THESIS

ASSESSMENT OF METHODS TO SCREEN FOR CAROTENOIDS IN YELLOW-FLESHED POTATO GERMPLASM

Submitted by

Jeremy Brandon Logrono

Department of Horticulture and Landscape Architecture

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2020

Master's Committee:

Advisor: David G. Holm
Co-Advisor: Sastry Jayanty

Adam Heuberger
Patrick Byrne

ABSTRACT

ASSESSMENT OF METHODS TO SCREEN FOR CAROTENOIDS IN YELLOW-FLESHED POTATO GERMPLASM

Rapid Evaporative Ionization Mass Spectrometry (REIMS) has the capability to rapidly perform tissue analysis without sample preparation, extractions or chromatography required. The study was conducted to evaluate REIMS as an efficient platform to identify carotenoids in yellow-fleshed potato germplasm (N = 60) from the Colorado Potato Breeding and Selection Program. The specific aim eventually is to improve selection efficiency and accelerate genetic gain in nutritional quality of potato cultivars. Phenotypic tuber flesh color (FC) rating (0 – 3), chroma values, and individual and total carotenoids data were collected, processed and combined with multivariate analyses to help in REIMS data interpretation. Results showed that orange-fleshed (FC 3) potato genotypes gave significantly higher overall carotenoid content ($P < 0.0001$) compared to the white-fleshed (FC 0), yellow-fleshed (FC 1) and dark yellow-fleshed (FC 2) genotypes. Zeaxanthin was the major carotenoid detected among the 60 selections/cultivars evaluated. The association between tuber flesh chroma and carotenoid content was analyzed. Results from Pearson correlation analysis revealed positive correlations overall. The correlation coefficient values *(r)* for lutein vs. chroma ($r = 0.56$, $P < 0.01$), zeaxanthin vs. chroma ($r = 0.60$, $P < 0.01$) and total carotenoid vs. chroma ($r = 0.63$, $P < 0.01$) were considered moderate. A metabolite mass fingerprint for each replicated sample was collected via REIMS to build a data matrix and processed to test the fit with prediction models. Multivariate methods of analysis (MVA) of principal component analysis (PCA), partial least

square (PLS) and orthogonal-PLS (OPLS) were created to determine any sample differentiation among the yellow germplasm. FC rating data (0 – 3) were integrated to MVA models as a covariate. Rep 3 samples were excluded in all MVA analyses due to high presence of noise in the raw data. PCA of Reps 1 and 2 (n = 95) showed a predictive power of 48.4% (Q^2). No apparent trends or separations based on flesh color was observed in the PCA model. PLS and OPLS supervised models illustrated better differentiation among sample components. OPLS model (n = 71) of high carotenoids (FC 3) vs. low carotenoids (FC 1 & 0) with a predictive power of 56.1% (Q^2) was considered the best model due to clear separations of high vs. low carotenoid samples. Loadings and variable importance score (VIP) data were also analyzed to rank metabolite masses that contributed to differentiation of samples, detecting mostly lipid class molecules. Precursor molecules of lutein and zeaxanthin were not detected from the REIMS analyses and carotenoid fragmentation products were most likely contributing to differentiation among samples. Further research is needed to verify identification of carotenoid fragmentation in REIMS as well as the use of more portable and cost-efficient devices.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Dr. David Holm for the opportunity to work for him in the Colorado Potato Breeding and Selection Program. I learned valuable potato breeding experience under his guidance. Thank you to Dr. Sastry Jayanty for all the guidance with carotenoid analysis and any laboratory related activities. Thank you to Dr. Adam Heuberger for his help with metabolomics and REIMS data interpretation. I would like to thank Dr. Patrick Byrne for his plant breeding and genetics courses that I enjoyed very much. I am also thankful for Dr. Corey Broeckling and Dr. Jessica Prenni for all the assistance with the REIMS instrument at PMF. Thank you to Caroline Gray, Mike Gray and Katie Gaudreau for their assistance to my research. To all SLVRC field crew, I would like to thank you for helping with all the field work. Thank you to the Horticulture and Landscape Architecture staff for providing a kind and helpful environment here in CSU. Thank you to my fellow graduate colleagues; Mansour Hamed, Esam Emragi, Tyler Mason, Sahar Toulabi and Andrew Katz for their friendship and support. Big shout out to my friends and family scattered across the globe for all their support. Mama, thank you for your love and support that gave me strength to strive for greatness. Dad, thank you for inspiring me to become the person I am today, and I will surely continue my career path in agriculture.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. LITERATURE REVIEW

1.1. Introduction

The potato plant (*Solanum tuberosum)* is a staple tuberous crop and the fourth most consumed food crop in the world (FAO 2008). The Colorado Potato Breeding and Selection Program aims to develop new varieties with increased yield, improved nutritional and health characteristics, improved quality, and resistance to biotic and abiotic stresses. The major classes of potatoes the breeding program is devoted to developing are russets, reds, and specialty varieties. Yellow-fleshed potatoes or "yellows" fall under the category specialty varieties due to their phenotype of having yellow-flesh. Carotenoids impart the yellow color of the potato flesh, and previous studies reported that carotenoids found in potatoes may prevent certain diseases including cardiovascular disease, eye disease, and certain cancers (Konschuh et al. 2005). However, current analytical methods to screen for carotenoids in potato germplasm is slow, tedious and arguably inefficient. Discovery of alternative analytical methods to improve the selection efficiency of high-carotenoid potato breeding lines is crucial to increase genetic gains.

1.2. Potato Domestication

1.2.1. Origin and History

Potatoes were domesticated approximately 8,000 – 10,000 years ago by South American indigenous people who have been cultivating the crop in the Andes mountains of Peru and Bolivia, and the origin of the plant *per se* (Hardigan et al. 2017). The domesticated environment in the Andean highlands was described as an arid region, 3,000 – 4,500 m above sea level, with cold temperatures and high solar radiation (Hardigan et al. 2017). The first written record of potato farming was dated back in 1553 in a journal by the Spanish explorer Pedro Cieza de Leon,

and it was believed that the Spanish introduced the plant in Europe between 1565 and 1580 (Acquaah 2012).

The first potatoes in Europe belonged to the sub-species *Solanum tuberosum subsp. andigena*, a short-day early maturing plant that suits the sub-tropic climate of the Spanish region of Sevilla. However, this sub-species was less adaptable for the northern regions of Europe (Tiemens-Hulscher et al. 2013). Historic records have shown that selection of the more adaptable long-day plants of *Solanum tuberosum* established its first breeding work in Europe (Tiemens-Hulscher et al. 2013). Starting from the seventeenth century, potatoes were taken from Europe and cultivated in many other parts of the world (Acquaah 2012).

Potato was introduced into North America by Irish immigrants in the early 1700's, and the first large-scale production began in Londonderry, New Hampshire, USA (Acquaah 2012). In 1851, a clergyman from New York named Chauncey Goodrich introduced a Chilean Tuberosum cultivar into the U.S. called Rough Purple Chili, which would become a valuable predecessor for many U.S. cultivars including the famous Russet Burbank, developed by the renowned plant breeder Luther Burbank in 1902 (Acquaah 2012; Spooner 2013). As of 2009, the *World Catalogue of Potato Varieties* was able to list more than 4,500 cultivars from 102 countries, showing the breeding success of potato as one of the most important crops globally (Pieterse & Hils 2009).

1.2.2. Genetic Diversity and Domesticated Traits

Potatoes are an autopolyploid species, and ploidy levels of wild and cultivated types can vary from diploid *(2n=2x=24)* up to hexaploid *(2n=6x=72)* including triploids, tetraploids and pentaploids (Spooner 2013). The major repository for potato germplasm can be found in the International Potato Center (CIP) in Lima, Peru. Potatoes have wide genetic diversity, with over 7,000 accessions of native landraces, wild accessions, and cultivars (Hijmans & Spooner 2001).

The genetic diversity of potato can be visually observed through the tubers' vast range of phenotypes of tuber size, shape, skin color and texture, and flesh color. Recent studies estimate that there are 100-110 wild species and four cultivated species, and most modern cultivars were sourced from the Chilean landraces (Spooner 2013). Wild relatives are native to south-central Chile, central Andes of Peru, Bolivia and southwestern United States (Spooner 2013). In the U.S., wild relatives can be found at Mesa Verde National Park in southern Colorado where *Solanum jamesii* is present; studies have claimed that it may possess frost and drought tolerance alleles (Bamberg 2010).

Although wild relatives of potato contain rich source of favorable traits, linkage drag is common and therefore, their introgression in modern cultivars is difficult. Moreover, the majority are not edible due to high concentrations of glycoalkaloids that makes the tubers bitter and toxic, and tuber size is often small similar to the size of a pea (Spooner 2013). By contrast, domesticated cultivars of *Solanum tuberosum* possess less toxic glycoalkaloids, increased tuber yield, adaptation to longer days, greater tuber dormancy and fixed quality traits for the processing and fresh market.

1.3. Potato Crop Production

1.3.1. General Botany

Most cultivars of potatoes (*Solanum tuberosum L.)* are tetraploids *(2n = 4x = 48)* within the nightshade family of *Solanaceae* and are annual crops with short (31 – 61 cm), erect, and branched stems (Acquaah 2012). The flowers are complete, with the stigma and anthers are intact and functional, and color range includes white, red, purple, and blue. Potatoes are primarily sel1f-pollinated, yet emasculation is common in breeding programs to control pollination to increase genetic diversity. The potato plant produces fruit berries when flowers are pollinated and if successful, the berries will contain true potato seeds (TPS). Grown TPS

seedlings are genetically unique and produce true seed tubers that can be replanted as seed tubers in the field.

Potato tubers are modified stems located underground, also known as the stolon, and the product of commercial interest. Tubers are produced by the plant as a storage compartment plus for vegetative or asexual propagation purposes, the main reproduction method for potatoes (Fernandez-Orozco et al. 2013). The "eyes" on the tuber are rudimentary leaf scars, and each eye contains at least three buds (Tiemens-Hulscher et al. 2013). The eyes are the growing points for sprout emergence, and the depth and number will depend on the variety. Some examples of phenotypic characteristics of potato tubers that may be used to confirm, and record identification include skin color, flesh color, texture, shape, eye depth and specific gravity (dry matter content).

1.3.2. Agronomic Practices

Potatoes are an annual cool-season crop and planting date varies from region to region, according to the cultivar, local climate conditions, and the targeted market use (Bohl and Johnson 2010). Typically, in Colorado, potatoes can be planted either as summer potatoes (planted in March, harvested in July or August) or winter potatoes (planted in April or May and harvested in September and October, McDonald et al. 2003). Potatoes can grow on a wide variety of soil types, but it is desirable for the soil to be well-drained. The optimal soil temperature for planting is between 7°C and 21°C during the frost-free period of the season (Bohl and Johnson 2010). Conventional tillage is used for land preparation, weed control, and to support crop growth. Crop rotation also supports accumulation of organic matter and nutrients in soils and helps control weeds and pests (Bohl and Johnson 2010). In Colorado, a two-year rotation of potatoes and malting barley is common. Proper fertilizer application is important to obtain high-yield and high-quality potatoes, and depending on the variety, adequate NPK should be applied. For example, a total rate of 200 lb N/A is recommended to cultivate Russet Burbank

for an expected yield of 400 cwt/A (Essah et al. 2014). It is also important to have a consistent planting depth and seed spacing as these are critical factors to improve plant uniformity, tuber yield and quality.

A planting technique called "hilling" is a common practice in which mounds of soils are formed around the plants to prevent tuber greening (exposure to sun), enhance tuber development, and facilitate harvest (Bohl and Johnson 2010). Use of center-pivot sprinkler irrigation is a common system for irrigation and the amount of water required depends on the cultivar. Potato tubers are harvested mechanically by harvesters, and it is crucial to handle harvested tubers with extreme care. Potato bruising during harvest is one of the major causes of economic losses for potatoes and needs to be controlled, as it reduces tuber quality and value significantly (Bohl and Johnson 2010). During post-harvest, tubers are stored at 50°F for two weeks, then held at 40°F with 96% humidity to support wound healing (McDonald et al. 2003).

1.3.3. Potato Production and Economic Importance

Potato is the fourth most important world food crop after wheat, rice, and maize due to its yield potential and nutritional value (Zaheer & Akhtar 2016). According to the *2019 Potato Statistical Yearbook*, total world production of potatoes in 2017 was 388 million metric tons. The world's largest potato producing country is China, followed by India, Russia, Ukraine and the United States (National Potato Council 2019).

Interestingly, production in developing countries has been increasing in the past number of years and accounts for about 55% of total world potato production as of 2010 (Zaheer & Akhtar 2016). This trend is associated with the significant increase of potato consumption in developing countries from 10 kg to 22 kg per capita between 1960 and 2008, in addition to the continual shift of socioeconomic statuses in these regions of the world (Avendano 2012). In India, it was estimated that total food demand for potato will increase from 20 to 30 million

metric tons as it is projected to be the world's most populous country before 2030 (Scott et al. 2019).

About 65% of the world potato production was processed into frozen products (mainly frozen fries) and 13% were used as fresh potatoes. The U.S. was the 5th global potato producer in 2018, with 1.03 million acres planted, producing 44 billion pounds with a value of $3.7 billion USD (National Potato Council 2019). Two-thirds of potato production in the U.S. is located in the western states, with Idaho and Washington accounting for more than 50% of the total production (USDA 2016). As of 2017, Idaho was the top producing state with 13.0 billion pounds while Colorado produced 2.1 billion pounds and ranked 6th nationally (National Potato Council 2019).

1.4. Potato Nutritional Attributes

1.4.1. Phytonutrients in Potatoes

Potato tubers are considered a carbohydrate-rich food. Carbohydrates comprise 75% of the total dry matter and starch is the most abundant component (Jansen et al. 2001). Potatoes are also a significant source of several essential nutrients that include protein, vitamin C and E, fiber, potassium, iron and magnesium (National Potato Council 2019). An average size potato contains about 115 calories, 1 mg of iron, 30 mg of vitamin C and 3.2 g of protein (Spooner 2013). In addition, tubers provide phytonutrients such as phenolics, anthocyanins, polyamines and carotenoids (Brown 2005). Potato is the most commonly consumed vegetable, and ideally its integration in the regular human diet can be beneficial (Lu et al. 2001).

Breeding efforts have improved current cultivars with increased antioxidants accumulating in tuber flesh and skin in response to selection of high pigment content in these parts of the tuber (Brown 2005). Carotenoids and anthocyanins in potatoes have antioxidant properties that are associated with human health benefits (Cao et al. 1999; Wang et al. 1999). In

terms of potato classes that represent each antioxidant, yellows contain the most carotenoids while purple and red-fleshed potatoes contain the most anthocyanins (Brown et al. 2004)

1.4.2. Carotenoids in Potatoes

Carotenoids are C_{40} isoprenoids (tetraterpenoids) lipophilic pigments that are synthesized by plants. They have a role in photosynthesis, signaling, defense against abiotic and biotic stresses, and attraction of animal pollinators (Tanaka et al. 2008). The biosynthesis pathway of carotenoids starts with the condensation of four C_5 isoprene unit (IPP) to form C_{20} geranylgeranylpyrophosphate (GGPP). The phytoene synthase (PSY) is responsible for coupling two C_{20} GGPP molecules to yield the first C_{40} tetraterpenoid, phytoene, which is eventually modified by the phytoene desaturase (PDS) and carotene desaturase (SDS) to yield the intermediate product of lycopene in the pathway (Tanaka et al. 2008). The key regulatory gene for biosynthesis of carotenoids is the Lycopene epsilon cyclase (*LcyE*), which initially acts as a "steering wheel" to flux either the α-carotene or β-carotene branch pathway (Chandler et al. 2013). Low expression of *LcyE* increases the flux towards the β-carotene branch and vice versa. Some examples of carotenoid hydroxylation products in the β-carotene branch are zeaxanthin, antheraxanthin, and violaxanthin, while in the α-carotene branch includes lutein (Chandler et al. 2013).

Study reports have shown that there are several health benefits of carotenoids when incorporated in the human diet. Lists of claimed health benefits include: some pro-vitamin A activity, enhanced immune functions and anti-inflammatory properties, reduction of chronic diseases of certain cancers due to the antioxidant properties that may protect cells from oxidative damage, and reduction of cardiovascular diseases and age-related macular degeneration (AMD), (Brown 2005; Liu 2003). The action of carotenoids against certain diseases has been attributed to the antioxidant capability to quench singlet oxygen and interact with free radicals that damage

DNA, and other parts of the cell in the human body (Palozza & Krinsky 1992). AMD is a progressive blinding disease and the number one cause of visual impairment in the aging population in the developed world and the third globally (DeAngelis et al. 2017). Since humans cannot produce carotenoids, incorporation of the specific carotenoids of lutein and zeaxanthin in the diet have been shown to reduce risks of developing AMD (Seddon et al. 1994) as both carotenoids are components of the human retina and must be sourced from food for proper eye health (Brown 2008). Furthermore, studies have shown that lutein and zeaxanthin carotenoids are more bioavailable than others, signifying the potential of these carotenoids to be absorbed in the human body (Saini et al. 2015). Bioaccessibility of lutein and zeaxanthin in yellows ranged from 76-82% and 24-55%, respectively, and studies have shown that addition of dietary fat can enhance bioaccessibility for both carotenoids (Lachman et al. 2013).

There are about 750 known carotenoids that occurs naturally in plants, and xanthophylls are the most abundant class found in potatoes (Britton et al. 2004). Previous studies showed that there are about six major carotenoids detected in most modern varieties: lutein, zeaxanthin, violaxanthin, neoxanthin, antheraxanthin and beta-cryptoxanthin (Lu et al. 2001; Brown et al 2007; Burgos et al. 2009; Kotíková et al. 2016). However, only a small trace of beta-carotene was detected, exposing the fact that potatoes are not a significant source of pro-vitamin A (Brown 2005; Burgos et al. 2009). However, through metabolic engineering using transgenic technologies, beta-carotene can be expressed with a 20-fold increase in so-called “golden potatoes” (Diretto et al. 2007). Lutein and violaxanthin are the primary carotenoids detected in yellow-fleshed potatoes; 54-93% of the total carotenoids predominantly detected were lutein (Brown et al. 2007; Lachman et al. 2016), while lutein and zeaxanthin were the primary carotenoids detected in dark yellow to orange-fleshed potatoes (Brown et al. 1993; Haynes et al. 2010).

1.5. Breeding for High Carotenoid Yellow Potatoes

1.5.1. Yellow Potato Germplasm

Yukon Gold variety was introduced in the 1980s and established the niche of yellows in North America. Since then, there has been a growing market for growers and consumers (Konschuh et al. 2005). From 2016 to 2018, the percent of potato acreage planted in Colorado to yellow types increased from 8% to 10%, showing a steady incentive for growers to produce yellow varieties (USDA-NASS 2019). Therefore, breeding efforts to improve highly nutritional yellow-flesh varieties with high concentrations of carotenoids is of research interest in potato breeding programs to meet these demands.

The majority of current white-fleshed commercial cultivars only have a range of total carotenoids of 50 to 100 μg/100 g fresh weight (FW, Brown 2008), whereas the total carotenoid content of current yellow cultivars can range as high as 560 μg/100 g FW (Lu et al. 2001). However, through integration of landrace germplasm to potato breeding programs, there is a potential to increase carotenoid levels drastically. Native cultivars of the *Solanum tuberosum subsp. andigena* group developed by farmers in the Andes have been reported to contain carotenoid levels ranging from 535 to 3,895 μg/100 g FW (Andre et al. 2007), and diploid *Solanum phureja* group of the *Papa Amarilla* yellow population of the same geographic region contain carotenoid levels ranging from 800 to 2,000 μg/100 g FW (Brown 2008). The potato genotype *Yema de Heuvo* from the *Solanum phureja* group is an example of a yellow landrace germplasm used as a parental material to develop highly yellow and high carotenoid breeding lines in potato breeding programs in the U.S. (Brown et al. 2006; Haynes et al. 2011).

1.5.2. Inheritance of the Yellow Flesh Trait

The yellow versus white flesh color trait is considered to be controlled by a single dominant gene *(Y/y)* mapped to chromosome 3, where inheritance of allele *Y* (yellow) is

dominant to allele *y* (white, Bonierbale et al. 1988). As previously mentioned, lutein and violaxanthin are the most abundant carotenoids in yellow-fleshed genotypes while zeaxanthin and lutein are highly concentrated in dark yellow to orange-fleshed genotypes. Studies by Brown and colleagues identified a distinct allele *Or* at the *Y* locus with an orange flesh color trait from a diploid hybrid population of *Solanum phureja x Solanum stenotomum,* claiming that *Or* is dominant to the alleles of *Y* and *y* (Brown et al. 1993). Once again, incorporation of diploid landrace germplasm is essential to produce dark yellow to orange-fleshed potato lines with high levels of zeaxanthin (Brown 2008; Brown et al. 1993).

1.6. Analytical Methods for Analysis of Carotenoids

1.6.1. Colorimeter and Spectrophotometer

Since many studies have indicated that consumption of high carotenoid potatoes can protect consumers from various diseases, considerable interest is being shown by the scientific community to screen and develop potato cultivars with increased concentrations of total and individual carotenoids. Advances in analytical techniques have made it possible to identify, quantify and understand the biological functions of carotenoids in potatoes. Use of colorimetric measurements on tuber flesh is a tool for researchers to collect data on color values associated with carotenoid levels (Rodriguez-Amaya & Kimura 2004; Brown et al. 1993). Reflectance colorimeters can collect color data by generating a set of Cartesian coordinates that pinpoint the measured color in three-dimensional color space (McGuire 1992). The interpretation of the color data will be described in more detail in the colorimetric methods and analysis section.

A spectrophotometer is another useful tool to collect and estimate organic compounds in potatoes like protein, starch, glucose and potentially carotenoids (Burgos et al. 2009). Unlike the colorimeter, spectrophotometers collect color-space coordinates in a more specific target by measuring the reflectance of the organic compound of interest throughout the visible spectrum

(typically at 380 ~ 780 nanometers), (Voss 1992). In particular, 450 nanometers (nm) is the standard wavelength to detect and quantify carotenoids absorbance by using spectrophotometers. (Bonierbale et al. 2009; Lu et al. 2001). Some of the advantages of application of colorimeters and spectrophotometers to estimate carotenoids in potatoes include their rapid assay, non-destructive, and relatively inexpensive approaches (Bonierbale et al. 2009).

1.6.2. High-Performance Liquid Chromatography

The xanthophyll carotenoids found in potatoes are C_{40} in length and contains a series of conjugated double bonds, ketones, aldehydes, and hydroxyl groups, making them more polar in nature and thus extractable into ethanol solvents (Fraser et al. 2007). The use of the analytical chemistry method of high-performance liquid chromatography (HPLC) is ideal to detect and quantify individual (specific) carotenoids (Burgos et al. 2009; Fraser et al. 2007). Using HPLC, individual carotenoids can be separated by their polarity, which causes them to elute at different retention times. The polarity of the liquid mobile phase will gradually be eluted from the column, a narrow tube that contains a polymer where analytes will be retained known as the stationary phase and the time of separation will be detected by the HPLC detector. Analytes are the modified biological metabolites of carotenoids prior to detection. The detector then provides quantification of analyte abundance under the chromatography peak area in arbitrary units (AU) and through external calibration of using authentic carotenoid standards, actual quantity (μg /mL) can be collected. Furthermore, carotenoids are pigments that have the characteristic of the ultraviolet-visible spectra, hence photo-diode array (PDA) detectors are common to be coupled to the HPLC apparatus in detection analysis (Fraser et al. 2007).

The analytical method flow to screen for carotenoids in HPLC consists of experimental design, extraction, detection and quantification, data analysis and processing. However, extraction can be the limiting step to screen for carotenoids in the HPLC workflow. Extraction

involves multistep sample preparation of lyophilization and grinding of potato flesh, adding a mixture of solvents, and as well as the need to control the environmental variables such as temperature and light that may affect the stability of carotenoids (Kotíková et al. 2016). Simply put, the extraction step is tedious, time-consuming and can be inefficient depending on the protocol. Therefore, this study explores an alternative analytical method to HPLC to skip the extraction step and improve efficiency for screening for carotenoids in yellow potato germplasm.

1.7. Rapid Evaporative Ionization Mass Spectrometry

Because of the tedious extraction process required for carotenoid analysis, this study was designed to evaluate the novel technology of Rapid Evaporative Ionization Mass Spectrometry (REIMS) developed by the Waters Corporation (Milford, MA). REIMS technology is relatively new in biological sciences, and its promising rapid tissue analysis is becoming an area of interest by researchers. Initially, the earliest application of REIMS was as a handheld "intelligent knife" (iknife) by surgeons as a detection tool that tells them whether the human tissue, they are removing is cancerous in real-time (Alexander et al. 2017). REIMS also demonstrated that the novel technology could classify and detect differentiation between meat species for beef and fish, providing meat authenticity at high precision and efficiency (Balog et al. 2016; Black et al. 2017). There is no current scientific publication on the use of REIMS in potato improvement, thus providing an exciting opportunity for this research field.

REIMS works by enabling high conductivity of the monopolar cutting electrode (iknife) that generates high temperature as it makes contact to flesh samples, producing vapor or smoke that is taken up by a quadrupole time-of-flight mass spectrometry (Q-TOF-MS). It takes only a few seconds to collect tissue for metabolomic identification (Paxton 2020). This is how the term rapid evaporation was derived, requiring samples to have a high moisture content to allow effective conductivity. In contrast to HPLC, REIMS coupled with Q-TOF collects and measures

a fingerprint of analyte mass/charge ratio (m/z) based on the time the ionized molecule takes to travel between two places within a tube, and separation and detection is based on molecule size (molecules with smaller m/z will travel more quickly than larger ones, Paxton 2020). Basically, REIMS fingerprint profiling provides *in situ,* real-time data by ionizing biological tissue samples with no sample preparation.

Identification and quantification can be challenging for xanthophyll carotenoids due to its sensitivity in the presence of light, heat or oxygen, narrowing the appropriate analytical methods for their analysis. The chemical structures of xanthophylls have been attributed to be oxidized much more easily due to the presence of oxygen with lone electron pairs, and addition of cations (positive charge) can be appropriate (Guaratini et al. 2005). Ionization techniques coupled with mass spectrometry have been applied before to analyze carotenoids such as the use of atmospheric pressure chemical ionization (APCI) and electrospray ionization (ESI), (Rezanka et al. 2009). Previous studies have shown that ESI-MS in positive-ionization mode was capable of protonation ($[M]^+$, $[M+H]^+$) to detect fragmentation and parent molecule products of xanthophylls such as lutein and zeaxanthin (Rezanka et al. 2009; Guaratini et al. 2005; Rivera et al. 2013). Therefore, REIMS technique in positive-ionization mode may potentially detect carotenoids.

There are two major advantages of the application of REIMS to potentially screen carotenoids in yellow potato germplasm. First, there will be no requirement for extraction, sample preparation or chromatography, creating an opportunity of eliminating the most time-consuming step and arguably the biggest incentive of using this novel technology by improving overall protocol efficiency. Second, data collection of metabolites that determine differences between samples is instant. REIMS as a direct analysis for rapid phenotyping in potato may

potentially screen hundreds of yellow breeding lines in a short period of time, improving the selection efficiency of high-carotenoid genotypes significantly.

1.8. Conclusion

Genetic gain from selection is defined as the improvement in average phenotypic value of the trait of interest per unit of time (Rutkoski 2019). In this case, it is the improvement of selection for high-carotenoid yellow-fleshed potato breeding lines within a population over cycles of breeding. In any plant breeding program, innovation such as incorporation of high-throughput phenotyping (HTP) technologies is always encouraged to increase the efficiency of phenotypic evaluation to increase genetic gains. REIMS may potentially allow a breeding program to significantly speed up evaluation and selection of a large number of potato genotypes. Essentially, the goal is to verify the hypothesis of this study, which was to assess REIMS's sensitivity and efficiency to detect lutein and zeaxanthin carotenoids in real-time among yellow-fleshed potato germplasm. Color and carotenoid information collected via analytical methods will be combined with multivariate analyses, with an aim to support REIMS data interpretation. Overall, this thesis will focus on the following objectives:

1. Evaluate Rapid Evaporative Ionization Mass Spectrometry (REIMS) as a novel method to screen for carotenoids in yellow-fleshed potato germplasm;
2. Evaluate and compare the tuber-flesh color of 60 selected yellow genotypes in the Colorado Potato Breeding and Selection Program;
3. Measure total and individual carotenoid content for the 60 selected yellow genotypes;
4. Integrate collected color and individual carotenoid data to multivariate analyses to help with REIMS data interpretation.

CHAPTER 2. PLANT MATERIALS

2.1. Introduction

Potato genotypes in this study were sourced from the 2018 growing season of the Colorado Potato Breeding and Selection Program at San Luis Valley Research Center (SLVRC), an Agriculture Experiment Station (AES) of Colorado State University (CSU). The SLVRC is located near Center in southern Colorado and is an ideal target production environment for a potato breeding program due to its high altitude at 7600 ft, mild summer temperatures and dry climate. Sixty potato genotypes were processed and analyzed to collect flesh color, total and individual carotenoids, and REIMS data.

2.2. Cultural Management and Tuber Preparations

Planting, vine kill, harvest dates and number of days grown from planting until vine kill are provided in Table 2.1. The soil type in SLVRC field plots is a sandy loam classified as Dunul cobbly sandy loam. Row plots were spaced 86 cm (34") apart with in-row plant spacing of 30 cm (12") between hills. Fertilizer (N-P-K-S-Zn) application rate of pounds per acre (lb /A) was 143-35-20-17-1, and center pivot irrigation was used with a 27 cm (10.7") gross application rate in 2018. Rainfall for the duration of the growing season was 8.8 cm (3.47").

Harvest bags were prepared and labeled with the clone name, plot field number, and number of tubers to be harvested. For this study, ten potato tubers per genotype were harvested, totaling 600 potato tubers. Harvested tubers were stored in a cold storage (4.4°C/40°F) with a relative humidity of 95% until further processing and analysis.

2.3. Plant Material Overview

The 60 selected potato genotypes consist of tubers at different stages of selection to represent the yellow population of the breeding program, in addition to selected released check varieties. Of the 60 selected potato lines, 53 yellow lines were identified from the preliminary, intermediate and advanced selections, four yellow-fleshed check cultivars of Yukon Gold, Masquerade, Agria and Inka Gold from the maintenance plots (seed tuber increase plots). Two check cultivars of Russet Nugget and Chipeta were selected for their white-flesh to serve as controls, and one purple-flesh cultivar of Purple Majesty was included to determine if REIMS can detect anthocyanins.

All genotypes were tetraploids, with the exception of two yellow lines of BDC701-4-3W/Y and BDC701-1-1W/Y which were diploids. Many of the selected parents from the yellow population have a pedigree that includes the *Solanum phureja* group of Inca Dawn, Inka Gold, and Mayan Gold. These yellow germplasms have been introduced to the breeding program to increase the genetic diversity of breeding high-carotenoid yellow potatoes.

2.4. Phenotypic Scoring of Flesh Color

Tuber flesh color for the selected genotypes were phenotypically scored based on the standard color chart ratings used in the Western Regional Potato Variety Trials. A qualitative score scale of 0 – 3 (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange) was implemented to rate tuber flesh color. Initially, Russet Nugget (FC 0) was used as the white flesh control, Yukon Gold (FC 1) as yellow, CO13033-2W/Y (FC 2) as dark yellow and CO13033-12W/Y (FC 3) as orange. Listed genotypes above were used as the standard guide to visually score the rest of the selected genotypes to collect tuber flesh color rating data. The application of this tuber flesh color (FC) rating system is exceptionally advantageous in the breeding program, as the method itself is simple, fast, and can be performed at the field during harvest by cutting tubers in half to

evaluate flesh color. Potato tuber flesh color data were collected among the 60 yellow germplasm: white (n = 2), yellow (n = 25), dark yellow (n = 16) and orange (n = 17, Table 2.2). In addition to flesh color and its rating, tuber shape, trial stage and parental information are shown in Table 2.2.

2.5. Chapter 2 Summary

Potato tuber genotypes were carefully selected based on uniformity of shape and size, and quality appearance of disease/damage free tubers. Phenotypic data of flesh color (FC) rating will be widely used as a descriptive covariate to diagrams and statistical analyses. Since tubers were obtained from different stages of the selection, there is an observable phenotypic variance of tuber flesh color, tuber shape and size, and skin color among selected genotypes.

Table 2.1. Growing season of the potato germplasm at SLVRC (Center, CO).

	2018
Planting Date	15-May
Vine Kill Date	16-Aug
Harvesting Date	11-Sept
Number of Days Grown	93

Table 2.2. Characteristics of the 60 potato genotypes harvested during the 2018 growing season at SLVRC.

Clone	Shape	Flesh Color	FC Rating[a]	Trial stage	Parents[b]
AC05175-3P/Y	Oval	Yellow	1	Intermediate	A99331-2R/Y x COA99261-1RY
AC06908-1W/Y	Oval	Dark yellow	2	Advance	Inka Gold x Reiche
AC08172-2W/Y	Round	Dark yellow	2	Advance	AH57-80 x EGAO9703-4Y
Agria	Oblong	Yellow	1	Maintenance	Quarta x Semlo
BDC701-4-3W/Y	Oblong	Orange	3	Advance	N/A
BDC704-1-1W/Y	Round	Orange	3	Advance	N/A
Chipeta	Round	White	0	Maintenance	WNC612-13 x Wischip
CO05037-2R/Y	Long	Dark yellow	2	Intermediate	Midnight Moon x CO97227-2P/PW
CO05037-3W/Y	Oval	Yellow	1	Intermediate	Midnight Moon x CO97227-2P/PW
CO07044-3RU/Y	Round	Yellow	1	Advance	AC00550-4RU x PA4X137-12
CO07131-1W/Y	Round	Orange	3	Advance	PA4X137-12 x 4X91E22
CO08155-2RU/Y	Oblong	Yellow	1	Advance	Fortress Russet x Innovator
CO09128-3W/Y	Oval	Dark yellow	2	Advance	A00293-2Y x CO03060-2W/Y
CO09128-5W/Y	Round	Dark yellow	2	Advance	A00293-2Y x CO03060-2W/Y
CO09218-4W/Y	Oval	Dark yellow	2	Advance	ATC00293 -1W/Y x CO00412-5W/Y
CO10098-5W/Y	Round	Dark yellow	2	Advance	CO04099-3W/Y x CO04099-4W/Y
CO11252-1W/Y	Oval	Dark yellow	2	Intermediate	CO99045-1W/Y x POR02PG37-2
CO11266-1W/Y	Oval	Yellow	1	Intermediate	CO03060-2W/Y x A99433-5Y
CO11324-2W/Y	Round	Dark yellow	2	Intermediate	CO04099-4W/Y x CO04029-5W/Y
CO13011-1RU/Y	Long	Yellow	1	Preliminary	Fortress Russet x CO07021-2RU/Y
CO13019-1W/Y	Long	Yellow	1	Advance	Fortress Russet x Inca Dawn
CO13019-2W/Y	Long	Yellow	1	Advance	Fortress Russet x Inca Dawn
CO13019-3W/Y	Oblong	Yellow	1	Advance	Fortress Russet x Inca Dawn
CO13033-10W/Y	Oblong	Orange	3	Preliminary	AC06358-2W/Y x Inca Dawn

CO13033-12W/Y	Round	Orange	3	Preliminary	AC06358-2W/Y x Inca Dawn
CO13033-1W/Y	Oval	Dark yellow	2	Preliminary	AC06358-2W/Y x Inca Dawn
CO13033-2W/Y	Oblong	Dark yellow	2	Preliminary	AC06358-2W/Y x Inca Dawn
CO13033-4W/Y	Oval	Dark yellow	2	Preliminary	AC06358-2W/Y x Inca Dawn
CO13033-5W/Y	Round	Orange	3	Preliminary	AC06358-2W/Y x Inca Dawn
CO13033-9W/Y	Round	Dark yellow	2	Preliminary	AC06358-2W/Y x Inca Dawn
CO13050-1RU/Y	Long	Yellow	1	Advance	CO05040-1RU x Inca Dawn
CO13069-8RU/Y	Long	Yellow	1	Preliminary	CO05068-1RU x Inca Dawn
CO13101-3RU/Y	Oval	Yellow	1	Advance	CO07039-3RU/Y x CO07044-3RU/Y
CO13101-5RU/Y	Oblong	Yellow	1	Advance	CO07039-3RU/Y x CO07044-3RU/Y
CO13101-7RU/Y	Oval	Yellow	1	Advance	CO07039-3RU/Y x CO07044-3RU/Y
CO13126-1RW/Y	Round	Dark yellow	2	Advance	CO07131-1W/Y x PA99P2-1
CO13127-2RW/Y	Round	Orange	3	Preliminary	CO07131-1W/Y x Inca Dawn
CO13127-6W/Y	Round	Orange	3	Preliminary	CO07131-1W/Y x Inca Dawn
CO13143-1RU/Y	Long	Yellow	1	Advance	CO07249-1RU/Y x CO07039-3RU/Y
CO13143-3RU/Y	Long	Yellow	1	Advance	CO07249-1RU/Y x CO07039-3RU/Y
CO13175-2RW/Y	Round	Dark yellow	2	Preliminary	PA99P2-1 x CO07131-1W/Y
CO13188-1RU/Y	Long	Orange	3	Advance	Mayan Gold x CO07131-1W/Y
CO13188-2RU/Y	Long	Orange	3	Advance	Mayan Gold x CO07131-1W/Y
CO14226-1W/Y	Round	Yellow	1	Preliminary	AC07315-1W/Y x BDC704-1-1W/Y
CO14226-2W/Y	Round	Yellow	1	Preliminary	AC07315-1W/Y x BDC704-1-1W/Y
CO14226-3W/Y	Round	Yellow	1	Preliminary	AC07315-1W/Y x BDC704-1-1W/Y
CO14274-4W/Y	Oval	Orange	3	Preliminary	BDC701-4-3W/Y x OR04198-1
CO14274-5W/Y	Round	Orange	3	Preliminary	BDC701-4-3W/Y x OR04198-1
CO14274-6W/Y	Round	Orange	3	Preliminary	BDC701-4-3W/Y x OR04198-1
CO14282-3RW/YP	Round	Orange	3	Preliminary	BDC704-1-1W/Y x CO07131-2W/Y
CO14282-4R/YR	Round	Orange	3	Preliminary	BDC704-1-1W/Y x CO07131-2W/Y
CO14369-1RU/Y	Round	Yellow	1	Advance	CO07044-3RU/Y x CO09128-5W/Y
CO14479-4W/Y	Round	Yellow	1	Preliminary	Yukon Nugget x CO09128-5W/Y
CO97237-5RU/Y	Oval	Yellow	1	Advance	CO94222-6 x CO94161-2
Inka Gold	Round	Dark yellow	2	Maintenance	89S104-4 x Mi Peru
Masquerade	Round	Yellow	1	Maintenance	Inka Gold x A91846-5R
OR04198-1	Round	Orange	3	Advance	POR02PG4-1 x PA99P2-1
Purple Majesty	Oval	Purple	0	Maintenance	ND2008-2 x All Blue
Russet Nugget	Long	White	0	Maintenance	Krantz x AND71609-1
Yukon Gold	Oblong	Yellow	1	Maintenance	N/A

[a]Flesh color (FC) ratings: 0 – 3 (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange).
[b]Parents that were crossed to produce listed progeny clones (female parent x male parent).

CHAPTER 3. COLOR AND CAROTENOID CONTENT ANALYSIS

3.1. Introduction

Potato tubers are one of the richest sources of antioxidants among staple food crops, consisting of the major antioxidants of ascorbic acid, anthocyanins, tocopherols, polyphenols and carotenoids (Perla et al. 2012). Previous research has shown that the yellow-flesh color in tubers are imparted by certain carotenoids, particularly the prevalence of the xanthophylls of lutein and zeaxanthin (Brown et al. 1993). Intensity of yellowness of tuber flesh color is moderately to highly associated with carotenoid content, signifying that colorimetric measurements of tuber flesh color can be useful for selection to develop high carotenoid yellow-fleshed potato breeding lines (Lu et al. 2001; Lachman et al. 2016). A reflectance colorimeter is an essential tool to collect quantitative flesh color data. Devices such as Konica Minolta chroma meters and HunterLab colorimeters are widely used to collect tissue color data not only in potatoes but also in other crops and vegetables.

Moreover, the research interest of increasing carotenoid content in potatoes to benefit consumers' health relies on analytical methods (De Ritter & Purcell 1981; Brown et al 1993). Identification and quantification of total and individual carotenoid content of the selected yellow potato germplasm in the breeding program can provide essential data to guide the decision-making of selection to advance yellow breeding lines. Use of near-infrared reflectance spectroscopy (NIRS) and microplate reader spectrophotometers are useful devices to estimate total carotenoids in potato germplasm (Burgos et al. 2009). Likewise, use of high-performance liquid chromatography (HPLC) is more applicable for estimating individual carotenoids (Lu et al. 2001; Lachman et al. 2016; Kotíková et al. 2016)

3.2. Colorimetric Methods and Analysis

3.2.1. HunterLab Colorimeter and Protocol

To collect quantitative tuber flesh color data of the 60 selected yellow germplasm, a HunterLab MiniScan XE colorimeter (HunterLab, Reston, VA) was used. Initially, three potato tuber replicates per genotype totaling to 180 samples were used for the colorimetric measurements sourced from the cold storage located at SLVRC. Potato tubers were removed from the cold storage on the day of data collection and placed at room temperature. The HunterLab MiniScan XE colorimeter was calibrated using the recommended white and black standard tiles prior to colorimetric scans on tuber flesh. Tubers were sliced in half and flesh color measurements were taken at the bud, center, and stem end of each tuber half. The darker pith area down the center of the tuber was avoided. Each tuber half had one random scan on the designated areas (three tuber replicates per genotype).

3.2.2. L, a*, b*, and Chroma Values*

The data output for each flesh color scan from the HunterLab colorimeter produces three-dimensional color space coordinate values of L*, a*, and b* (Voss 1992; HunterLab 2003). The L – axis measures lightness and values range from 100 for white, and 0 for black. For the a – axis, positive values are red, negative values are green and 0 is neutral. For the b – axis, the positive values are yellow, negative values are blue and 0 is neutral (HunterLab 2003). Chroma implies the intensity or saturation of a color, and a* and b* coordinate values can reflect the intensity of the yellow color when calculated (Brown et al. 1993, McGuire 1992). The formula to calculate chroma is $[a^2+b^2]^{1/2}$ and has been used as the quantitative colorimeter data for yellow-fleshed genotypes to correlate with carotenoid content (Brown et al. 1993; Lu et al. 2001).

3.2.3. Summary of the Mean of the Colorimeter Measurements

As seen in Table 3.1, mean and standard errors of the colorimeter measurements of the white, yellow, dark yellow and orange flesh types were calculated. White-fleshed genotypes (FC 0) have a relatively higher mean L* value of 62.07 ± 0.56 than the rest of the more yellow flesh types. Orange-fleshed genotypes (FC 3) on the other hand had the highest mean positive a* and b* values of 5.84 ± 0.60 and 33.17 ± 0.23 respectively, leaning on the color associated to more red and yellow. In addition, dark yellow (FC 2) and orange flesh types (FC 3) showed the highest mean chroma values of 28.55 ± 0.76 and 34.57 ± 0.25 respectively, demonstrating the high yellow intensity in flesh color on the designated genotypes. By contrast, white flesh types only showed a mean value of 15.94 ± 0.67 for chroma. Figure 3.1 shows the complete chroma value data associated to each genotype in this study.

Table 3.1. Colorimeter measurements[a] of L* a* b* and chroma values among the four flesh types.

Flesh Color	FC rating[b]	n[c]	L*	a*	b*	Chroma[d]
White	0	6	62.07 ± 0.56	-1.08 ± 0.13	15.90 ± 0.11	15.94 ± 0.67
Yellow	1	75	58.16 ± 0.60	-1.43 ± 0.34	26.25 ± 0.63	26.29 ± 0.85
Dark Yellow	2	48	56.43 ± 0.45	0.07 ± 0.03	28.52 ± 0.32	28.55 ± 0.76
Orange	3	51	55.60 ± 0.71	5.84 ± 0.32	33.17 ± 0.23	34.57 ± 0.25

[a]Mean and standard error (SE) of the colorimeter measurements data of L* (lightness), a* (positive = red, negative = green), b* (positive = yellow, negative = blue) and chroma (intensity of yellowness).
[b]Flesh color (FC) rating: 0 – 3 (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange).
[c]Sample size (n) of the replicated tubers (3 rep tubers per genotype, totaling to 180 samples) with the designated flesh color and FC rating.
[d]Formula to calculate chroma values = $[a^2+b^2]^{1/2}$

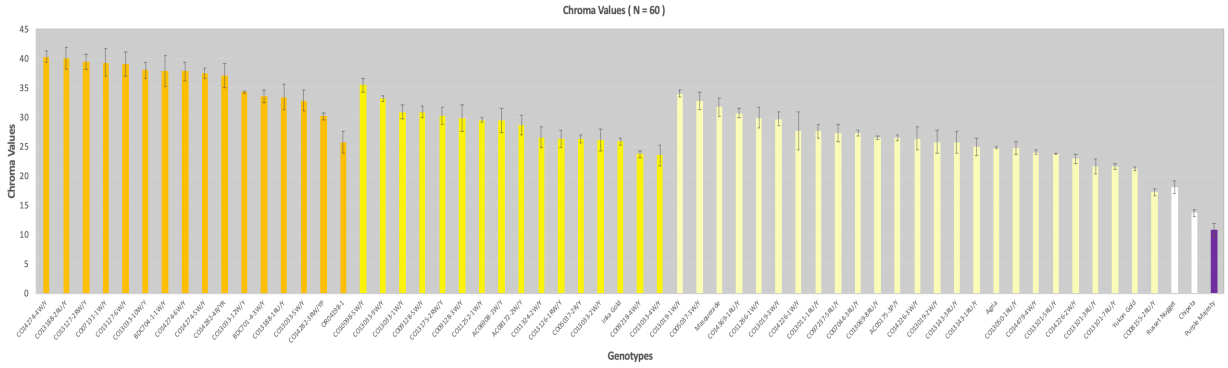

Figure 3.1. Bar chart of chroma values among 60 genotypes. Bins colored based on the FC rating score 0 – 3 (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange). Error bars represent SE (standard error of the mean, 3 reps).

3.3. Carotenoid Content Analysis

3.3.1. Experimental Design and Extraction Protocols

The analytical method workflow to screen for individual carotenoids in HPLC starts with the experimental design of randomized complete block design (RCBD) with three tuber replicates per potato genotype (180 total tuber samples). Also, all HPLC analysis was performed with three injection replicates (RCBD) per genotype (180 total injections). Pre-cut potato tuber samples were sourced from the -80°C freezer with labeled replication number. The tubers used for the HPLC analysis were also the same tubers from the REIMS analysis, which will be discussed more in detail in the REIMS materials and methods section.

The carotenoid extraction protocol was based on Kotíková et al. (2016). Since carotenoids are sensitive to light, it was critical to minimize the light environment during extraction protocols which was achieved with a low light setting and wrapping plastic falcon tubes with foil. Potato tubers were cut to smaller cubes, then put in a freeze dryer (Labconco FreeZone 6, Kansas City, MO) for lyophilization for 72 hours until samples were dry. Freeze-dried tuber cubes were ground to a powdered form using a coffee blender. In brief, 1.0 g of freeze-dried potato powder was weighed into a 50 mL falcon tube labeled with the replication number and was extracted with 10 mL of acetone overnight in a dark refrigerator (at 4°C).

Samples were then centrifuged (Beckman J2-21 Centrifuge, Palo Alto, CA) for 7000 rpm, 4°C, for 30 minutes. Ten mL of the supernatant was collected and transferred to a new 50-mL plastic falcon tube, and the extraction procedure with 10 mL of acetone overnight was performed again. Use of internal standard normalization of trans-b-apo-8-carotenal (Sigma-Aldrich Corporation, St. Louis, MO) was also added to all extract samples to minimize bias and variation among samples by accounting for the variation in presence tuber flesh tissues (Brown et al. 1993; Konschuh et al. 2005). The combined volume of 20 mL supernatant with internal

standard concentration of 10 µg/mL was then evaporated under vacuum using the Rotavapor R-200 (Büchi Labortechnik, AG, Flawil, Switzerland) for 120 rpm at 40°C water bath. The final extract was then reconstituted with 2 mL ethanol: acetone (3:2) mixture with the addition of 0.2% butylated hydroxytoluene (BHT). The samples were filtered through a syringe filter (PVDF, 0.45 µm) and transferred into 2 mL Eppendorf tubes and stored at -20°C freezer until injection.

3.3.2. HPLC Materials, Methods and Chromatographic Separation

To identify and quantify individual carotenoids of the listed yellow germplasm, a Water 2695 HPLC equipped with Water 996 PDA detector (Waters Corporation, Milford, MA) was used. The analytes were separated by a YMC C30 carotenoid column (250 mm x 3.0 mm, S-3 µm, YMC, Wilmington, USA) with a column heater set to 25°C and a flow rate of 0.6 mL/min. Vials were labeled based on the RCBD, and a 10 µL sample was injected into the HPLC.

The spectral acquisition was set to 300-700 nm with detection at $\lambda = 445$ nm, and total run time per injection of 15 minutes. Mobile phase consisted of methanol (A), and tert-butyl methyl ether (B). For carotenoid analysis, the gradient elution of the mobile phase consists of 80:20 (methanol: tert-butyl methyl ether) and has the following conditions: initial conditions of 80% A, and 20% B were kept constant for 1 minute, then gradually increasing to 70% A, 30% B at 5 minutes, reaching 40% A and 60% B at 15 minutes (Kotíková et al. 2016). Column condition calibration and wet prime flush in the HPLC apparatus were implemented every ten sample injections.

3.3.3. Individual Carotenoid Content Analysis and Results

Because HPLC analysis is capable of targeted metabolomics for individual carotenoids, authentic external standards were used to develop standard curves to quantify actual carotenoid

content. Lutein standard was obtained from Cayman Chemical Company (Ann Arbor, MI), zeaxanthin and violaxanthin from Sigma-Aldrich Corporation (St. Louis, MO), and antheraxanthin, neoxanthin, beta-cryptoxanthin from DHI Water & Environment (Hørsholm, Denmark). All standards used in this study had a purity of (≥ 95%) and were diluted in ethanol as recommended by the manufacturers.

In the carotenoid detection analysis for all extracted samples, lutein, zeaxanthin, violaxanthin and the internal standard of trans-b-apo-8-carotenal were detected. Antheraxanthin, neoxanthin and beta-cryptoxanthin were not detected in this study. This implies the variation and inconsistency of detected individual carotenoids among other publications with different protocols (Lu et al. 2001; Brown et al 2007; Burgos et al. 2009; Kotíková et al. 2016). Table 3.2 shows sample retention times of each detected carotenoid standards. Individual carotenoid content was calculated based on the external standard curves of the following equations:

1. lutein ($y = 41451x + 4134.4$, $R^2 = 0.9993$),
2. zeaxanthin ($y = 53368x – 1059$, $R^2 = 0.9982$),
3. violaxanthin ($y = 13482x – 3620.3$, $R^2 = 0.9932$),
4. trans-b-apo-8-carotenal ($y = 13482x – 3620.3$, $R^2 = 0.9973$).

In addition, individual carotenoid concentrations were adjusted and calculated relative to the internal standard concentration with the following ratio:

$$\frac{\text{Peak area (AU) of detected carotenoid}}{\text{Peak area (AU) of internal standard.}}$$

Quantified individual carotenoid content of the yellow potato germplasm has been expressed in micrograms (µg) of the particular analyte per 1.0 g of dry matter (µg/g DM). Detected carotenoid contents of lutein and zeaxanthin among 59 genotypes (excluding Purple Majesty) were illustrated in Figures 3.2 and 3.4, respectively. For violaxanthin content, 53

genotypes were detected and illustrated in Figure 3.6. Bar chart bins have been colored based on the FC rating score (0-3) for each genotype. By visual estimation from the bar charts, there was a consistent trend of FC rating of (FC 3) genotypes to have the highest concentrations of lutein and zeaxanthin and vice versa for FC ratings of (FC 1) and (FC 0) genotypes (Figure 3.2 & Figure 3.4). However, violaxanthin showed inconsistent concentrations among FC ratings (Figure 3.6).

For both lutein and zeaxanthin content, the top and bottom ten yellow potato lines were listed in Figures 3.3 and 3.5 respectively. Lutein content ranged from 1.03 µg/g DM (Russet Nugget) to 25.20 µg/g DM (CO13127-6W/Y) while zeaxanthin content ranged from 5.31 µg/g DM (Russet Nugget) to 745.89 µg/g DM (OR04198-1). Zeaxanthin was the major carotenoid detected among the yellow clones analyzed. In addition, orange-fleshed genotypes of CO13127-6W/Y and OR04198-1 showed the top two highest lutein and zeaxanthin carotenoid content yellow lines. In contrast, the white-fleshed control cultivar of Russet Nugget had the lowest carotenoid content.

Table 3.2. Sample chromatography peaks (AU) and retention times of carotenoid standards.

Carotenoid Standards[a]	AU[b]	Retention time (min)[c]
Violaxanthin	145785	3.73
Lutein	417124	4.45
Zeaxanthin	529537	4.94
Trans-b-apo-8-carotenal	93065	7.48

[a]External standards of violaxanthin, lutein, and zeaxanthin. Internal standard of trans-b-apo-8-carotenal.
[b]Detected chromatography peak area in arbitrary units (AU) of standard concentration of 10 μg/mL.
[c]Retention time of the analyte eluted from the column (15 minutes HPLC run time per injection).

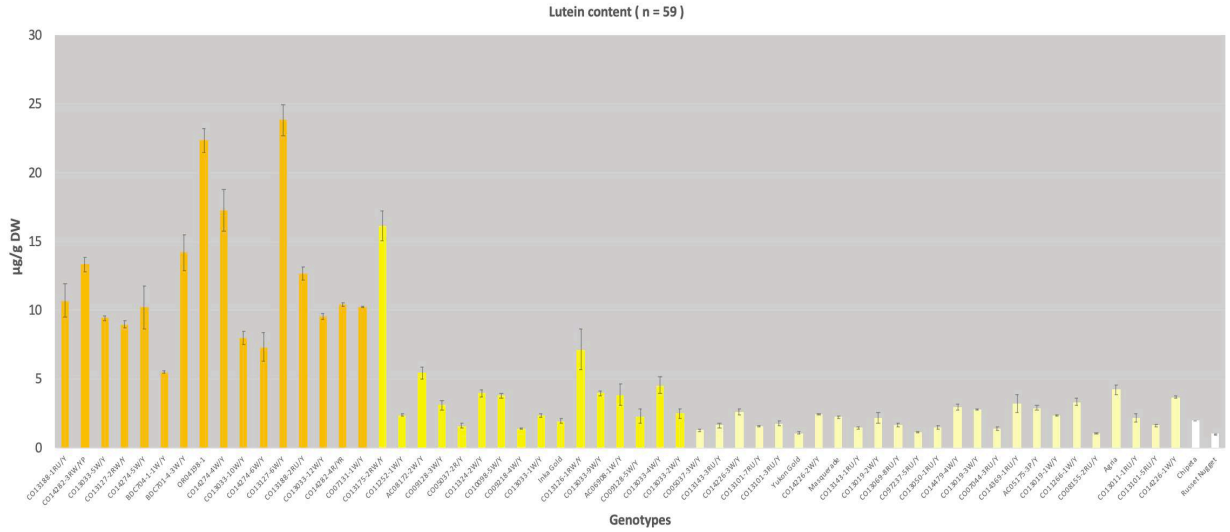

Figure 3.2. Bar chart of lutein content among 59 genotypes. Bins colored based on the FC rating score 0 – 3 (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange). Error bars represent SE (standard error of the mean, 3 reps).

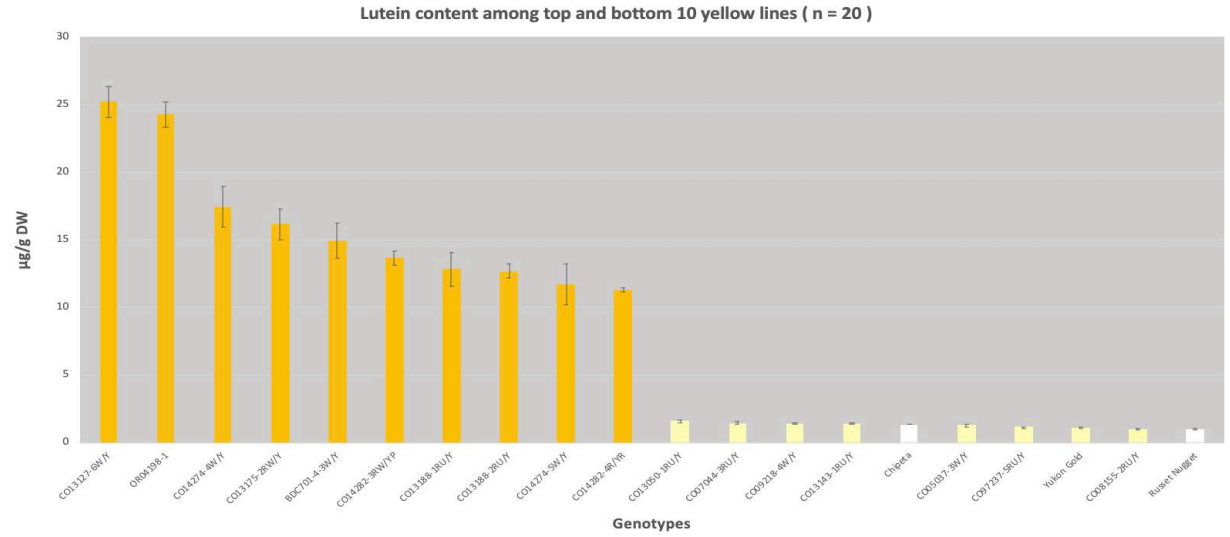

Figure 3.3. Bar chart of lutein content among top and bottom 10 yellow lines. Bins colored based on the FC rating score 0 – 3 (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange). Error bars represent SE (standard error of the mean, 3 reps).

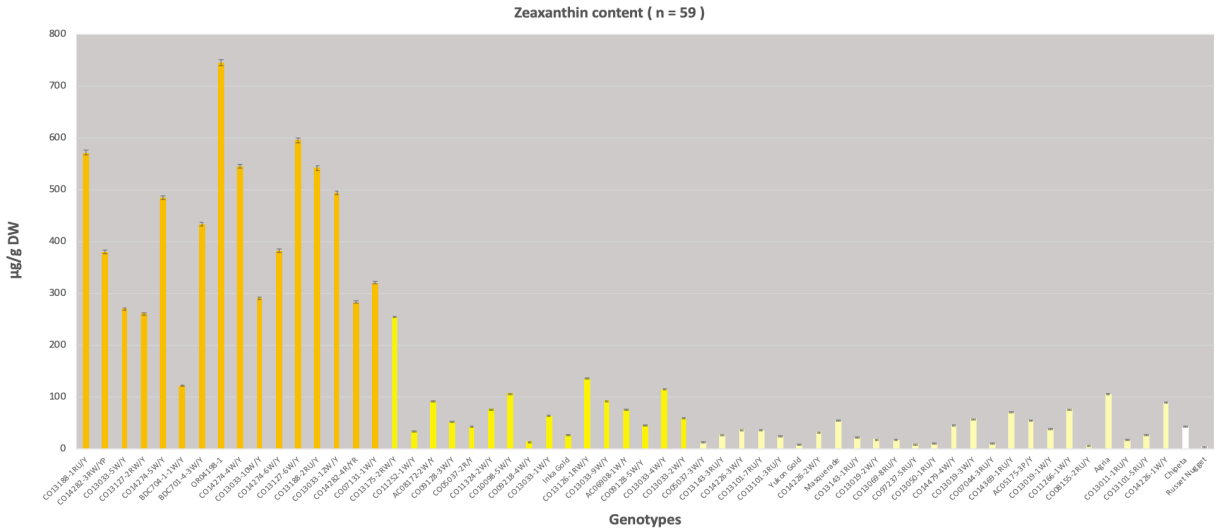

Figure 3.4. Bar chart of zeaxanthin content among 59 genotypes. Bins colored based on the FC rating score 0 – 3 (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange). Error bars represent SE (standard error of the mean, 3 reps).

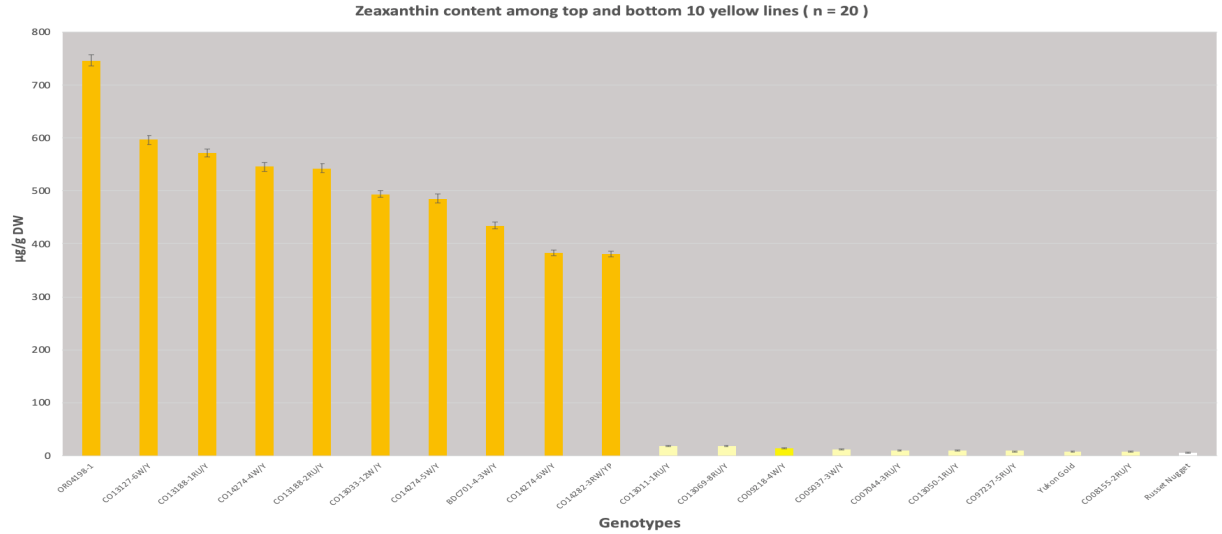

Figure 3.5. Bar chart of zeaxanthin content among top and bottom 10 yellow lines. Bins colored based on the FC rating score 0 – 3 (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange). Error bars represent SE (standard error of the mean, 3 reps).

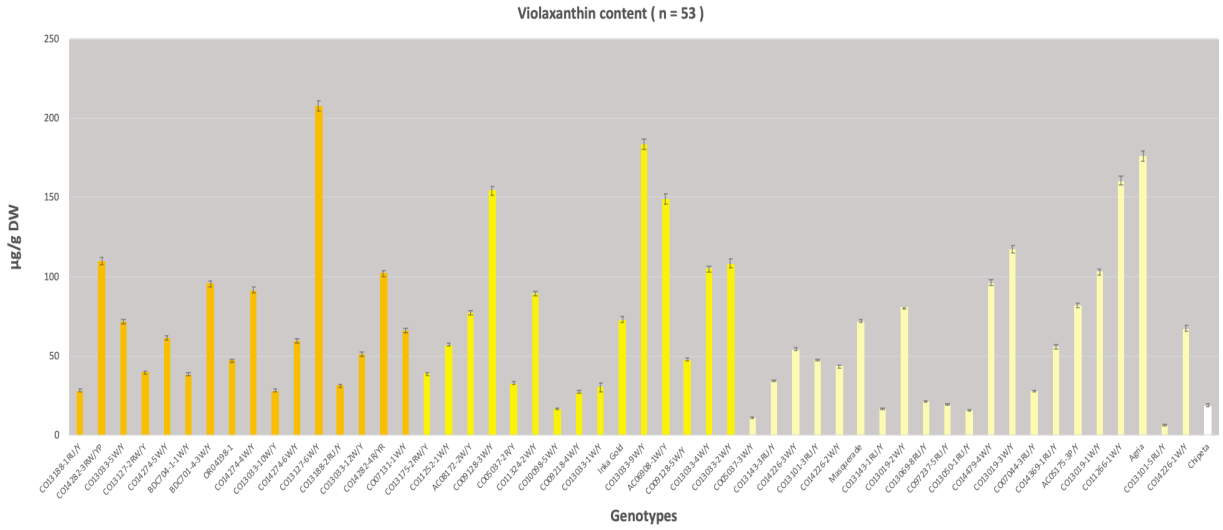

Figure 3.6. Bar chart of violaxanthin content among 53 genotypes. Bins colored based on the FC rating score 0 – 3 (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange). Error bars represent SE (standard error of the mean, 3 reps).

3.3.4. Total Carotenoid Content Analysis

A 96-well Bio-Rad 680 Microplate Reader Spectrophotometer (Bio-Rad Laboratories Inc. Hercules, CA) was used to determine the total carotenoid content. Extracted samples from the -20°C freezer were diluted by a factor of 10. For the dilution, 20 µL of the samples were dispensed into the 96-well plate and 180 µL of ethanol solvents were added (200 µL total volume per well). Total carotenoid content was determined by the absorbance values of the sample extracts by using a 450 nm filter (Burgos et al. 2009). Since zeaxanthin was the major carotenoid detected among the yellow clones analyzed by HPLC, total carotenoid content was calculated using a zeaxanthin standard curve. Total carotenoid content was expressed as micrograms of zeaxanthin equivalents per 1.0 g of dry matter (µg of ZEA/g DM).

The zeaxanthin curve was prepared by determining the absorbance of zeaxanthin standard solution ranging in concentrations from 0, 5, 10, 15, 20 and 25 (µg/mL). Microplate spectrophotometer readings for the samples at 450 nm were converted into zeaxanthin equivalents based on the following equation: $y = 0.00565x - 0.0218$, $R^2 = 0.993$, where x = absorbance at 450 nm and y = µg zeaxanthin equivalents/g DM. Like Figures 3.2 and 3.4, there was a consistent trend of FC rating (FC 3) genotypes to have the highest concentrations of total carotenoids. Figure 3.7 illustrates the calculated total carotenoids among the 59 yellow clones.

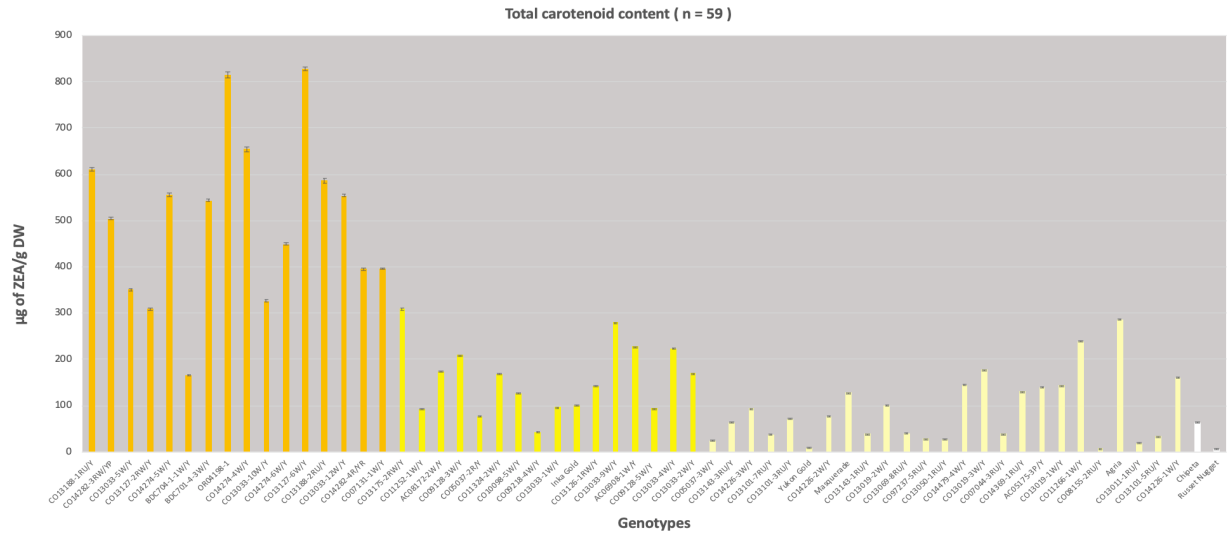

Figure 3.7. Bar chart of total carotenoid content among 59 genotypes. Bins colored based on the FC rating score 0 – 3 (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange). Error bars represent SE (standard error of the mean, 3 reps).

3.4. Flesh Color and Carotenoid Content Statistical Analysis

3.4.1. Flesh Color Rating vs. Carotenoid Content

Tukey adjusted pairwise comparison (Tukey's test) was performed to determine any significant differences between the treatment means of individual and total carotenoids. Analysis was done using RStudio version 3.4.3 (RStudio, Inc., Boston, MA). A one-way analysis of variance (ANOVA) was fit into the model, having the response variables as lutein, zeaxanthin, violaxanthin and total carotenoid content. Flesh color (FC) rating was included in the model as a covariate. By using the emmeans (estimated marginal means) package in RStudio, we can calculate p-values at significance level of $\alpha = 0.05$.

After the analysis of Tukey adjustment pairwise comparison shown in Table 3.3, it was concluded that any comparison of a FC rating of (FC 3) had a significantly higher mean response for individual carotenoids of lutein, zeaxanthin and total carotenoids with ($P < 0.0001$), indicating that orange-fleshed genotypes (FC 3) have higher overall carotenoid content as compared to the white-fleshed (FC 0), yellow-fleshed (FC 1) and dark yellow-fleshed (FC 2) genotypes. These results could be visually validated from Figures 3.2, 3.4 and 3.7. Interestingly, there were no significant p-values among other pairwise comparisons between (FC 0), (FC 1) and (FC 2) flesh colors and no significance identified in violaxanthin as a response variable (Figure 3.6 & Table 3.3). Previous studies have shown that the combined total of lutein and zeaxanthin in orange-fleshed genotypes are four to five times higher than in lighter yellow-fleshed genotypes, generally supporting the results in Table 3.3 (Brown et al. 1993; Lu et al. 2001; Haynes et al. 2011).

3.4.2. Chroma vs. Carotenoid Content

The yellow and orange color of the tuber flesh is imparted by lutein and zeaxanthin respectively, while the intensity of the yellowness is moderately to highly correlated with

individual and total carotenoid content (Lu et al. 2001; Lachman et al. 2016). Pearson correlation coefficients *(r)* were analyzed using RStudio version 3.4.3 (RStudio, Inc., Boston, MA) to determine the association of yellow intensity (chroma) versus individual and total carotenoid among the yellow germplasm. As shown in Table 3.4, lutein vs. chroma ($r = 0.56$, $P < 0.01$), zeaxanthin vs. chroma ($r = 0.60$, $P < 0.01$) and total carotenoid vs. chroma ($r = 0.63$, $P < 0.01$) indicated moderate positive correlations overall. Conversely, violaxanthin vs. chroma ($r = 0.32$, $P < 0.05$) show a low correlation. Lutein vs. zeaxanthin ($r = 0.92$, $P < 0.01$) have a high positive correlation, indicated both individual carotenoids may co-elute together. Likewise, total carotenoids vs. lutein and zeaxanthin ($r = 0.93$, $P < 0.01$), ($r = 0.97$, $P < 0.01$) respectively showed high positive correlations (Table 3.4).

Table 3.3. Emmeans output of Tukey adjusted pairwise comparison.

FC[b] Contrast	P-values[a]			
	Lut	**Zea**	**Vio**	**Total**
0 - 1	0.9730	0.9963	0.3753	0.8254
0 - 2	0.5193	0.7580	0.5530	0.2847
0 - 3	<0.0001	<0.0001	0.5240	<0.0001
1 - 2	0.2470	0.4318	0.9219	0.2422
1 - 3	<0.0001	<0.0001	0.9631	<0.0001
2 - 3	<0.0001	<0.0001	0.9994	<0.0001

[a]P-values of treatment means of carotenoids ($\alpha = 0.05$) via emmeans package: Lut = lutein, Zea = zeaxanthin, Vio = Violaxanthin, Total = total carotenoids.
[b]Flesh color (FC) rating (0 - 3) contrasts. For example, pairwise comparison of white-flesh types (FC 0) vs. yellow-flesh types (FC 1) with a response variable of a given carotenoid above.

Table 3.4. Pearson correlation coefficients[a] *(r)* between individual[b] and total carotenoid content and chroma among yellow genotypes.

	Lut	**Zea**	**Vio**	**Total[c]**	**Chroma**
Lut	-	0.92**	0.30*	0.93**	0.56**
Zea		-	0.20	0.97**	0.60**
Vio			-	0.43*	0.32*
Total[c]				-	0.63**
Chroma					-

[a]Significance: $*P < 0.05$, $**P < 0.01$
[b]Lut = lutein, Zea = zeaxanthin, Vio = violaxanthin.
[c]Total = total carotenoids

3.5. Chapter 3 Summary

Through the use of a reflectance colorimeter, chroma can be calculated and be used as a quantitative colorimetric variable for the yellow germplasm in this study. Although chroma is an indirect measure of carotenoid content in yellows, and despite the moderate correlation coefficient results, colorimeters can still be an acceptable tool to aid selection of high-carotenoid yellow potato breeding lines due to its simplicity and relatively inexpensive application. Microplate reader spectrophotometers and the HPLC apparatus are the standard methods to quantify and identify total and individual carotenoids respectively.

The details of the inefficient extraction protocols demonstrated it as the limiting step for individual carotenoid content analysis. Therefore, REIMS's capability as a novel technology that does not require sample preparation and extraction to improve overall selection efficiency was assessed. The objective was to identify and quantify the prevalence of lutein and zeaxanthin content in the yellow potato germplasm as the main parameters.

CHAPTER 4. REIMS ANALYSIS

4.1. Introduction

Rapid Evaporative Ionization Mass Spectrometry (REIMS) is a technique that allows *in situ* analysis and real-time identification of biological tissues (Schäfer et al. 2009). REIMS was originally developed for surgical operations of human tissue identification by combining mass spectrometry with electrosurgical tools but since then, the application has expanded from the medical fields to cover the characterization and identification of other organisms such as the analysis of food and environmental samples (Paxton 2020). Currently, Waters Corporation is commercializing the REIMS system for research use only and it is not intended for diagnostic or therapeutic purposes (Waters Corporation 2015).

As a novel technology, only a few research publications have been released and the majority were in animal sciences. Based on food authenticity testing results, REIMS could differentiate between fish species with large differences in price, e.g., between the more expensive cod and the cheaper whiting (Black et al. 2017). In addition to fish, REIMS has been applied to differentiate raw meat products of bovine and horse breeds, providing 97% accuracy at the breed level and 100% accuracy at the species level. The study was made in response to the 2013 horse meat scandal in the United Kingdom where beef labeled products were found to contain horse meat (Balog et al. 2016).

REIMS was applied in crop science to help understand how to deal with weed grass species with emerging herbicide-resistance traits (Stead et al. 2016). Its rapid leaf tissue analysis was able to differentiate wild type black-grass, annual rye-grass and multi-herbicide resistant populations of black-grass, demonstrating that REIMS have the potential for rapid phenotyping in crops (Stead et al. 2016). The incentive of rapid phenotyping for identification of high-

carotenoid content of lutein and zeaxanthin breeding lines in yellow-fleshed potato germplasm has generated strong interest in the use of REIMS that may ultimately benefit the Colorado Potato Breeding and Selection Program.

4.2. REIMS Materials and Methods

4.2.1. REIMS Instrumentation Overview

The experimental setup for REIMS analysis consists of a monopolar cutting electrode (iknife) as the ion source to apply high-frequency electric current directly to the biological tissue, and a return electrode where samples are placed to couple the generated electric current (Paxton 2020). As the iknife makes contact with flesh tissues, smoke containing both positive and negative ions from the rapid heating and evaporation of the samples are then transferred through a 2 m long polytetrafluoroethylene (PTFE) tubing coupled to the inlet tube to the mass spectrometer. A venture gas jet pump inside the apparatus is then used to clear the site of charged smoke and drive it toward the entry of a quadrupole time-of-flight mass spectrometer (Q-TOF-MS, Paxton 2020). Data acquisition of the tissue's metabolomic fingerprint is instant, collecting REIMS mass spectra (denoted by m/z and each m/z's peak intensity) in real-time.

4.2.2. REIMS Advantages in Individual Carotenoid Screening

REIMS's promising rapid, direct-from-tissue analysis would eliminate sample preparations and extraction protocols, potentially making the technology a preferred analytical method to screen for carotenoids in potato germplasm. In addition, REIMS is claimed to be an effective instrument for analyzing small molecules, lipids (particularly phospholipids with mass range between 600 – 900 m/z) and organics (Waters Corporation 2015; Paxton 2020). Both lutein and zeaxanthin carotenoids are small organic pigments with relatively low molecular weight (568.87 g/mol, Sigma-Aldrich Corporation) and theoretically, REIMS may detect both carotenoids.

In addition, REIMS analysis can be performed via positive-ionization or negative-ionization mode (Paxton, 2020). The rational of positive-ionization mode as default for all samples was based on previous studies that detected carotenoid fragmentation and parent protonated products (Rezanka et al. 2009; Guaratini et al. 2005; Rivera et al. 2013). Sample expected m/z signals for both lutein and zeaxanthin in positive-ionization mode via different analytical platforms are provided in Table 4.1. Parental ion m/z for lutein and zeaxanthin are provided, ranging from 551.5 – 569.5 m/z ($[M+H-18]^+$, $[M+H]^+$) and 568.9 – 569.5 m/z ($[M+H]^+$), respectively (Table 4.1). Also, fragment motifs for both carotenoids can be observed, ranging from 283.2 – 537.409 m/z ($[M+H]^+$, $[M]^+$, Table 4.1).

4.2.3. Pilot REIMS Experiment in Potato

Since there are no released publications conducted as a reference on the use of REIMS on potato improvement, a pilot experiment was performed to become familiar with the system prior to the actual data collection of the 60 yellow potato genotypes. In this pilot study, three genotypes with varying levels of FC rating of Chipeta (FC 0), Inka Gold (FC 2) and CO07131-1W/Y (FC 3) were assessed. Samples were analyzed using a Synapt G2 Si Q-TOF, fitted with a REIMS ionization source coupled with a monopolar cutting electrode (iknife, Waters Corporation, Milford, MA).

There were technical factors of the REIMS apparatus explored in this preliminary study, with an objective to develop the most practical protocol to collect reasonable data output. First, the method of applying the iknife to tuber flesh was investigated, and lightly slicing or "burning" the flesh for about 1-inch (2.54 cm) length produced the most stable spectra peaks. It was important to note that one burn in tuber flesh yields one REIMS spectra peak in real-time and waiting for the peak to go back to the baseline allowed consistent repeatability of data acquisitions.

Second, using the "dry cut" mode at varying power levels between 10 to 40 Watts was assessed. A dry cut mode at power of 40 W generated the best burns in tuber flesh tissues with adequate smoke to be taken up by REIMS, producing the most stable spectra peaks. The preliminary raw spectra data of the three genotypes were compared; however, no observable differences were detected (Figure 4.1). Converting raw data to data matrix and using multivariate statistical analyses was processed to help understand REIMS data.

4.2.4. REIMS Experimental Design and General Protocol

REIMS method workflow to screen for lutein and zeaxanthin carotenoids among the 60 potato selections started with the experimental design of randomized complete block design (RCBD) with three tuber replicates per genotype (180 total tuber samples). Days was applied as a blocking factor in the experimental design (Rep 1 = 1st day, Rep 2 = 2nd day, Rep 3 = 3rd day), assaying 60 sample data acquisitions per day (total of 180 sample data acquisitions). Samples were analyzed using a Synapt G2 Si Q-TOF, fitted with a REIMS ionization source coupled with a monopolar cutting electrode (iknife, Waters Corporation, Milford, MA) powered by an Erbotom ICC 300 electrosurgical generator (Erbe Elektromedizin GmbH, Tubingen, Germany).

Potato tubers approximately two months after the harvest date were removed from the cold storage on the day of data collection and then held at room temperature. Tubers were cut in half longitudinally exposing the flesh area and placed into a return electrode under a hood system. Dry cut mode was applied at a power of 40 Watts with the cone voltage set to 40 Volts. To enhance the response for lipids in the REIMS source, a continual flow (200 µL/min) of 2 ng/mL leucine-enkephalin internal standard (556.277 m/z) was injected directly to the REIMS apparatus during the duration of sampling.

REIMS spectra were collected in positive-ionization mode and set data acquisition for 100 – 1000 m/z for 3 minutes duration per sample. At least six tuber flesh burns were collected

for each sample for about 1-inch (2.54 cm) length in longitude direction, with each burn lasting approximately two seconds for consistency. The darker pith area down the center of the tuber flesh was avoided if possible. Fresh cut tuber samples contained high moisture content, generating adequate smoke. This yielded six REIMS spectra per sample which were observed in real-time using the LiveID software (Waters Corporation, Milford, MA). REIMS inlet was cleaned daily prior to data collection. Once data collection of a sample was completed, processed tuber halves were wrapped in foil and put into a labeled Ziploc bag and stored in a cooler with dry ice. These were later transferred to a -80°C freezer so they could be the same tubers processed for the HPLC analysis with the purpose of reducing biological variation.

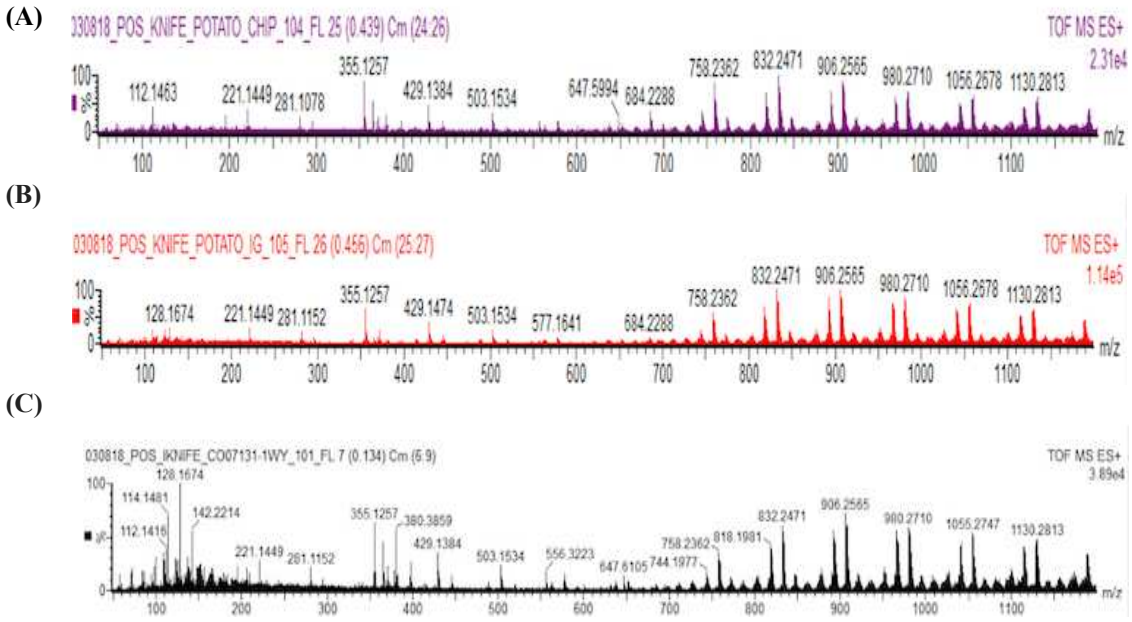

Figure 4.1. Sample raw REIMS spectra (positive-ionization mode) obtained from the pilot experiment via LiveID software. Intensity (0 – 100% relative scale) peaks in the (y) range vs. accurate mass m/z in (x) range. Preliminary potato genotypes of Chipeta (A), Inca Gold (B) and CO07131-1W/Y (C).

Table 4.1. Sample expected m/z signals for lutein and zeaxanthin (parent and fragment ions)

Carotenoid	MS Technique[a]	Molecular Ion	m/z[b]	Adduct(s)	Reference
Lutein	LC-MS/APCI	Parent	551.5	$[M+H-18]^+$	Rezanka et al. 2009
Lutein	LC-MS/MS	Parent	569.4359	$[M+H]^+$	HMDB version. 4.0
Lutein	ESI-MS	Parent	569	$[M+H-18]^+$	Rivera et al. 2013
Zeaxanthin	LC-MS/APCI	Parent	569.5	$[M+H]^+$	Rezanka et al. 2009
Zeaxanthin	LC-MS/MS	Parent	569.4359	$[M+H]^+$	HMDB version. 4.0
Zeaxanthin	ESI-MS	Parent	568.9	$[M+H]^+$	Rivera et al. 2013
Lutein	LC-MS/MS	Fragment	425.3208	$[M+H]^+$	HMDB version. 4.0
Lutein	LC-MS/MS	Fragment	395.2739	$[M+H]^+$	HMDB version. 4.0
Lutein	ESI-MS	Fragment	476.6	$[M]^+$	Rivera et al. 2013
Zeaxanthin	LC-MS/MS	Fragment	385.2895	$[M+H]^+$	HMDB version. 4.0
Zeaxanthin	LC-MS/MS	Fragment	537.4096	$[M+H]^+$	HMDB version. 4.0
Zeaxanthin	ESI-MS	Fragment	283.2	$[M]^+$	Rivera et al. 2013

[a]Mass signal data (m/z) derived from liquid chromatography mass spectrometry coupled with atmospheric pressure chemical ionization (LC-MS/APCI), electrospray ionization mass spectrometry (ESI-MS) and liquid chromatography tandem mass spectrometry (LC-MS/MS).
[b]Expected/predicted m/z value detected for the corresponding carotenoid molecule. Source provided above.

4.3. REIMS Multivariate Statistical Analysis

4.3.1. Conversion of Raw Data to Data Matrix

REIMS spectra derived from different FC rating tuber tissues were highly specific; however, the differences were not immediately distinguishable as observed from the pilot experiment (Figure 4.1) and differentiation of tissue spectra would rely on multivariate statistical tools such as principal component analysis (PCA) and partial least square (PLS), (Balog et al. 2016, Stead et al. 2016, Paxton 2020). A study about optimal machine learning algorithm in REIMS analysis has also been conducted to increase prediction accuracy to classify quality attributes in beef (Gredell et al. 2019). A total of eight machine learning algorithms were compared, and results showed that dimension reduction and feature selection of REIMS data in the processing and analysis method stage improved all predictive models (Gredell et al. 2019). REIMS data is complex, and application of machine learning is a powerful tool for REIMS analyses to rapidly detect and characterize sample patterns and differentiation.

The first step in multivariate analyses was to convert raw data generated by REIMS into a data matrix using a prototype Abstract Model Builder software, AMX version 1.0.1581.0 (Waters Research Centre, Budapest, Hungary). Raw data were pre-processed and were background subtracted setting a lock mass correction using leucine-enkephalin (556.277 m/z) and normalized to TIC (total ion current) before exposure to multivariate analysis. Peak binning was performed at intervals of 0.5 m/z to increase the signal intensity to capture more bin information that may potentially cover carotenoids. Bins from the six burn spectra were summed for each sample. Mass bins ranging from 556 – 559 were excluded from the data matrix to remove the internal standard signal (leucine-enkephalin, 556.277 m/z). The data matrix model was generated using the bin mass regions of 560 – 900 m/z to specifically detect parent molecules of lutein and zeaxanthin as provided in Table 4.1.

4.3.2. Multivariate Statistical Tools

The processed data matrix was exported to SIMCA version 15.0.2 (Umetrics, Umea, Sweden) for multivariate methods of analysis (MVA). Moreover, collected color and individual carotenoid data were integrated to multivariate models to help with REIMS data interpretation. FC rating (0 – 3), chroma values, lutein and zeaxanthin (µg/g DM) data were added to MVA statistical models. FC ratings (0 – 3) were incorporated as the categorical covariate to all models, while chroma, lutein and zeaxanthin were treated as the numerical response variables.

Three MVA were incorporated for the REIMS data interpretation. First, principal component analysis (PCA) which was an unbiased model, was purely based on sample variations of the REIMS data matrix. Second, partial least square (PLS) which was a biased model, where data matrix (x) were regressed against a response variable (y). Lastly, orthogonal-PLS (OPLS) which was a PLS model with an addition of component of variation to PLS that was not associated with the specified (y), and the preferred model for data fusion from different analytical platforms (Bylesjö et al. 2007). Cross-validation coefficient (Q^2) from the MVA analyses was used to represent the predictive power of the model (50 – 100% show good model fit) while (R^2, cumulative) indicated the variation described by all components in the model. All data were subjected to unit variance scaling (UV).

4.3.3. PCA Analysis of All Replicated Samples

The RCBD with three replicates of the 60 yellow genotypes for REIMS analysis produced 180 total data samples that were converted into a data matrix. PCA models were created via SIMCA version 15.0.2 (Umetrics, Umea, Sweden), generating nine components. Figure 4.2 shows the PCA of all replicated samples minus obvious outliers (n = 171), explaining 92.2% (R^2) of sample variation in the model. A cross-validation coefficient of 83.1% (Q^2) implied a good model fit. While the PCA model illustrated an observable separation of

replicated samples, it was unusual that almost all of Rep 3 (301 – 360) score values (PC1 vs. PC2 scores) are concentrated in the first quadrant of the PCA plot, suggesting that separation is based by reps and not metabolite differences (Figure 4.2).

All REIMS spectra data were re-evaluated to determine the occurrence using the Abstract Model Builder software, AMX version 1.0.1581.0 (Waters Research Centre, Budapest, Hungary). Interestingly, it was determined that all Rep 3 spectra peaks contained high noise signal abundance. Since days was established as blocking for the RCBD, it was speculated that in Rep 3 (3rd day), the REIMS instrument accumulated a fair amount of starch aggregates in the system. When Rep 3 samples were excluded in multivariate models regardless of PCA, PLS or OPLS, there was an improvement in sample differentiation. As a result, all Rep 3 samples were excluded in all future MVA in addition of selected replicated samples that had similar high noise abundance. Re-processing of the data matrix with Rep 3 exclusion was done via Abstract Model Builder software, AMX version 1.0.1581.0 (Waters Research Centre, Budapest, Hungary).

PCA of Rep 1 and Rep 2 was re-analyzed using SIMCA version 15.0.2 (Umetrics, Umea, Sweden). FC rating (0 – 3) was included as a categorical covariate in the PCA model (Figure 4.3). PCA of Reps 1 and 2 minus obvious outliers and replicated samples with noisy spectra (n = 95) explain 67.4% (R^2) of sample variation in the model with a moderate cross-validation coefficient of 48.4% (Q^2, Figure 4.3). Although Figure 4.3 showed a more normal distribution of score sample plots compared to Figure 4.2, there were no apparent trends or separations based on flesh color after analyses of all nine components. Hence, the next series of MVA included partial least square (PLS) and orthogonal-PLS (OPLS) biased models to determine any differentiation between samples and individual carotenoids.

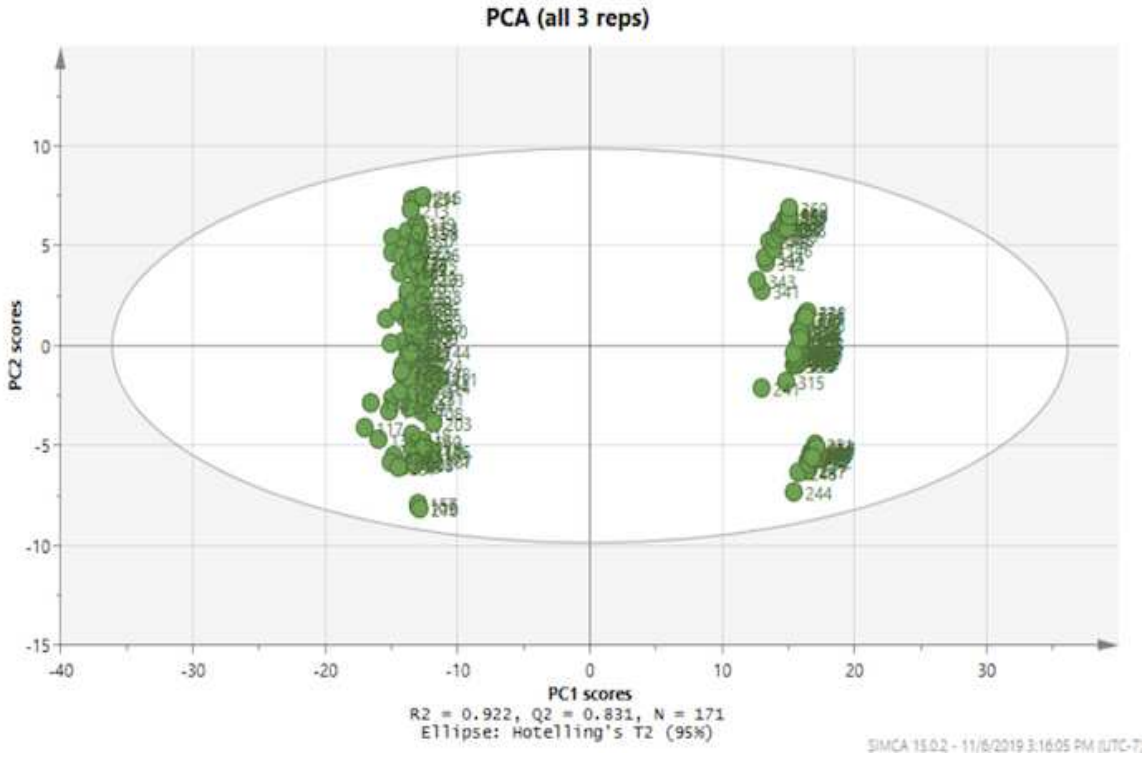

Figure 4.2. PCA of all replicated samples (n^a = 171). Outliers were excluded in the model outside the extreme range of the elliptical region (95% confidence interval).

[a]Rep_1 (101-160), rep_2 (201-260), rep_3 (301-360), excluding outliers.

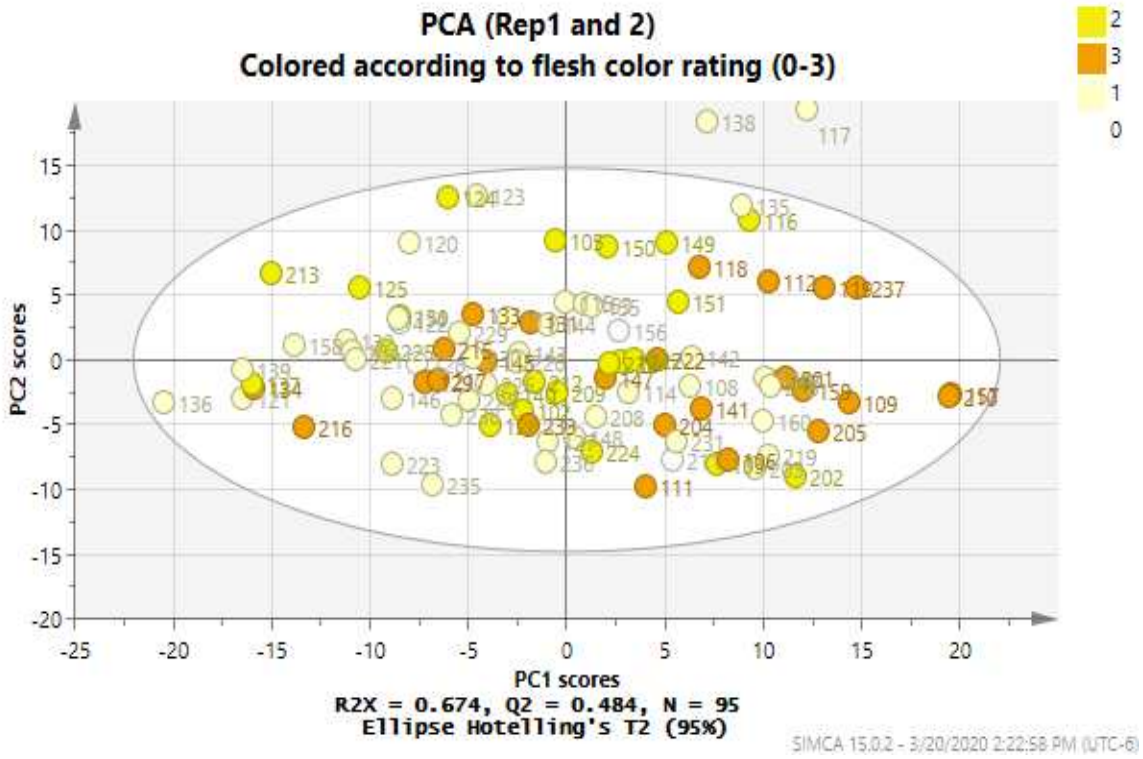

Figure 4.3. PCA[a] of rep_1 and rep_2 (n[b] = 95). Outliers were excluded in the model outside the extreme range of the elliptical region (95% confidence interval).

[a]PCA scores colored according to flesh color rating (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange).
[b]Rep_1 (101-160), rep_2 (201-260). Excluding outliers, rep_3 (301-360) and replicated samples with noisy spectra.

4.3.4. PLS and OPLS Analysis and Results

To determine multivariate trends for the REIMS analysis, the data matrix (x) was regressed against response variables (y) of lutein, zeaxanthin and chroma content. PCA of Rep 1 and Rep 2 (n = 95, Figure 4.3) was modified to partial least square analysis (PLS) using SIMCA version 15.0.2 (Umetrics, Umea, Sweden), generating five components. PLS was fitted into a model as lutein content as the response variable that explained 94.8% (R^2Y) of component variation in the model with a cross-validation coefficient of 67.8% (Q^2) implying good model fit (Figure 4.4). Similarly, PLS was fitted into a model as zeaxanthin content as the response variable that explained 80.5% (R^2Y) of component variation in the model with a cross-validation coefficient of 56.4% (Q^2) implying good model fit (Figure 4.5).

Orthogonal-PLS (OPLS) models were also generated to determine any improvement in REIMS data interpretation. An OPLS was fitted into a model as lutein content as the response variable that explained 75.1% (R^2Y) of orthogonal component variation in the model with a cross-validation coefficient of 63.8% (Q^2) implying good model fit (Figure 4.6). Likewise, an OPLS was fitted into a model as zeaxanthin content as the response variable that explain 80.5% (R^2Y) of orthogonal component variation in the model with a cross-validation coefficient of 59% (Q^2) implying good model fit (Figure 4.7). Moreover, an OPLS was fitted into a model as chroma values as the response variable that explained 53.3% (R^2Y) of orthogonal component variation in the model with a cross-validation coefficient of 29.4% (Q^2) implying poor model fit (Figure 4.8).

Finally, an OPLS was fitted into a model with all the response variables of lutein, zeaxanthin and chroma that explain 65.6% (R^2Y) of orthogonal component variation in the model with a cross-validation coefficient of 48.7% (Q^2) implying a moderate model fit (Figure

4.9). Overall, there were no significant differences in predictive powers (Q^2) between PLS and OPLS models showing the same individual carotenoid response variables.

In addition, it can be observed from all PLS and OPLS models that a FC rating (FC 3) sample components were more separated from other FC ratings of 0-2. Therefore, the next MVA analyses consisted of OPLS models of high carotenoids (FC 3) vs. low carotenoids (FC 1 & 0), and medium carotenoids (FC 2) vs. low carotenoids (FC 1& 0) with an objective to illustrate any clear separation between sample populations. The OPLS model with all the response variables (n = 95, Figure 4.9) was modified to exclude samples that have FC rating (FC 2). Figure 4.10 (n = 71) show an OPLS model of high carotenoids (FC 3) vs. low carotenoids (FC 1 & 0) that explained 87.3% (R^2Y) of orthogonal component variation in the model with a cross-validation coefficient of 56.1% (Q^2) implying a good model fit and for the first time, there was a clear observable separation. In contrast, Figure 4.11 (n = 68) shows an OPLS model of medium carotenoids (FC 2) vs. low carotenoids (FC 1 & 0) that explained only 23.6% (R^2Y) of orthogonal component variation in the model with a low cross-validation coefficient of 13.7% (Q^2) implying a very poor fit model and displayed no separation.

Fundamentally, replicates of samples are expected to have near identical metabolite profiles and should be very close distance together in the orthogonal component plot. For example, replicates of (131) and (233) located in the lower fourth quadrant of the OPLS plot was designated as the OR04198-1 orange-fleshed clone and could be identified clearly next to each other (Figure 4.10). Because of the criteria of clear separations of high vs. low carotenoid samples, good model fit of 56.1% (Q^2) and near distance of replicated samples, Figure 4.10 was considered the best MVA model of REIMS differentiation among the screened yellow germplasm in this study.

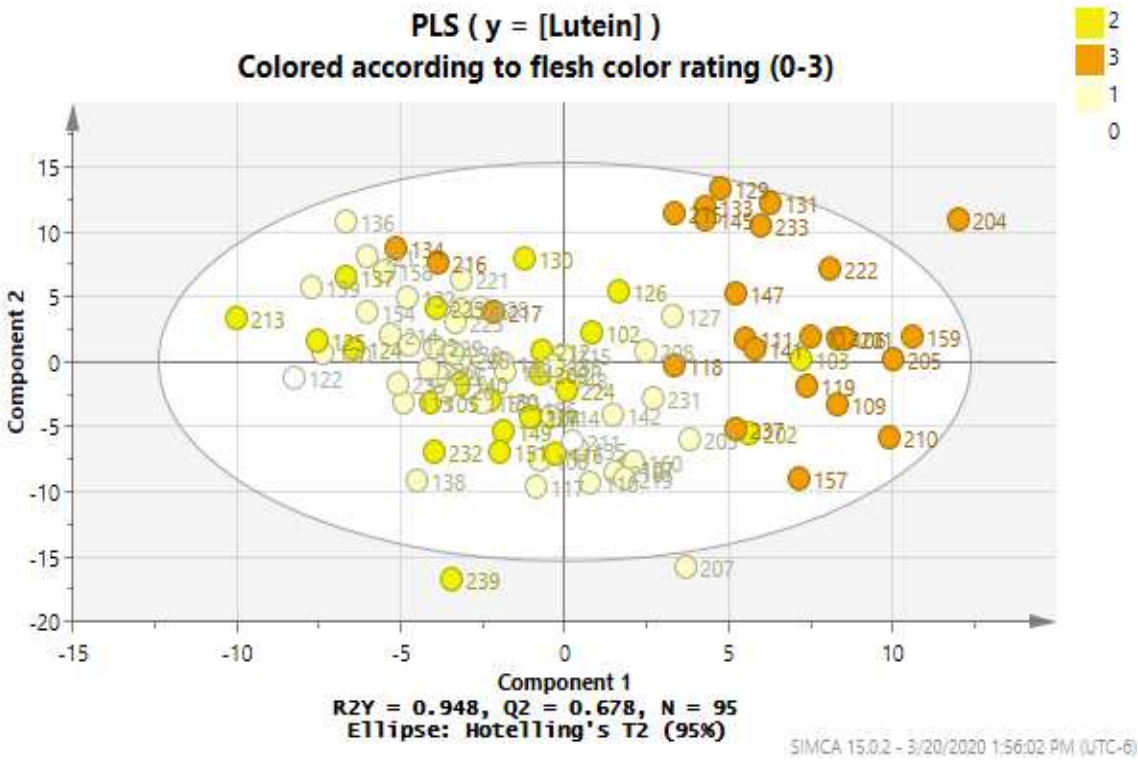

Figure 4.4. PLS[a] of lutein content as a response variable (n[b] = 95). Outliers were excluded in the model outside the extreme range of the elliptical region (95% confidence interval).

[a]PLS components colored according to flesh color rating (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange).
[b]Rep_1 (101-160), rep_2 (201-260). Excluding outliers, rep_3 (301-360) and replicated samples with noisy spectra.

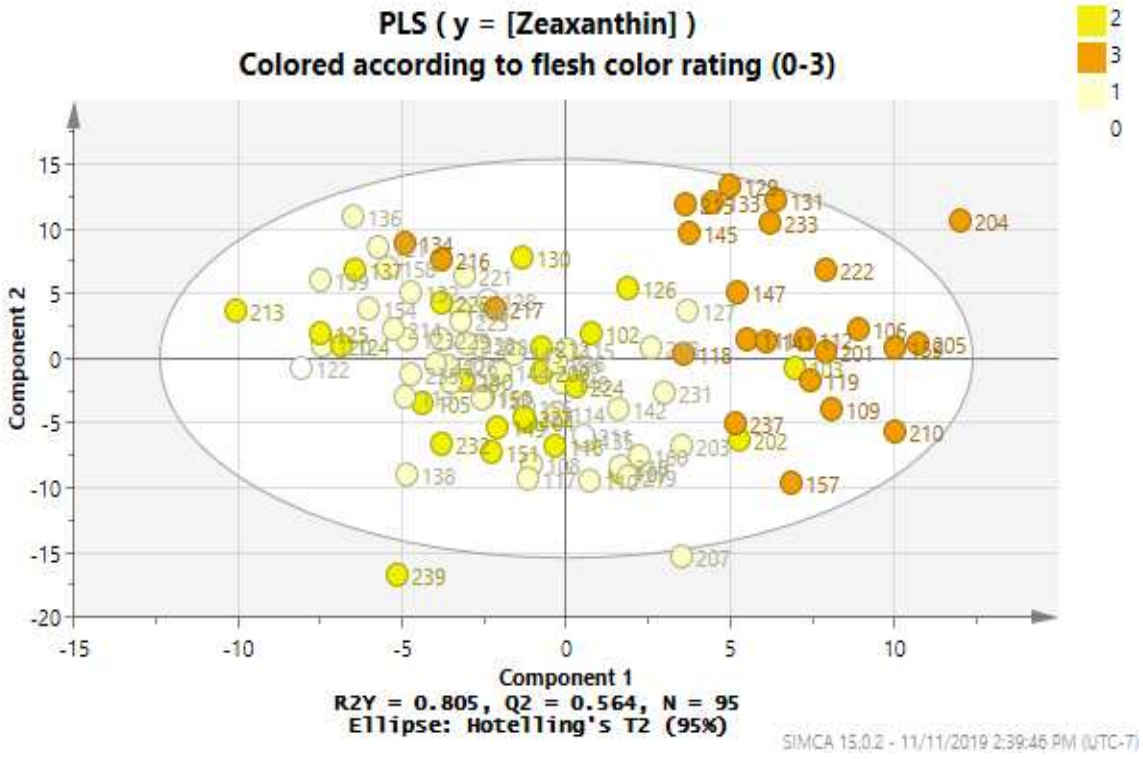

Figure 4.5. PLS[a] of zeaxanthin content as a response variable (n[b] = 95). Outliers were excluded in the model outside the extreme range of the elliptical region (95% confidence interval).

[a]PLS components colored according to flesh color rating (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange).
[b]Rep_1 (101-160), rep_2 (201-260). Excluding outliers, rep_3 (301-360) and replicated samples with noisy spectra.

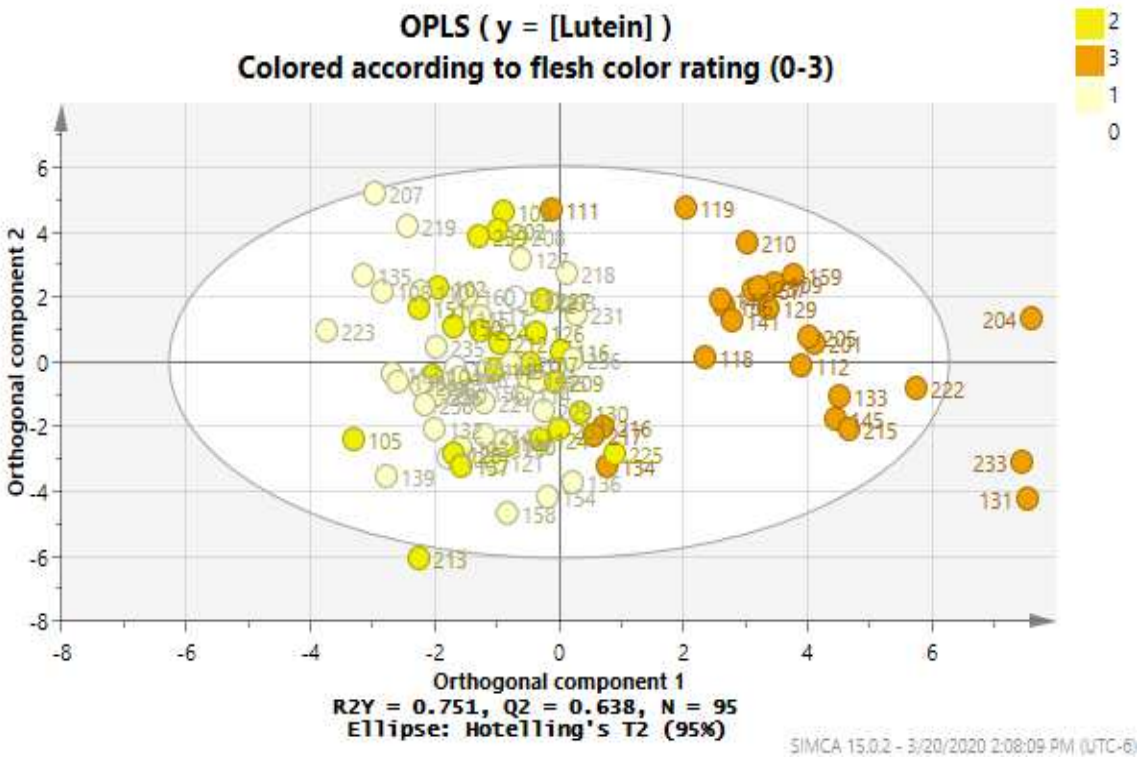

Figure 4.6. OPLS[a] of lutein content as a response variable (n[b] = 95). Outliers were excluded in the model outside the extreme range of the elliptical region (95% confidence interval).

[a]OPLS orthogonal components colored according to flesh color rating (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange).
[b]Rep_1 (101-160), rep_2 (201-260). Excluding outliers, rep_3 (301-360) and replicated samples with noisy spectra.

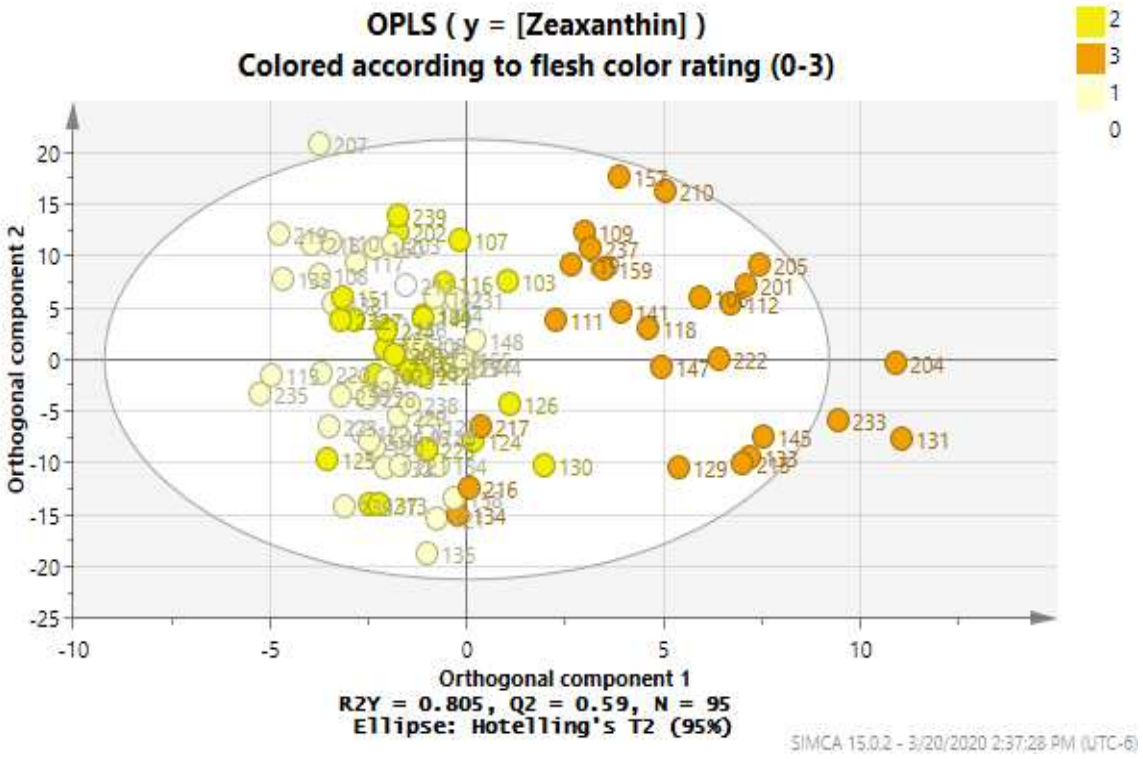

Figure 4.7. OPLS[a] of zeaxanthin content as a response variable (n[b] = 95). Outliers were excluded in the model outside the extreme range of the elliptical region (95% confidence interval).

[a]OPLS orthogonal components colored according to flesh color rating (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange).
[b]Rep_1 (101-160), rep_2 (201-260). Excluding outliers, rep_3 (301-360) and replicated samples with noisy spectra.

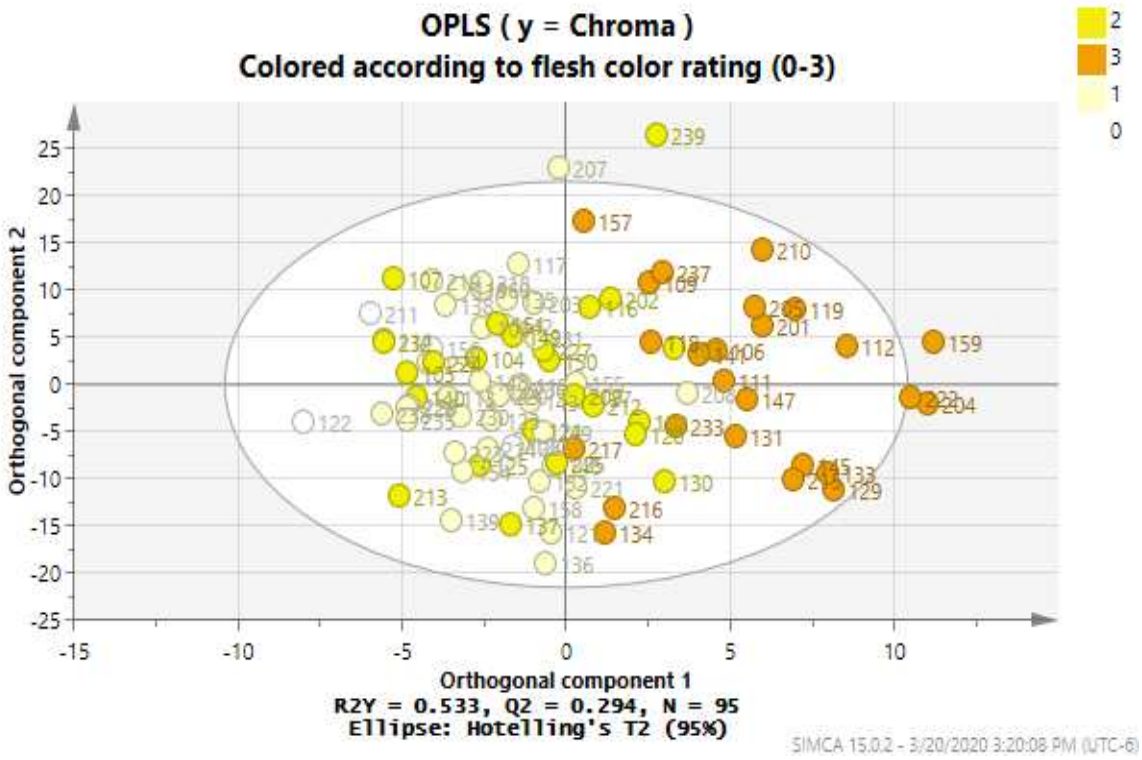

Figure 4.8. OPLS[a] of chroma values as a response variable (n[b] = 95). Outliers were excluded in the model outside the extreme range of the elliptical region (95% confidence interval).

[a]OPLS orthogonal components colored according to flesh color rating (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange).
[b]Rep_1 (101-160), rep_2 (201-260). Excluding outliers, rep_3 (301-360) and replicated samples with noisy spectra.

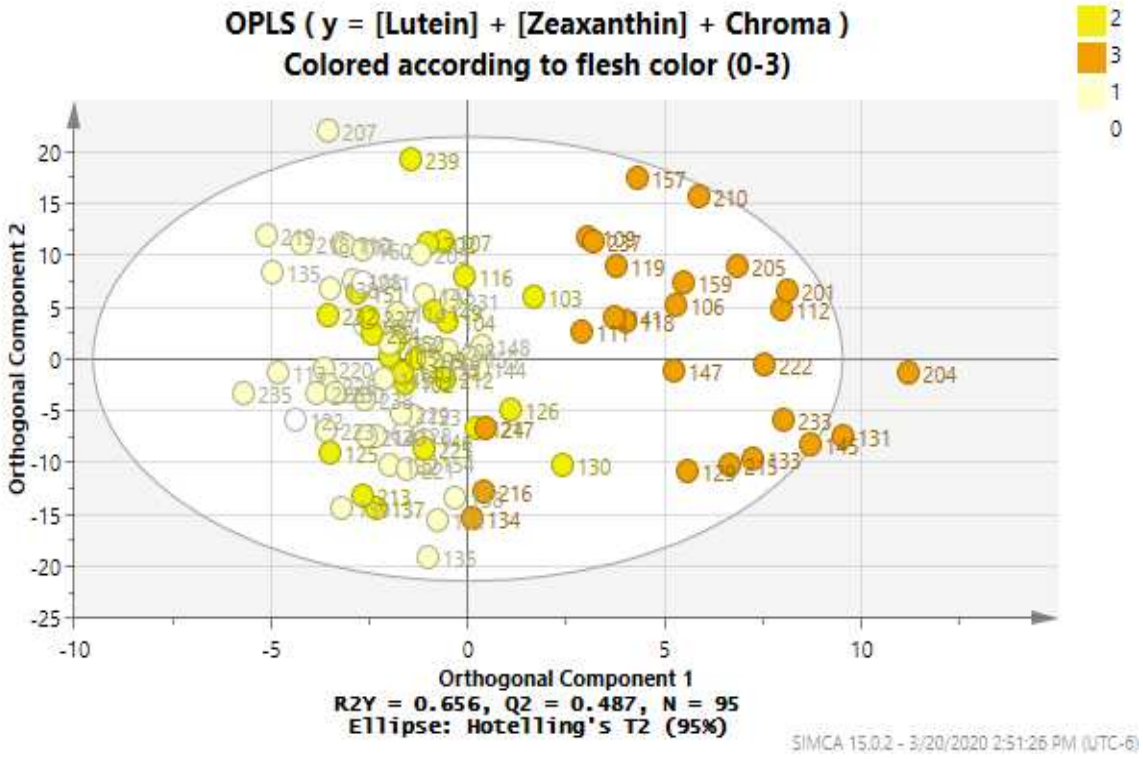

Figure 4.9. OPLS[a] of lutein, zeaxanthin and chroma as response variables (n[b] = 95). Outliers were excluded in the model outside the extreme range of the elliptical region (95% confidence interval).

[a]OPLS orthogonal components colored according to flesh color rating (0 = white, 1 = yellow, 2 = dark yellow, 3 = orange).
[b]Rep_1 (101-160), rep_2 (201-260). Excluding outliers, rep_3 (301-360) and replicated samples with noisy spectra.

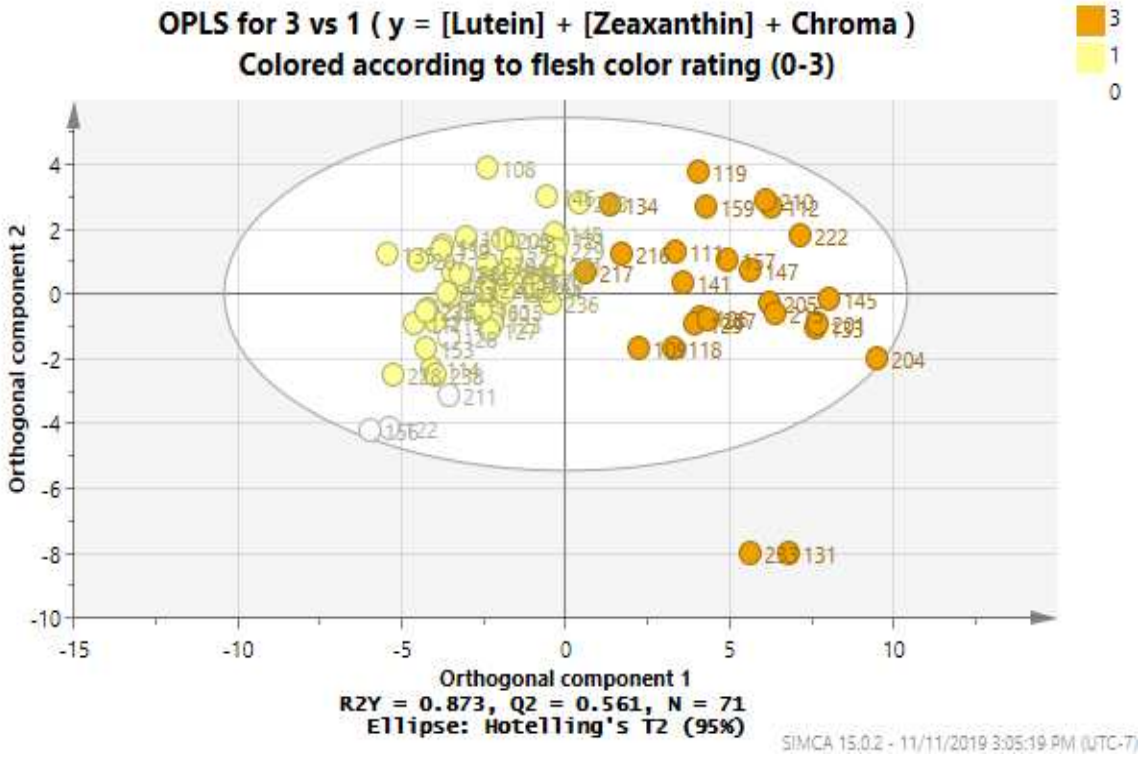

Figure 4.10. OPLS[a] of high carotenoids (FC 3) vs. low carotenoids (FC 1 & 0) as lutein, zeaxanthin and chroma as response variables (n[b] = 71). Outliers were excluded in the model outside the extreme range of the elliptical region (95% confidence interval).

[a]OPLS orthogonal components colored according to flesh color rating (0 = white, 1 = yellow, 3 = orange).
[b]Rep_1 (101-160), rep_2 (201-260). Excluding outliers, rep_3 (301-360), replicated samples with noisy spectra and samples with (FC 2) ratings

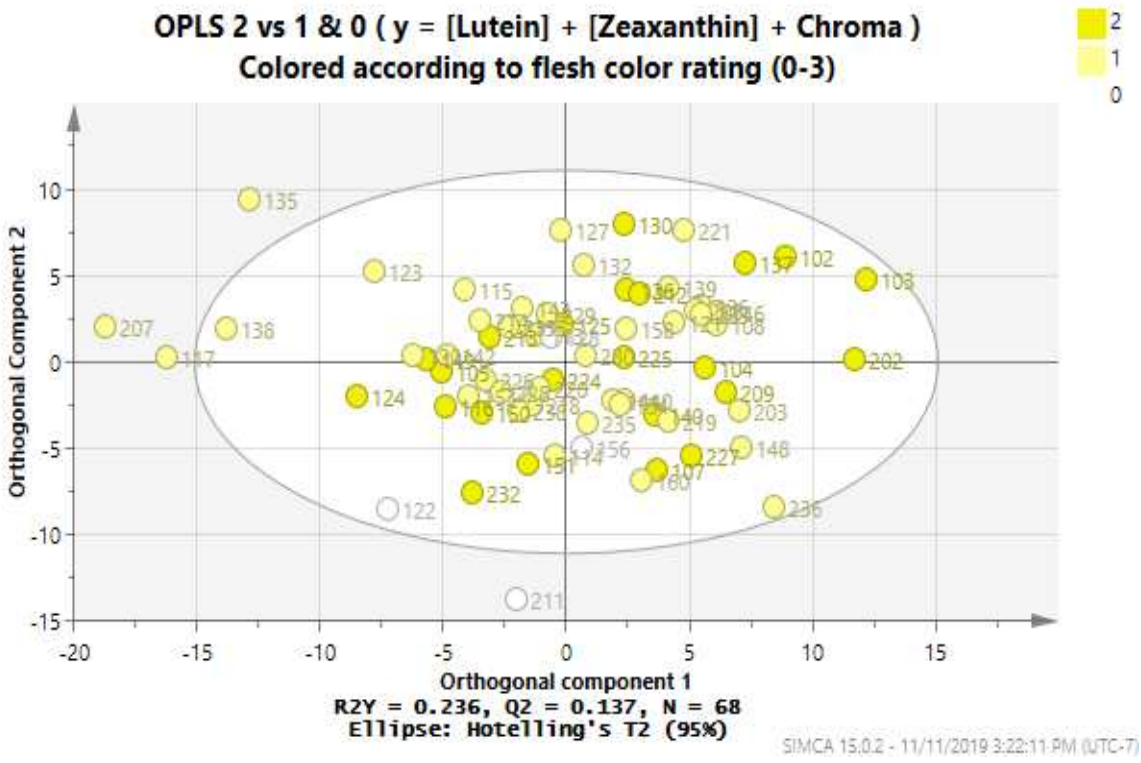

Figure 4.11. OPLS[a] of medium carotenoids (FC 2) vs. low carotenoids (FC 1 & 0) as lutein, zeaxanthin and chroma as response variables (n[b] = 68). Outliers were excluded in the model outside the extreme range of the elliptical region (95% confidence interval).

[a]OPLS orthogonal components colored according to flesh color rating (0 = white, 1 = yellow, 2 = Dark Yellow).
[b]Rep_1 (101-160), rep_2 (201-260). Excluding outliers, rep_3 (301-360), replicated samples with noisy spectra and samples with (FC 3) ratings.

4.4. REIMS Loadings Interpretation

4.4.1. Loadings and VIP Results

So far, multivariate analysis of OPLS has demonstrated that REIMS has some potential to differentiate samples of high carotenoids (FC 3) vs. low carotenoids (FC 1 & 0) among the yellow germplasm, but it did not give information about which metabolites are contributing to the alleged differentiation (Figure 4.10). REIMS loadings data was therefore analyzed to determine which metabolite masses differentiate samples with the aim of specifically identifying individual carotenoids of lutein and zeaxanthin. In addition, loadings were transformed into variable importance scores (VIP) to rank important metabolite masses, simplifying loadings analysis and helping with feature selection.

Loadings and VIP analyses were processed in SIMCA version 15.0.2 (Umetrics, Umea, Sweden). The best OPLS model of high carotenoids (FC 3) vs. low carotenoids (FC 1 & 0) was processed to extract loadings and VIP data (Figure 4.10). Since there were hundreds of mass bins associated to a theoretical metabolite that differentiates samples, only the top 20 variable importance scores (VIP > 0.8, Figure 4.12) were extracted to fit into an OPLS loadings model (Figure 4.13). It was observed that for the regression analysis for lutein and zeaxanthin carotenoids (y), mass bins such as 818.5, 817.5, and 739.5 (near-distance clustering) were among the highest contributors that may explain metabolite mass differentiation among samples (Figures 4.12 & 4.13).

4.4.2. Identification of Accurate and True Metabolite Mass

Mass bins reported were total of mass signals in the detected ion bin category (e.g., 818.5 m/z) and were not accurate masses. To report accurate mass data, the top five VIP mass bins ranks were analyzed from the raw REIMS spectra using the prototype Abstract Model Builder software, AMX version 1.0.1581.0 (Waters Research Centre, Budapest, Hungary). Manual

search of assigned peak accurate mass values (m/z) were processed from the software and were registered in the METLIN metabolite database (Scripps Research, San Diego, CA) to match with a true mass value (m/z) with a mass accuracy threshold of 30 ppm and protonated charge adducts ($[M+H]^+$ & $[M+Na]^+$, Table 4.2). Identified metabolites were registered to HMDB metabolome database (HMDB version 4.0, Alberta, Canada) to obtain organic compound class information. Four major organic compound classes of glycerophospholipid (GP), phosphatidylcholine (PC), phosphatidylglycerol (PG), and flavonoid-3-o-glycosides described the high-ranking metabolites from the loadings analysis (Table 4.2). These were mostly lipid molecules for which REIMS was claimed to be an effective instrument for detection (Waters Corporation 2015; Paxton 2020).

4.4.3. Identification of Theoretical Carotenoid Fragments

Unfortunately, precursor molecules of lutein and zeaxanthin (569.4359 m/z $[M+H]^+$, HMDB version 4.0, Alberta, Canada) were not detected from REIMS loadings and VIP analyses (Table 4.2, Figures 4.12 & 4.13). There was no detected mass ion bin category in the range of (569 ± 1.0 m/z) that was assigned for the target of interest of individual carotenoids. This was even though the data matrix model was explicitly created expending in the bin mass regions of 560 – 900 m/z to specifically detect lutein and zeaxanthin molecules. Instead, REIMS sensitivity to detect lipids were validated and might have contributed to the reason for the differentiation among samples (Table 4.2, Figures 4.10, 4.12 & 4.13).

It was most likely that the REIMS source was detecting fragmentation products of lutein and zeaxanthin. Previous studies have also reported the collection of theoretical metabolite fragments via REIMS source, and apparently a common phenomenon in positive or negative ionization mode (Stead et al. 2016; Black et al. 2019, Paxton 2020). In order to identify any theoretical carotenoid fragments of lutein and zeaxanthin, a raw REIMS spectrum of high

carotenoid genotype (OR04198-1) with a FC rating of (FC 3) was analyzed using the prototype Abstract Model Builder software, AMX version 1.0.1581.0 (Waters Research Centre, Budapest, Hungary).

Theoretical carotenoid fragments of lutein and zeaxanthin predicted spectra data were sourced from the HMDB metabolome database (HMDB version 4.0, Alberta, Canada), setting expected mass range of 100 – 550 m/z in positive-ionization mode with collision energy of 40V collected in a Q-TOF-MS instrument. In the predicted lutein and zeaxanthin spectra from the HMDB metabolome database, three carotenoid fragment motifs (385.2895, 425.3208, 537.4906 m/z, $[M+H]^+$) have approximate spectral peak matches with the raw REIMS spectrum of high carotenoid genotype (OR04198-1, Figure 4.14). Carotenoid fragments with mass range below (280 m/z) were excluded in the analysis due to high abundance of noise in the raw REIMS spectrum (Figure 4.14). Even though there were detected spectral peaks in the raw REIMS data that could be potential candidates for lutein or zeaxanthin fragments, further processing and analyses are needed to justify such claims, and the possibility of re-collecting raw REIMS fingerprint data among the 60 yellow germplasm to specifically identify carotenoid fragments.

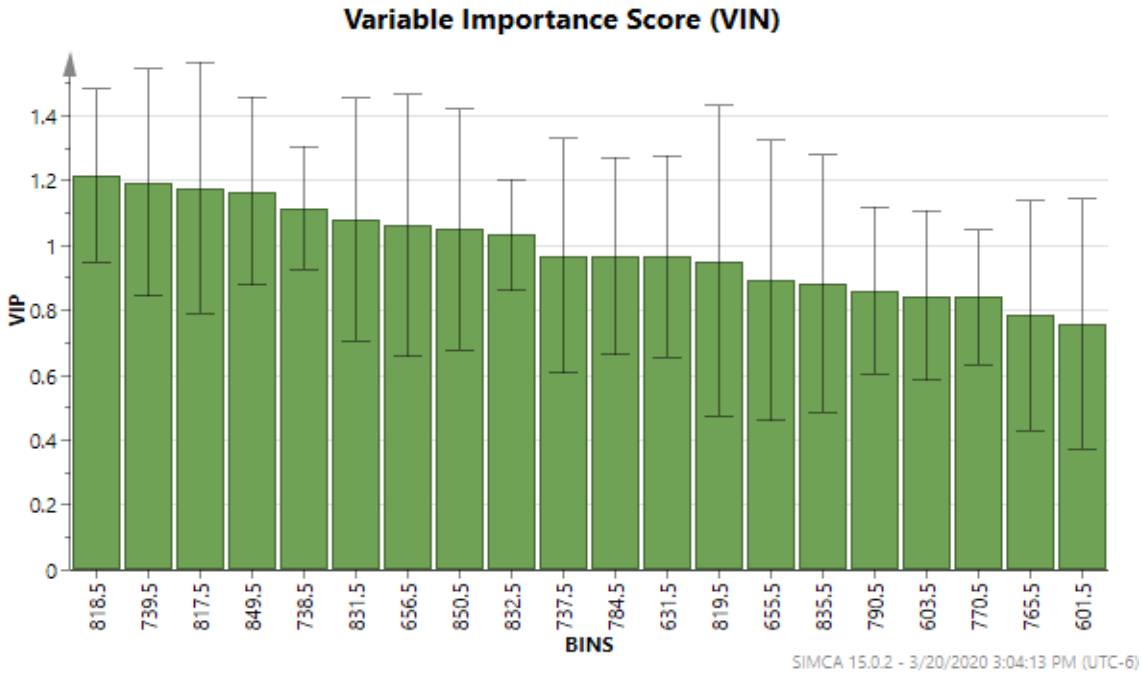

Figure 4.12. Top 20 variable importance scores (VIP > 0.8) that represent high-ranking metabolite ion mass bin categories extracted via best OPLS model (Figure 4.10). Mass bins reported as (m/z). Error bars represent SE (standard error of the mean, 2 reps).

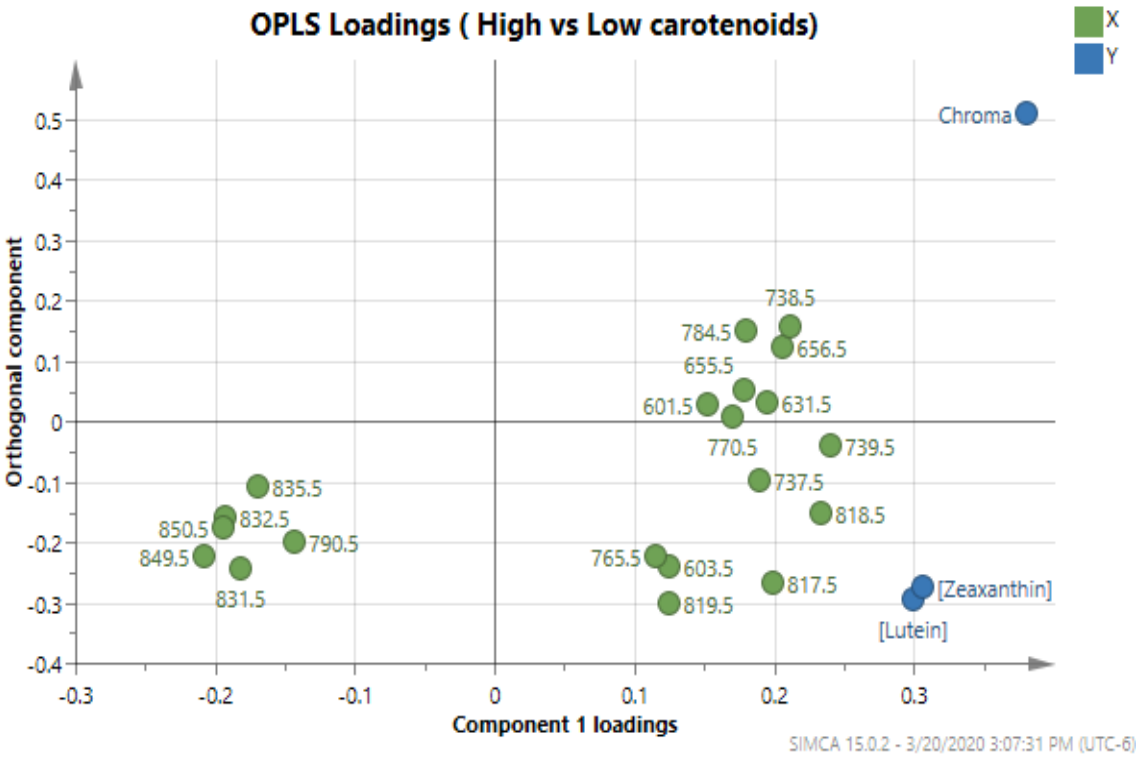

Figure 4.13. OPLS loadings[a] of high carotenoids (FC 3) vs. low carotenoids (FC 1 & 0). Top 20 VIP represent mass bins above. Mass bins in near-distance and clustering together will mostly co-vary with lutein and zeaxanthin (y).

[a]X = detected mass bins (m/z). Y = numerical response variables of lutein, zeaxanthin and chroma content.

Table 4.2. List of organic compounds classified based on REIMS loadings and VIP data.

Ion bin category[a]	Accurate mass (m/z)[b]	True mass (m/z)[c]	Δ ppm[d]	Organic compound class[e]	Adduct(s)
819.5	818.658	818.6633	6	Glycerophospholipid (GP)	$[M+H]^+$
818.5	818.552	818.5670	18	Phosphatidylcholine (PC)	$[M+Na]^+$
739.5	739.452	739.4521	0	Phosphatidylglycerol (PG)	$[M+Na]^+$
817.5	817.552	817.5460	7	Phosphatidylglycerol (PG)	$[M+H]^+$
849.5	849.252	849.2622	12	Flavonoid-3-o-glycosides	$[M+H]^+$

[a]Detected top five bin masses (Figures 12 & 13).
[b]Accurate mass detected via Abstract Model Builder software, AMX version 1.0.1581.0.
[c]True mass data sourced from METLIN metabolomics database.
[d]Calculated ppm error sourced from METLIN metabolomics database (threshold = 30 ppm).
[e]Listed organic compound classes of detected metabolites via HMDB metabolome database v. 4.0.

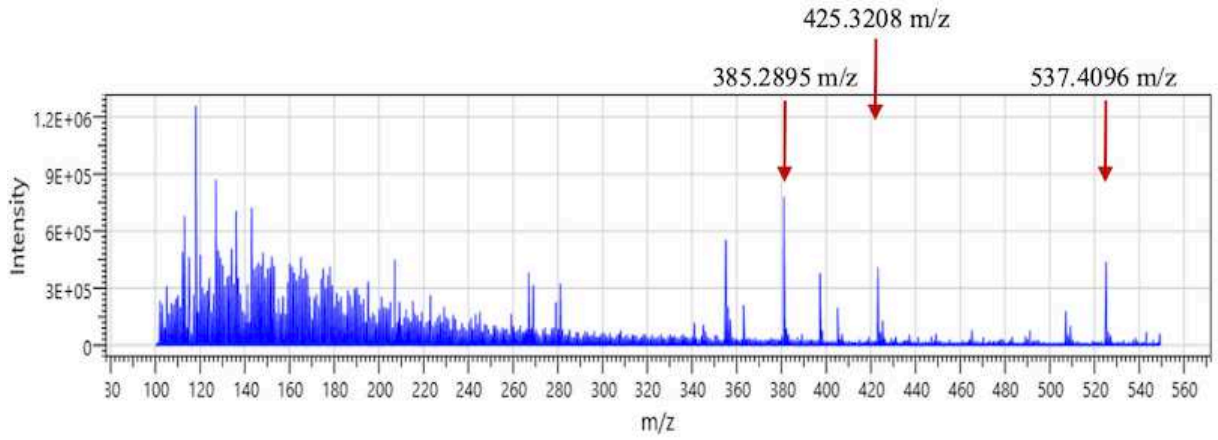

Figure 4.14. Raw REIMS spectrum (100 – 550 m/z, positive-ionization mode) of orange-fleshed (FC 3) genotype (OR04198-1) with potential carotenoid fragment peaks. Obtained via Abstract Model Builder software, AMX version 1.0.1581.0. Reported theoretical lutein and zeaxanthin fragment mass (m/z, $[M+H]^+$) above sourced from HMDB metabolome database v. 4.0.

CHAPTER 5. CONCLUSIONS AND DISCUSSION

One of the missions of the Colorado Potato Breeding and Selection Program is to develop potato cultivars with improved nutritional and health characteristics to benefit consumers. Biofortifying potato cultivars with increased lutein and zeaxanthin carotenoids can be achieved through plant breeding efforts. Efficient screening for high-carotenoid yellow breeding lines is therefore essential to develop new yellow cultivars with improved nutritional traits. The application of REIMS as a rapid phenotyping tool without the need for sample preparation and extraction protocols was explored for improving selection efficiency.

Phenotypic flesh color rating (0 – 3) technique was shown to be a handy method to evaluate tuber flesh color that would aid indirect selection for high-carotenoid potato lines. FC ratings of orange-fleshed genotypes (FC 3) have exhibited significantly higher overall carotenoid content from the HPLC analysis compared to white-fleshed (FC 0), yellow-fleshed (FC 1) and dark yellow-fleshed (FC 2) genotypes. Pearson correlation results suggested that chroma values measured by a reflectance colorimeter were moderately correlated with individual and total carotenoid content among the selected 60 yellow germplasm. However, the current reflectance colorimeter used in the breeding program did not precisely capture chroma in relative to individual and total carotenoid content. Multivariate analyses of PLS and OPLS models facilitated REIMS data interpretation. The OPLS model of high carotenoids (FC 3) vs. low carotenoids (FC 1 & 0) was considered the best model from this study due to the observable differentiation between sample replicates.

The primary objective of this study was to evaluate REIMS as a novel method to screen for carotenoids in yellow-fleshed potato germplasm. Consequently, the measurement focus was REIMS's sensitive capability to detect lutein and zeaxanthin carotenoids in real-time among the

60 selected yellow-fleshed potato germplasm. In the given outlined REIMS protocols in this study, the REIMS source did not detect parent molecules of lutein and zeaxanthin carotenoids. At most, REIMS detected carotenoid fragmentation products that might have contributed to sample separations observed from the best OPLS model, in addition to the identified and validated lipid molecules. REIMS could be considered as a good platform for an indirect technique to identify lutein and zeaxanthin carotenoids based on the results of this study.

There are assumptions of why REIMS's sensitivity to carotenoids was limited. Carotenoids did not ionize effectively in the presence of high energy electric current and high temperature, which the REIMS instrument basically did to potato flesh tissues. REIMS as a hard ionization method most likely degraded carotenoid molecules and merely collected fragmentation products. In future analyses, modifying the data matrix to detect carotenoid fragmentation products will be implemented. This can be achieved by setting the acquisition in the bin mass regions of 200 – 570 m/z with an aim to detect both parent and fragment ions as provided in Table 4.1. It is also possible to tweak the REIMS protocols that it could potentially allow detection of precursor carotenoids. For example, the energy electric current applied to tuber flesh could be lowered; however, there are tradeoffs, such as the possibility of collecting unstable spectra peaks due to less smoke signals generated by the instrument. Lipid molecules on the other hand ionized well with REIMS source. Furthermore, REIMS spectra were collected in positive-ionization mode, and while there is no guarantee that collection in negative mode will yield better results, the action is a possible protocol modification for any future analyses.

Generally, REIMS technology is relatively new to biological science and understanding how target ionization for a specific compound, in this case carotenoids, is still challenging. REIMS method optimization to specifically detect carotenoids in potatoes can be achieved. Since it has been shown that REIMS have potential to differentiate between high carotenoids (FC

3) vs. low carotenoids (FC 1 and 0), one approach for future method development is to only include white-fleshed vs. orange-fleshed potato genotypes with a relatively smaller population size (e.g., N = 20). Another possible method development is to spike or inject white-fleshed tubers or potato powder with known authentic carotenoid standards with an aim to analyze individual carotenoids while maintaining high moisture content on all samples. In addition, in a potato breeding model perspective, rank order analysis of yellow breeding lines can be informative to guide selection of high carotenoid genotypes. This can be done by applying individual carotenoid data via HPLC to rank yellow lines with the highest to lowest carotenoid content and regress the data matrix in a PLS model. Although, PLS analysis have the limitation of including all metabolites in the model such as the inclusion of those that are not relevant to sample differentiation. Based on the findings of Gredell and colleagues, incorporation of machine learning algorithms of dimension reduction and feature selection can potentially improve the predictive power (Q^2) of all models in this study and will be part of the protocol development in future REIMS analyses (Gredell et al. 2019).

There are also other factors that need to be considered in the application of REIMS technology in potato improvement. Currently, the REIMS apparatus is estimated to cost about $500,000 USD (Waters Corporation 2015) and for the use of the instrument provided by the Proteomics and Metabolomics Facility at Colorado State University for $64 per hour, the application itself is impractically expensive. Moreover, REIMS raw spectra contain high abundance of noise. In REIMS MVA analyses, all Rep 3 samples in addition to replicated samples with noisy spectra were excluded just to attain meaningful MVA models. Potato tubers are comprised mostly of starchy compounds, possibly aggregating inside the REIMS instrumentation resulting in noise abundance. Also, the method of applying the iknife to tuber flesh was manual. If iknife burns were automated in some way using robotics, it may contribute

to more consistent collection of fingerprint mass spectra data in a high-throughput manner and possibly with lower noise. REIMS data analyses can arguably be complex, relying on multiple software platforms, pre-processing protocols and color and carotenoid data from different analytical platforms to improve MVA analyses. Finally, REIMS is not a portable instrument that can be transported to the field and evaluate carotenoid content among hundreds of yellow breeding lines. Use of direct analysis in real time mass spectrometry (DART-MS) is a prospective efficient platform for carotenoid detection and quantification (Gross 2014). Similar to REIMS technology, DART-MS is capable of *in situ* analysis on biological samples, and collecting rapid mass fingerprint data in real-time, but it relies on gas-phase ionization mechanisms (Gross 2014). DART-MS has been proven to detect xanthophyll carotenoids, particularly beta-cryptoxanthin, making it an incentive future project for screening carotenoids in yellow potato germplasm (Ma et al. 2017).

Ideally, having a portable device that is fairly non-destructive, no extractions approach to tuber flesh tissues that directly measures color in high precision and carotenoid content in real-time is the preferred tool for selection of high-carotenoid potato genotypes. Fortunately, such portable devices that can potentially measure color and carotenoid content in potato tuber flesh without sample preparations and extractions exist in the market today. A handheld Konica Minolta Spectrophotometer (CM-700d) is a practical candidate tool to screen for chroma that may potentially have high correlations with carotenoid content in yellow potato germplasm. According to the brochure information, the Konica Minolta CM-700d instrument is a spectrophotometer that is designed to evaluate the color of objects producing three-dimensional color space coordinate values of L*, a*, and b* at high precision and measure wavelengths from 400-700 nm which may potentially cover carotenoid (450 nm) compounds (Konica Minolta 2007). One unit costs approximately $10,000 USD, a fraction of the cost compared to REIMS.

In terms of future directions for carotenoid screening in yellow potato germplasm, the Colorado Potato Breeding and Selection Program should consider exploring more practical and relatively lower cost tools such as the Konica Minolta CM-700d spectrophotometer with an objective to improve overall selection efficiency.

LITERATURE CITED

Acquaah, G. 2012. Breeding Potato. *Principles of plant genetics and breeding (2nd ed.)*. John Wiley & Sons. (pp. 647 – 656).

Alexander, J., Gildea, L., Balog, J., Speller, A., McKenzie, J., Muirhead, L., & Hoare, J. 2017. A novel methodology for in vivo endoscopic phenotyping of colorectal cancer based on real-time analysis of the mucosal lipidome: a prospective observational study of the iKnife. *Surgical endoscopy*, *31*(3), 1361-1370.

Andre, C. M., Ghislain, M., Bertin, P., Oufir, M., del Rosario Herrera, M., Hoffmann, L., & Evers, D. 2007. Andean potato cultivars (Solanum tuberosum L.) as a source of antioxidant and mineral micronutrients. *Journal of agricultural and food chemistry*, *55*(2), 366-378.

Avendano, M. 2012. Correlation or causation? Income inequality and infant mortality in fixed effects models in the period 1960–2008 in 34 OECD countries. *Social science & medicine*, *75*(4), 754-760.

Balog, J., Perenyi, D., Guallar-Hoyas, C., Egri, A., Pringle, S. D., Stead, S., & Takats, Z. 2016. Identification of the species of origin for meat products by rapid evaporative ionization mass spectrometry. *Journal of agricultural and food chemistry*, *64*(23), 4793-4800.

Bamberg, J. B. 2010. Tuber dormancy lasting eight years in the wild potato Solanum jamesii. *American journal of potato research*, *87*(2), 226-228.

Black, C., Chevallier, O. P., Haughey, S. A., Balog, J., Stead, S., Pringle, S. D., & Nikolopoulos, D. S. 2017. A real time metabolomic profiling approach to detecting fish fraud using rapid evaporative ionisation mass spectrometry. *Metabolomics*, *13*(12), 153.

Black, C., Chevallier, O. P., Cooper, K. M., Haughey, S. A., Balog, J., Takats, Z., & Cavin, C. 2019. Rapid detection and specific identification of offals within minced beef samples utilising ambient mass spectrometry. *Scientific Reports*, *9*(1), 1-9.

Bohl, W. and S. Johnson. 2010. Commerical potato production in North America. *Potato Association of America handbook* 1-85.

Bonierbale, M.W., R.L. Plaisted, and S.D. Tanksley. 1988. RFLP maps based on a common set of entries reveal modes of chromosomal evolution in potato and tomato. *Genetics* 120:1095-1103.

Bonierbale, M., Grüneberg, W., Amoros, W., Burgos, G., Salas, E., Porras, E., & zum Felde, T. 2009. Total and individual carotenoid profiles in Solanum phureja cultivated potatoes: II. Development and application of near-infrared reflectance spectroscopy (NIRS)

calibrations for germplasm characterization. *Journal of Food Composition and Analysis*, *22*(6), 509-516.

Britton, G., Liaaen-Jensen, S. and Pfander, H. 2004. *Carotenoids Handbook*. Basel, Switzerland: Birkhauser.

Brown, C., C. Edwards, C. Yang, and B. Dean. 1993. Orange flesh trait in potato: inheritance and carotenoid content. *Journal of the American Society for Horticultural Science* 118(1):145-150.

Brown C., C. Yang, D. Navarre, and D. Culley. 2004. Carotenoid and anthocyanin concentrations and associated antioxidant values in high pigment potatoes. American *Journal of Potato Research* 81:48.

Brown, C. 2005. Antioxidants in potato. *American Journal of Potato Research* 82:163-172.

Brown, C., T. Kim, Z. Ganga, K. Haynes, D. De Jong, M. Jahn, I. Paran, and W. De Jong. 2006. Segregation of total carotenoid in high level potato germplasm and its relationship to beta-carotene hydroxylase polymporphism. *American Journal of Potato Research* 83:365-372.

Brown, C., D. Culley, M. Bonierbale, W. Amoros, and ARS, USDA. 2007. Anthocyanin, carotenoid content, and antioxidant values in native South American potato cultivars. *HortScience* 42(7):1733-1736.

Brown, C. 2008. Breeding for phytonutrient enhancement of potato. American Journal of Potato Research 85:298-307.

Burgos, G., Salas, E., Amoros, W., Auqui, M., Munoa, L., Kimura, M., & Bonierbale, M. 2009. Total and individual carotenoid profiles in Solanum phureja of cultivated potatoes: I. Concentrations and relationships as determined by spectrophotometry and HPLC. *Journal of Food Composition and Analysis*, *22*(6), 503-508.

Bylesjö, M., Eriksson, D., Kusano, M., Moritz, T., & Trygg, J. 2007. Data integration in plant biology: the O2PLS method for combined modeling of transcript and metabolite data. *The Plant Journal*, *52*(6), 1181-1191.

Cao G, B Shukitt-Hale, PC Bickford, JA Joseph, J McEwen, and RL Prior. 1999. Hyperoxia-induced changes in antioxidant capacity and the effect of dietary antioxidants. J Appl Physio186:1817-1822.

Chandler, K., Lipka, A. E., Owens, B. F., Li, H., Buckler, E. S., Rocheford, T., & Gore, M. A. 2013. Genetic analysis of visually scored orange kernel color in maize. *Crop Science*, *53*(1), 189-200.

DeAngelis, M. M., Owen, L. A., Morrison, M. A., Morgan, D. J., Li, M., Shakoor, A., & Farrer, L. A. 2017. Genetics of age-related macular degeneration (AMD). *Human molecular genetics*, *26*(R1), R45-R50.

De Ritter, E., & Purcell, A. E. 1981. Carotenoid analytical methods. In *Carotenoids as colorants and vitamin A precursors* (pp. 815-923). Academic Press.

Diretto, G., Al-Babili, S., Tavazza, R., Papacchioli, V., Beyer, P., & Giuliano, G. 2007. Metabolic engineering of potato carotenoid content through tuber-specific overexpression of a bacterial mini-pathway. *PLoS One*, *2*(4).

Essah, S.Y.C., Davidson, R.D., Davis, J.D. 2014. Fertilizing Potatoes. Colorado State University Extenstion. Fact Sheet No. 0.541. Retrieved from: http://extension.colostate.edu/docs/pubs/crops/00541.pdf

Fernandez-Orozco, R., Gallardo-Guerrero, L., & Hornero-Méndez, D. 2013. Carotenoid profiling in tubers of different potato (Solanum sp) cultivars: Accumulation of carotenoids mediated by xanthophyll esterification. *Food chemistry*, *141*(3), 2864-2872.

Food and Agriculture Organization of the United Nations (FAO). 2008. The potato plant. Retrieved from: http://www.fao.org/potato-2008/en/potato/index.html

Fraser, P. D., Enfissi, E. M., Goodfellow, M., Eguchi, T., & Bramley, P. M. 2007. Metabolite profiling of plant carotenoids using the matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *The Plant Journal*, *49*(3), 552-564.

Gredell, D. A., Schroeder, A. R., Belk, K. E., Broeckling, C. D., Heuberger, A. L., Kim, S. Y., & Woerner, D. R. 2019. Comparison of machine learning algorithms for predictive modeling of beef attributes using rapid evaporative ionization mass spectrometry (REIMS) data. *Scientific reports*, *9*(1), 1-9.

Gross, J. H. 2014. Direct analysis in real time—a critical review on DART-MS. *Analytical and bioanalytical chemistry*, *406*(1), 63-80.

Guaratini, T., Vessecchi, R., Pinto, E., Colepicolo, P., & Lopes, N. P. 2005. Balance of xanthophylls molecular and protonated molecular ions in electrospray ionization. *Journal of mass spectrometry*, *40*(7), 963-968.

Hardigan, M. A., Laimbeer, F. P. E., Newton, L., Crisovan, E., Hamilton, J. P., Vaillancourt, B., & Veilleux, R. E. 2017. Genome diversity of tuber-bearing Solanum uncovers complex evolutionary history and targets of domestication in the cultivated potato. *Proceedings of the National Academy of Sciences*, *114*(46), E9999-E10008.

Haynes, K. G., Clevidence, B. A., Rao, D., & Vinyard, B. T. 2011. Inheritance of carotenoid content in tetraploid× diploid potato crosses. *Journal of the American Society for Horticultural Science*, *136*(4), 265-272.

Haynes, K., B. Clevidence, D. Rao, B. Vineyard, and J. White. 2010. Entry x environment interactions for potato tuber carotenoid content. Journal of the American Society for Horticultural Science 135(3):250-258.

Hijmans, R. J., & Spooner, D. M. 2001. Geographic distribution of wild potato species. *American Journal of Botany*, *88*(11), 2101-2112.

HunterLab. 2003. MiniScan XE plus user's guide: measurement values. 10-1.

Jansen G, Flamme W, Schüler K, Vandrey M. 2001. Tuber and starch quality of wild and cultivated potato species and cultivars. Potato Res 44:137–146.

Konica Minolta. 2007. Spectrophotometer CM-700d/600d Brochure. Retrieved from: https://sensing.konicaminolta.us/uploads/cm-700d_600d_catalog-5v0252564m.pdf

Konschuh, M., McAllister, T., DalpØ, S., & Lewis, T. 2005. Assessment of Carotenoid Content of Yellow-Fleshed Potato Varieties Grown in Alberta to Determine Potential Nutritional Benefits, Final Report Project: 2004-008, Alberta Agricultural. *Food and Rural Development, Alberta*, 1-19.

Kotíková, Z., Šulc, M., Lachman, J., Pivec, V., Orsák, M., & Hamouz, K. 2016. Carotenoid profile and retention in yellow-, purple-and red-fleshed potatoes after thermal processing. *Food chemistry*, *197*, 992-1001.

Lachman, J., Hamouz, K., Musilová, J., Hejtmánková, K., Kotíková, Z., Pazderů, K., & Cimr, J. 2013. Effect of peeling and three cooking methods on the content of selected phytochemicals in potato tubers with various colour of flesh. *Food Chemistry*, *138*(2-3), 1189-1197.

Lachman, J., Hamouz, K., Orsák, M., & Kotíková, Z. 2016. Carotenoids in potatoes–a short overview. *Plant, Soil and Environment*, *62*(10), 474-481.

Liu, R. H. 2003. Health benefits of fruit and vegetables are from additive and synergistic combinations of phytochemicals. American Journal of Clinical Nutrition 78:517S-520S.

Lu, W., K. Haynes, E. Wiley, and B. Clevidence. 2001. Carotenoid content and color in diploid potatoes. Journal of the American Society for Horticultural Science 126(6):722–726.

Ma, G., Zhang, L., Iida, K., Madono, Y., Yungyuen, W., Yahata, M., & Kato, M. 2017. Identification and quantitative analysis of β-cryptoxanthin and β-citraurin esters in Satsuma mandarin fruit during the ripening process. *Food chemistry*, *234*, 356-364.

McDonald, S. K., Hofsteen, L. Downey, L. 2003. Crop Profile for Potatoes in Colorado. Retrieved from: http://www.ipmcenters.org/cropprofiles/docs/copotatoes.pdf

McGuire, R.G. 1992. Reporting of objective color measurements. HortScience 27(12):1254-1255.

National Potato Council. 2019. *Potato statistical yearbook*. Washington, DC: National Potato Council.

Palozza, P., & Krinsky, N. I. 1992. Antioxidant effects of carotenoids in Vivo and in Vitro: An overview. In *Methods in enzymology* (Vol. 213, pp. 403-420). Academic Press.

Paxton, T. 2020. Rapid evaporative ionization mass spectrometry. In *Ambient Ionization Mass Spectrometry in Life Sciences* (pp. 241-270). Elsevier.

Perla, V., Holm, D. G., & Jayanty, S. S. 2012. Effects of cooking methods on polyphenols, pigments and antioxidant activity in potato tubers. *LWT-Food Science and Technology*, *45*(2), 161-171.

Pieterse, L., & Hils, U. 2009. *World catalogue of potato varieties*. Agrimedia.

Rezanka, T., Olsovska, J., Sobotka, M., & Sigler, K. 2009. The use of APCI-MS with HPLC and other separation techniques for identification of carotenoids and related compounds. *Current Analytical Chemistry*, *5*(1), 1-25.

Rivera, S. M., Christou, P., & Canela-Garayoa, R. 2014. Identification of carotenoids using mass spectrometry. *Mass Spectrometry Reviews*, *33*(5), 353-372.

Rodriguez-Amaya, D. B., & Kimura, M. 2004. *HarvestPlus handbook for carotenoid analysis* (Vol. 2). Washington, DC: International Food Policy Research Institute (IFPRI).

Rutkoski, J. 2019. A practical guide to genetic gain. *Advances in agronomy*, *157*, 217-249.

Saini, R. K., Nile, S. H., & Park, S. W. 2015. Carotenoids from fruits and vegetables: Chemistry, analysis, occurrence, bioavailability and biological activities. *Food Research International*, *76*, 735-750.

Schäfer, K. C., Dénes, J., Albrecht, K., Szaniszló, T., Balog, J., Skoumal, R., & Takáts, Z. 2009. In vivo, in situ tissue analysis using rapid evaporative ionization mass spectrometry. *Angewandte Chemie International Edition*, *48*(44), 8240-8242.

Scott, G. J., Petsakos, A., & Suarez, V. 2019. Not by Bread Alone: Estimating Potato Demand in India in 2030. *Potato Research*, *62*(3), 281-304.

Seddon, J. M., Ajani, U. A., Sperduto, R. D., Hiller, R., Blair, N., Burton, T. C., & Yannuzzi, L. A. 1994. Dietary carotenoids, vitamins A, C, and E, and advanced age-related macular degeneration. *Jama*, *272*(18), 1413-1420.

Spooner, D. M. 2013. Solanum tuberosum (Potatoes). *Brenner's Encyclopedia of Genetics. Maloy and K. Hughes (eds.), 2nd ed. Academic Press. Waltham, MA*, 481-483.

Stead, S. J. Pope, R. Edwards, M. Brazier-Hicks, C. Tetard-Jones, Z. Takats. 2016. REIMS (rapid evaporative ionization mass spectrometry) and multivariant statistics, two tools in

support of weed grass speciation and phenotype characterization, Poster Session presented at: 64th Conference on Mass Spectrometry and Allied Topics (ASMS)

Tanaka, Y., Sasaki, N., & Ohmiya, A. 2008. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *The Plant Journal*, *54*(4), 733-749.

Tiemens-Hulscher, M., Delleman, J., Eisinger, E., & Van Bueren, E. L. 2013. Potato breeding: a practical manual for the potato chain. (pp. 10 – 11, 42)

United States Department of Agriculture (USDA). 2016. United States Department of Agriculture Economic Research Service. Vegetable & Pulses. *Potatoes.* Retrieved from: https://www.ers.usda.gov/topics/crops/vegetables-pulses/potatoes/

United States Department of Agriculture National Agricultural Statistics Service (USDA-NASS). 2019. *Potatoes 2018 Summary.* ISSN: 1949-1514

Voss, D.H. 1992. Relating colorimeter measurement of plant color to the Royal Horticultural Society Colour Chart. HortScience 27(12):1256-1260.

Wang H, MG Nair, GM Strasburg, YC Chang, AM Booren, JI Gray, and DL DeWitt. 1999. Antioxidant and anti-inflammatory activities of anthocyanins and their aglycon, cyanidin, from tartcherries. JNat Prod 62:294-296.

Waters Corporation. 2015. REIMS Research System With iKnife Sampling Brochure. Retrieved from: https://www.waters.com/waters/library.htm?lid.134846772.

Zaheer, K., & Akhtar, M. H. 2016. Potato production, usage, and nutrition—a review. *Critical reviews in food science and nutrition*, *56*(5), 711-721.