BMC genomics, 11*(1), 468.

Diapari, M., Sindhu, A., Bett, K., Deokar, A., Warkentin, T. D., & Tar'an, B. (2014). Genetic diversity and association mapping of iron and zinc concentrations in chickpea (Cicer arietinum L.). *Genome, 57*(8), 459-468.

Diapari, M., Sindhu, A., Warkentin, T. D., Bett, K., & Tar'an, B. (2015). Population structure and marker-trait association studies of iron, zinc and selenium concentrations in seed of field pea (Pisum sativum L.). *Molecular breeding, 35*(1), 30.

Dotaniya, M. L., & Meena, V. D. (2015). Rhizosphere effect on nutrient availability in soil and its uptake by plants: a review. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences, 85(1)*, 1-12.

Easton, D. F., Pooley, K. A., Dunning, A. M., Pharoah, P. D., Thompson, D., Ballinger, D. G., ... & Wareham, N. (2007). Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature, 447(7148)*, 1087-1093.

Ellis, THN., Turner, L., Hellens, R.P., Lee, D., Harker, C.L., Enars, C., Domoney, C., Davies, D.R. (1992) Linkage maps in pea. *Genetics 130:*649–663

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

Elwell, W. T., & Gidley, J. A. F. (2013). Atomic-Absorption Spectrophotometry: *International Series of Monographs in Analytical Chemistry (Vol. 6).* Elsevier.

Enamorado-Báez, S. M., Abril, J. M., & Gómez-Guzmán, J. M. (2013). Determination of 25 trace element concentrations in biological reference materials by ICP-MS following

different microwave-assisted acid digestion methods based on scaling masses of digested samples. *ISRN Analytical Chemistry*, 2013.

Fageria, N. K. (2016). The use of nutrients in crop plants. CRC press.

FAO (2010) Available from http://www.fao.org/ag/agn/nutrition/household_micronutrients_en.stm [Accessed on 5th April 2016]

FAOSTAT (2016) Available from http://faostat.fao.org [Accessed on 7th September 2018]

Ferrari, B., Romani, M., Aubert, G., Boucherot, K., Burstin, J., Pecetti, L., ... & Annicchiarico, P. (2016). Association of SNP markers with agronomic and quality traits of field pea in Italy. Czech J. Genet. *Plant Breed, 52(3),* 83-93.

French, R. J. (2016) Field Pea: Agronomy. *Reference Module in Food Science*.

Gali, K. K., Liu, Y., Sindhu, A., Diapari, M., Shunmugam, A. S., Arganosa, G., ... & Warkentin, T. D. (2018). Construction of high-density linkage maps for mapping quantitative trait loci for multiple traits in field pea (Pisum sativum L.). *BMC plant biology*, *18*(1), 172. Garcia-Oliveira, A. L., Tan, L., Fu, Y., & Sun, C. (2009). Genetic identification of quantitative trait loci for contents of mineral nutrients in rice grain. *Journal of Integrative Plant Biology, 51(1),* 84-92.

Gei, M. G., & Powers, J. S. (2015). The influence of seasonality and species effects on surface fine roots and nodulation in tropical legume tree plantations. *Plant and soil, 388(1-2),* 187-196.

George, A. W., & Cavanagh, C. (2015). Genome-wide association mapping in plants. *Theoretical and Applied Genetics*, *128*(6), 1163-1174.

Gilroy, S., & Jones, D. L. (2000). Through form to function: root hair development and nutrient uptake. *Trends in plant science, 5(2),* 56-60.

Gowda, M., Das, B., Makumbi, D., Babu, R., Semagn, K., Mahuku, G., ... & Prasanna, B. M. (2015). Genome-wide association and genomic prediction of resistance to maize lethal necrosis disease in tropical maize germplasm. *Theoretical and Applied Genetics, 128(10),* 1957-1968.

Graves, H., Rayburn, A. L., Gonzalez-Hernandez, J. L., Nah, G., Kim, D. S., & Lee, D. K. (2016). Validating DNA polymorphisms using KASP assay in prairie cordgrass (Spartina pectinata Link) populations in the US. *Frontiers in plant science*, *6*, 1271.

Gupta, A. K. (2004). Origin of agriculture and domestication of plants and animals linked to early Holocene climate amelioration. *CURRENT SCIENCE-BANGALORE-*, *87*, 54-59.

Gupta, M., & Gupta, S. (2017). An overview of selenium uptake, metabolism, and toxicity in plants. *Frontiers in plant science, 7,* 2074.

Gupta, N., Ram, H., & Kumar, B. (2016). Mechanism of Zinc absorption in plants: uptake, transport, translocation and accumulation. *Reviews in Environmental Science and Bio/Technology, 15(1),* 89-109.

Gupta, P. K., Rustgi, S., & Mir, R. R. (2008). Array-based high-throughput DNA markers for crop improvement. *Heredity, 101(1),* 5-18. 20

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

Ho, E. (2004). Zinc deficiency, DNA damage, and cancer risk. *The Journal of Nutritional Biochemistry, 15(10),* 572-578.

Huang, X., & Han, B. (2014). Natural variations and genome-wide association studies in crop plants. *Annual review of plant biology, 65,* 531-551.

Hussain, B., Lucas, S. J., Ozturk, L., & Budak, H. (2017). Mapping QTLs conferring salt tolerance and micronutrient concentrations at seedling stage in wheat. *Scientific reports, 7*(1), 15662.

Illumina, Inc. (2006) Available from https://www.illumina.com/documents/products/workflows/workflow_goldengate_assay.pdf [Accessed on 10th April 2018]

Jain, M., Misra, G., Patel, R. K., Priya, P., Jhanwar, S., Khan, A. W., ... & Yadav, M. (2013). A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). *The Plant Journal, 74(5),* 715-729.

Jha, A.B., Ashokkumar, K., Diapari, M., Ambrose, S.J., Zhang, H., Tar‟an, B., Bett, K.E., Vandenberg, A., Warkentin, T.D., & Purves, R.W. (2015). Genetic diversity of folate profiles in seeds of common bean, lentil, chickpea and pea. *Journal of Food Composition and Analysis 42*:134-140.

Jha, A. B., Gali, K. K., Tar'an, B., & Warkentin, T. D. (2017). Fine mapping of QTLs for ascochyta blight resistance in pea using heterogeneous inbred families. *Frontiers in plant science, 8,* 765.

Jha, A. B., Tar'an, B., Stonehouse, R., & Warkentin, T. D. (2016). Identification of QTLs associated with improved resistance to ascochyta blight in an interspecific pea recombinant inbred line population. *Crop Science, 56*(6), 2926-2939.

Jiang, G. L. (2013). Molecular markers and marker-assisted breeding in plants. *Plant Breeding from Laboratories to Fields* ed. Andersen, S. B. InTech (Available from https://www.intechopen.com/books/plant-breeding-from-laboratories-to-fields/molecular-markers-and-marker-assisted-breeding-in-plants).

Jobbágy, E. G., & Jackson, R. B. (2001). The distribution of soil nutrients with depth: global patterns and the imprint of plants. *Biogeochemistry, 53*(1), 51-77.

Kawai, S., Takagi, S. I., & Sato, Y. (1988). Mugineic acid-family phytosiderophores in root-secretions of barley, corn and sorghum varieties. *Journal of Plant Nutrition, 11*(6-11), 633-642.

Khazaei, H., Podder, R., Caron, C. T., Kundu, S. S., Diapari, M., Vandenberg, A., & Bett, K. E. (2017). Marker–trait association analysis of iron and zinc concentration in lentil (Lens culinaris Medik.) seeds. *The plant genome,* 10(2).

Kim, C., Guo, H., Kong, W., Chandnani, R., Shuang, L. S., & Paterson, A. H. (2016). Application of genotyping by sequencing technology to a variety of crop breeding programs. *Plant Science*, *242*, 14-22.

King, K. E., Peiffer, G. A., Reddy, M., Lauter, N., Lin, S. F., Cianzio, S., & Shoemaker, R. C. (2013). Mapping of iron and zinc quantitative trait loci in soybean for association to iron deficiency chlorosis resistance. *Journal of plant nutrition*, *36*(14), 2132-2153.

Klasen, J. R., Barbez, E., Meier, L., Meinshausen, N., Bühlmann, P., Koornneef, M., ... & Schneeberger, K. (2016). A multi-marker association method for genome-wide association studies without the need for population structure correction. *Nature communications*, 7.

Korte, A., & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant methods*, *9(1)*, 29.

Kujur, A., Bajaj, D., Upadhyaya, H. D., Das, S., Ranjan, R., Shree, T., ... & Gowda, C. L. L. (2015). Employing genome-wide SNP discovery and genotyping strategy to extrapolate the natural allelic diversity and domestication patterns in chickpea. *Frontiers in plant science*, *6*, 162.

Kumar, A., Choudhary, A. K., Pooniya, V., Suri, V. K., & Singh, U. (2016). Soil factors associated with micronutrient acquisition in crops-biofortification perspective. In *Biofortification of food crops* (pp. 159-176). Springer, New Delhi.

Kumar, J., Choudhary, A. K., Solanki, R. K., & Pratap, A. (2011). Towards marker-assisted selection in pulses: a review. *Plant Breeding*, *130*(3), 297-313.

Kwon, S. J., Brown, A.F., Hu, J., McGee, R., Watt, C., Kisha, T., Timmerman-Vaughan, G., Grusak, M., McPhee, K.E., & Coyne, C.J. (2012). Genetic diversity, population structure and genome-wide marker-trait association analysis emphasizing seed nutrients of the USDA pea (Pisum sativum L.) core collection. *Genes & Genomics. 1;34(3):*305-20.

Leonforte, A., Sudheesh, S., Cogan, N. O., Salisbury, P. A., Nicolas, M. E., Materne, M., ... & Kaur, S. (2013). SNP marker discovery, linkage map construction and identification of QTLs for enhanced salinity tolerance in field pea (Pisum sativum L.). *BMC plant biology*, *13(1)*, 161.

LGC Limited, Available from https://www.lgcgroup.com/kasp/ [Accessed on 5th July 2018]

Li, H., Peng, Z., Yang, X., Wang, W., Fu, J., Wang, J., ... & Liu, J. (2013). Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nature genetics*, *45(1)*, 43-50.

Li, Y. X., Li, C., Bradbury, P. J., Liu, X., Lu, F., Romay, C. M., ... & Song, Y. (2016). Identification of genetic variants associated with maize flowering time using an extremely large multi-genetic background population. *The Plant Journal*, *86(5)*, 391-402.

Lozoff, B., Clark, K. M., Jing, Y., Armony-Sivan, R., Angelilli, M. L., & Jacobson, S. W. (2008). Dose-response relationships between iron deficiency with or without anemia and infant social-emotional behavior. *The Journal of pediatrics*, *152*(5), 696-702.

Ma, H., Evangelou, M. W., Vontobel, P., & Schulin, R. (2016). Root growth and Zn uptake of three common crop plants in response to heterogeneity in soil texture and Zn distribution. *Environmental and Experimental Botany, 127*, 45-54.

Ma, Y., Coyne, C. J., Grusak, M. A., Mazourek, M., Cheng, P., Main, D., & McGee, R. J. (2017). Genome-wide SNP identification, linkage map construction and QTL mapping for seed mineral concentrations and contents in pea (*Pisum sativum* L.). *BMC plant biology, 17(1),* 43.

MacDonald, R. S. (2000). The role of zinc in growth and cell proliferation. *The Journal of Nutrition, 130(5),* 1500S-1508S.

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

Magallanes-López, A. M., Hernandez-Espinosa, N., Velu, G., Posadas-Romano, G., Ordoñez-Villegas, V. M. G., Crossa, J., Karim, A., & Guzmán, C. (2017). Variability in iron, zinc and phytic acid content in a worldwide collection of commercial durum wheat cultivars and the effect of reduced irrigation on these traits. *Food chemistry, 237*, 499-505.

Mammadov, J. A., Chen, W., Ren, R., Pai, R., Marchione, W., Yalçin, F., ... & Kumpatla, S. P. (2010). Development of highly polymorphic SNP markers from the complexity reduced portion of maize [Zea mays L.] genome for use in marker-assisted breeding. *Theoretical and applied genetics, 121*(3), 577-588.

Mao, H., Wang, H., Liu, S., Li, Z., Yang, X., Yan, J., ... & Qin, F. (2015). A transposable element in a NAC gene is associated with drought tolerance in maize seedlings. *Nature communications, 6*, 8326.

McCouch, S. R., Wright, M. H., Tung, C. W., Maron, L. G., McNally, K. L., Fitzgerald, M., ... & Greenberg, A. J. (2016). Open access resources for genome-wide association mapping in rice. *Nature communications, 7,* 10532.

McKay, K., Schatz, B., & Enders, G. (2003). *Field pea production*. North Dakota State University and US Department of Agriculture.

McLaren, T. I., Guppy, C. N., & Tighe, M. K. (2012). A rapid and nondestructive plant nutrient analysis using portable x-ray fluorescence. *Soil Science Society of America Journal, 76*(4), 1446-1453..

Meenakshi, J.V., Johnson, N.L., Manyong, V.M., DeGroote, H., Javelosa, J., Yanggen, D.R., Naher, F., Gonzalez, C., Garcia, J., & Meng, E. (2010) *World Development 38*, 64-75

Morgounov, A., Gómez-Becerra, H. F., Abugalieva, A., Dzhunusova, M., Yessimbekova, M., Muminjanov, H., Zelenskiy, Y., Ozturk, L., & Cakmak, I. (2007). Iron and zinc grain density in common wheat grown in Central Asia. *Euphytica, 155*(1-2), 193-203.

Morrissey, J., & Guerinot, M. L. (2009). Iron uptake and transport in plants: the good, the bad, and the ionome. *Chemical reviews, 109(10),* 4553-4567.

Nalls, M. A., Pankratz, N., Lill, C. M., Do, C. B., Hernandez, D. G., Saad, M., ... & Schulte, C. (2014). Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nature genetics, 46(9),* 989-993.

National Climate Report, USA 2014 Avilable from https://www.ncdc.noaa.gov [Accessed on 10th July 2018]

NIH Available from https://www.nih.gov [Accessed on 18th April 2016]

Oliveira, N. P., Faquin, V., Costa, A. L. D., Livramento, K. G. D., Pinho, P. J. D., & Guilherme, L. R. G. (2016). Genotypic variation of agronomic traits as well as concentrations of Fe, Zn, P and phytate in soybean cultivars. *Revista Ceres*, *63*(3), 403-411.

Olukolu, B. A., Tracy, W. F., Wisser, R., De Vries, B., & Balint-Kurti, P. J. (2016). A genome-wide association study for partial resistance to maize common rust. *Phytopathology, 106(7),* 745-751.

Paltridge, N. G., Palmer, L. J., Milham, P. J., Guild, G. E., & Stangoulis, J. C. (2012). Energy-dispersive X-ray fluorescence analysis of zinc and iron concentration in rice and pearl millet grain. *Plant and Soil, 361(1-2),* 251-260.

Patterson, C. A., Maskus, H., & Dupasquier, C. (2009). Pulse crops for health. *Cereal Foods World (CFW), 54(3),* 108.

Perseguini, J. M. K. C., Oblessuc, P. R., Rosa, J. R. B. F., Gomes, K. A., Chiorato, A. F., Carbonell, S. A. M., ... & Benchimol-Reis, L. L. (2016). Genome-wide association studies of anthracnose and angular leaf spot resistance in common bean (*Phaseolus vulgaris* L.). *PloS one, 11(3),* e0150506.

Pfeiffer, W. H., & McClafferty, B. (2007). HarvestPlus: breeding crops for better nutrition. *Crop Science*, *47*(Supplement_3), S-88.

Phillips, D. A. (1980). Efficiency of symbiotic nitrogen fixation in legumes. *Annual Review of Plant Physiology*, *31*(1), 29-49.

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PloS one*, *7*(2), e32253.h

Prasad, A. S. (1996). Zinc deficiency in women, infants, and children. *Journal of the American College of Nutrition, 15(2),* 113-120. 21

Prasad, A. S., Bao, B., Beck, F. W., Kucuk, O., & Sarkar, F. H. (2004). Antioxidant effect of zinc in humans. *Free Radical Biology and Medicine, 37(8),* 1182-1190.

Rakitsch, B., Lippert, C., Stegle, O., & Borgwardt, K. (2012). A Lasso multi-marker mixed model for association mapping with population structure correction. *Bioinformatics*, *29*(2), 206-214.

Ray, H., Bett, K., Tar‟an, B., Vandenberg, A., Thavarajah, D., & Warkentin, T.D. (2014) *Crop Science 54,* 1698-1708

Rayman, M. P. (2005). Selenium in cancer prevention: a review of the evidence and mechanism of action. *Proceedings of the Nutrition Society, 64(04),* 527-542.

Retka, J., Maksymowicz, A., & Karmasz, D. (2010). Determination of Cu, Ni, Zn, Pb, Cd by ICP-MS and Hg by AAS in Plant Samples. Department of Analytical Chemistry, Chemical Faculty, Gdansk University of Technology, september, 19-23.

Rogers, E. D., & Benfey, P. N. (2015). Regulation of plant root system architecture: implications for crop advancement. *Current opinion in biotechnology, 32,* 93-98.

Roohani, N., Hurrell, R., Kelishadi, R., & Schulin, R. (2013). Zinc and its importance for human health: An integrative review. *Journal of Research in Medical Sciences, 18(2),* 144-157.

Rostan, E. F., DeBuys, H. V., Madey, D. L., & Pinnell, S. R. (2002). Evidence supporting zinc as an important antioxidant for skin. *International Journal of Dermatology, 41(9),* 606-611.

Roy, F., Boye, J. I., & Simpson, B. K. (2010). Bioactive proteins and peptides in pulse crops: Pea, chickpea, and lentil. *Food Research International, 43(2),* 432-442.

RSC Available from http://www.liskeard.cornwall.sch.uk/images/Liskeard-Sixth-Form/Atomic-Absorption-Spectrometry.pdf [Accessed on 12th April 2016]

Schenkeveld, W. D. C., Oburger, E., Gruber, B., Schindlegger, Y., Hann, S., Puschenreiter, M., & Kraemer, S. M. (2014). Metal mobilization from soils by phytosiderophores–experiment and equilibrium modeling. *Plant and soil*, *383*(1-2), 59-71.

Selby-Pham, J., Lutz, A., Moreno-Moyano, L. T., Boughton, B. A., Roessner, U., & Johnson, A. A. (2017). Diurnal changes in transcript and metabolite levels during the iron deficiency response of rice. *Rice*, *10*(1), 14.

Semagn, K., Babu, R., Hearne, S., & Olsen, M. (2014). Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. *Molecular Breeding*, *33*(1), 1-14.

Shenkin, A., (2006). The key role of micronutrients. *Clinical Nutrition, 25(1),* pp.1-13.

Singh, U., Praharaj, C. S., Singh, S. S., Bohra, A., & Shivay, Y. S. (2015). Biofortification of Pulses: Strategies and Challenges. *In Proceedings of The Second International Conference on Bio-Resource and Stress Management, Hyderabad, India* (pp. 50-55).

Singh, U., Williams, P. C., & Petterson, D. S. (2000). Processing and grain quality to meet market demands. In *Linking Research and Marketing Opportunities for Pulses in the 21st Century* (pp. 155-166). Springer, Dordrecht.

Singh, V. K., Khan, A. W., Jaganathan, D., Thudi, M., Roorkiwal, M., Takagi, H., ... & Sutton, T. (2016). QTL-seq for rapid identification of candidate genes for 100-seed weight and root/total plant dry weight ratio under rainfed conditions in chickpea. *Plant biotechnology journal, 14*(11), 2110-2119.

Smartt, J. (1984). Evolution of grain legumes. I. Mediterranean pulses. *Experimental Agriculture*, *20*(4), 275-296.

Smýkal, P., Aubert, G., Burstin, J., Coyne, C. J., Ellis, N. T., Flavell, A. J., & McPhee, K. E. (2012). Pea (Pisum sativum L.) in the genomic era. *Agronomy, 2(2),* 74-115.

Smýkal, P., Šafářová, D., Navrátil, M., & Dostalová, R. (2010). Marker assisted pea breeding: eIF4E allele specific markers to pea seed-borne mosaic virus (PSbMV) resistance. *Molecular breeding, 26*(3), 425-438.

Smýkal, P., Varshney, R. K., Singh, V. K., Coyne, C. J., Domoney, C., Kejnovský, E., & Warkentin, T. (2016). From Mendel‟s discovery on pea to today‟s plant genetics and breeding. *Theoretical and Applied Genetics, 129(12),* 2267-2280.

Sonah, H., Bastien, M., Iquira, E., Tardivel, A., Légaré, G., Boyle, B., ... & Belzile, F. (2013). An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. *PloS one, 8*(1), e54603.

Stangoulis, J. C. R., Huynh, B. L., Welch, R. M., Choi, E. Y., & Graham, R. D. (2007). Quantitative trait loci for phytate in rice grain and their relationship with grain micronutrient content. *Euphytica, 154(3),* 289-294.

Statistics Canada (2017) available on http://www.statcan.gc.ca/daily-quotidien/170421/dq170421b-eng.pdf [Accessed on 22nd of May 2018]

Stich, B., & Melchinger, A. E. (2010). An introduction to association mapping in plants. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources, 5, 1-9*. 22

Stuth, J., Jama, A., & Tolleson, D. (2003). Direct and indirect means of predicting forage quality through near infrared reflectance spectroscopy. *Field Crops Research, 84(1),* 45-56.

Sudheesh, S., Lombardi, M., Leonforte, A., Cogan, N. O., Materne, M., Forster, J. W., & Kaur, S. (2015). Consensus genetic map construction for field pea (Pisum sativum L.), trait dissection of biotic and abiotic stress tolerance and development of a diagnostic marker for the er1 powdery mildew resistance gene. *Plant molecular biology reporter, 33(5),* 1391-1403.

Sukumaran, S., Yu, J., (2013) Association mapping of genetic resources. *Genomics of plant genetic resources, 1,* 209-210

Tar'an, B., Warkentin, T., Somers, D. J., Miranda, D., Vandenberg, A., Blade, S., ... & 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*(8), 1482-1491.

Tar'an, B., Warkentin, T., Somers, D. J., Miranda, D., Vandenberg, A., Blade, S., & Bing, D. (2004). Identification of quantitative trait loci for grain yield, seed protein concentration and maturity in field pea (Pisum sativum L.). *Euphytica, 136*(3), 297-306.

Tayeh, N., Aubert, G., Pilet-Nayel, M. L., Lejeune-Hénaut, I., Warkentin, T. D., & Burstin, J. (2015). Genomic tools in pea breeding programs: status and perspectives. *Frontiers in Plant Science, 6,* 1037.

Teumer, A., Tin, A., Sorice, R., Gorski, M., Yeo, N. C., Chu, A. Y., ... & Taliun, D. (2015). Genome-wide association studies identify genetic loci associated with albuminuria in diabetes. *Diabetes,* db151313.

Thavarajah, D., Vandenberg, A., George, G.N., & Pickering, I.J., (2007) *J Agric Food Chem 55*:7337–7341

Thavarajah, D., Warkentin, T., & Vandenberg, A. (2010). Natural enrichment of selenium in Saskatchewan field peas (Pisum sativum L.). Canadian Journal of Plant Science, 90(4), 383-389.

Thavarajah, P. (2012). Evaluation of chickpea (Cicer arietinum L.) micronutrient composition: Biofortification opportunities to combat global micronutrient malnutrition. *Food research international*, *49*(1), 99-104.

Thavarajah, P., Wejesuriya, A., Rutzke, M., Glahn, R. P., Combs, G. F., & Vandenberg, A. (2011). The potential of lentil (Lens culinaris L.) as a whole food for increased selenium, iron, and zinc intake: preliminary results from a 3-year study. *Euphytica, 180(1),* 123-128.

Thudi, M., Upadhyaya, H. D., Rathore, A., Gaur, P. M., Krishnamurthy, L., Roorkiwal, M., ... & Fikre, A. (2014). Genetic dissection of drought and heat tolerance in chickpea through genome-wide and candidate gene-based association mapping approaches. *Plos one, 9(5),* e96758.

Timmerman-Vaughan, G. M., Moya, L., Frew, T. J., Murray, S. R., & Crowhurst, R. (2016). Ascochyta blight disease of pea (Pisum sativum L.): defence-related candidate genes associated with QTL regions and identification of epistatic QTL. *Theoretical and applied genetics, 129(5),* 879-896.

Tinggi, U. (2008). Selenium: its role as antioxidant in human health. *Environmental Health and Preventive Medicine, 13(2),* 102-108.

Tiwari, V.K., Rawat, N., Chhuneja, P., Neelam, K., Aggarwal, R., Randhawa, G.S., Dhaliwal, H.S., Keller, B., & Singh, K. (2009). Mapping of quantitative trait loci for grain iron and zinc concentration in diploid A genome wheat. *Journal of Heredity.* esp030.

Trebbi, D., Maccaferri, M., de Heer, P., Sørensen, A., Giuliani, S., Salvi, S., ... & Tuberosa, R. (2011). High-throughput SNP discovery and genotyping in durum wheat (Triticum durum Desf.). *Theoretical and applied genetics*, *123*(4), 555-569.

Tulchinsky, T. H. (2010). Micronutrient deficiency conditions: global health issues. *Public Health Reviews*, *32*(1), 243.

Ubayasena, L., Bett, K., Tar'an, B., & Warkentin, T. (2011). Genetic control and identification of QTLs associated with visual quality traits of field pea (Pisum sativum L.). *Genome*, *54*(4), 261-272.

Upadhyaya, H. D., Bajaj, D., Das, S., Kumar, V., Gowda, C. L. L., Sharma, S., ... & Parida, S. K. (2016). Genetic dissection of seed-iron and zinc concentrations in chickpea. *Scientific reports,* 6.

Vandemark, G. J., Grusak, M. A., & McGee, R. J. (2018). Mineral concentrations of chickpea and lentil cultivars and breeding lines grown in the US Pacific Northwest. *The Crop Journal, 6*(3), 253-262.

Varshney, R. K., Song, C., Saxena, R. K., Azam, S., Yu, S., Sharpe, A. G., ... & Millan, T. (2013). Draft genome sequence of chickpea (Cicer arietinum) provides a resource for trait improvement. *Nature biotechnology, 31(3),* 240-246.

Vos, P. G., Paulo, M. J., Voorrips, R. E., Visser, R. G., van Eck, H. J., & van Eeuwijk, F. A. (2017). Evaluation of LD decay and various LD-decay estimators in simulated and SNP-array data of tetraploid potato. *Theoretical and Applied Genetics, 130*(1), 123-135.

Wakeel, A., Farooq, M., Bashir, K., & Ozturk, L. (2018). Micronutrient Malnutrition and Biofortification: Recent Advances and Future Perspectives. In *Plant Micronutrient Use Efficiency* (pp. 225-243).

Wang, N., Wang, Z. P., Liang, X. L., Weng, J. F., Lv, X. L., Zhang, D. G., ... & Jiang, L. Y. (2016). Identification of loci contributing to maize drought tolerance in a genome-wide association study. *Euphytica, 210(2),* 165-179.

Warkentin, T. D., Sloan, A. G., & Ali-Khan, S. T. (1997). Proximate and mineral composition of field peas. *Canadian Journal of Plant Science, 77(1),* 101-103.

Warkentin, T.D., Smykal, P., Coyne, C.J., Weeden, N., Domoney, C., Bing, D., Leonforte, A., Xuxiao, Z., Dixit, G., Boros, L., McPhee, K., McGee, R., Burstin, J., and Ellis, N. (2015) Pea (Pisum sativum L.). pp. 37-83, In: Antonio M. De Ron (Ed.) *Grain Legumes, Series Handbook of Plant Breeding.* Springer Science+Business Media, New York.

Waters, B. M., & Grusak, M. A. (2008). Quantitative trait locus mapping for seed mineral concentrations in two Arabidopsis thaliana recombinant inbred populations. *New Phytologist, 179,* 1033-1047.

Weeden, N. F., Swiecicki, W. K., Timmerman-Vaughan, G. M., Ellis, T. H. N., & Ambrose, M. (1996). The current pea linkage map. *Pisum Genet, 28,* 1-4.

Wen, Z., Tan, R., Yuan, J., Bales, C., Du, W., Zhang, S., .., & Wang, D. (2014). Genome-wide association mapping of quantitative resistance to sudden death syndrome in soybean. *BMC genomics, 15(1),* 809.

White, P. J., Bradshaw, J. E., Finlay, M., Dale, B., Ramsay, G., Hammond, J. P., & Broadley, M. R. (2009). Relationships between yield and mineral concentrations in potato tubers. *HortScience, 44*(1), 6-11.

White, P. J., & Broadley, M. R. (2011). Physiological limits to zinc biofortification of edible crops. *Frontiers in Plant Science, 2,* 80. Wiltgren, A. R., Booth, A. O., Kaur, G., Cicerale, S., Lacy, K. E., Thorpe, M. G., ... & Riddell, L. J. (2015). Micronutrient supplement use and diet quality in university students. Nutrients, 7(2), 1094-1107.

Wong, M. M., Gujaria-Verma, N., Ramsay, L., Yuan, H. Y., Caron, C., Diapari, M., ... & Bett, K. E. (2015). Classification and characterization of species within the genus Lens using genotyping-by-sequencing (GBS). *PLoS One, 10*(3), e0122025.

World Health Organization, (n.d.) Available on http://www.who.int/nutrition/topics/micronutrients/en/ [Accessed on 20th April 2016]

Wu, C., DeWan, A., Hoh, J., & Wang, Z. (2011). A comparison of association methods correcting for population stratification in case–control studies. *Annals of human genetics, 75(3),* 418-427.

Xu, Y., Li, P., Yang, Z., & Xu, C., (2017) Genetic mapping of quantitative trait loci in crops. *The Crop Journal, 5(2),* 175-184.

Yan, J., Yang, X., Shah, T., Sánchez-Villeda, H., Li, J., Warburton, M., ... & Xu, Y. (2010). High-throughput SNP genotyping with the GoldenGate assay in maize. *Molecular Breeding*, *25*(3), 441-451.

Zargar, S. M., Raatz, B., Sonah, H., Bhat, J. A., Dar, Z. A., Agrawal, G. K., & Rakwal, R. (2015). Recent advances in molecular marker techniques: insight into QTL mapping, GWAS and genomic selection in plants. *Journal of Crop Science and Biotechnology, 18(5),* 293-308.

Zeng, H., & Combs, G. F. (2008). Selenium as an anticancer nutrient: roles in cell proliferation and tumor cell invasion. *The Journal of Nutritional Biochemistry, 19(1),* 1-7.

Zhu, C., Gore, M., Buckler, E. S., & Yu, J. (2008). Status and prospects of association mapping in plants. *The Plant Genome, 1(1),* 5-20

Appendix

Appendix A: Pea genome wide association study (GWAS) panel from University of Saskatchewan pea breeding program

Entry	Accession name	Seed source	Country of origin
1	PS05ND0232	NDSU, Fargo, ND	USA
2	PS07ND0164	NDSU, Fargo, ND	USA
3	PS07ND0190	NDSU, Fargo, ND	USA
4	NDP080111	NDSU, Fargo, ND	USA
5	NDP080138	NDSU, Fargo, ND	USA
6	Bohatyr	NDSU, Fargo, ND	Czech Republic
7	Shawnee	NDSU, Fargo, ND	USA
8	PS05ND327	NDSU, Fargo, ND	USA
9	PS05ND330	NDSU, Fargo, ND	USA
10	PS05ND0434	NDSU, Fargo, ND	USA
11	Audit	Limagrain, Netherlands	Netherlands
12	Avantgarde	Limagrain, Netherlands	Netherlands
13	Lasso	Limagrain, Netherlands	Netherlands
14	Quadril	Limagrain, Netherlands	Netherlands
15	Rebel	Limagrain, Netherlands	Netherlands
16	Satelit	Limagrain, Netherlands	Netherlands
17	Nette	Limagrain, Netherlands	Netherlands
18	Emerald	Limagrain, Netherlands	Netherlands
19	Evergreen	Limagrain, Netherlands	Netherlands
20	Abarth	Limagrain, Netherlands	Netherlands
21	Sorento	Limagrain, Netherlands	Netherlands
22	Neon	Limagrain, Netherlands	Netherlands
23	Strada	Limagrain, Netherlands	Netherlands
24	Prophet	Limagrain, Netherlands	Netherlands
25	Garde	Limagrain, Netherlands	Netherlands
26	Camry	Limagrain, Netherlands	Netherlands
27	Matrix	Limagrain, Netherlands	Netherlands
28	CEB-Montech 4152	Limagrain, Netherlands	Netherlands
29	Spider	Limagrain, Netherlands	Netherlands
30	Aukland	Limagrain, Netherlands	Netherlands
31	Lifter	USDA, Pullman, WA	USA
32	Serge	USDA, Pullman, WA	USA
33	Medora	USDA, Pullman, WA	USA
34	Melrose	USDA, Pullman, WA	USA
35	Dove HR	USDA, Pullman, WA	USA
36	GRAY'S	USDA, Pullman, WA	USA
37	No. 9292	USDA, Pullman, WA	USA
38	G 9173	USDA, Pullman, WA	USA
39	No. 8120	USDA, Pullman, WA	USA
40	AMPLISSIMO ZAZERSKIJ	USDA, Pullman, WA	Russia
41	PLP 105A	USDA, Pullman, WA	India

42	22778	USDA, Pullman, WA	Ethiopia
43	22791	USDA, Pullman, WA	Ethiopia
44	Klatovsky zeleny	Palacký University, Czech Republic	Czech Republic
45	Milion zeleny	Palacký University, Czech Republic	Czech Republic
46	Slovensky expres	Palacký University, Czech Republic	Czech Republic
47	Moravsky Hrotovicky krajovy	Palacký University, Czech Republic	Czech Republic
48	Dalibor	Palacký University, Czech Republic	Czech Republic
49	Dick Trom	Palacký University, Czech Republic	Czech Republic
50	Prebohatyr	Palacký University, Czech Republic	Czech Republic
51	Odeon	Palacký University, Czech Republic	Czech Republic
52	Kamelot	Palacký University, Czech Republic	Czech Republic
53	B 99/108	Palacký University, Czech Republic	Czech Republic
54	Sponsor	Palacký University, Czech Republic	Czech Republic
55	Terno	Palacký University, Czech Republic	Czech Republic
56	Moravska krajova	Palacký University, Czech Republic	Czech Republic
57	Purpurviolett Schottige Nero	Palacký University, Czech Republic	Czech Republic
58	Stupicka jarni	Palacký University, Czech Republic	Czech Republic
59	Kapucin	Palacký University, Czech Republic	Czech Republic
60	Carouby de Maussane	INRA, Dijon, France	France
61	Champagne	INRA, Dijon, France	France
62	Chemin Long	INRA, Dijon, France	France
63	Cote D'or	INRA, Dijon, France	France
64	Fin de la Bievre	INRA, Dijon, France	France
65	Gloire de Correze	INRA, Dijon, France	France
66	Merveille D'etampes	INRA, Dijon, France	France
67	Normand	INRA, Dijon, France	France
68	Picar	INRA, Dijon, France	France
69	Piver	INRA, Dijon, France	France
70	Serpette D'auvergne	INRA, Dijon, France	France
71	Morgan	DPI, Horsham, Australia	Australia
72	EXCELL	DPI, Horsham, Australia	Australia
73	KASPA	DPI, Horsham, Australia	Australia
74	OZP0805	DPI, Horsham, Australia	Australia
75	OZP0819	DPI, Horsham, Australia	Australia
76	OZP0902	DPI, Horsham, Australia	Australia
77	OZP0903	DPI, Horsham, Australia	Australia
78	OZP1001	DPI, Horsham, Australia	Australia
79	OZP1002	DPI, Horsham, Australia	Australia
80	OZP1101	DPI, Horsham, Australia	Australia
81	OZP1004	DPI, Horsham, Australia	Australia
82	OZP1103	DPI, Horsham, Australia	Australia
83	OZP1104	DPI, Horsham, Australia	Australia
84	OZP1102	DPI, Horsham, Australia	Australia
85	PARAFIELD	DPI, Horsham, Australia	Australia
86	PBA GUNYAH	DPI, Horsham, Australia	Australia

87	PBA OURA	DPI, Horsham, Australia	Australia
88	PBA PERCY	DPI, Horsham, Australia	Australia
89	PBA TWILIGHT	DPI, Horsham, Australia	Australia
90	STURT	DPI, Horsham, Australia	Australia
91	BK12	John Innes Centre, Norwich, UK	UK
92	BK72	John Innes Centre, Norwich, UK	UK
93	BK83	John Innes Centre, Norwich, UK	UK
94	BK88	John Innes Centre, Norwich, UK	UK
95	KB115	John Innes Centre, Norwich, UK	UK
96	KB117	John Innes Centre, Norwich, UK	UK
97	KB121	John Innes Centre, Norwich, UK	UK
98	KB124	John Innes Centre, Norwich, UK	UK
99	KB138	John Innes Centre, Norwich, UK	UK
100	KB139	John Innes Centre, Norwich, UK	UK
101	KB144	John Innes Centre, Norwich, UK	UK
102	KB151	John Innes Centre, Norwich, UK	UK
103	KB168	John Innes Centre, Norwich, UK	UK
104	BE15	John Innes Centre, Norwich, UK	UK
105	BE47	John Innes Centre, Norwich, UK	UK
106	BE99	John Innes Centre, Norwich, UK	UK
107	EB125	John Innes Centre, Norwich, UK	UK
108	EB157	John Innes Centre, Norwich, UK	UK
109	EB173	John Innes Centre, Norwich, UK	UK
110	EB188	John Innes Centre, Norwich, UK	UK
111	EB192	John Innes Centre, Norwich, UK	UK
112	EB193	John Innes Centre, Norwich, UK	UK
113	EB208	John Innes Centre, Norwich, UK	UK
114	BE218	John Innes Centre, Norwich, UK	UK
115	BE220	John Innes Centre, Norwich, UK	UK
116	EK3	John Innes Centre, Norwich, UK	UK
117	EK9	John Innes Centre, Norwich, UK	UK
118	EK11	John Innes Centre, Norwich, UK	UK
119	EK25	John Innes Centre, Norwich, UK	UK
120	EK34	John Innes Centre, Norwich, UK	UK
121	EK69	John Innes Centre, Norwich, UK	UK
122	EK73	John Innes Centre, Norwich, UK	UK
123	EK80	John Innes Centre, Norwich, UK	UK
124	KE106	John Innes Centre, Norwich, UK	UK
125	KE109	John Innes Centre, Norwich, UK	UK
126	KE134	John Innes Centre, Norwich, UK	UK
127	KE160	John Innes Centre, Norwich, UK	UK
128	KE206	John Innes Centre, Norwich, UK	UK
129	Kahuna-PGRO	John Innes Centre, Norwich, UK	UK
130	Enigma-PGRO	John Innes Centre, Norwich, UK	UK
131	EB181	John Innes Centre, Norwich, UK	UK
132	Brutus	John Innes Centre, Norwich, UK	UK

133	Kahuna-NIAB	John Innes Centre, Norwich, UK	UK
134	Enigma-NIAB	John Innes Centre, Norwich, UK	UK
135	Cameor	INRA, Dijon, France	France
136	TMP 15116	CDC, University of Saskatchewan	Ukraine
137	TMP 15133	CDC, University of Saskatchewan	Lithuania
138	TMP 15159	CDC, University of Saskatchewan	Russia
139	TMP 15202	CDC, University of Saskatchewan	Russia
140	TMP 15206	CDC, University of Saskatchewan	Russia
141	MPG87	CDC, University of Saskatchewan	Canada
142	TMP 15213	CDC, University of Saskatchewan	Belarus
143	Cutlass	CDC, University of Saskatchewan	Canada
144	CDC Bronco	CDC, University of Saskatchewan	Canada
145	CDC Centennial	CDC, University of Saskatchewan	Canada
146	CDC Golden	CDC, University of Saskatchewan	Canada
147	CDC Meadow	CDC, University of Saskatchewan	Canada
148	DS Admiral	CDC, University of Saskatchewan	Denmark
149	Eclipse	CDC, University of Saskatchewan	Netherlands
150	CDC Striker	CDC, University of Saskatchewan	Canada
151	Cooper	CDC, University of Saskatchewan	Netherlands
152	Nitouche	CDC, University of Saskatchewan	Denmark
153	Matar	CDC, University of Saskatchewan	India
154	Carneval	CDC, University of Saskatchewan	Sweden
155	MP1401	CDC, University of Saskatchewan	Canada
156	Orb	CDC, University of Saskatchewan	UK
157	Alfetta	CDC, University of Saskatchewan	Netherlands
158	CDC Sage	CDC, University of Saskatchewan	Canada
159	Espace	CDC, University of Saskatchewan	Netherlands
160	Terese	CDC, University of Saskatchewan	France
161	Torsdag	CDC, University of Saskatchewan	Russia
162	Delta	CDC, University of Saskatchewan	Netherlands
163	CDC Acer	CDC, University of Saskatchewan	Canada
164	Naparnyk	CDC, University of Saskatchewan	Russia
165	Trapper	CDC, University of Saskatchewan	Canada
166	Radley	CDC, University of Saskatchewan	UK
167	CDC Vienna	CDC, University of Saskatchewan	Canada
168	Highlight	CDC, University of Saskatchewan	Sweden
169	CDC 1-150-81	CDC, University of Saskatchewan	Canada
170	CDC 1-2347-144	CDC, University of Saskatchewan	Canada
171	Agassiz	CDC, University of Saskatchewan	Canada
172	Hardy	CDC, University of Saskatchewan	France
173	Cartouche	CDC, University of Saskatchewan	France
174	Superscout	CDC, University of Saskatchewan	USA
175	Lido	CDC, University of Saskatchewan	Denmark
176	Aragorn	CDC, University of Saskatchewan	USA
177	CDC Dakota	CDC, University of Saskatchewan	Canada

Appendix B: Preliminary study to determine appropriate sample size for AAS evaluations

Methodology:

In order to determine the appropriate sample size for AAS evaluations, seed samples of three selected cultivars, i.e., CDC Striker, CDC Dakota, and CDC Meadow, obtained from two locations in 2016, Rosthern and Saskatoon, with two biological replicates per location were analyzed in three sample sizes of 10 seeds, 60 seeds, and 120 seeds following the methodology described in Chapter 3. Factorial analysis of variance was performed for obtained concentration using a general linear model in SAS 9.3.

Results:

Table B.1: Fe, Zn, and Se concentration (µg/g) in solid material for all the samples analyzed in the preliminary study

No	Accession	Location	Replicate	sample size	Technical replicate	Se (µg/g) in solid material	Zn (µg/g) in solid material	Fe (µg/g) in solid material
1	CDC Striker	Rosthern	R1	10	T1	0.761	32.49	47.48
2	CDC Striker	Rosthern	R1	10	T2	0.711	33.87	47.86
3	CDC Striker	Rosthern	R1	10	T3	0.717	33.02	44.16
4	CDC Striker	Rosthern	R1	60	T1	0.549	31.96	46.23
5	CDC Striker	Rosthern	R1	60	T2	0.537	32.04	46.51
6	CDC Striker	Rosthern	R1	60	T3	0.535	32.17	46.27
7	CDC Striker	Rosthern	R1	120	T1	0.569	33.89	52.42
8	CDC Striker	Rosthern	R1	120	T2	0.510	33.45	54.31
9	CDC Striker	Rosthern	R1	120	T3	0.545	32.98	48.58
10	CDC Striker	Rosthern	R2	10	T1	0.645	29.10	45.23
11	CDC Striker	Rosthern	R2	10	T2	0.657	29.41	44.79
12	CDC Striker	Rosthern	R2	10	T3	0.650	30.17	46.50
13	CDC Striker	Rosthern	R2	60	T1	0.752	28.98	45.93
14	CDC Striker	Rosthern	R2	60	T2	0.780	29.11	48.03
15	CDC Striker	Rosthern	R2	60	T3	0.765	29.66	49.43
16	CDC Striker	Rosthern	R2	120	T1	0.754	30.79	51.40
17	CDC Striker	Rosthern	R2	120	T2	0.710	29.71	46.78
18	CDC Striker	Rosthern	R2	120	T3	0.748	30.29	47.19
19	CDC Striker	Saskatoon	R1	10	T1	0.585	30.43	43.68
20	CDC Striker	Saskatoon	R1	10	T2	0.577	30.42	44.73
21	CDC Striker	Saskatoon	R1	10	T3	0.581	29.85	43.33
22	CDC Striker	Saskatoon	R1	60	T1	0.609	27.74	42.48
23	CDC Striker	Saskatoon	R1	60	T2	0.643	28.18	49.05
24	CDC Striker	Saskatoon	R1	60	T3	0.622	27.90	46.68
25	CDC Striker	Saskatoon	R1	120	T1	0.657	30.29	46.12

26	CDC Striker	Saskatoon	R1	120	T2	0.649	30.78	53.53
27	CDC Striker	Saskatoon	R1	120	T3	0.657	30.67	45.72
28	CDC Striker	Saskatoon	R2	10	T1	0.685	27.40	47.44
29	CDC Striker	Saskatoon	R2	10	T2	0.704	27.89	49.69
30	CDC Striker	Saskatoon	R2	10	T3	0.782	27.58	48.94
31	CDC Striker	Saskatoon	R2	60	T1	0.823	26.85	41.44
32	CDC Striker	Saskatoon	R2	60	T2	0.742	26.04	47.03
33	CDC Striker	Saskatoon	R2	60	T3	0.761	27.28	45.97
34	CDC Striker	Saskatoon	R2	120	T1	0.770	29.06	46.80
35	CDC Striker	Saskatoon	R2	120	T2	0.747	28.14	46.85
36	CDC Striker	Saskatoon	R2	120	T3	0.761	28.74	48.08
37	CDC Dakota	Rosthern	R1	10	T1	0.416	29.54	50.59
38	CDC Dakota	Rosthern	R1	10	T2	0.397	28.24	46.16
39	CDC Dakota	Rosthern	R1	10	T3	0.403	28.99	49.93
40	CDC Dakota	Rosthern	R1	60	T1	0.386	33.43	51.88
41	CDC Dakota	Rosthern	R1	60	T2	0.360	34.71	52.12
42	CDC Dakota	Rosthern	R1	60	T3	0.365	33.93	54.68
43	CDC Dakota	Rosthern	R1	120	T1	0.452	35.23	56.77
44	CDC Dakota	Rosthern	R1	120	T2	0.441	32.60	51.80
45	CDC Dakota	Rosthern	R1	120	T3	0.453	33.06	53.94
46	CDC Dakota	Rosthern	R2	10	T1	0.140	32.23	48.68
47	CDC Dakota	Rosthern	R2	10	T2	0.144	34.24	49.89
48	CDC Dakota	Rosthern	R2	10	T3	0.141	34.07	53.11
49	CDC Dakota	Rosthern	R2	60	T1	0.251	29.18	46.79
50	CDC Dakota	Rosthern	R2	60	T2	0.290	30.02	46.27
51	CDC Dakota	Rosthern	R2	60	T3	0.263	30.16	50.10
52	CDC Dakota	Rosthern	R2	120	T1	0.254	31.78	48.25
53	CDC Dakota	Rosthern	R2	120	T2	0.270	32.50	48.49
54	CDC Dakota	Rosthern	R2	120	T3	0.248	32.48	48.61
55	CDC Dakota	Saskatoon	R1	10	T1	3.475	26.82	49.74
56	CDC Dakota	Saskatoon	R1	10	T2	3.523	26.16	45.78
57	CDC Dakota	Saskatoon	R1	10	T3	3.610	25.31	46.26
58	CDC Dakota	Saskatoon	R1	60	T1	3.615	28.91	50.72
59	CDC Dakota	Saskatoon	R1	60	T2	2.791	26.68	54.16
60	CDC Dakota	Saskatoon	R1	60	T3	2.835	28.21	51.90
61	CDC Dakota	Saskatoon	R1	120	T1	2.815	25.83	49.57
62	CDC Dakota	Saskatoon	R1	120	T2	3.081	26.06	39.27
63	CDC Dakota	Saskatoon	R1	120	T3	2.951	28.00	56.63
64	CDC Dakota	Saskatoon	R2	10	T1	0.870	27.47	50.25
65	CDC Dakota	Saskatoon	R2	10	T2	0.860	27.20	49.21
66	CDC Dakota	Saskatoon	R2	10	T3	0.870	26.83	52.31
67	CDC Dakota	Saskatoon	R2	60	T1	1.043	26.52	49.00
68	CDC Dakota	Saskatoon	R2	60	T2	1.042	26.58	46.32
69	CDC Dakota	Saskatoon	R2	60	T3	1.054	26.92	49.22
70	CDC Dakota	Saskatoon	R2	120	T1	1.073	26.47	51.98
71	CDC Dakota	Saskatoon	R2	120	T2	1.017	26.01	38.60

72	CDC Dakota	Saskatoon	R2	120	T3	1.033	25.91	38.67
73	CDC Meadow	Rosthern	R1	10	T1	0.839	33.66	51.11
74	CDC Meadow	Rosthern	R1	10	T2	0.814	33.16	49.05
75	CDC Meadow	Rosthern	R1	10	T3	0.826	33.13	51.59
76	CDC Meadow	Rosthern	R1	60	T1	0.796	32.99	53.94
77	CDC Meadow	Rosthern	R1	60	T2	0.792	33.19	49.99
78	CDC Meadow	Rosthern	R1	60	T3	0.778	31.19	51.60
79	CDC Meadow	Rosthern	R1	120	T1	0.815	33.64	51.97
80	CDC Meadow	Rosthern	R1	120	T2	0.814	32.54	50.40
81	CDC Meadow	Rosthern	R1	120	T3	0.810	32.52	52.92
82	CDC Meadow	Rosthern	R2	10	T1	1.900	27.13	60.53
83	CDC Meadow	Rosthern	R2	10	T2	1.800	27.16	57.90
84	CDC Meadow	Rosthern	R2	10	T3	1.797	27.12	58.46
85	CDC Meadow	Rosthern	R2	60	T1	1.919	23.37	52.04
86	CDC Meadow	Rosthern	R2	60	T2	1.925	22.70	48.95
87	CDC Meadow	Rosthern	R2	60	T3	1.991	25.18	51.19
88	CDC Meadow	Rosthern	R2	120	T1	1.768	25.86	56.98
89	CDC Meadow	Rosthern	R2	120	T2	1.761	25.20	57.89
90	CDC Meadow	Rosthern	R2	120	T3	1.723	25.72	54.75
91	CDC Meadow	Saskatoon	R1	10	T1	0.494	26.01	43.06
92	CDC Meadow	Saskatoon	R1	10	T2	0.490	25.90	44.37
93	CDC Meadow	Saskatoon	R1	10	T3	0.503	25.77	43.02
94	CDC Meadow	Saskatoon	R1	60	T1	0.498	26.82	48.71
95	CDC Meadow	Saskatoon	R1	60	T2	0.499	27.45	43.21
96	CDC Meadow	Saskatoon	R1	60	T3	0.491	26.64	47.68
97	CDC Meadow	Saskatoon	R1	120	T1	0.455	27.64	46.68
98	CDC Meadow	Saskatoon	R1	120	T2	0.494	28.41	51.37
99	CDC Meadow	Saskatoon	R1	120	T3	0.525	28.85	42.28
100	CDC Meadow	Saskatoon	R2	10	T1	1.128	29.00	44.87
101	CDC Meadow	Saskatoon	R2	10	T2	1.117	29.79	46.84
102	CDC Meadow	Saskatoon	R2	10	T3	1.147	30.13	50.03
103	CDC Meadow	Saskatoon	R2	60	T1	1.311	30.81	51.22
104	CDC Meadow	Saskatoon	R2	60	T2	1.321	31.29	54.03
105	CDC Meadow	Saskatoon	R2	60	T3	1.306	31.11	50.26
106	CDC Meadow	Saskatoon	R2	120	T1	1.213	30.28	48.46
107	CDC Meadow	Saskatoon	R2	120	T2	1.216	30.22	48.99
108	CDC Meadow	Saskatoon	R2	120	T3	1.224	30.50	59.41

Table B.2: Least square means in µg/g (LSM), analysis of variance (ANOVA), mean, and coefficient of variation from the factorial analysis of Fe, Zn and Se concentrations for three sample sizes (10 seeds, 60 seeds, 120 seeds) in cultivars of CDC Striker, CDC Dakota and CDC Meadow obtained from two locations of Rosthern and Saskatoon with two biological replicates per location in 2016.

	Se	**Zn**	**Fe**
LSM_10	1.0	29.4	48.5
LSM_60	1.0	29.1	48.9
LSM_120	1.0	29.9	49.8
Mean	1.0	29.4	49.1
ANOVA (F value)	1.18^{ns}	1.21^{ns}	1.32^{ns}
Sample_size	0.01^{ns}	0.91^{ns}	0.97^{ns}
CV%	76.3	9.1	8.1

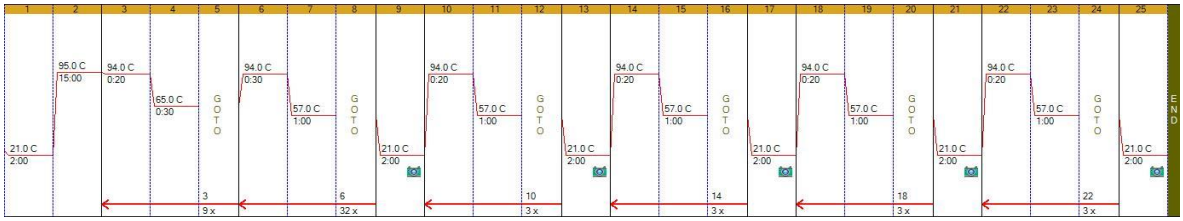
(ns - not significant)

Discussion:

No significant differences were detected for each of Se, Zn and Fe concentration for all the factors considered (genotype, location, and sample size) at the 5 % level of significance (Table B.2). Thus, the sample size did not significantly affect the concentration of the particular minerals, and 10, 60 or 120 seeds could be used as the standard sample size.

Of the three sample sizes, using 120 seeds for the main study was considered an unnecessary use of resources and time. Using 10 seeds would be a suitable sample size, as fewer resources are needed and it would be efficient in terms of counting seeds and grinding. However, in order to reduce any practical errors and to increase accuracy, it was decided to use 60 seeds for the main study. Also, 60 seeds allowed for reserve ground sample for further use if required.

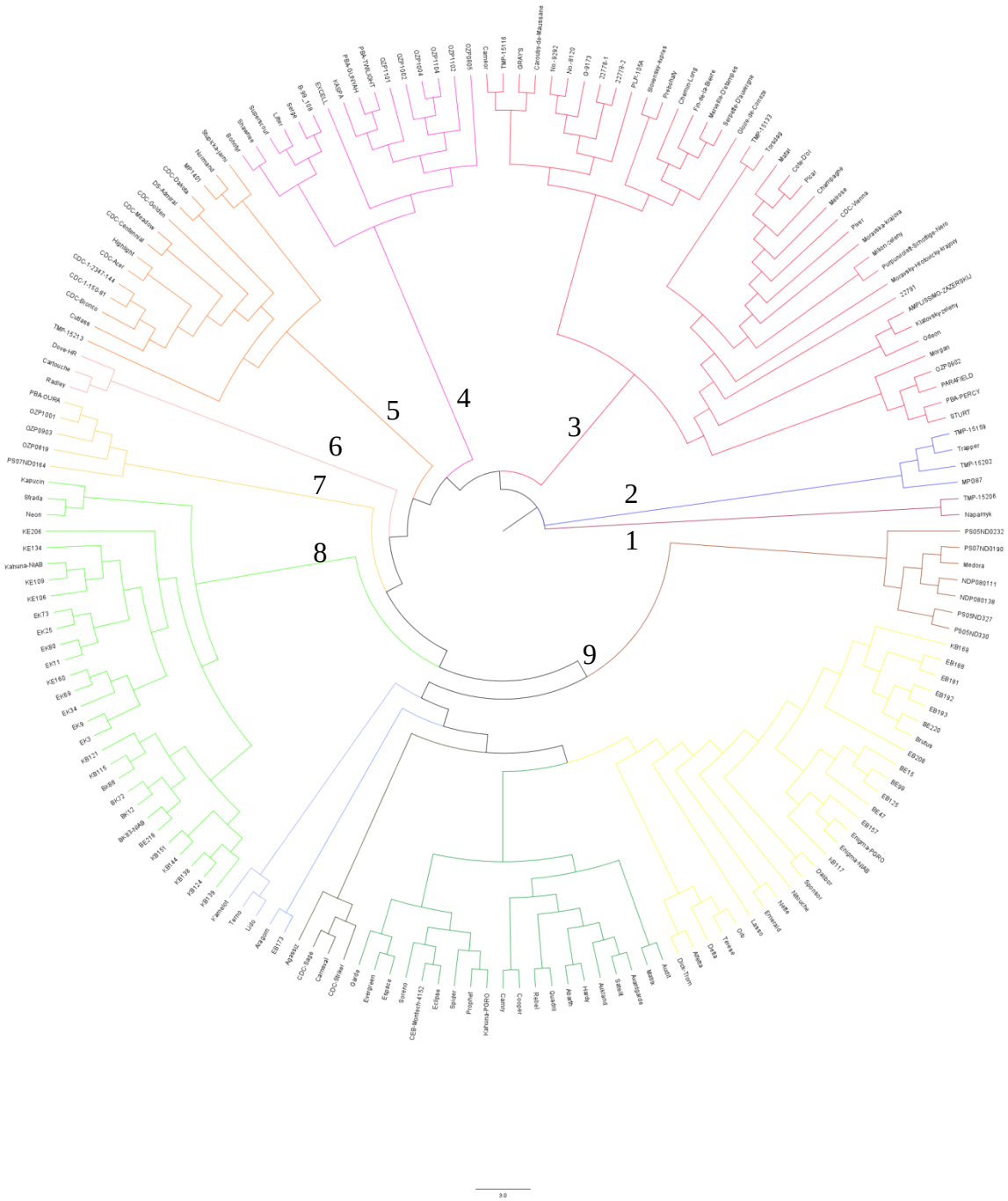
Appendix C: Thermal cycling protocol in KASP genotyping



1. 21.0 C for 2:00
2. 95.0 C for 15:00
3. 94.0 C for 0:20
4. 65.0 C for 0:30
   Decrement temperature by -0.8 C per cycle
5. GOTO 3, 9 more times
6. 94.0 C for 0:30
7. 57.0 C for 1:00
8. GOTO 6, 32 more times
9. 21.0 C for 2:00
   + Plate Read
10. 94.0 C for 0:20
11. 57.0 C for 1:00
12. GOTO 10, 3 more times
13. 21.0 C for 2:00
    + Plate Read
14. 94.0 C for 0:20
15. 57.0 C for 1:00
16. GOTO 14, 3 more times
17. 21.0 C for 2:00
    + Plate Read
18. 94.0 C for 0:20
19. 57.0 C for 1:00
20. GOTO 18, 3 more times
21. 21.0 C for 2:00
    + Plate Read
22. 94.0 C for 0:20
23. 57.0 C for 1:00
24. GOTO 22, 3 more times
25. 21.0 C for 2:00
    + Plate Read
    END

Appendix D: Phylogenetic clustering of the GWAS panel (MAF at 0.05)

Source: Gali et al. (in prep)



Appendix E: Accession origin and Fe, Zn, and Se concentration based on the different clusters

Cluster	Name	Entry	Country of origin	Fe	Zn	Se
1	TMP 15206	140	Russia	50.55	31.03	1.02
1	Naparnyk	164	Russia	47.70	30.98	0.90
2	TMP 15159	138	Russia	57.10	34.89	0.98
2	Trapper	165	Canada	58.49	34.12	1.02
2	MPG87	141	Canada	63.66	34.43	1.02
2	TMP 15202	139	Russia	50.57	28.14	0.74
3	OZP0902	76	Australia	60.83	31.59	1.15
3	Torsdag	161	Russia	48.62	31.47	0.90
3	22778	42	Ethiopia	57.34	31.37	1.41
3	TMP 15116	136	Ukraine	55.23	27.83	0.78
3	Moravska krajova	56	Czech Republic	62.66	29.78	1.10
3	Prebohaty	50	Czech Republic	53.80	32.04	0.97
3	AMPLISSIMO ZAZERSKIJ	40	Russia	63.69	36.92	0.93
3	Matar	153	India	62.94	33.34	0.80
3	Cameor	135	France	50.37	26.63	0.71
3	Fin de la Bievre	64	France	50.55	31.85	0.78
3	No. 9292	37	USA	57.39	29.64	1.07
3	Chemin Long	62	France	50.63	29.30	0.91
3	Gloire de Correze	65	France	50.73	33.12	0.94
3	Odeon	51	Czech Republic	58.97	33.23	0.95
3	STURT	90	Australia	54.27	29.21	1.11
3	Champagne	61	France	61.51	36.46	0.99
3	TMP 15133	137	Lithuania	57.27	32.46	1.32
3	No. 8120	39	USA	57.83	34.86	0.89
3	GRAY'S	36	USA	52.87	30.17	0.65
3	G 9173	38	USA	48.63	29.47	0.72
3	Klatovsky zeleny	44	Czech Republic	52.14	34.31	1.07
3	Cote D'or	63	France	63.13	34.61	0.89
3	Picar	68	France	59.39	35.12	0.97
3	Melrose	34	USA	60.56	34.88	1.12
3	Slovensky expres	46	Czech Republic	53.27	31.53	1.36
3	Serpette D'auvergne	70	France	58.77	34.78	1.72
3	Merveille D'etampes	66	France	56.56	31.00	1.13
3	Purpurviolett Schottige Nero	57	Czech Republic	57.07	30.71	1.56
3	CDC Vienna	167	Canada	53.32	28.30	1.47
3	PARAFIELD	85	Australia	56.78	30.75	1.41
3	Piver	69	France	56.69	32.02	0.90
3	PLP 105A	41	India	67.03	32.01	1.20
3	Moravsky Hrotovicky krajovy	47	Czech Republic	57.33	32.39	1.26
3	22791	43	Ethiopia	52.66	27.93	1.15
3	PBA PERCY	88	Australia	48.96	28.29	0.98

3	Milion zeleny	45	Czech Republic	50.37	27.41	1.02
3	Morgan	71	Australia	54.87	32.22	0.93
3	Carouby de Maussane	60	France	59.54	36.41	1.05
4	PBA GUNYAH	86	Australia	48.87	31.44	0.85
4	Bohatyr	6	Czech Republic	49.85	29.22	0.75
4	Serge	32	USA	49.21	30.59	0.86
4	EXCELL	72	Australia	43.90	32.40	0.81
4	KASPA	73	Australia	54.99	32.58	1.07
4	OZP1102	84	Australia	49.83	29.26	0.78
4	Superscout	174	USA	49.38	30.58	0.93
4	OZP1002	79	Australia	43.46	27.54	0.89
4	B 99/108	53	Czech Republic	51.90	29.02	0.88
4	OZP1101	80	Australia	47.79	30.02	0.79
4	OZP1004	81	Australia	47.27	29.24	0.73
4	PBA TWILIGHT	89	Australia	51.89	32.95	0.96
4	OZP1104	83	Australia	44.61	27.74	0.79
4	Lifter	31	USA	48.65	29.85	1.16
4	Shawnee	7	USA	47.19	31.32	0.80
4	OZP0805	74	Australia	48.02	31.02	1.05
5	CDC 1-2347-144	170	Canada	46.75	26.89	1.01
5	CDC Bronco	144	Canada	43.36	24.39	0.80
5	Normand	67	France	55.59	32.02	0.98
5	DS Admiral	148	Denmark	42.43	25.29	1.05
5	CDC Acer	163	Canada	49.39	30.45	0.99
5	Cutlass	143	Canada	46.41	25.54	0.93
5	CDC Golden	146	Canada	47.46	28.30	0.88
5	Stupicka jarni	58	Czech Republic	59.31	35.62	0.78
5	Highlight	168	Sweden	44.51	26.83	0.95
5	CDC Dakota	177	Canada	44.27	25.76	0.89
5	CDC Meadow	147	Canada	46.90	26.87	0.72
5	TMP 15213	142	Belarus	48.81	24.69	0.64
5	MP1401	155	Canada	48.18	30.13	0.87
5	CDC Centennial	145	Canada	48.96	28.25	1.01
5	CDC 1-150-81	169	Canada	45.05	26.68	1.13
6	Radley	166	UK	56.47	31.00	0.99
6	Dove HR	35	USA	56.01	30.45	0.78
6	Cartouche	173	France	47.33	27.78	0.94
7	PS07ND0164	2	USA	48.92	31.24	0.99
7	OZP1001	78	Australia	47.18	28.79	0.88
7	PBA OURA	87	Australia	53.33	32.22	0.75
7	OZP0903	77	Australia	44.51	27.87	1.09
7	OZP0819	75	Australia	51.26	28.33	0.86
8	Kapucin	59	Czech Republic	53.35	36.95	0.97
8	Kahuna-NIAB	133	UK	53.59	34.30	1.23
8	KE109	125	UK	50.48	29.62	1.07
8	BK12	91	UK	49.24	30.48	0.97
8	EK34	120	UK	49.92	31.32	0.83

8	Strada	23	Netherlands	51.61	32.87	0.78
8	KE160	127	UK	49.89	31.94	0.84
8	EK25	119	UK	47.80	30.59	0.82
8	KB151	102	UK	51.41	33.82	0.87
8	EK73	122	UK	53.09	35.43	0.93
8	KE106	124	UK	50.91	31.66	0.88
8	KB121	97	UK	51.33	35.32	0.79
8	KB139	100	UK	48.89	28.35	1.23
8	BK83	93	UK	51.29	32.40	0.76
8	BK72	92	UK	48.35	28.73	1.04
8	EK9	117	UK	49.10	27.19	0.78
8	KE206	128	UK	51.21	30.85	1.03
8	KB138	99	UK	52.46	30.49	1.10
8	BE218	114	UK	54.34	31.73	1.01
8	EK80	123	UK	54.69	31.05	0.91
8	BK88	94	UK	50.45	30.81	0.91
8	Neon	22	Netherlands	53.24	32.73	1.00
8	EK3	116	UK	46.39	27.99	0.91
8	KB144	101	UK	50.55	28.38	0.82
8	KB115	95	UK	50.42	31.31	1.17
8	EK11	118	UK	52.84	31.56	1.02
8	KB124	98	UK	45.92	30.95	0.82
8	KE134	126	UK	54.97	33.08	0.76
8	EK69	121	UK	47.37	31.42	1.00
9	Garde	25	Netherlands	44.61	26.06	0.82
9	CEB-Montech 4152	28	Netherlands	50.39	30.91	1.27
9	EB125	107	UK	47.59	25.94	1.02
9	CDC Striker	150	Canada	46.12	26.93	0.90
9	Hardy	172	France	44.76	28.06	0.85
9	Agassiz	171	Canada	44.99	26.23	0.84
9	Alfetta	157	Netherlands	44.71	28.60	0.91
9	Dalibor	48	Czech Republic	46.67	25.98	0.72
9	Satelit	16	Netherlands	48.28	26.96	1.11
9	Nette	17	Netherlands	44.01	23.15	1.07
9	KB168	103	UK	47.66	27.61	0.78
9	BE15	104	UK	47.68	29.39	1.01
9	Soreno	21	Netherlands	47.54	29.10	0.80
9	EB208	113	UK	47.05	27.08	1.00
9	Aukland	30	Netherlands	44.95	25.17	0.85
9	BE99	106	UK	44.07	25.94	0.88
9	Audit	11	Netherlands	52.16	25.98	0.79
9	Aragorn	176	USA	47.92	29.34	1.12
9	Enigma-PGRO	130	UK	44.26	24.19	0.87
9	NDP080138	5	USA	48.69	25.48	1.22
9	PS05ND327	8	USA	54.38	27.97	1.21
9	KB117	96	UK	48.19	30.34	1.01
9	EB188	110	UK	48.86	27.47	0.65

9	Rebel	15	Netherlands	47.66	27.63	0.88
9	Abarth	20	Netherlands	47.35	25.20	0.83
9	Terno	55	Czech Republic	51.49	26.07	0.80
9	EB192	111	UK	48.70	29.09	0.94
9	Dick Trom	49	Czech Republic	47.80	29.66	1.12
9	Kamelot	52	Czech Republic	45.94	26.55	0.73
9	Terese	160	France	48.99	30.13	0.72
9	EB157	108	UK	48.86	27.46	1.01
9	EB173	109	UK	49.15	26.14	0.71
9	EB193	112	UK	47.36	26.24	0.97
9	Orb	156	UK	46.49	27.80	0.75
9	Enigma-NIAB	134	UK	45.52	25.39	0.76
9	Eclipse	149	Netherlands	49.71	29.34	0.72
9	Nitouche	152	Denmark	47.40	28.23	0.86
9	Avantgarde	12	Netherlands	46.55	25.22	0.78
9	BE220	115	UK	47.86	27.13	0.80
9	Sponsor	54	Czech Republic	47.40	26.42	1.12
9	OZP1103	82	Australia	49.43	31.79	1.10
9	Carneval	154	Sweden	44.30	26.00	0.92
9	Brutus	132	UK	46.58	25.94	0.88
9	Spider	29	Netherlands	47.33	27.07	0.90
9	Kahuna-PGRO	129	UK	55.00	36.56	1.06
9	Matrix	27	Netherlands	48.06	30.21	0.90
9	Medora	33	USA	48.56	26.11	1.13
9	Emerald	18	Netherlands	45.59	26.91	1.25
9	PS05ND0434	10	USA	45.99	24.59	0.81
9	CDC Sage	158	Canada	43.87	26.65	0.70
9	Cooper	151	Netherlands	44.14	28.81	0.76
9	PS05ND0232	1	USA	47.99	26.64	0.90
9	Evergreen	19	Netherlands	48.29	23.28	0.86
9	Prophet	24	Netherlands	46.66	25.91	1.04
9	NDP080111	4	USA	45.70	24.18	1.11
9	Quadril	14	Netherlands	50.31	26.41	1.09
9	Camry	26	Netherlands	46.36	29.31	0.76
9	PS07ND0190	3	USA	48.32	27.11	1.12
9	PS05ND330	9	USA	51.30	29.46	1.12
9	Espace	159	Netherlands	48.03	25.79	1.08
9	Lido	175	Denmark	43.22	27.25	1.23
9	EB181	131	UK	48.43	26.59	0.85
9	BE47	105	UK	47.43	25.84	0.76
9	Delta	162	Netherlands	48.89	26.83	0.93
9	Lasso	13	Netherlands	49.63	27.63	0.94

Appendix F: SNP markers associated with Se concentration in seeds of the pea GWAS panel over five location-years.

Location-year	Marker	Scaffold	*P* value	-log$_{10}$	Other location-years that the marker is present (-log$_{10}$ *P* value ≥ 3.16)
2013S	Sc_1139_105036	1139	1.03E-06	5.98*	none
2013S	Sc_3049_27354	5765	2.43E-06	5.61*	none
2015S	Sc_5765_134167	6064	8.22E-09	8.09***	none
2015S	Sc_6064_83250	6064	6.77E-08	7.17***	none
2015S	Sc_6064_83570	6064	6.77E-08	7.17***	none
2015S	Sc_6064_83588	6064	6.77E-08	7.17***	none
2015S	Sc_6064_83592	6064	6.77E-08	7.17***	none
2015S	Sc_3542_172841	3542	1.25E-07	6.90**	none
2015S	Sc_6393_64958	6393	1.51E-07	6.82**	none
2015S	Sc_6393_64973	6393	1.51E-07	6.82**	none
2015S	Sc_4211_152482	4211	3.35E-07	6.47**	none
2015S	Sc_5765_134136	5765	3.82E-07	6.42**	none
2015S	Sc_5765_134139	5765	3.82E-07	6.42**	none
2015S	Sc_5765_134315	5765	3.90E-07	6.41**	none
2015S	Sc_1_37135	1	4.87E-07	6.31**	none
2015S	Sc_4643_173162	4643	5.09E-07	6.29**	none
2015S	Sc_2158_60008	2158	5.75E-07	6.24**	none
2015S	Sc_2158_60179	2158	5.75E-07	6.24**	none
2015S	Sc_6086_72254	6086	6.05E-07	6.22**	none
2015S	Sc_393_898089	393	6.11E-07	6.21**	none
2015S	Sc_5247_82400	5247	6.17E-07	6.21**	none
2015S	Sc_2997_81595	2997	6.17E-07	6.21**	none
2015S	Sc_5765_134168	5765	6.22E-07	6.21**	none
2015S	Sc_1512_627644	1512	6.26E-07	6.20**	none
2015S	Sc_2783_213981	2783	6.79E-07	6.17**	none
2015S	Sc_4407_37120	4407	9.17E-07	6.04*	none
2015S	Sc_1622_628173	1622	1.03E-06	5.99*	none
2015S	Sc_2552_354124	2552	1.26E-06	5.90*	none

2015S	Sc_266_87261	266	1.46E-06	5.84*	none
2015S	Sc_5765_134154	5765	1.60E-06	5.80*	none
2015S	Sc_3685_291242	3685	1.66E-06	5.78*	none
2015S	Sc_1417_25038	1417	2.01E-06	5.70*	none
2015S	Sc_2158_60010	2158	2.11E-06	5.68*	none
2015S	Sc_6830_59544	6830	2.24E-06	5.65*	none
2015S	Sc_788_50361	788	2.26E-06	5.65*	none
2015S	Sc_2252_490227	2252	2.45E-06	5.61*	none
2015S	Sc_3984_38749	3984	2.54E-06	5.60*	none
2015S	Sc_3878_244774	3878	2.54E-06	5.60*	none
2015S	Sc_3984_39060	3984	2.63E-06	5.58*	none
2015S	Sc_3984_39069	3984	2.63E-06	5.58*	none
2015S	Sc_3984_38833	3984	2.64E-06	5.58*	none
2015S	Sc_449_149728	449	2.64E-06	5.58*	none
2015S	Sc_5765_134095	5765	2.65E-06	5.58*	none
2015S	Sc_2084_228433	2084	3.42E-06	5.47*	none

2013S - 2013 Saskatoon, 2014F - 2014 Fargo, 2015S - 2015 Saskatoon, 2016R - 2016 Rosthern, 2016S - 2016 Saskatoon. *** $P \leq 0.001$ ($-\log_{10} P$ value $\geq$ 7.16), ** $P \leq 0.01$ ($-\log_{10} P$ value $\geq$ 6.16), * $P \leq 0.05$ ($-\log_{10} P$ value $\geq$ 5.46)

Appendix G: Average yield with Se, Zn, and Fe concentration

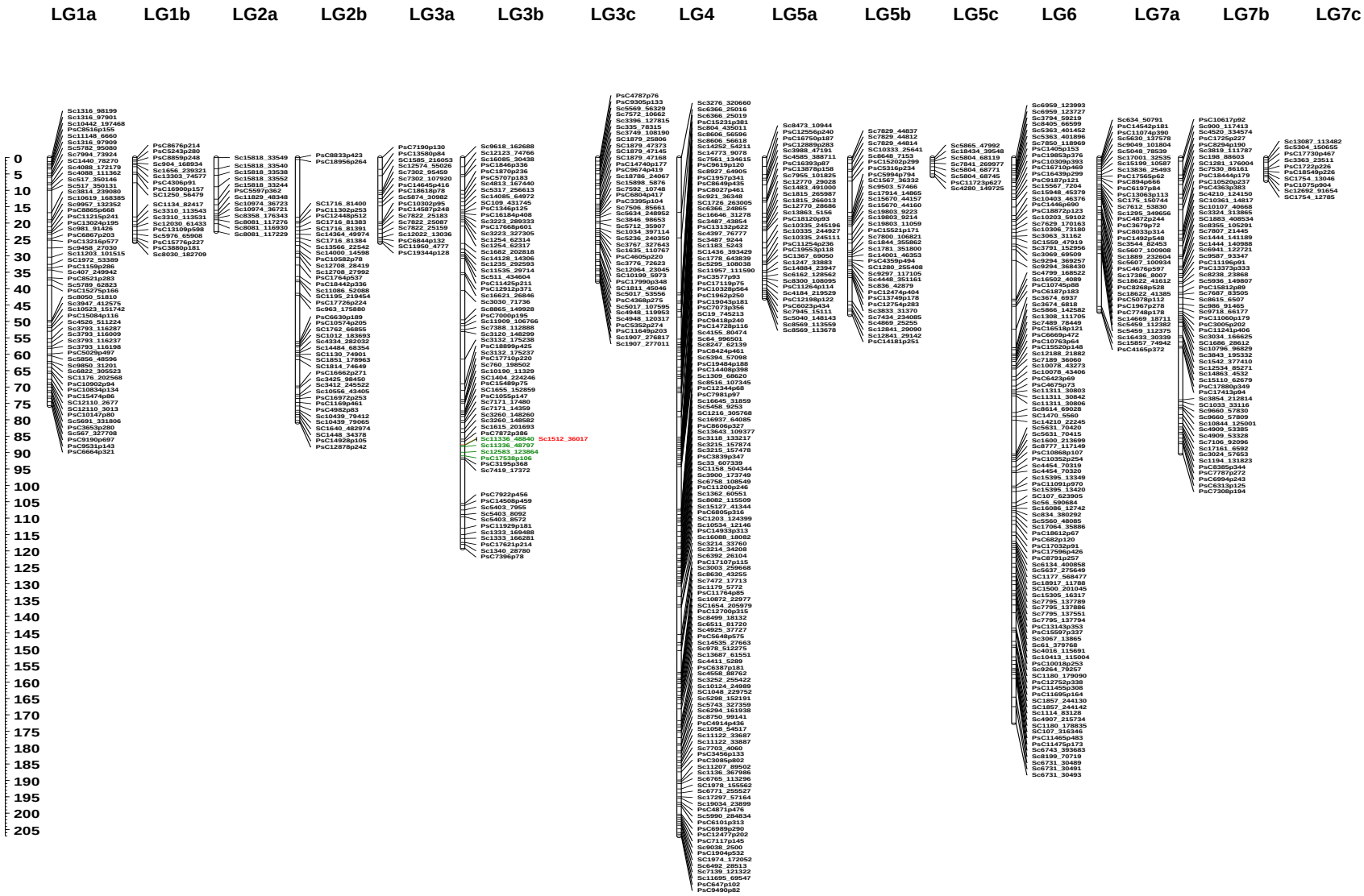
Entry	Name	Yield	Se		Zn		Fe	
		kg/ha	µg/g	g/ha	µg/g	g/ha	µg/g	g/ha
1	PS05ND0232	2622.04	0.90	2.37	26.64	69.86	47.99	125.83
2	PS07ND0164	2924.79	0.99	2.89	31.24	91.36	48.92	143.07
3	PS07ND0190	2575.59	1.12	2.88	27.11	69.83	48.32	124.44
4	NDP080111	3051.95	1.11	3.39	24.18	73.78	45.70	139.45
5	NDP080138	2807.90	1.22	3.42	25.48	71.56	48.69	136.71
6	Bohatyr	2053.20	0.75	1.54	29.22	59.99	49.85	102.35
7	Shawnee	2553.56	0.80	2.06	31.32	79.99	47.19	120.49
8	PS05ND327	2326.21	1.21	2.81	27.97	65.05	54.38	126.48
9	PS05ND330	2727.03	1.12	3.07	29.46	80.33	51.30	139.89
10	PS05ND0434	2766.96	0.81	2.24	24.59	68.05	45.99	127.24
11	Audit	2722.75	0.79	2.15	25.98	70.75	52.16	142.02
12	Avantgarde	2545.96	0.78	1.99	25.22	64.22	46.55	118.50
13	Lasso	2959.16	0.94	2.79	27.63	81.76	49.63	146.86
14	Quadril	2903.32	1.09	3.16	26.41	76.68	50.31	146.07
15	Rebel	2564.70	0.88	2.26	27.63	70.86	47.66	122.22
16	Satelit	2469.18	1.11	2.74	26.96	66.57	48.28	119.20
17	Nette	2635.57	1.07	2.83	23.15	61.02	44.01	115.99
18	Emerald	2189.00	1.25	2.75	26.91	58.91	45.59	99.78
19	Evergreen	2509.36	0.86	2.16	23.28	58.42	48.29	121.16
20	Abarth	2849.76	0.83	2.38	25.20	71.81	47.35	134.94
21	Soreno	2633.80	0.80	2.11	29.10	76.65	47.54	125.21
22	Neon	2478.12	1.00	2.47	32.73	81.12	53.24	131.92
23	Strada	2520.60	0.78	1.98	32.87	82.86	51.61	130.09
24	Prophet	2421.63	1.04	2.52	25.91	62.75	46.66	112.99
25	Garde	2612.81	0.82	2.15	26.06	68.10	44.61	116.56
26	Camry	2580.82	0.76	1.97	29.31	75.64	46.36	119.63
27	Matrix	2388.68	0.90	2.16	30.21	72.16	48.06	114.79
28	CEB-Montech 4152	2921.53	1.27	3.72	30.91	90.29	50.39	147.21
29	Spider	2688.37	0.90	2.43	27.07	72.78	47.33	127.22
30	Aukland	2291.56	0.85	1.96	25.17	57.68	44.95	102.99
31	Lifter	2809.91	1.16	3.27	29.85	83.89	48.65	136.69
32	Serge	1873.32	0.86	1.61	30.59	57.30	49.21	92.17
33	Medora	2668.56	1.13	3.02	26.11	69.68	48.56	129.57
34	Melrose	2039.43	1.12	2.29	34.88	71.13	60.56	123.50
35	Dove HR	2636.56	0.78	2.06	30.45	80.29	56.01	147.66
36	GRAY'S	2709.90	0.65	1.76	30.17	81.75	52.87	143.26
37	No. 9292	2604.20	1.07	2.80	29.64	77.18	57.39	149.46
38	G 9173	2307.32	0.72	1.68	29.47	67.99	48.63	112.21
39	No. 8120	2015.86	0.89	1.79	34.86	70.28	57.83	116.57
40	AMPLISSIMO ZAZERSKIJ	2177.77	0.93	2.03	36.92	80.41	63.69	138.70
41	PLP 105A	2218.68	1.20	2.67	32.01	71.03	67.03	148.70
42	22778	2216.54	1.41	3.13	31.37	69.54	57.34	127.09

43	22791	2600.25	1.15	3.00	27.93	72.61	52.66	136.92
44	Klatovsky zeleny	2477.91	1.07	2.66	34.31	85.02	52.14	129.19
45	Milion zeleny	2285.25	1.02	2.33	27.41	62.64	50.37	115.09
46	Slovensky expres	2139.69	1.36	2.91	31.53	67.46	53.27	113.96
47	Moravsky Hrotovicky krajovy	2318.52	1.26	2.92	32.39	75.09	57.33	132.92
48	Dalibor	2624.32	0.72	1.89	25.98	68.18	46.67	122.47
49	Dick Trom	1768.59	1.12	1.98	29.66	52.45	47.8	84.54
50	Prebohaty	2027.84	0.97	1.98	32.04	64.97	53.8	109.09
51	Odeon	2542.76	0.95	2.41	33.23	84.49	58.97	149.93
52	Kamelot	2948.50	0.73	2.16	26.55	78.28	45.94	135.44
53	B 99/108	1934.03	0.88	1.71	29.02	56.13	51.9	100.37
54	Sponsor	2418.96	1.12	2.71	26.42	63.91	47.4	114.64
55	Terno	2306.79	0.80	1.86	26.07	60.13	51.49	118.76
56	Moravska krajova	2443.39	1.10	2.69	29.78	72.75	62.66	153.09
57	Purpurviolett Schottige Nero	2087.54	1.56	3.27	30.71	64.11	57.07	119.14
58	Stupicka jarni	2458.17	0.78	1.93	35.62	87.55	59.31	145.78
59	Kapucin	2033.41	0.97	1.98	36.95	75.14	53.35	108.48
60	Carouby de Maussane	1989.31	1.05	2.10	36.41	72.42	59.54	118.44
61	Champagne	1895.40	0.99	1.89	36.46	69.10	61.51	116.57
62	Chemin Long	2417.43	0.91	2.20	29.3	70.84	50.63	122.39
63	Cote D'or	2158.66	0.89	1.92	34.61	74.72	63.13	136.28
64	Fin de la Bievre	2115.64	0.78	1.65	31.85	67.39	50.55	106.93
65	Gloire de Correze	1940.08	0.94	1.82	33.12	64.26	50.73	98.41
66	Merveille D'etampes	1458.35	1.13	1.65	31.00	45.20	56.56	82.48
67	Normand	2294.19	0.98	2.26	32.02	73.47	55.59	127.52
68	Picar	2285.61	0.97	2.21	35.12	80.27	59.39	135.73
69	Piver	2076.69	0.90	1.88	32.02	66.49	56.69	117.72
70	Serpette D'auvergne	1895.17	1.72	3.26	34.78	65.92	58.77	111.37
71	Morgan	2489.39	0.93	2.31	32.22	80.21	54.87	136.59
72	EXCELL	2633.45	0.81	2.15	32.4	85.32	43.9	115.60
73	KASPA	2725.72	1.07	2.92	32.58	88.79	54.99	149.88
74	OZP0805	2488.75	1.05	2.62	31.02	77.19	48.02	119.49
75	OZP0819	2594.61	0.86	2.24	28.33	73.49	51.26	133.00
76	OZP0902	1913.92	1.15	2.20	31.59	60.47	60.83	116.42
77	OZP0903	2830.10	1.09	3.09	27.87	78.87	44.51	125.96
78	OZP1001	3102.49	0.88	2.74	28.79	89.33	47.18	146.37
79	OZP1002	2918.29	0.89	2.61	27.54	80.38	43.46	126.81
80	OZP1101	2598.63	0.79	2.05	30.02	78.00	47.79	124.18
81	OZP1004	2807.86	0.73	2.06	29.24	82.11	47.27	132.73
82	OZP1103	2504.74	1.10	2.77	31.79	79.63	49.43	123.80
83	OZP1104	2807.07	0.79	2.23	27.74	77.86	44.61	125.21
84	OZP1102	2764.14	0.78	2.17	29.26	80.87	49.83	137.73
85	PARAFIELD	2342.09	1.41	3.30	30.75	72.01	56.78	132.97
86	PBA GUNYAH	3042.40	0.85	2.58	31.44	95.64	48.87	148.69
87	PBA OURA	2810.63	0.75	2.12	32.22	90.56	53.33	149.90
88	PBA PERCY	2765.45	0.98	2.72	28.29	78.22	48.96	135.39

89	PBA TWILIGHT	2837.57	0.96	2.72	32.95	93.49	51.89	147.22
90	STURT	2963.81	1.11	3.31	29.21	86.58	54.27	160.83
91	BK12	2609.29	0.97	2.54	30.48	79.52	49.24	128.48
92	BK72	3086.66	1.04	3.20	28.73	88.68	48.35	149.23
93	BK83	2388.33	0.76	1.81	32.4	77.39	51.29	122.50
94	BK88	2578.01	0.91	2.35	30.81	79.44	50.45	130.05
95	KB115	2746.21	1.17	3.21	31.31	85.99	50.42	138.47
96	KB117	2841.24	1.01	2.88	30.34	86.20	48.19	136.92
97	KB121	2722.93	0.79	2.17	35.32	96.17	51.33	139.76
98	KB124	2582.57	0.82	2.12	30.95	79.92	45.92	118.58
99	KB138	2382.10	1.10	2.62	30.49	72.64	52.46	124.96
100	KB139	2954.81	1.23	3.65	28.35	83.76	48.89	144.46
101	KB144	2879.82	0.82	2.37	28.38	81.73	50.55	145.56
102	KB151	2420.00	0.87	2.10	33.82	81.84	51.41	124.40
103	KB168	3058.74	0.78	2.38	27.61	84.45	47.66	145.79
104	BE15	2315.37	1.01	2.34	29.39	68.04	47.68	110.38
105	BE47	3059.52	0.76	2.33	25.84	79.05	47.43	145.11
106	BE99	2823.54	0.88	2.50	25.94	73.24	44.07	124.43
107	EB125	2627.43	1.02	2.84	25.94	68.14	47.59	125.04
108	EB157	2693.21	1.01	2.73	27.46	73.96	48.86	131.58
109	EB173	2693.00	0.71	1.91	26.14	70.41	49.15	132.34
110	EB188	2833.26	0.65	1.84	27.47	77.83	48.86	138.42
111	EB192	2773.65	0.94	2.61	29.09	80.69	48.7	135.08
112	EB193	2757.12	0.97	2.69	26.24	72.34	47.36	130.57
113	EB208	2727.73	1.00	2.75	27.08	73.88	47.05	128.33
114	BE218	2756.63	1.01	2.79	31.73	87.46	54.34	149.77
115	BE220	2685.13	0.80	2.15	27.13	72.85	47.86	128.50
116	EK3	2920.82	0.91	2.68	27.99	81.75	46.39	135.48
117	EK9	3165.96	0.78	2.48	27.19	86.08	49.10	155.44
118	EK11	2566.27	1.02	2.61	31.56	81.00	52.84	135.59
119	EK25	2661.11	0.82	2.19	30.59	81.40	47.80	127.18
120	EK34	2775.45	0.83	2.31	31.32	86.94	49.92	138.53
121	EK69	2738.16	1.00	2.75	31.42	86.02	47.37	129.71
122	EK73	2710.39	0.93	2.52	35.43	96.02	53.09	143.88
123	EK80	2869.61	0.91	2.63	31.05	89.11	54.69	156.92
124	KE106	2809.39	0.88	2.47	31.66	88.93	50.91	143.01
125	KE109	2977.74	1.07	3.20	29.62	88.21	50.48	150.30
126	KE134	3121.29	0.76	2.38	33.08	103.25	54.97	171.56
127	KE160	2625.26	0.84	2.21	31.94	83.86	49.89	130.96
128	KE206	2644.50	1.03	2.72	30.85	81.58	51.21	135.42
129	Kahuna-PGRO	2820.18	1.06	3.10	36.56	103.12	55.00	155.11
130	Enigma-PGRO	3059.09	0.87	2.65	24.19	74.01	44.26	135.41
131	EB181	2475.13	0.85	2.10	26.59	65.81	48.43	119.86
132	Brutus	3074.99	0.88	2.72	25.94	79.78	46.58	143.22
133	Kahuna-NIAB	2732.21	1.23	3.37	34.3	93.70	53.59	146.41
134	Enigma-NIAB	2764.62	0.76	2.12	25.39	70.20	45.52	125.85

135	Cameor	1927.40	0.71	1.38	26.63	51.33	50.37	97.09
136	TMP 15116	2595.59	0.78	2.03	27.83	72.23	55.23	143.34
137	TMP 15133	2634.58	1.32	3.49	32.46	85.51	57.27	150.88
138	TMP 15159	2462.79	0.98	2.42	34.89	85.92	57.10	140.63
139	TMP 15202	2455.01	0.74	1.81	28.14	69.07	50.57	124.15
140	TMP 15206	2444.58	1.02	2.51	31.03	75.86	50.55	123.58
141	MPG87	2005.79	1.02	2.04	34.43	69.06	63.66	127.68
142	TMP 15213	3224.66	0.64	2.08	24.69	79.63	48.81	157.38
143	Cutlass	2768.67	0.93	2.57	25.54	70.71	46.41	128.49
144	CDC Bronco	3666.52	0.80	2.93	24.39	89.41	43.36	158.98
145	CDC Centennial	3071.26	1.01	3.10	28.25	86.78	48.96	150.37
146	CDC Golden	3269.86	0.88	2.89	28.3	92.53	47.46	155.17
147	CDC Meadow	3247.00	0.72	2.35	26.87	87.25	46.9	152.26
148	DS Admiral	2715.09	1.05	2.85	25.29	68.66	42.43	115.20
149	Eclipse	3249.71	0.72	2.35	29.34	95.35	49.71	161.55
150	CDC Striker	2609.80	0.90	2.34	26.93	70.29	46.12	120.37
151	Cooper	3009.37	0.76	2.31	28.81	86.69	44.14	132.84
152	Nitouche	2786.94	0.86	2.39	28.23	78.69	47.4	132.10
153	Matar	2337.50	0.80	1.87	33.34	77.94	62.94	147.13
154	Carneval	2861.28	0.92	2.63	26.00	74.40	44.30	126.74
155	MP1401	2492.91	0.87	2.16	30.13	75.12	48.18	120.09
156	Orb	2494.59	0.75	1.88	27.80	69.34	46.49	115.96
157	Alfetta	2566.27	0.91	2.35	28.60	73.41	44.71	114.73
158	CDC Sage	2635.23	0.70	1.84	26.65	70.23	43.87	115.59
159	Espace	2710.53	1.08	2.93	25.79	69.90	48.03	130.17
160	Terese	2339.52	0.72	1.68	30.13	70.49	48.99	114.62
161	Torsdag	2407.84	0.90	2.17	31.47	75.77	48.62	117.07
162	Delta	2523.17	0.93	2.35	26.83	67.70	48.89	123.36
163	CDC Acer	3267.69	0.99	3.24	30.45	99.49	49.39	161.36
164	Naparnyk	2441.32	0.90	2.20	30.98	75.63	47.70	116.46
165	Trapper	2498.90	1.02	2.56	34.12	85.26	58.49	146.16
166	Radley	2474.18	0.99	2.45	31.00	76.70	56.47	139.70
167	CDC Vienna	2785.60	1.47	4.11	28.30	78.84	53.32	148.52
168	Highlight	2967.60	0.95	2.84	26.83	79.62	44.51	132.07
169	CDC 1-150-81	2812.97	1.13	3.18	26.68	75.06	45.05	126.73
170	CDC 1-2347-144	2918.10	1.01	2.96	26.89	78.48	46.75	136.40
171	Agassiz	2949.98	0.84	2.47	26.23	77.39	44.99	132.71
172	Hardy	2581.56	0.85	2.19	28.06	72.44	44.76	115.56
173	Cartouche	2737.33	0.94	2.57	27.78	76.05	47.33	129.55
174	Superscout	1917.01	0.93	1.78	30.58	58.63	49.38	94.65
175	Lido	2852.59	1.23	3.52	27.25	77.74	43.22	123.28
176	Aragorn	2681.47	1.12	3.01	29.34	78.67	47.92	128.49
177	CDC Dakota	3246.58	0.89	2.90	25.76	83.63	44.27	143.70

Appendix H: Linkage map of PR-07 with the SNP marker Sc_1512_36017



Appendix I: Analysis of the relationship between soil and seed micronutrient concentration

Background:

When evaluating the GWAS panel accessions for Se, Zn and Fe concentration, in some cases, biological replications in the same location varied substantially, particularly for Se. Soil analysis was conducted to evaluate the correlation between soil and seed micronutrient concentration. Understanding this relationship could inform biofortification strategies.

Objective:

Observing the correlation of micronutrient concentration of soil with the particular micronutrient concentration of field pea seeds.

Methodology:

Ten varieties were selected randomly from the GWAS panel for this study (NDP080138, CAROUBY DE MAUSSANE, SERPETTE D'AUVERGNE, BK12, KB139, EB157, EK3, CDC EK25, TMP 15133, and CDC DAKOTA). Soil samples were collected from two locations, i.e., Saskatoon and Rosthern from two blocks in each location for three depths (0-15cm, 15-30cm, 30-60cm) in the harvest year 2017. Collected soil samples were submitted to the ALS Company for chemical analysis (CRC ICPMS for Se, available micronutrient method for Fe and Zn, available nitrate for N, available phosphate for P and available K). The seed samples were extracted from the corresponding plots that were selected for the soil collection washed, dried, ground and analyzed using atomic absorption spectrometry according to Thavarajah et al. (2007). Statistical analysis of the obtained data was performed using SAS 9.4 software.

Results:

Table I.1: Pearson's correlation coefficient (r) between the micronutrient concentration of soil and seeds for the three depths in Rosthern and Saskatoon, 2017

Depth	Rosthern			Saskatoon		
	Se	Zn	Fe	Se	Zn	Fe
0-15 cm	-0.30	0.03	0.25	-0.07	0.50*	0.62**
15-30 cm	0.36	0.43	0.26	-0.05	0.44	0.49*
30-60 cm	0.57*	-0.26	0.21	0.76***	-0.13	0.40

Significance: * at 0.05, ** at 0.01, *** at 0.001

Table I.2: Pearson's correlation coefficient (r) between Se, Zn, and Fe concentration of seeds and each N, P, and K in the soil for the three depths in Rosthern and Saskatoon, 2017.

Depth		Rosthern			Saskatoon		
		Se_seed	Zn_seed	Fe_seed	Se_seed	Zn_seed	Fe_seed
0-15 cm	N	-0.06	0.00	-0.04	0.27	-0.29	-0.11
	P	-0.30	0.22	0.29	-0.10	0.12	0.02
	K	-0.49*	0.16	0.27	-0.06	-0.05	-0.01
15-30 cm	N	-0.39	0.44	0.47	0.67**	0.03	0.14
	P	-0.46*	0.15	0.27	0.14	0.42	0.51*
	K	-0.45*	0.14	0.21	-0.15	0.25	0.17
30-60cm	N	-0.06	-0.16	-0.11	0.35	-0.37	-0.26
	P	0.40	0.21	0.26	0.00	0.07	0.16
	K	-0.34	0.10	0.19	-0.05	0.01	0.03

Significance: * at 0.05, ** at 0.01, *** at 0.001

Table I.3: Descriptive statistics of Se, Zn, and Fe concentration (µg/g) of the seeds for two blocks in Saskatoon and Rosthern, 2017.

Field	Block	Variable	Minimum	Maximum	Mean	SD	CV%
Rosthern	Block 1	Se_seed	0.05	0.35	0.15	0.09	58.7
		Zn_seed	29.28	47.43	36.23	6.01	16.6
		Fe_seed	46.97	69.86	55.83	7.15	12.8
	Block 2	Se_seed	0.16	1.48	0.56	0.39	68.9
		Zn_seed	29.96	45.37	35.75	5.97	16.7
		Fe_seed	47.23	69.56	55.79	7.56	13.6
Saskatoon	Block 1	Se_seed	0.51	3.86	1.30	0.98	75.9
		Zn_seed	26.76	38.49	31.65	3.85	12.2
		Fe_seed	37.11	52.9	43.14	4.76	11.0
	Block 2	Se_seed	0.52	1.91	0.88	0.39	44.1
		Zn_seed	26.65	35.44	32.63	3.06	9.4
		Fe_seed	36.45	57.07	44.47	6.10	13.7

Table I.4: Descriptive statistics of Se, Zn, and Fe concentration (µg/g) of the seeds in Saskatoon and Rosthern, 2017 (regardless of the blocks).

Field	Variable	Minimum	Maximum	Mean	SD	CV%
Rosthern	Se_seed	0.05	1.48	0.36	0.35	97.7
	Zn_seed	29.28	47.43	35.99	5.84	16.2
	Fe_seed	46.97	69.86	55.81	7.16	12.8
Saskatoon	Se_seed	0.51	3.86	1.09	0.76	69.5
	Zn_seed	26.65	38.49	32.14	3.42	10.6
	Fe_seed	36.45	57.07	43.80	5.37	12.3

Table I.5: Descriptive statistics of Se, Zn, and Fe concentration (µg/g) of the soil for two blocks in Saskatoon and Rosthern, 2017.

Location	Block	Depth	Variable	Min	Max	Mean	SD	CV
Rosthern	Block 1	0-15 cm	Se_soil	0.31	0.37	0.35	0.02	5.7
			Zn_soil	1.50	2.82	2.21	0.36	16.2
			Fe_soil	90.60	289.00	195.83	71.76	36.6
		15-30 cm	Se_soil	0.22	0.34	0.28	0.04	16.0
			Zn_soil	0.22	1.70	0.61	0.47	77.5
			Fe_soil	23.30	167.00	98.10	52.74	53.8
		30-60 cm	Se_soil	0.21	0.28	0.24	0.03	12.0
			Zn_soil	0.20	0.33	0.25	0.05	20.5
			Fe_soil	15.90	91.40	47.52	23.51	49.5
	Block 2	0-15 cm	Se_soil	0.32	0.38	0.35	0.02	7.0
			Zn_soil	0.94	2.40	1.48	0.43	28.8
			Fe_soil	60.10	83.20	69.42	7.25	10.4
		15-30 cm	Se_soil	0.22	0.85	0.39	0.19	49.8
			Zn_soil	0.21	0.30	0.25	0.03	12.9
			Fe_soil	13.10	19.50	16.59	2.36	14.2
		30-60 cm	Se_soil	0.21	0.58	0.36	0.14	37.8
			Zn_soil	0.20	0.61	0.31	0.17	54.60
			Fe_soil	14.10	22.40	17.05	2.47	14.46
Saskatoon	Block 1	0-15 cm	Se_soil	0.50	0.73	0.58	0.07	12.2
			Zn_soil	0.74	2.03	1.24	0.44	35.9
			Fe_soil	32.90	65.30	48.45	10.43	21.5
		15-30 cm	Se_soil	0.29	0.65	0.36	0.11	29.7
			Zn_soil	0.38	1.82	0.63	0.42	67.2
			Fe_soil	12.40	70.90	23.61	17.21	72.9
		30-60 cm	Se_soil	0.22	0.62	0.31	0.12	39.6

		Zn_soil	0.22	0.42	0.33	0.06	18.4
		Fe_soil	12.20	23.40	16.30	3.72	22.8
Block 2	0-15 cm	Se_soil	0.39	0.82	0.64	0.11	17.1
		Zn_soil	0.55	1.95	1.42	0.46	32.1
		Fe_soil	30.20	73.70	56.18	13.67	24.3
	15-30 cm	Se_soil	0.24	0.65	0.42	0.15	35.0
		Zn_soil	0.24	2.12	0.68	0.54	79.1
		Fe_soil	14.30	68.70	28.45	16.21	57.0
	30-60 cm	Se_soil	0.25	0.34	0.28	0.03	11.7
		Zn_soil	0.21	0.53	0.32	0.11	34.7
		Fe_soil	10.10	36.80	20.69	8.21	39.7

Table I.6: Descriptive statistics of Se, Zn, and Fe concentration (µg/g) of the soil for three depths in Saskatoon and Rosthern, 2017 (regardless of the blocks).

Location	Depth	Variable	Min	Max	Mean	SD	CV
Rosthern	0-15 cm	Se_soil	0.31	0.38	0.35	0.02	6.2
		Zn_soil	0.94	2.82	1.84	0.54	29.2
		Fe_soil	60.1	289	132.63	81.67	61.6
	15-30 cm	Se_soil	0.22	0.85	0.33	0.15	43.9
		Zn_soil	0.21	1.7	0.47	0.41	85.4
		Fe_soil	13.1	167	57.35	55.4	96.6
	30-60 cm	Se_soil	0.21	0.58	0.3	0.11	37.8
		Zn_soil	0.2	0.61	0.28	0.12	42.3
		Fe_soil	14.1	91.4	32.29	22.56	69.9
Saskatoon	0-15 cm	Se_soil	0.39	0.82	0.61	0.09	15.4
		Zn_soil	0.55	2.03	1.33	0.45	33.7
		Fe_soil	30.2	73.7	52.32	12.48	23.9
	15-30 cm	Se_soil	0.24	0.65	0.39	0.13	32.9
		Zn_soil	0.24	2.12	0.66	0.47	72.0
		Fe_soil	12.4	70.9	26.03	16.46	63.2
	30-60 cm	Se_soil	0.22	0.62	0.3	0.09	31.5
		Zn_soil	0.21	0.53	0.32	0.09	26.4
		Fe_soil	10.1	36.8	18.5	6.6	35.7

Discussion:

No significant correlation between soil-seed concentration could be observed for Fe and Zn in Rosthern for all the three depths. Fe correlated significantly in 0-15 cm and 15-30 cm depths in Saskatoon. Zn showed a significant correlation only in 0-15 cm depth in Saskatoon.

For Se, soil-seed concentration was significant only in the 30-60 cm depth in both locations (Table I.1).

Fe_seed concentration significantly correlated with P in 15-30 cm but not in any other combination. Zn_seed concentration did not correlate significantly with N, P, and K in soil for all the three depths in both the locations. Se_seed showed significant correlations with N, P, and K in several combinations. No significant correlation could be observed in 30-60 cm depth between N, P, and K in soil and Se, Zn, and Fe in seeds (Table I.2).

In terms of Zn and Fe concentration in seeds, the mean and CV% are more or less similar. Even if the CV% of the Se concentration in Rosthern is higher than Saskatoon (97.7% > 69.5%) (I.4), when considering the blocks, seeds grown in Saskatoon show a considerable difference between blocks (75.9% and 44.1% in block 1 and block 2, respectively) compared to Rosthern (Table I.3).

Generally, the Fe and Zn concentration in the soil decreases from top to bottom depths. Se concentration in soil between the three depths is more similar in Rosthern but not in Saskatoon (Table I.5).

However, in the soil, Zn and Fe showed much more variation across the plots than Se regardless of the block effect (Table I.6). Even if the block factor is considered as 'random' factor in general, it seems like block factor has a significant effect on the particular fixed (treatment) factor, i.e., the micronutrient concentration. However, there was not sufficient evidence to prove that fact and to identify any particular trend across blocks. Levene's test was significant for the two locations and therefore a combined analysis was not performed for the data. In conclusion, this soil study did not suggest a strong relationship between the micronutrients in seeds and the micronutrients in the soil.