THESIS

GENETIC AND METABOLOMIC ANALYSES OF BARLEY AND COWPEA: IMPLICATIONS ON QUALITY AND NUTRITION OF FINISHED FOODS

Submitted by

Brooke Sayre-Chavez

Department of Soil and Crop Science

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2021

Master's committee:

Advisor: María Muñoz-Amatriaín

Elizabeth P. Ryan

Adam L. Heuberger

ABSTRACT

GENETIC AND METABOLOMIC ANALYSES OF BARLEY AND COWPEA: IMPLICATIONS ON QUALITY AND NUTRITION OF FINISHED FOODS

The finished foods of a cereal (barley) and a legume (cowpea) were subjects of this thesis and analyzed in two separate studies. High quality barley that meets malting standards, is economically worth billions each year to the malting and brewing industry. The prevalence of craft brewing has been on the rise and with that, an increased interest in understanding the basis of beer flavor. Malt has been the subject of most research on beer flavor, and currently there is a lack of understanding on the contributions that the barley variety has to product flavor. The second crop that was subject of this thesis is cowpea. Cowpea is a nutritious food, that grows well in sub-Saharan Africa, where malnutrition is prevalent. It is well adapted to the resource-poor farming practices common in these regions, and highly valued as a food security crop. Despite the known health benefits, potential to alleviate malnutrition, and use in nutritional studies, there are no biomarkers identified for cowpea and its metabolic profile is currently not well characterized. The research goals of this thesis are broken down by crop. Regarding barley, the goals were to 1) test the hypothesis that barley genotype contributes to beer flavor, 2) to identify regions of the genome that control traits associated with flavor, and 3) identify candidate genes that control traits associated with flavor. Regarding cowpea, the goals were to 1) characterize the metabolic profile of three cowpea flours from varieties commonly consumed in sub-Saharan Africa (Ghana, specifically), and 2) test the hypothesis that there are metabolites unique to cowpea (and cowpea varieties). Metabolomic approaches were applied to both crops as

finished foods, with additional genetic analysis of barley. We concluded that barley genotype does contribute to beer flavor, and that cowpea has distinct and characterizable metabolomic differences from other legumes. In barley, QTLs (quantitative trait loci) for malt quality, beer sensory, and metabolite traits were mapped, and candidate genes identified. The results of this study set a foundation for future genetic and breeding efforts surrounding barley and beer flavor, allowing for integration of various quality attributes. In cowpea, comparisons were made between cowpea, pigeon pea (another legume common to sub-Saharan Africa), and common bean on two non-targeted mass spectrometry platforms. Comparisons between the legumes illuminated metabolites that were either common to, or unique to each legume type or variety. The annotated metabolites from both analyses were from a diverse set of classifications and metabolic pathways, many with numerous known nutritional benefits. The metabolomic profiling of cowpea (and cowpea varieties) will allow for easier identification of nutritional biomarkers in future feeding studies.

ACKNOWLEDGMENTS

I would like to extend my deepest gratitude to my advisor, Dr. María Muñoz-Amatriaín for choosing me as your first graduate student. Learning from you was a true pleasure. All your guidance, advice, and encouragement has been incredibly valuable to me.

I would also like to express my gratitude towards the rest of my committee, Dr. Elizabeth P. Ryan and Dr. Adam L. Heuberger, for your time, support, and mentorship.

A special thanks to Dr. Patrick Hayes as well, who although wasn't a part of my committee, was a true mentor and was always incredibly enthusiastic.

I also owe thanks to the many professors I've had the pleasure of learning from during my time at CSU.

And a particularly special thanks to my family and friends whose excitement and support made this journey all the more enjoyable.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Finished foods from a cereal (barley) and a legume (cowpea) are subjects of this thesis. Barley, one of the most important temperate-climate cereals, is the most used grain in the malting and brewing industry, which contributed $116 billion to the U.S. economy in 2019 (Brewers Association). Cowpea, a warm-season legume, is a major food security crop that provides food, nutrition, and a livelihood to millions in sub-Saharan Africa (Gómez, 2004). Genetic and metabolomic analyses were performed on finished beers to get a better understanding of the impact of the barley genotype on beer flavor. Metabolic profiles from flours of cowpea varieties commonly consumed in Ghana were also generated and analyzed to identify potential biomarkers for nutritional studies.

Barley

Barley (*Hordeum vulgare* L.) is a cereal grain, belonging to the Poaceae family and Triticeae tribe, first domesticated in the Fertile Crescent over 10,000 years ago (Badr et al., 2000; von Bothmer et al., 2003). Barley is highly adaptable to a wide range of climates and soil conditions, more so than other cereals, making it the fourth most important cereal in the world in terms of production and cultivated area (Shewry and Ullrich, 2014; Tricase et al., 2018). Other major crops belonging to the Triticeae tribe include wheat (*Triticum aestivum* L.) and rye (*Secale cereale* L.) (von Bothmer et al., 2003). Barley is a monocot, diploid (2n = 14), self-pollinating species with a genome size of about 5.1 Gb (Kumlehn and Stein, 2014; Sato, 2020). Ample genetic and genomic resources exist for barley, including a reference genome sequence (Mascher et al., 2017), making it ideal for genetic studies (Colmsee et al., 2015; Sato, 2020).

A diversity of barley varieties exists, which can be classified as winter, spring or facultative types, and two-row or six-row barleys. Winter barley requires vernalization (prolonged exposure to cold temperatures) to flower, and it is sown in the fall, while spring barley does not require vernalization and can be sown in spring. Facultative varieties do not require vernalization, but they are tolerant of cold temperatures, so they can be planted in the fall or in the spring (Kolar et al., 1991). These can exist as either six or two-rowed varieties based on their spike morphology. Barley heads have spikelets in alternating sets of three (triplets) but in two-rowed barley the side florets are sterile and do not develop a seed (Kumlehn and Stein, 2014; von Bothmer et al., 2003). The difference in number of kernels developed impacts the grain size, as well as protein and starch contents (Magliano et al., 2014). Most breeding efforts to date have been focused on agronomic traits such as lodging and grain yield, disease resistance traits, and biotic and abiotic stress tolerance (Briggs, 1998; Friedt et al., 2011; von Bothmer et al., 2003). Efforts have also been made to breed for barley with increased malt quality, via lower protein content, higher malt extract, and plumper kernels (Anderson and Reinsbergs, 1985).

Importance of barley to the malting and brewing industry

Although some of the main uses of barley include animal feeding and human consumption, the greatest economic impact of barley is associated with the malting and brewing industries. To be accepted for malt production, barley grain must meet strict quality standards (Ash and Hoffman, 1990; Shewry and Ullrich, 2014). Many organizations exist around the world that set standards and approve new malting varieties [American Malting Barley Association (AMBA), the Brewing and Malting Barley Research Institute (BMBRI), Barley Australia, the Canadian Malting Barley Technical Centre (CMBTC), and the European Brewery Convention (EBC) to name a few]. Strict guidelines and lengthy testing requirements are in place to test new

varieties, and the genetics of these malting quality traits have been largely explored (Fang et al., 2019; Han et al., 1997; Igartua et al., 2000; Mather et al., 1997; Mohammadi et al., 2015; Muñoz-Amatriaín et al., 2010; Potokina et al., 2004; Szűcs et al., 2009). Typically, malting barley will be lower in protein and higher in starch content than feed barley (Paynter, 2015; Shewry and Ullrich, 2014). Two-rowed barley tends to be preferred by the brewing industry because of its greater extract value and lower protein content compared to six-rowed barley. This is because two-row kernels tend to be larger, and large seeds are associated with lower protein content (Magliano et al., 2014). Some of the recommended guidelines put forth by AMBA for quality two-row barley include: >98% germinative energy, >90% kernels retained on a 6/64" sieve, and <12-13% protein content (American Malting Barley Association, 2019). Individual components of malting quality include kernel plumpness, kernel weight, barley color, barley protein, malt extract, wort color, wort protein, soluble/total protein (S/T), diastatic power (DP), β-glucan (BG), α-amylase (AA), and free amino nitrogen (FAN). These are further explained in the section below. Targets for these traits vary depending on the intended use of the malt (all-malt brewing, adjunct brewing, or distilling).

The first evidence of brewing and beer consumption predates the invention of bread (Tremblay and Tremblay, 2005). For as long as grains have been cultivated, they have been stored in some way. Natural fermentation can occur during storage under the right conditions, which is what most likely initiated the first discovery of "beer" making (Kumlehn and Stein, 2014). Since becoming an important part of many cultures, the brewing industry has become more formal and economically impactful on a global scale. In the U.S. alone, Americans drank $101.5 billion worth of beer in 2014 (Reid and Gatrell, 2015). As the industry grew, so did the demand for a diversity of styles and flavors (Withers, 2017). This led to the rise of craft brewing.

In 1981 there were only five established craft breweries in the U. S., but by 2014, there were over 3,400. This correlated to $19.6 billion worth of beer bought from designated craft breweries in the U.S. that year (Reid and Gatrell, 2015). Since then, the industry has only grown. Craft brewing is smaller scale in production than commercial brewing and tends to be limited to local markets. There is also more of an emphasis on distinct flavors, although it is acknowledged that there is a knowledge gap in the understanding the basis of these flavors (Brewers Association, 2015).

Malting and brewing

Malt is a product of the controlled germination of a cereal grain. The process of malting involves three steps: steeping, germination, and kilning. Briefly, steeping initiates germination of the grain by alternations of wet and dry cycles, germination involves chemical changes to the grain, and kilning halts germination through drying of the grain (Mallet, 2014; Paynter, 2015; Paynter and Young, 1996). Heavily relied upon in the brewing and distilling industries, malt is a one of the four main ingredients in beer (along with water, yeast, and hops) that provides critical starches and enzymes needed for fermentation (Briggs, 1998; Paynter and Young, 1996). The malting process breaks down the internal cell walls of the grain, stimulates the production of diastatic enzymes (which convert starch to malt extract), and contributes to color and flavor of the malt (Paynter and Young, 1996). Kilning also contributes heavily to the flavor and color of the malt through alterations in temperature and time (Mallet, 2014; Paynter, 2015; Paynter and Young, 1996). Many malting methods exist, contributing to the wide range of beer styles made from various base and specialty malts.

Brewing is the process of converting a starchy grain into an alcoholic beverage by fermentation. It involves mashing, lautering, boiling, cooling, fermentation, and maturing/aging

steps. Overall, the process breaks down starches provided by the malted barley grain, to create a sugary solution called wort, where yeast consumes the sugars and releases ethanol along with other metabolic byproducts during fermentation (Willaert, 2007). The barley grains provides important enzymes that catalyze major biological changes during the brewing process, a few of these enzymes include α-amylases, β-amylases, β-glucanases, β-glucosidases, and limit dextrinases (Briggs, 1998; Evans et al., 2010). To describe the brewing process briefly, mashing starts the process by mixing the malted grain with water and heating to create wort and start degradation of the cell wall, starches, and proteins; lautering filters the insoluble material out from the wort; boiling initiates a variety of complex physical and chemical reactions (including enzyme inactivation and fixation of carbohydrate composition) that prepare the wort for fermentation; cooling removes excess protein and clarifies the beer, fermentation allows yeast to consume available sugars in exchange for ethanol and carbon dioxide production; and maturing prepares beer for packaging by reduction of off-flavors and removal of yeast (Lipnizki, 2015a, 2015b; Masschelein, 1986; Willaert, 2007).

Defining malt quality traits

Some of the most important malt quality traits include: 1) kernel plumpness, 2) kernel weight, 3) barley color, 4) barley protein, 5) malt extract, 6) wort color, 7) wort protein, 8) soluble/total protein (S/T), 9) diastatic power (DP), 10) β-glucan (BG), 11) α-amylase (AA), and 12) free amino nitrogen (FAN). These are described below.

1. *Kernel plumpness*

Kernel plumpness, sometimes referred to as percent on 6/64", is a measurement of seed size. Kernels that are retained on a 6/64" sieve when sifted are counted in the

percentage. AMBA recommends >90% retention of kernels on a 6/64” sieve (American Malting Barley Association, 2019). Plump kernels are preferred because they tend to contain more starch, which correlates to more beer per weight from the malt. Two-rowed varieties tend to be plumper and more utilized in the malting and brewing industry than six-rowed ones (Magliano et al., 2014; Wrigley et al., 2016).

2. *Kernel weight*

Kernel weight is a factor correlated to kernel plumpness and itis another way to measure potential quality. Typically, higher kernel weights mean more malt extract and higher quality malt (Mather et al., 1997). Kernel weight is measured as a thousand seed weight, reported in grams and representative of the dry weight.

3. *Barley color*

Dark or stained barley can indicate the presence of unwanted bacteria or incomplete germination due to wet weather at harvest. This typically leads to undesirable flavors, so bright, light yellow barley is desired. Barley color also varies based on the growing environment. Wet climates tend to have duller, darker colored grains than dry climates (Skarsaune et al., 1971). Barley color is measured spectrophotometrically (American Society of Brewing Chemists, 1992).

4. *Barley protein*

The amount of protein in the barley grain directly affects the enzyme composition and chemical makeup of the malt. Too much protein limits malt extract and makes malt modification difficult, but too little protein leaves insufficient enzyme levels for brewing. With high protein levels, steeping takes longer and germination can be erratic, which can

lead to significant malt losses. High protein levels also contribute to darker wort and beer colors, as well as hazier beers (Burger and LaBerge, 1985). Environment can have an impact on protein levels; hot and dry environment result in higher grain protein content than cool and wet environments (Wrigley et al., 2016). Barley protein is measured by either the Kjeldahl method or NIR (near-infrared) spectrometry (American Society of Brewing Chemists, 1992; Burger and LaBerge, 1985). It is typically reported as a percentage. There are varying recommendations for protein content but they are generally between 9% and 12% (American Malting Barley Association, 2019; Fang et al., 2019; Mather et al., 1997).

5. *Malt extract*

Many consider malt extract to be the most important malt quality parameter. Malt extract is a quantitative measurement of sugar extracted from the grain (Fang et al., 2019). Malt extract determines the amount of alcohol that can be made from the grain based on the amount of fermentable sugar available (Paynter and Young, 1996). Higher levels (>80% DBFG) of malt extract are preferred. Malt extract is typically reported as a percentage (American Society of Brewing Chemists, 1992). The percentage units are either in % DBFG or % DBCG, which stands for dry basis, and either fine or course grind.

6. *Wort color*

Wort is an intermediate product of beer created during the mashing process (Willaert, 2007). Wort color increases during the boiling process as an effect of the Maillard reactions. The Maillard reaction, a non-enzymatic browning reaction, occurs between

amino acids and sugars and gives rise to “toasted” flavors (Ellis, 1959; Tamanna and Mahmood, 2015). The longer wort boils, the darker the color, and the more toasted the taste. Wort color is measured spectrophotometrically (American Society of Brewing Chemists, 1992).

7. *Wort protein*

Like barley protein, wort protein requires a fine balance. Too much protein in the wort will result in opaque and viscous beer, while too little protein will result in a thin beer with little body. These proteins also effect the amount of foam, which is another trait valued by most brewers. Wort protein can also be measured by either the Kjeldahl method or by spectrophotometry (American Society of Brewing Chemists, 1992; Burger and LaBerge, 1985). It is typically reported as a percentage, and AMBA recommends 4.8-5.6% wort protein (American Malting Barley Association, 2019).

8. *Soluble/Total protein (S/T)*

The ratio of soluble to total protein, which can also be referred to as the kolbach index, gives insight into the amount of protein degradation that took place during germination. Protein degradation effects yeast growth, and low ratios of protein degradation correlate to lower enzymatic activity, lower malt extract, and more difficult filtration. Yeast also ages too quickly with a low S/T, leading to a thin tasting beer. This parameter is the best indicator of how long germination was allowed to proceed (degree of malt modification). More soluble proteins mean a higher S/T (Fang et al., 2019). The Kjeldahl method is used to measure the amount of each protein and then it is presented as a percentage of the soluble / total protein content (Burger and LaBerge, 1985). Consensus

target ranges from 35-48% according to Brewers Association and Fang et al. (Brewers Association, 2015; Fang et al., 2019).

9. Diastatic power (DP)

Diastatic power is a measurement of the total activity of malt enzymes (α-amylase, β-amylase, limit dextrinase, and α-glucosidase) (Fang et al., 2019). Diastatic enzymes convert starch into soluble sugars (malt extract). High diastatic power means a higher potential for more malt extract (Paynter and Young, 1996). Although, higher diastatic power can be problematic for batch consistency, specifically during the mash phase if not controlled carefully (Brewers Association, 2015). Diastatic power can be calculated by either Fehlings or Ferricyanide titration methods (American Society of Brewing Chemists, 1992). Consensus target for diastatic power is <150 Lintner according to Brewers Association (Brewers Association, 2015).

10. *Alpha-amylase (AA)*

α-amylase is arguably the most important enzyme in the brewing process. It converts starch (a large, insoluble molecule) into soluble sugars that are utilized by yeast in fermentation. If there is an inadequate amount of starch converted, yeast will not be able to properly ferment. α-amylase is measured as a unit of dextrinization time, or the time it takes to breakdown the sugars (The Institute of Brewing, 1969). AMBA recommends >50 DU (dextrinizing units) (American Malting Barley Association, 2019).

11. *Beta-glucan (BG)*

β-glucan is a polysaccharide. It accounts for 75% of the endosperm cell wall in barley. During the malting process, the barley cell wall gets broken down, releasing β-

glucan. Incomplete breakdown of the cell wall during malting, leads to an excess of β-glucan. This decreases the amount of malt extract and lowers the quality of the beer. Excess β-glucan can also increase the viscosity of wort and beer, which makes filtration more difficult (Fang et al., 2019). β-glucan can be measured fluorometrically with dyes such as Calcofluor, or by NIR (near-infrared) spectrometry (American Society of Brewing Chemists, 1992). Consensus target is <140 ppm according to the Brewers Association (Brewers Association, 2015).

12. *Free amino nitrogen (FAN)*

Free amino nitrogen is the exclusive nitrogen source for yeast, making it crucial to ensure desired yeast growth, synthesis, and metabolic changes. It also contributes directly to beer flavor, or off-flavors. Too much or too little FAN can be a problem so the aim is generally 180-220 mg/L (Fang et al., 2019). FAN can be measured with colorimetry and the ninhydrin method (American Society of Brewing Chemists, 1992).

Flavor as an attribute of interest

Off-flavors of malt and beer have been well studied and chemically annotated, but desired flavors are mostly understood at the anecdotal level, since malt is approved and chosen based on the lack of off-flavors, as opposed to the presence of desired flavors. Over 1,000 flavor compounds of malt and beer have been chemically identified by mass spectrometry and, although many of these compounds may be under our taste detection threshold, they still contribute to the overall flavor profile (Kamimura and Kaneda, 1992). Whether desired or undesired, these small compounds (also referred to as metabolites), are the basis of beer flavor and are a significant factor in determining overall product quality.

Recently, some genetic work has been done on the contributions of barley genotype on beer flavor. Using a small sample set, Herb et al. (2017a and 2017b) showed that barley genotype impacts beer flavor, regardless of degree of malt modification or barley growing environment (Herb et al., 2017a; Herb et al., 2017). Bettenhausen et al. (2020) expanded on the flavor assessment with metabolomics data to define the chemical basis for differences in beer flavor from genetically distinct malt sources (Bettenhausen et al., 2020). Sensory assessments of hot steep (malt sensory evaluation) and beer on a limited number of barley varieties have also been performed and showed that the genotype leads to differences in flavor profiles (Windes et al., 2020). Flavor is one of the most important attributes when it comes to consumer preference of beer, and with the rise in popularity of craft brewing, the interest in understanding flavor is becoming increasingly important.

Cowpea

Cowpea (*Vigna unguiculata* L. Walp) is a warm-season legume native to Africa that is well adapted to sandy soils and low input farming practices which are common throughout sub-Saharan Africa (Faris, 1965; Ji et al., 2019; Rawal, 1975; Vaillancourt and Weeden, 1992). Cowpea is a major food security crop for a large part of the world (da Silvia et al., 2018). Approximately 5.8 million tons (dry weight) are produced each year, mostly in Africa (da Silvia et al., 2018; Dugje et al., 2009). Smallholder farmers are responsible for the majority (around 95%) of cowpea production in rural parts of the continent (Baoua et al., 2021).

Cowpea for food security and health in Africa

Malnutrition, or an imbalance of various nutrients consumed, continues to be a serious problem in many low-income countries. Globally, one in three children under the age of five is malnourished and 340 million children under the age of five suffer from micronutrient deficiencies (UNICEF, 2019). All forms of malnutrition are linked to an increased risk of illness and death (UNICEF, 2019). Africa (broken down into West and Central, Eastern and Southern) accounts for the some of the highest percentages (39.4% and 42.1%, respectively) of children suffering from various forms of malnutrition, such as stunting (low height for age) and wasting (low weight for height) (UNICEF, 2019; USAID, 2018). Food insecurity, combined with early motherhood makes it especially hard to break the cycle of malnutrition (Fink et al., 2014; USAID, 2018). In Ghana specifically, two in three children (age six months - two years) are not fed the foods they need for healthy development (USAID, 2018). Although nutritious food, including legumes such as cowpea, are grown throughout Africa, they only make up a minor portion of a child's diet because the crop is also used for animal feed or sold for income (Dugje et al., 2009; Gómez, 2004; UNICEF, 2019). The majority of diets in rural Africa consist mostly of starches from cassava, yam, plantain, millet, sorghum, and maize but the addition of even small amounts of cowpea could have tremendous impacts in balancing and enhancing nutrition in these areas (Singh et al., 2002).

Cowpea can help alleviate malnutrition when consumed, due to its abundance of proteins, vitamins, trace minerals, antioxidants, amino acids, fibers, lipids, and phytochemicals which also contribute to an assortment of health benefits. Consuming cowpea also provides preventative health benefits. It can lower cholesterol and blood pressure, reduce inflammation, and even help prevent diseases such as diabetes and cancer (Abizari et al., 2013; Awika and Duodu, 2017;

Ayogu et al., 2016; da Silvia et al., 2018; Dakora and Belane, 2019; Jayathilake et al., 2018). Despite established benefits of consuming cowpea, the crop is often sold or used as animal feed instead. Livestock production makes up a large part of farmers income in sub-Saharan Africa and in some areas, premiums are given for certain cowpea varieties so there is a strong incentive to sell instead of consume them (Dugje et al., 2009; Langyintuo et al., 2003).

Importance of biomarkers to nutritional studies

Nutritional studies involving other legumes exist, but because cowpeas are so prevalent and well-adapted to areas where malnutrition is also prevalent, more detailed characterization of their metabolic profile and candidate nutritional biomarkers are needed. Though there are efforts to promote legume and cowpea consumption, there is no reliable way to measure compliance in cowpea feeding studies yet. Nutritional biomarkers are useful for indication and validation of consumption, as well as quantitatively measuring consumption (Potischman and Freudenheim, 2003). Nutritional biomarkers have been identified for many other legume species such as soybeans, green peas, chickpeas, lentils, pinto beans, navy beans, kidney beans, lima beans, and black beans (Borresen et al., 2017; Hofinger et al., 1975; Lu et al., 2010; Madrid-Gambin et al., 2017; Perera et al., 2015; Sri Harsha et al., 2018; Tsopmo and Muir, 2010; Zarei et al., 2021).

Methodology

The studies involving barley and cowpea had different aims, and therefore different approaches and methodology were used. Below, the concepts of QTL mapping (used for studying the genetics of barley contribution to beer flavor) and mass spectrometry (used both to investigate the metabolic profiling of cowpea flours and barley beers) are described.

QTL mapping

A quantitative trait locus (QTL) is a region of the genome that is responsible for the variation of a particular trait. Development of mapping populations is needed to genetically map QTLs, and bi-parental populations are among the most common types. Bi-parental populations are generally derived from parents that are highly homozygous and that differ in the phenotypes of the particular trait(s) of interest (Collard et al., 2005). Commonly used bi-parental populations include recombinant-inbred lines (RILs), and doubled haploids (DHs). A doubled haploid (DH) is a plant derived from a haploid cell (e.g. immature pollen grain) that underwent chromosomal doubling. Chromosomal doubling can either happen spontaneously or be induced with treatments such as colchicine (Ohnoutkova et al., 2019). By developing DHs from the F_1 plants of a cross, a population of completely homozygous lines can be developed in just one generation instead of the 6-8 generations of self-pollination needed to develop RIL populations. The use of DH populations has become fundamental for genetic analyses such as linkage mapping and quantitative trait loci (QTL) mapping in many crops (Cistué et al., 2003). Two main protocols exist in barley for creating doubled haploids; anther culture, and hybridization with *Hordeum bulbosum* (Cistué et al., 2003; Maluszynski et al., 2003).

Once a relatively large bi-parental population is generated, the lines and parents are usually genotyped to identify single nucleotide polymorphisms (SNPs), the most common type of genetic marker. SNPs are base pair substitutions that are frequent in a population. Because of their high frequency and wide distribution in the genome, SNPs help locate the QTL region or gene(s) that encode traits of interest (Collard et al., 2005). SNPs can be identified through DNA sequencing of a set of individuals and used to develop high-density SNP arrays (Bayer et al., 2017). Several SNP chips exists for barley (Close et al., 2009; Comadran et al., 2012; Szűcs et

al., 2009), including the barley 50k Illumina Infinium iSelect array that makes high-throughput genotyping more cost effective than before (Bayer et al., 2017). Due to the availability of a reference genome sequence for barley, the physical positions for most of those SNPs are known, and a direct link with barley genes can be established (Mascher et al., 2017).

Prior to QTL mapping, a genetic map of the bi-parental population is generated, which positions those SNPs in linkage groups based on recombination frequencies (Collard et al., 2005). Recombination frequencies allow for inference of genetic distances between the markers, although they are not equal to the physical positions. The recombination frequencies are converted to centiMorgans (cM), and supplementation with physical positions allows for chromosome assignment and orientation of the linkage groups (Collard et al., 2005). Linkage groups tend to correlate to chromosomes but sometimes more than one linkage group will belong to the same chromosome if marker density is insufficient (Y. Wu et al., 2008).

The linkage map and SNP information of the DHs is used in QTL analysis to identify the genomic regions associated with the measured phenotypes (Collard et al., 2005). QTL analysis works by detecting differences in means of the phenotype(s) of interest and correlating them to specific marker loci (Collard et al., 2005; Hackett, 2002). Since barley has a high-quality reference genome with SNP and gene annotations, specific genes within marker regions can be identified as candidates that control or contribute to the phenotype (Hackett, 2002; Monat et al., 2019). QTL mapping is a widely used method that has helped improve our understanding of the genetic control of important agronomic, disease resistance, and quality traits (Fang et al., 2019; Pauli et al., 2014; Singh et al., 2019).

Mass spectrometry

Mass spectrometry is arguably one of the most versatile analytical techniques that provides precise chemical information on a sample. It has broad uses in many fields including physics, chemistry, medicine, pharmacology, geology, nuclear science, forensic science, and many more (Dass, 2007). Small compounds (<1,500 Da) called metabolites are detected based on their molecular mass and charge (calculated as an m/z ratio) after ionization (Gross, 2006). Mass spectrometry also provides incredibly detailed quantitative and structural information on the metabolites in each sample (Dass, 2007). Mass spectrometry is one of the most common platforms used in metabolomics, and many types exist based on the analysis method and sample type (volatile or non-volatile; polar or non-polar; as well as solid, liquid, or gas) (Dass, 2007; Putri et al., 2013). Mass spectrometers have three main components, this includes an ionizer which applies a charge (positive or negative) to the system, a mass analyzer which reports a detected mass, and a mass detector which calculates the quantity of each metabolite (Gross, 2006). The mass spectrometer reports data as a mass spectrum. This is a two-dimensional representation of mass-to-charge ratio (m/z) versus peak intensity, which reflects relative abundance (Gross, 2006).

The general metabolomics workflow involves extraction of the metabolites, separation, detection, annotation, and quantitation. Extraction is typically done by liquid solvent such as methanol, acetonitrile, ethanol, chloroform, or perchloric acid (H. Wu et al., 2008). Although, some samples such as beverages or biofluids do not require an extraction solvent. Metabolite separation occurs by chromatography, which yields a chromatogram (retention time versus metabolite quantity) (Lu et al., 2017). The two most common forms of chromatography include gas chromatography (GC), which separates molecules based on volatility, and liquid

chromatography (LC), which separates molecules based on polarity (Lu et al., 2017). Metabolite detection occurs by mass spectrometry (MS), as described above (Dass, 2007; Gross, 2006). Various systems exist for mass spectrometry. The two main ones are quadrupole (Q) and time of flight (TOF) mass spectrometry. These can also be coupled together in tandem mass spectrometry. Quadrupole systems measure mass from resonance through a magnetic field, and time of flight systems measure mass by time it takes molecules to travel between two places (Dass, 2007). Metabolites are annotated by reference to a database using computer software based on metabolite retention time and mass spectra. Without the inclusion of internal standards, annotation confidence will vary based on the information gathered from mass spectra and chromatogram. Internal standards also allow for absolute quantitation, whereas with no internal standards, data can be reported in relative quantitation after data normalization occurs (Lu et al., 2017). The use of internal standards is definitive of a targeted metabolomics approach. Non-targeted metabolomics is a more global approach where there is no a priori knowledge of what metabolites are of interest in the sample (Zhang et al., 2016). Metabolomics and the use of mass spectrometry have become indispensable tools in many fields of science for a variety of applications.

Research goals and hypotheses

Two crop systems were used in this work. The overall goals and hypotheses of this thesis are broken down by crop and described below.

Barley

Flavor is an attribute of increasing interest in the malting and brewing industry. Chapter 2 of this thesis, titled Genetic Basis of Barley Contributions to Beer Flavor, aims to 1) test the hypothesis that barley genotype contributes to beer flavor, 2) identify regions of the genome that control traits associated with flavor, and 3) identify candidate genes that control traits associated with flavor.

Cowpea

Cowpea is a highly valued food for its nutritional content and contribution to food security in many developing parts of the world. The need of biomarkers for nutritional studies is undisputed, but currently there is a lack of work done specific to cowpea. Chapter 3 of this thesis, titled Non-targeted Metabolomics of Cooked Cowpeas (*Vigna unguiculata*) and Pigeon Pea (*Cajanus cajan*) from Ghana, aims to 1) characterize the metabolic profiles of three cowpea flours, and 2) test the hypothesis that there are metabolites unique to cowpea (and cowpea variety) that can be potential biomarkers for nutritional studies.

REFERENCES

Abizari, A.R., Pilime, N., Armar-Klemesu, M., Brouwer, I.D., 2013. Cowpeas in Northern Ghana and the Factors that Predict Caregivers' Intention to Give Them to Schoolchildren. PLoS One 8, 1–8.

American Malting Barley Association, 2019. Two-row Variety Development Update.

American Society of Brewing Chemists, 1992. Methods of Analysis. In: ASBC. St. Paul, MN.

Anderson, M.K., Reinsbergs, E., 1985. Barley Breeding. In: Barley. pp. 231–268.

Ash, M., Hoffman, L., 1990. Barley: Background for 1990 Farm Legislation.

Awika, J.M., Duodu, K.G., 2017. Bioactive polyphenols and peptides in cowpea (Vigna unguiculata) and their health promoting properties: A review. J. Funct. Foods 38, 686–697.

Ayogu, R.N.B., Nnam, N.M., Mbah, M., 2016. Evaluation of two local cowpea species for nutrient, antinutrient, and phytochemical compositions and organoleptic attributes of their wheat-based cookies. Food Nutr. Res. 60, 1–8.

Badr, A., Müller, K., Schäfer-Pregl, R., El Rabey, H., Effgen, S., Ibrahim, H.H., Pozzi, C., Rohde, W., Salamini, F., 2000. On the origin and domestication history of barley (Hordeum vulgare). Mol. Biol. Evol. 17, 499–510.

Baoua, I., Rabé, M.M., Murdock, L.L., Baributsa, D., 2021. Cowpea production constraints on smallholders' farms in Maradi and Zinder regions, Niger. Crop Prot. 142.

Bayer, M.M., Rapazote-Flores, P., Ganal, M., Hedley, P.E., Macaulay, M., Plieske, J., Ramsay, L., Russell, J., Shaw, P.D., Thomas, W., Waugh, R., 2017. Development and evaluation of a barley 50k iSelect SNP array. Front. Plant Sci. 8, 1–10.

Bettenhausen, H.M., Benson, A., Fisk, S., Herb, D., Hernandez, J., Lim, J., Queisser, S.H., Shellhammer, T.H., Vega, V., Yao, L., Heuberger, A.L., Hayes, P.M., 2020. Variation in Sensory Attributes and Volatile Compounds in Beers Brewed from Genetically Distinct Malts: An Integrated Sensory and Non-Targeted Metabolomics Approach. J. Am. Soc. Brew. Chem. 78, 136–152.

Borresen, E.C., Zhang, L., Trehan, I., Nealon, N.J., Maleta, K.M., Manary, M.J., Ryan, E.P., 2017. The nutrient and metabolite profile of 3 complementary legume foods with potential to improve gut health in rural Malawian children. Curr. Dev. Nutr. 1, 1–11.

Brewers Association, 2015. Malting Barley Characteristics for Craft Brewers Executive Summary 1–6.

Briggs, D.E., 1998. Malts and Malting.

Burger, W.C., LaBerge, D.E., 1985. Malting and Brewing Quality. In: Rasmusson, D.C. (Ed.), Barley, Volume 26.

Cistué, L., Vallés, M.P., Echávarri, B., Sanz, J.M., Castillo, A., 2003. Barley anther culture. Doubled Haploid Prod. Crop Plants 29–34.

Close, T.J., Bhat, P.R., Lonardi, S., Wu, Y., Rostoks, N., Ramsay, L., Druka, A., Stein, N., Svensson, J.T., Wanamaker, S., Bozdag, S., Roose, M.L., Moscou, M.J., Chao, S., Varshney, R.K., Szucs, P., Sato, K., Hayes, P.M., Matthews, D.E., Kleinhofs, A., Muehlbauer, G.J., DeYoung, J., Marshall, D.F., Madishetty, K., Fenton, R.D., Condamine, P., Graner, A., Waugh, R., 2009. Development and implementation of high-throughput SNP genotyping in barley. BMC Genomics 10, 1–13.

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

Colmsee, C., Beier, S., Himmelbach, A., Schmutzer, T., Stein, N., Scholz, U., Mascher, M., 2015. BARLEX - the Barley Draft Genome Explorer. Mol. Pant 8, 964–966.

Comadran, J., Kilian, B., Russell, J., Ramsay, L., Stein, N., Ganal, M., Shaw, P., Bayer, M., Thomas, W., Marshall, D., Hedley, P., Tondelli, A., Pecchioni, N., Francia, E., Korzun, V., Walther, A., Waugh, R., 2012. Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. Nat. Genet. 44, 1388–1391.

da Silvia, A.C., da Costa Santos, D., Teixeira Jr., D.L., da Silvia, P.B., dos Santos, R.C., Siviero, A., 2018. Cowpea: A Strategic Legume Species for Food Security and Health. In: Legume Seed Nutraceutical Research. pp. 47–65.

Dakora, F.D., Belane, A.K., 2019. Evaluation of Protein and Micronutrient Levels in Edible Cowpea (Vigna Unguiculata L. Walp.) Leaves and Seeds. Front. Sustain. Food Syst. 3, 1–10.

Dass, C., 2007. Fundamentals of Contemporary Mass Spectrometry.

Dugje, I.Y., Omoigui, L.O., Ekeleme, F., Kamara, A. Y., Ajeigbe, H., 2009. Farmers' Guide to Cowpea Production in West Africa, IITA, Ibadan, Nigeria. Ibadan, Niger. IITA 20.

Ellis, G.P., 1959. The Maillard Reaction. Adv. Carbohydr. Chem. 14, 63–134.

Evans, D.E., Li, C., Eglinton, J.K., 2010. The Properties and Genetics of Barley Malt Starch Degrading Enzymes. In: Zhang, G., Li, Chengdao (Eds.), Genetics and Improvement of Barley Malt Quality. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 143–189.

Fang, Y., Zhang, X., Xue, D., 2019. Genetic analysis and molecular breeding applications of malting quality QTLs in barley. Front. Genet. 10, 1–11.

Faris, D.G., 1965. The Origin and Evolution of the Cultivated Forms of Vigna sinensis. Can. J. Genet. Cytol. 7, 433–452.

Fink, G., Sudfeld, C.R., Danaei, G., Ezzati, M., Fawzi, W.W., 2014. Scaling-up access to family planning may improve linear growth and child development in low and middle income countries. PLoS One 9, 1–9.

Friedt, W., Horsley, R.D., Harvey, B.L., Poulsen, D.M.E., Lance, R.C.M., Ceccarelli, S., Grando, S., Capettini, F., 2011. Barley Breeding History, Progress, Objectives, and Technology. In: Barley: Production, Improvement, and Uses. Wiley-Blackwell, pp. 160–

220.

Gómez, C., 2004. Introduction: Economic and Social Impact, COWPEA: Post-Harvest Operations.

Gross, J.H., 2006. Mass Spectrometry: A Textbook.

Hackett, C.A., 2002. Statistical methods for QTL mapping in cereals. Plant Mol. Biol. 48, 585–599.

Han, F., Ullrich, S.E., Kleinhofs, A., Jones, B.L., Hayes, P.M., Wesenberg, D.M., 1997. Fine structure mapping of the barley chromosome-1 centromere region containing malting-quality QTLs. Theor. Appl. Genet. 95, 903–910.

Herb, D., Filichkin, T., Fisk, S., Helgerson, L., Hayes, P., Benson, A., Vega, V., Carey, D., Thiel, R., Cistue, L., Jennings, R., Monsour, R., Tynan, S., Vinkemeier, K., Li, Y., Nguygen, A., Onio, A., Meints, B., Moscou, M., Romagosa, I., Thomas, W., 2017a. Malt modification and its effects on the contributions of barley genotype to beer flavor. J. Am. Soc. Brew. Chem. 75, 354–362.

Herb, D., Filichkin, T., Fisk, S., Helgerson, L., Hayes, P., Meints, B., Jennings, R., Monsour, R., Tynan, S., Vinkemeier, K., Romagosa, I., Moscou, M., Carey, D., Thiel, R., Cistue, L., Martens, C., Thomas, W., 2017b. Effects of barley (Hordeum vulgare L.) variety and growing environment on beer flavor. J. Am. Soc. Brew. Chem. 75, 345–353.

Hofinger, M., Monseur, X., Pais, M., Jarreau, F.X., 1975. Further confirmation of the presence of indolylacrylic acid in lentil seedlings and identification of hypaphorine as its precursor. Phytochemistry 14, 475–477.

Igartua, E., Edney, M., Rossnagel, B.G., Spaner, D., Legge, W.G., Scoles, G.J., Eckstein, P.E., Penner, G.A., Tinker, N.A., Briggs, K.G., Falk, D.E., Mather, D.E., 2000. Marker-based selection of QTL affecting grain and malt quality in two-row barley. Crop Sci. 40, 1426–1433.

Jayathilake, C., Visvanathan, R., Deen, A., Bangamuwage, R., Jayawardana, B.C., Nammi, S., Liyanage, R., 2018. Cowpea: an overview on its nutritional facts and health benefits. J. Sci. Food Agric. 98, 4793–4806.

Ji, J., Zhang, C., Sun, Z., Wang, L., Duanmu, D., Fan, Q., 2019. Genome Editing in Cowpea Vigna unguiculata Using CRISPR-Cas9. Int. J. Mol. Sci. 20, 1–13.

Kamimura, M., Kaneda, H., 1992. Off-flavors in beer, Developments in Food Science. Elsevier B.V.

Kolar, S.C., Hayes, P.M., Chen, T.H.H., Linderman, R.G., 1991. Genotypic variation for cold tolerance in winter and facultative barley. Crop Sci. 31, 1149–1152.

Kumlehn, J., Stein, N., 2014. Biotechnological Aproaches to Barley, Biotechnology in Agriculture and Forestry.

Langyintuo, A.S., Lowenberg-DeBoer, J., Faye, M., Lambert, D., Ibro, G., Moussa, B., Kergna, A., Kushwaha, S., Musa, S., Ntoukam, G., 2003. Cowpea supply and demand in West and

Central Africa. F. Crop. Res. 82, 215–231.

Lipnizki, F., 2015a. Beer Maturation. In: Drioli, E., Giorno, L. (Eds.), Encyclopedia of Membranes. Springer Berlin Heidelberg, Berlin, Heidelberg, p. 1.

Lipnizki, F., 2015b. Beer Clarification. In: Drioli, E., Giorno, L. (Eds.), Encyclopedia of Membranes. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–2.

Lu, C.T., Tang, H.F., Sun, X.L., Wen, A.D., Zhang, W., Ma, N., 2010. Indole alkaloids from chickpea seeds (Cicer arietinum L.). Biochem. Syst. Ecol. 38, 441–443.

Lu, W., Su, X., Klein, M.S., Lewis, I.A., Fiehn, O., Rabinowitz, J.D., 2017. Metabolite Measurement: Pitfalls to Avoid and Practices to Follow. Physiol. Behav. 176, 139–148.

Madrid-Gambin, F., Llorach, R., Vazquez-Fresno, R., Urpi-Sarda, M., Almanza-Aguilera, E., Garcia-Aloy, M., Estruch, R., Corella, D., Andres-Lacueva, C., 2017. Urinary 1H Nuclear Magnetic Resonance Metabolomic Fingerprinting Reveals Biomarkers of Pulse Consumption Related to Energy-Metabolism Modulation in a Subcohort from the PREDIMED study. J. Proteome Res. 16, 1493–1491.

Magliano, P.N., Prystupa, P., Herman Gutierrez Boem, F., 2014. Protein content of grains of different size fractions in malting barley. J. Inst. Brew. 120.

Mallet, J., 2014. Malt: A Practical Guide from Field to Brewhouse. Brewers Publications, Boulder, Colorado.

Maluszynski, M., Kasha, K.J., Forster, B.P., Szarejko, I., 2003. Doubled Haploid Production in Crop Plants.

Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S.O., Wicker, T., Radchuk, V., Dockter, C., Hedley, P.E., Russell, J., Bayer, M., Ramsay, L., Liu, H., Haberer, G., Zhang, X.Q., Zhang, Q., Barrero, R.A., Li, L., Taudien, S., Groth, M., Felder, M., Hastie, A., Šimková, H., Stanková, H., Vrána, J., Chan, S., Munõz-Amatriaín, M., Ounit, R., Wanamaker, S., Bolser, D., Colmsee, C., Schmutzer, T., Aliyeva-Schnorr, L., Grasso, S., Tanskanen, J., Chailyan, A., Sampath, D., Heavens, D., Clissold, L., Cao, S., Chapman, B., Dai, F., Han, Y., Li, H., Li, X., Lin, C., McCooke, J.K., Tan, C., Wang, P., Wang, S., Yin, S., Zhou, G., Poland, J.A., Bellgard, M.I., Borisjuk, L., Houben, A., Doleael, J., Ayling, S., Lonardi, S., Kersey, P., Langridge, P., Muehlbauer, G.J., Clark, M.D., Caccamo, M., Schulman, A.H., Mayer, K.F.X., Platzer, M., Close, T.J., Scholz, U., Hansson, M., Zhang, G., Braumann, I., Spannagl, M., Li, C., Waugh, R., Stein, N., 2017. A chromosome conformation capture ordered sequence of the barley genome. Nature 544, 427–433.

Masschelein, C.A., 1986. The Biochemistry of Maturation. J. Inst. Brew. 92, 213–219.

Mather, D.E., Tinker, N.A., LaBerge, D.E., Edney, M., Jones, B.L., Rossnagel, B.G., Legge, W.G., Briggs, K.G., Irvine, R.B., Falk, D.E., Kasha, K.J., 1997. Regions of the genome that affect grain and malt quality in a North American two-row Barley Cross. Crop Sci. 37, 544–554.

Mohammadi, M., Blake, T.K., Budde, A.D., Chao, S., Hayes, P.M., Horsley, R.D., Obert, D.E., Ullrich, S.E., Smith, K.P., 2015. A genome-wide association study of malting quality across eight U.S. barley breeding programs. Theor. Appl. Genet. 128, 705–721.

Monat, C., Padmarasu, S., Lux, T., Wicker, T., Gundlach, H., Himmelbach, A., Ens, J., Li, C., Muehlbauer, G.J., Schulman, A.H., Waugh, R., Braumann, I., Pozniak, C., Scholz, U., Mayer, K.F.X., Spannagl, M., Stein, N., Mascher, M., 2019. TRITEX: Chromosome-scale sequence assembly of Triticeae genomes with open-source tools. bioRxiv 1–18.

Muñoz-Amatriaín, M., Xiong, Y., Schmitt, M.R., Bilgic, H., Budde, A.D., Chao, S., Smith, K.P., Muehlbauer, G.J., 2010. Transcriptome analysis of a barley breeding program examines gene expression diversity and reveals target genes for malting quality improvement. BMC Genomics 11.

Ohnoutkova, L., Vlcko, T., Ayalew, M., 2019. Barley Anther Culture. In: Harwood, W.A. (Ed.), Barley: Methods and Protocols. Springer New York, New York, NY, pp. 37–52.

Pauli, D., Muehlbauer, G.J., Smith, K.P., Cooper, B., Hole, D., Obert, D.E., Ullrich, S.E., Blake, T.K., 2014. Association Mapping of Agronomic QTLs in U.S. Spring Barley Breeding Germplasm. Plant Genome 7.

Paynter, B., 2015. Grain protein and malting quality in barley.

Paynter, B., Young, K., 1996. Managing grain protein in malting barley.

Perera, T., Young, M.R., Zhang, Z., Murphy, G., Colburn, N.H., Lanza, E., Hartman, T.J., Cross, A.J., Bobe, G., 2015. Identification and monitoring of metabolite markers of dry bean consumption in parallel human and mouse studies. Mol. Nutr. Food Res. 59, 795–806.

Potischman, N., Freudenheim, J.L., 2003. Biomarkers of nutritional exposure and nutritional status: An overview. J. Nutr. 133, 873–874.

Potokina, E., Caspers, M., Prasad, M., Kota, R., Zhang, H., Sreenivasulu, N., Wang, M., Graner, A., 2004. Functional association between malting quality trait components and cDNA array based expression patterns in barley (Hordeum vulgare L.). Mol. Breed. 14, 153–170.

Putri, S.P., Yamamoto, S., Tsugawa, H., Fukusaki, E., 2013. Current metabolomics: Technological advances. J. Biosci. Bioeng. 116, 9–16.

Rawal, K.M., 1975. Natural Hybridization Among Wild, Weedy and Cultivated Vigna unguiculata (L.) Walp. Euphytica 24, 699–707.

Reid, N., Gatrell, J.D., 2015. Brewing Growth, Economic Development Journal.

Sato, K., 2020. History and future perspectives of barley genomics. DNA Res. 27, 1–8.

Shewry, P.R., Ullrich, S.E., 2014. Barley: Chemistry and Technology, Second Edition. AACC International.

Singh, B., Mehta, S., Aggarwal, S.K., Tiwari, M., Bhuyan, S.I., Bhatia, S., Islam, M.A., 2019. Barley, Disease Resistance, and Molecular Breeding Approaches. In: Wani, S.H. (Ed.), Disease Resistance in Crop Plants: Molecular, Genetic and Genomic Perspectives. Springer International Publishing, Cham, pp. 261–299.

Singh, B.B., Ehlers, J.D., Sharma, B., Freire Filho, F.R., 2002. Recent Progress in cowpea breeding. In: Challenges and Opportunities for Enhancing Sustainable Cowpea Production: Proceedings of the World Cowpea Conference III Held at IITA, 4-8 September 2000. pp. 1–

424.

Skarsaune, S.K., Banasik, O.J., Watson, C.A., 1971. The Measurement of Barley Kernel Color. Proceedings. Annu. Meet. - Am. Soc. Brew. Chem. 29, 5–8.

Sri Harsha, P.S.C., Wahab, R.A., Garcia-Aloy, M., Madrid-Gambin, F., Estruel-Amades, S., Watzl, B., Andrés-Lacueva, C., Brennan, L., 2018. Biomarkers of legume intake in human intervention and observational studies: A systematic review. Genes Nutr. 13, 1–16.

Szűcs, P., Blake, V.C., Bhat, P.R., Chao, S., Close, T.J., Cuesta-Marcos, A., Muehlbauer, G.J., Ramsay, L., Waugh, R., Hayes, P.M., 2009. An Integrated Resource for Barley Linkage Map and Malting Quality QTL Alignment. Plant Genome 2.

Tamanna, N., Mahmood, N., 2015. Food processing and maillard reaction products: Effect on human health and nutrition. Int. J. Food Sci. 2015.

The Institute of Brewing, 1969. Method for Determination of Alpha-amylase Activity. J. Inst. Brew. 75, 430–432.

Tremblay, V.J., Tremblay, C.H., 2005. The U.S. Brewing Industry: Data and Economic Analysis. MIT Press.

Tricase, C., Amicarelli, V., Lamonaca, E., Rana, R.L., 2018. Economic Analysis of the Barley Market and Related Uses. In: Tadele, Z. (Ed.), Grasses as Food and Feed. IntechOpen.

Tsopmo, A., Muir, A.D., 2010. Chemical profiling of lentil (Lens culinaris Medik.) cultivars and isolation of compounds. J. Agric. Food Chem. 58, 8715–8721.

UNICEF, 2019. Children, Food and Nutrition 24.

USAID, 2018. Ghana: Nutrition Profile. USAID 1–4.

Vaillancourt, R.E., Weeden, N.F., 1992. Chloroplast DNA Polymorphism Suggests Nigerian Center of Domestication for the Cowpea, Vigna Unguiculata (Leguminosae). Am. J. Bot. 79, 1194–1199.

von Bothmer, R., van Hintum, T., Knupffer, K., Sato, K., 2003. Diversity in Barley. Elsevier.

Willaert, R., 2007. The Beer Brewing Process: Wort Production and Beer Fermentation. In: Handbook of Food Products Manufacturing. pp. 443–490.

Windes, S., Bettenhausen, H.M., Simaeys, K.R.V., Clawson, J., Fisk, S., Heuberger, A.L., Lim, J., Queisser, S.H., Shellhammer, T.H., Hayes, P.M., 2020. Comprehensive Analysis of Different Contemporary Barley Genotypes Enhances and Expands the Scope of Barley Contributions to Beer Flavor. J. Am. Soc. Brew. Chem. 0, 1–25.

Withers, E.T., 2017. The Impact and Implications of Craft Beer Research: An Interdisciplinary Literature Review. In: Kline, C., Slocum, S.L., Cavaliere, C.T. (Eds.), Craft Beverages and Tourism, Volume 1. pp. 11–24.

Wrigley, C., Batey, I.L., Miskelly, D., 2016. Cereal Grains: Assessing and Managing Quality.

Wu, H., Southam, A.D., Hines, A., Viant, M.R., 2008. High-throughput tissue extraction protocol for NMR- and MS-based metabolomics. Anal. Biochem. 372, 204–212.

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

Zarei, I., Baxter, B.A., Oppel, R.C., Borresen, E.C., Brown, R.J., Ryan, E.P., 2021. Plasma and urine metabolite profiles impacted by increased dietary navy bean intake in colorectal cancer survivors: A randomized-controlled trial. Cancer Prev. Res. 14, 497–508.

Zhang, X., Zhu, X., Wang, C., Zhang, H., Cai, Z., 2016. Non-targeted and targeted metabolomics approaches to diagnosing lung cancer and predicting patient prognosis. Oncotarget 7, 63437–63448.

CHAPTER 2 – GENETIC BASIS OF BARLEY CONTRIBUTIONS TO BEER FLAVOR[1]

Overview

Barley malt is critical for the malting, brewing, and distilling industries, as it is one of the main ingredients of beer and some types of spirits. There is growing evidence that barley genotype - via malt - can impact the flavors of beers and spirits. However, information on the barley genes involved in these flavors is lacking. Therefore, we used quantitative trait locus (QTL) mapping of malt quality traits, beer sensory descriptors and metabolic compounds on a biparental population of doubled haploids derived from the cross of the cultivars Golden Promise and Full Pint. Candidate genes for QTLs were identified by alignment with the reference barley genome sequence. There were thirty-seven QTLs across all chromosomes except 4H, with three QTL clusters located on 3H (1 cluster) and 5H (2 clusters: mid-5H and end-5H). Those "hotspots" contained QTLs for multiple phenotypes. Several candidate genes that regulate plant metabolism were identified within the QTLs and included *HvAlaAT*, *HvDep1*, *HvMKK3*, *HvGA20ox1* and *HvGA20ox2*. These genes are involved in seed dormancy and plant height. Alleles at these loci, and perhaps at physically linked loci, can have key downstream effects on malting quality, beer flavor, and abundance of volatile metabolites.

---

[1] This chapter is a modified form of a manuscript submitted to the Journal of Cereal Science:

Brooke Sayre-Chavez, Harmonie Bettenhausen, Sarah Windes, Patricia Aron, Luis Cistué, Scott Fisk, Laura Helgerson, Adam L. Heuberger, Sean Tynan, Patrick Hayes4, and María Muñoz-Amatriaín

Introduction

Barley (*Hordeum vulgare L.*) is the main cereal grain used in the malting and brewing industries, as well as in distilling. Barley malt provides critical starches and enzymes to the brewing process, which in turn provides the necessary sugar and nutrients to yeast for fermentation in order create the end-products, which are typically beer (Paynter and Young, 1996) and/or spirits. Modifications to any step of the malting process can alter the overall malt flavor, with the largest driver being kilning, resulting in the wide range of base and specialty malts used to make different styles of beer. A growing body of evidence – based on base malts and pale lager/ale style malt-forward beers brewed from them – indicates that barley genotype can make significant contributions to beer flavor (Bettenhausen et al., 2020, 2018; Craine et al., 2021; Herb et al., 2017a, 2017a; Kyraleou et al., 2021; Morrissy et al., 2021; Windes et al., 2020). The environment can modulate the effects of these genes – a source of barley variety "terroir". Kyraleou et al., (2021) for example, reported differences in flavor of spirits attributed to where the barley was grown.

The assessment of contributions of genotype and environment to flavor is an area of recent research; historically, the suitability of barley varieties for brewing has been based on a suite of malt quality parameters. Organizations around the world set the acceptable standards and approve new malting varieties. Notable examples include the American Malting Barley Association (AMBA), the Brewing and Malting Barley Research Institute (BMBRI), Barley Australia, the Canadian Malting Barley Technical Centre (CMBTC), and the European Brewery Convention (EBC). These organizations have strict guidelines and lengthy testing requirements for barley varieties to be approved and recommended for malting and brewing. AMBA, for example, uses pilot malting evaluations followed by plant (commercial) scale malting and

brewing trials. A barley variety is recommended only if ratings are satisfactory at all stages. However, a barley may have satisfactory malt specifications but fail at later stage, commercial scale brewing trials due to negative flavors. Therefore, to date, selection for barley flavor has been a defect elimination process, rather than a process designed to identify and promote positive flavors. Negative attributes may be due to flavor components of barley grain, flavors developed during malting that are not associated with current malting quality attributes, and/or flavors developed in brewing as a result of interactions with hops, yeast and other beer ingredients. The defect elimination strategy does provide evidence for flavor contributions from the barley genotype and/or the production environment. If these factors can contribute negative flavor attributes, could they potentially contribute positive flavor attributes? Ultimately, all causes of differences in beer flavor - be they positive or negative - between barley genotypes, and production environments, will have a genetic basis. Identifying these genes will require systematic assessment of all possible phenotypes that could contribute to beer flavor. A starting point, for barley, is the suite of malting quality parameters, such as kernel weight and plumpness, malt extract, wort color and protein, barley protein, and various enzyme parameters.

Barley malting quality traits, as well as their regulatory genetics, have been the subjects of intense study (Fang et al., 2019; Han et al., 1997; Igartua et al., 2000; Mather et al., 1997; Mohammadi et al., 2015; Muñoz-Amatriaín et al., 2010; Potokina et al., 2004; Szűcs et al., 2009). Malt quality per se is a complex meta-phenotype, and target values for critical component attributes will vary depending on the intended use of the malt – all-malt brewing, adjunct brewing, or distilling. In general, well-modified malts with high levels of malt extract are desirable for all end uses. Modification is a term widely used in malting and brewing to describe a malt with the optimum balance of starch and protein-related factors for the intended end-

product. Enzyme-related parameters – such as α-amylase activity and diastatic power - and free amino nitrogen (FAN) will vary between end uses. Lower enzyme levels are desired for all-malt brewing; higher enzyme levels are required for adjunct brewing; and even higher levels are targeted for grain distillers' malts. Genes encoding key enzymes have been cloned and the bases of allelic variation described (reviewed in Shewry and Darlington, 2002). These include loci encoding α-amylases, β-amylases, β-glucanases, β-glucosidases, and limit dextrinases (Bamforth, 2009; Evans and Eglinton, 2009; Knox et al., 1987; Muthukrishnan et al., 1984; Erkkila et al., 1998; Clark et al., 2003; Kreis et al., 2009; Han et al., 1995; Litts et al., 1990; Tibbot et al., 1996; Burton et al., 1999). These enzymes are important in the breakdown of starches which are utilized during fermentation and brewing. However, known-function and candidate genes are lacking for many other important malting quality traits – and these are typically only reported as quantitative trait loci (Mohammadi et al., 2015).

The degree of modification, and resulting malting quality profile, is driven by the grain's ability to germinate, and to germinate uniformly. Germination characteristics - which can be further broken down as germinative energy, capacity, and water sensitivity - are key gateway characteristics that are assessed on a sample of prospective malting barley prior to malting (Briggs, 1978). Seed dormancy lies at one extreme of the germination continuum. Pre-harvest sprouting (PHS) susceptibility lies at the other end of this continuum. Between these extremes lie the degrees of dormancy, which may be key drivers of traits affecting malting. The hormones abscisic acid, gibberellin, ethylene, and auxin play key roles in seed dormancy (Corbineau et al., 2014; Ishibashi et al., 2017; Li et al., 2004; Liu et al., 2013). Briefly, abscisic acid induces and maintains dormancy (Corbineau et al., 2014; Ishibashi et al., 2017); gibberellins coordinate the release from dormancy so the seed can germinate (Corbineau et al., 2014; Ishibashi et al., 2017;

Li et al., 2004); ethylene inhibits abscisic acid while also stimulating the biosynthesis of gibberellin (Corbineau et al., 2014); and auxin plays a role in seed dormancy by stimulating abscisic acid action, therefore promoting dormancy (Liu et al., 2013).

Dormancy, and the degree dormancy, have been areas of intensive and extensive research in barley, leading the identification of qualitative and quantitative genetic determinants (Edney and Mather, 2004; Hori et al., 2007; Li et al., 2004; Nakamura et al., 2017, 2016; Prada et al., 2004). Two of the most important QTLs are *SD1* and *SD2*, located, respectively, in the centromeric region (mid) and long arm (end) of chromosome 5H (Nakamura et al., 2017). *Alanine aminotransferase* (*AlaAT*) has been identified as the causal gene for *SD1* (Sato et al., 2016). *Mitogen Activated Protein Kinase Kinase 3* (*MKK3*) was the first gene reported to be responsible for *SD2* (Nakamura et al., 2016), and this gene has been validated in subsequent reports (Mao et al., 2019; Shorinola et al., 2017; Vetch et al., 2020, 2019). Nagel et al. (2019) reported that a gibberellin oxidase gene involved in dormancy alleviation (*HvGA20ox1*) maps to the *SD2* region as well. *MKK3* and *HvGA20ox1* are located ~1,600 kb apart based on the reference genome Morex V2 (Monat et al., 2019). The distal end of the long arm of chromosome 5H is one of the key "hotspots" for barley malting quality QTLs (Fang et al., 2019; Igartua et al., 2000; Mather et al., 1997; Mohammadi et al., 2015). Therefore, it is tempting to speculate that the candidate gene(s) for *SD2* are involved in dormancy, degree of dormancy, and PHS.

Whereas tremendous progress has been made in elucidating the genetic basis of both agronomic and malting quality traits, information on the genes determining the contributions of barley to beer flavor is lacking. There are strong anecdotal opinions in the malting and brewing community – both for and against the contributions of barley to beer flavor: a key piece of evidence in flavor is the persistence of older varieties in the market due to their perceived unique

contributions to beer flavor. These varieties range from heirlooms, such as Chevalier and Bere to more recent varieties that do not have competitive agronomic and malting quality profiles, such as Golden Promise, Maris Otter, and Klages (Mallett, 2014). Genetic analysis requires harder evidence than opinions and sales figures. Recent experimental work has established the necessary foundation. Herb et al. (2017b) provided the first clear evidence that there is a genetic basis for the barley contribution to beer flavor. Sensory descriptors were notably different between parental varieties (Golden Promise and Full Pint) and variation for sensory attributes was observed in a sample of 34 doubled haploids derived from the cross (referred to as the Oregon Promise population). The conclusion that there is a genetic basis to flavor was based on estimates of heritability and preliminary estimates of marker:trait relationships. The malts upon which this research was based were generally under-modified – an unavoidable confounding factor when experimental genotypes and varieties of historical interest are micro-malted in batches using protocols designed for assessing contemporary and future malting varieties. Therefore, Herb et al. (2017a) specifically addressed the impact of degree of modification on barley genotype contributions to beer flavor and concluded that even with intentional under-modification and over-modification, there is a genetic contribution of the barley to beer sensory attributes. Bettenhausen et al. (2020) extended this flavor assessment to larger-scale malting and brewing on a subset of three Oregon Promise doubled haploids, confirmed differences in flavor, and identified the top-rated doubled haploid for release as the variety “Oregon Promise” in 2020.

The Bettenhausen et al. (2020) work also included metabolomics, a powerful tool that is used to better understand the chemical composition of a sample. Since the focus is on the sensory attributes of the beer, in this case volatile metabolites are of the most interest. Volatiles are the aromatic compounds contributing to flavor perception. Bettenhausen et al. (2018) first applied

metabolomics to answer questions about the effect of malt source on beer flavor and flavor stability. They demonstrated that malt sources (location grown, maltster) did have an impact on beer flavor/flavor stability and metabolite variation that could account for flavor differences among beers. Subsequently, metabolomics has been applied to the assessment of beers made from Oregon Promise selections (Bettenhausen et al., 2020) as well as beers and hot steeps made from currently available winter malting varieties and experimental spring varieties (Windes et al., 2020). Morrissy et al. (2021) extended the analysis pipeline of pilot malting, brewing, sensory, and metabolomics of hot steeps and beer to assess contributions to beer flavor in doubled haploids derived from crosses of Maris Otter with contemporary varieties.

There is evidence, therefore, that (i) barley genotypes can differ in their contributions to beer flavor, (ii) there is a genetic basis to these contributions, and (iii) differences in sensory attributes and metabolite profiles are not simply due to the degree of modification of malt and/or differences in beer analytics. In this report, we build on these findings by providing an integrated and comprehensive analysis of the genetic basis of malting quality, beer sensory traits, and beer metabolites. Specifically, we expand the scope of inference on the Oregon Promise population (Herb et al., 2017b) via genetically characterizing 236 doubled haploids from the Oregon Promise population, using a high-density genotyping array (Barley 50k iSelect SNP array; Bayer et al., 2017) and integrating - via biparental QTL mapping - the genotype data with phenotype data on malting quality, beer sensory, and beer metabolomics. This integration of new findings with a review of the literature on the topic provides a platform for identifying next steps in this exciting area of research.

Materials and Methods

Plant material and micro-malting

The development of the Oregon Promise population was described in detail by Herb et al. (2017b). Briefly, the entire population consists of 236 doubled haploids (DHs), developed using the anther culture protocol of Cistué et al. (2003). The full mapping population was grown at Corvallis, Oregon USA in 2013 and 2014. Grain samples from the 2013 crop were malted at the USDA Cereal Crops Research Unit (CCRU), Madison Wisconsin, USA, following the procedures described by Mahalingam et al. (2017). Analysis methods are per American Society of Brewing Chemists Methods, except for quality score and overall rank. Quality score is a weighted measure of all quality parameters (C. Martens, personal communication). The higher the value, the more suitable the malt for adjunct brewing. The overall rank is the inverse of quality score, where the top ranked sample (1) has the highest quality score. The malting quality data used for quantitative trait locus (QTL) mapping in this report trace to these samples. Sufficient malt from these samples was not available for nano-brewing (see Materials and Methods). Therefore, grain from the 2014 crop was malted at the CCRU but no malt analyses were conducted in order to have sufficient grain for nano-brewing. One hundred and sixty-two DH lines from the 2014 crop had sufficient grain for malting and subsequent brewing. Samples from the 2014 crop were submitted for malting in August 2015 – approximately one year post-harvest. For the parents, there was not sufficient residual grain from the Corvallis 2013 or 2014 crops for micro-malting and subsequent nano-brewing. Therefore, residual malt from a 2016 experiment grown at Lebanon, Oregon USA (described by Herb et al., 2017b) was used for nano-brewing of Golden Promise and Full Pint.

Nano-brewing

Nano-brewing was performed at Rahr Malting Co. (Shakopee, MN, USA) using a beer recipe developed by Rahr Malting Co. for the purpose of this project. The method was developed to accommodate the large number of samples and limited amount of malted barley available per sample. Each sample of micro-malt was milled according to the ASBC Coarse Grind Extract method (ASBC Extract Method, Malt-4) on Bühler Universal Laboratory Disk Mill DLFU (Bühler AG, Uzwil, Switzerland). 150 g of milled micro-malt of each sample were divided equally into each of two mash cans, which were prepared to yield 0.47 L of beer per sample. The strike water was prepared with gypsum and CaCl using reverse osmosis (RO) water. 0.45 L of strike water, heated to 65°C, were added to each mash can with malt. A single infusion mash was employed; the mash cans were maintained at 65°C for 60 min and stirred using magnetic stir-bars. At the end of the 60-min mash, the mash cans were removed from the bath and weighed. Reverse osmosis water was added to each mash can to standardize the volume of all mashes to 0.45 L before filtering. The contents of each of the two mash cans per sample were poured over Goldtone Reusable Basket Coffee Filters (GoldTone, Pompano Beach, FL, USA) into a single beaker, to separate the wort from grist. Sparge water was prepared by heating two beakers of RO water at 82°C. The sparge water was then cooled to 77°C before 0.2 l was added to the grist. Approximately 1 L of wort was collected from each beaker, covered with parchment paper, and heated to 204°C. Once boiling, the parchment paper was removed and 0.9 g of Fuggle hops, with 4.9% α-acid, were added to each beaker. The aim was ~20 IBUs in the wort post-boil, translating to 10-12 IBUs in the final beer. The beakers were boiled for approximately 60 min each. To clarify the beer, 0.1 g of Irish Moss were added during the final seven minutes of boiling. Beakers were then removed from the hot plate, covered with a sterile aluminum foil lid, and

transferred to an ice bath for 15-20 min. Samples were swirled periodically until the wort cooled to 18-20°C. Beakers were then sanitized and placed in a biohood for hot trub settling. After the hot trub settled, wort was poured into autoclaved 1000 mL VWR media bottles (Avantor, Radnor, PA, USA). Specific gravity was recorded and adjusted to 11°P using autoclaved reverse osmosis water. Yeast was pitched, using serological pipets, directly from White Labs Pure Pitch packets (White Labs, San Diego, CA, USA) with a goal of ~7-10 × 106 cells/mL. Pitched wort was mixed in the media bottle and incubated at 20°C for 6-7 days until fermentation was complete. The media bottles of beer were incubated at 1°C for 24 hours and then the beer was poured into 1 L SodaStream bottles (SodaStream, Kefar Sava, Israel) and carbonated. These beers were then ready for laboratory analyses. 11 DH nano-brews were brewed each day, along with a Golden Promise control.

Laboratory analyses included specific gravity, pH, color, alcohol by volume (ABV), and international bitterness units (IBUs). Data were collected throughout the brewing process at mash (pH), pre-boil (specific gravity, pH, and color), post-boil (specific gravity, pH, and color), fermentation (specific gravity and pH), pre-bottle (specific gravity, pH, and color), and bottle (specific gravity, pH, color, ABV, GC-MS, and IBUs) steps. These data are available upon request. The GC-MS used was a Thermo Scientific GENESYS 10S UV-Vis spectrophotometer (Thermo Scientific, Waltham, MA, USA). All analyses, excluding sensory, were done following ASBC methods (American Society of Brewing Chemists, 1992).

Sensory evaluation

Sensory evaluations were conducted at Rahr Malting Co. using a trained sensory panel. Sensory panelist candidates were chosen based on robust brewing knowledge and previous sensory panel experience. These panelists were trained on 30 common beer flavors, including

off-flavors, by beers spiked with ~3x the average threshold concentration for each flavor compound. Panelists that were able to correctly identify all the compounds more than 70% of the time were selected for the final panel. In total there were 20 panelists that participated in the sensory data collection for this project. Due to the limited and varying amounts of micro-malt available, 162 DH beers were tasted between 6 and 13 times each. On average, each beer was tasted 12 times.

Sensory assessment was based on the comparison to a reference method. Descriptors were provided for the research beers (beer color, sweet flavor, cereal flavor, malty flavor, honey flavor, caramel/toffee flavor, grassy flavor, fruity flavor, floral flavor, and toasted flavor). All research beers along with a randomly distributed Golden Promise (brewing control) and Miller High Life (Miller Brewing Company, Milwaukee, WI, USA) (sensory control) were presented to panelists with a unique 3-digit code so they would not know which samples they were receiving. Miller High Life was used as the reference/control beer because the trained panel had the consensus that it contains a relatively neutral flavor profile within the style and has consistent quality control. Panelists were aware of the controls but not their identities. All beers brewed at Rahr were carbonated to similar levels using a SodaStream (SodaStream, Kefar Sava, Israel) and stored at 4 °C in a capped 5 oz cup prior to distribution to the sensory panel. The DH population research beers and the Golden Promise research beers were all tasted the same day they completed brewing. A maximum of 13 samples were tasted each day (11 DHs, 1 Golden Promise, and 1 Miller High Life).

Sensory descriptors were scored on a scale of +4 to -4, where positive numbers indicate more intensity than the sensory reference (Miller High Life) and negative numbers indicate less intensity than the reference, and 0 being the same as the reference. We also transformed the

sensory data to a simplified +1/0/-1 scale (more than/similar to/less than the reference) to check if additional QTLs could be detected. Using that scale, we were able to identify significant QTLs for malty, honey, and grassy flavors. The original +4 to -4 scale was used for QTL mapping of beer color, cereal flavor, and toasted flavor.

Metabolomics analysis and data processing

After sensory was completed, the remaining beers were shipped frozen in 50 mL Falcon Tubes from Rahr Malting Company (Shakopee, MN) to the Analytical Resource Core – Bioanalysis and Omics laboratory at Colorado State University (ARC-BIO, Fort Collins, CO). Of the 162 original beers, 155 had enough sample to conduct further testing, including the Golden Promise and Full Pint controls.

The contents were then pipetted to 20mL vials. Headspace Solid Phase Microextraction gas chromatography mass spectrometry (HS/SPME-GC-MS) was used to detect volatile compounds such as ketones, aldehydes, and esters using methods previously described (Bettenhausen et al., 2018, 2020). For instrumental analysis, the samples were first incubated at 65°C for 5 min, and then the headspace volatiles were extracted at the same temperature by a SPME fiber (DVB/PDMS/CAR 50/30 µm, Stableflex, Sigma-Aldrich) for 20 min, and injected into a DBWAXUI column (30 m x 0.25 mm x 0.25 µm, Agilent) in a Trace1310 GC (Thermo) coupled to an ISQLT MS (Thermo). The SPME fiber was desorbed at injection port (250°C) for 3 min, and then at fiber conditioning port (270°C) for 5 min. The GC inlet was operated under splitless mode during fiber desorption. The oven program started at 40°C for 4 min, ramped to 240°C at a rate of 5°C/min, and a final hold at 240°C for 0.5 min. Data were acquired under electron impact mode, with full scan of 40-500 amu at a rate of 5 scans/second. Transfer line and source temperatures were held at 250°C. Samples were not provided in replicates. Pooled QC

samples were run every six samples. Data were processed as described by Bettenhausen et al., (2018). Briefly, each sample resulted in a matrix of molecular features generated using XCMS software in R. Samples were normalized, relative abundances were calculated, mass spectra was deconvoluted using RamClust (Broeckling et al., 2014), and then metabolites were annotated by searching clustered features against in house and external libraries (NIST [http://www.nist.gov], Metlin [Tautenhahn et al., 2012; Zhu et al., 2013], the Human Metabolome Database [Wishart et al., 2013], and the Golm Metabolome Database [Hummel et al., 2013]).

SNP Genotyping and linkage map construction

The entire population of 236 DH lines and the two parents were genotyped with the Barley 50k iSelect SNP array (Bayer et al., 2017). Genotyping was performed by the Neogen GeneSeek laboratory (Scotland, UK: https://www.neogen.com/). SNPs were called using the GenomeStudio 2.0 software (Illumina Inc, San Diego, CA, USA). SNPs with >20% missing and/or heterozygous calls were removed, together with monomorphic SNPs. We did not identify highly distorted SNPs in the dataset. Data for each DH line were inspected; seven duplicated individuals were identified and eliminated from the population as well as one line having a high percentage of missing and/or heterozygous calls (22.5%).

The resulting 228 DH lines and 12,458 SNPs were used for linkage mapping. MSTmap (Wu et al., 2008); http://www.mstmap.org/mstmap_online.html) was used, with the following parameters: grouping LOD criteria = 10; population type = DH; no mapping size threshold = 2; no mapping distance threshold: 15 cM; try to detect genotyping errors = no; and genetic mapping function = kosambi. Physical coordinates of iSelect SNPs on the barley reference genome (Morex v2; Monat et al., 2019) were retrieved from BARLEX (Colmsee et al., 2005; http://apex.ipk-gatersleben.de/) and used to name and orient linkage groups.

QTL analysis and candidate gene identification

QTL mapping of the malting quality, beer sensory, and metabolomics traits was conducted using the mixed model method of Xu (2013) that was implemented in R by Lo et al. (2018). $-\log_{10}$ (*p*-values) were generated for each SNP. A genome-wide significance cutoff value was calculated based on false discovery rate (FDR) correction (Benjamini and Hochberg, 1995) at $\alpha = 0.05$ and used to identify significant QTLs. The percentage of phenotypic variation attributed to each QTL was calculated as described in Lo et al. (2018).

QTLs were displayed on linkage groups using MapChart version 2.32 software. QTL naming was based on the format described in Szűcs et al. (2009). Briefly, QTL names start with a "Q", followed by abbreviations for the trait and the population, as well as the chromosome number. Multiple QTLs on the same chromosome are indicated by a period and then the QTL number. Metabolite QTLs have the metabolite abbreviations following the "Q", and then follow the same format for population abbreviation and chromosome numbering.

The physical region of each QTL was determined based on the barley reference genome (Morex V2 [Monat et al., 2019]) and used to identify genes underlying the QTL intervals. Physical positions in Morex V1 (Mascher et al., 2017) were also retrieved.

Results

Phenotypic variation for malting quality, sensory descriptors, and metabolites

All phenotypic data generated in the DH population and the parents are provided in Supplementary Table 2.1, and a summary of phenotypes can be seen in Table 2.1. The malting

quality traits barley color, wort β-glucan (BG) and overall rank - together with the sensory traits cereal flavor, malty flavor, and grassy flavor and the metabolites linalool, ethyl hexonate-like, and oxalic acid dibutyl ester - all had higher values in Golden Promise than in Full Pint. Conversely, Full Pint had higher values for the remaining malting quality traits (kernel weight, kernel plumpness, malt extract, wort color, barley protein, wort protein, the ratio of soluble to total protein (S/T), diastatic power (DP), α-amylase (AA), free amino acid (FAN) and quality score), the sensory traits beer color, honey flavor, and toasted flavor, and the metabolites 2-methoxy-4-vinylphenol, and acetic acid, 2-phenylethyl ester. Transgressive segregation was observed for all traits except kernel plumpness.

Overall, the malts are under-modified, as evidenced by the low malt extract and S/T values and the high β-glucan values. There was limited variation for beer sensory descriptors, for which significant QTLs were detected, compared to the sensory reference (Miller High Life). Overall, Golden Promise was rated higher for cereal, malty, and grassy flavors. Full Pint was rated higher for beer color, and honey and toasted flavors.

A total of 543 volatile compounds were detected in the beer samples, 144 of which could be annotated as metabolites. The description of the beer flavor metabolome generated in this study is reported in Supplementary Table 2.1. Metabolite variation among the beers was detected and is visualized in Supplementary Figure 2.1. Of the 144, only 5 metabolites could be associated to a QTL ($p < 0.05$, Table 2.1). Those metabolites are: 2-methoxy-4-vinylphenol (MVP), acetic acid, 2-phenylethyl ester (PEA), linalool (LOO), a compound with similar, but slightly modified structure as ethyl hexonate (ethyl hexonate-like, EHEXL), and oxalic acid dibutyl ester (DBOA).

All possible pairwise correlations (n = 325) among malt quality, sensory and metabolic traits as well as their respective *p*-values are shown in Supplementary Table 2.2. Correlations (positive and negative) between traits ranged from 0 to 0.9; of these, 119 were significant ($p < 0.05$). Most of the significant correlations were between malting quality traits. The highest positive correlation was found between FAN and wort protein (0.92; $p = 0.000$), while the highest negative correlation was between overall rank and quality score (-0.98; $p = 0.000$).

Focusing on correlations >0.5 or <-0.5, there were expected patterns of trait relationships for most malting quality traits. For example, barley protein was positively correlated with diastatic power and negatively correlated with malt extract; wort β-glucan was negatively correlated with wort protein and S/T; FAN, wort color, wort protein, S/T, AA, and quality score were all positively correlated with each other; malt exact, S/T, AA, and quality score were also all positively correlated with each other. Quality score was positively correlated with many of the traits listed above (malt extract, wort color, wort protein, S/T, AA, and FAN) and negatively correlated with overall rank. Conversely, overall rank was negatively correlated with the same traits listed for quality score. Of the sensory traits, there were no correlations >0.5 or <-0.5 between sensory traits. Of the five flavor metabolites, the content of DBOA was positively correlated with LOO content.

No correlations were found which included traits from each of the three categories, but beer color was positively correlated with wort protein, S/T, AA, FAN, quality score; and negatively correlated with overall rank. Also, PEA content was negatively correlated with BG (Supplementary Table 2.2).

Development of the Golden Promise genetic map

A total of 12,458 polymorphic SNPs and 228 DHs were used to generate a genetic map of the Oregon Promise population. MSTmap (Wu et al., 2008) was the software of choice for genetic map construction, which mapped 12,453 SNPs into 1,073 bins across the seven linkage groups representing each of the seven barley chromosomes (Table 2.2; Supplementary Table 2.3). The genetic map spanned 1,221.76 cM and had an average density of 1 bin per 0.88 cM. There were two large gaps of 31.1 and 30.5 cM on chromosomes 1H and 6H, respectively (Supplementary Table 2.3). Linkage groups ranged in size from 134.77 cM for chromosome 4H to 193.94 cM for chromosome 3H. The genetic map of the Oregon Promise population together with the SNP information used for its construction is available in Supplementary Table 2.3.

Identification of QTLs and candidate genes

QTLs were identified for malt quality, sensory, and metabolite traits using the mixed model for QTL mapping of Xu et al. (2013) implemented in R (Lo et al., 2018). These QTLs were distributed across all chromosomes except 4H and include: 21 QTLs for 14 malting quality traits, eight QTLs for six sensory descriptors, and eight QTLs for five metabolic compounds (Figure 2.1; Table 2.3; Supplementary Table 2.4). The percentage of phenotypic variation accounted for by individual QTLs ranged from low (6.25%) for one of the DP QTLs (*QDp.GpFp-1H*) to substantial (48.3%) for overall rank (*QOr.GpFp-5H*) (Table 2.3). QTL clustering for traits belonging to different categories were identified on chromosomes 3H and 5H (mid-5H and end-5H), the end of 5H being the largest QTL hotspot (Figure 2.1). Annotated barley genes that fell within each QTL region were identified for all QTLs and are provided in Supplementary Table 2.5. Details of QTLs and candidate genes within each phenotypic category are provided below.

Malt quality

Twenty-one QTLs for 14 malt quality traits were identified on chromosomes 1H, 2H, 3H, 5H, and 7H (Table 2.3; Figure 2.1) and their $-\log_{10}$ (*p*-values) ranged from 3.40 for one of the kernel plumpness QTLs (*QKp.GpFp-3H*) to 16.00 for wort protein, S/T, quality score, and overall rank (Table 2.3). The percentage of phenotypic variation accounted for by each QTL ranged from 6.25% for a DP QTL (*QDp.GpFp-1H*) to 48.27 % for overall rank (*QOr.GpFp-5H*). There were overlapping QTLs for kernel plumpness and barley protein on 3H, which were in close proximity to the kernel weight QTL (Figure 2.1). Kernel plumpness, barley color, and barley protein QTLs overlapped at the mid-5H hotspot, while malt extract, wort color, wort protein, S/T, AA, BG, FAN, quality score, and overall rank QTLs clustered at the end-5H hotspot (Table 2.3; Figure 2.1). It is expected that the quality score and overall rank would coincide, and that they would coincide with the QTLs for the traits used to calculate the score.

QTLs contained between 24 and 3,903 annotated genes for *QKp.GpFp-3H* and *QBc.GpFp-5H*, respectively, with an average of 540 genes (Supplementary Table 2.5). Barley genes *HORVU.MOREX.r2.5HG0398940* and *HORVU.MOREX.r2.5HG0397930*, which correspond to *Alanine aminotransferase* (*HvAlaAT*; Sato et al., 2016) and *Dense and erect panicle 1* (*HvDep1*; Wendt et al., 2016), respectively, were identified within the kernel plumpness (*HvAlaAT* and *HvDep1*) and barley color (*HvDep1*) QTLs at the 5H-mid QTL cluster. *HvAlaAT* has been shown to control the length of dormancy, while *HvDep1* is involved in culm elongation and grain size in barley. Gene models *HORVU.MOREX.r2.5HG0447180* and *HORVU.MOREX.r2.5HG0446540*, corresponding to *Mitogen-Activated Kinase Kinase 3* (*HvMKK3*; Nakamura et al., 2016) and *Gibberellin 20-oxidase 1* (*HvGA20ox1*; Nagel et al., 2019), respectively, were contained within the overlapping region of all QTLs except BG at the

5H-end hotspot. Both *HvMKK3* and *HvGA20ox1* are reported to be involved in the regulation of seed dormancy in barley. The relationships of these genes with QTLs for malting quality, and the other two categories of data (sensory and metabolite), are explored in the Discussion. The *Sdw1/Denso* locus, where the determinant gene is *HvGA20ox2* (Xu et al., 2017; Jia et al., 2009), is on chromosome 3H and coincides with the barley grain protein QTL. Two amino acid permease genes (*HORVU.MOREX.r2.3HG0256690* and *HORVU.MOREX.r2.3HG0256700*), with roles in nitrogen remobilization (Kohl et al., 2012), were also identified among the annotated genes in the kernel plumpness and barley protein QTLs (Supplementary Table 2.5). Candidate genes for the singleton malt quality QTLs (Figure 2.1; Table 2.3) were not explored in detail.

Sensory

A total of eight QTLs for six sensory traits were identified on chromosomes 2H, 3H, 5H, and 7H (Table 2.3; Figure 2.1), with $-\log_{10}$ (*p*-values) ranging from 3.52 for cereal flavor to 14.00 for beer color (Table 2.3). The percentage of phenotypic variation accounted for by each QTL ranged from 6.88 % for cereal flavor (*QCe.GpFp-7H*) to 21.36 % for beer color (*QCo.GpFp-5H*). Overlapping QTLs for this category were located on chromosome 3H (malty flavor [*QMa.GpFp-3H*] and toasted flavor [*QTo.GpFp-3H*]), which also overlapped with the malt quality QTL for kernel weight (*QKw.GpFp-3H*), and chromosome 5H (beer color [*QCo.GpFp-5H*] and toasted flavor [*QTo.GpFp-5H*]), which were located on the 5H-end hotspot for malt quality and metabolic traits (Figure 2.1). The QTL for honey flavor on 5H (*QHo.GpFp-5H*), which does not overlap with any other sensory QTL, is coincident with the malt quality QTLs for kernel plumpness, barley color, and barley protein at the mid-5H cluster (Figure 2.1).

Sensory QTLs contained between 1 (*QCe.GpFp-7H*) and 5,467 (*QHo.GpFp-5H*) genes, with an average of 792 genes (Supplementary Table 2.5). Candidate genes for the 5H QTLs included those mentioned above for malting quality: *HvAlaAT* (Sato et al., 2016) and *HvDep1* (Wendt et al., 2016) for the honey flavor QTL on the mid-5H cluster, and *HvMKK3* (Nakamura et al., 2016) and *HvGA20ox1* (Nagel et al., 2019) for the beer color QTL located on the 5H-end hotspot (Figure 2.1). Among the genes located within the 3H hotspot we can highlight *HORVU.MOREX.r2.3HG0259410*, which encodes an ethylene-responsive transcription factor (ERF). ERFs play crucial roles in plant developmental processes and have been associated with kernel size (Zhang et al., 2020). This region is ~5,000 kb from *HvGA20ox2* (Xu et al., 2017). Candidate genes for the remaining singleton sensory QTLs were not explored in detail.

Flavor metabolites

Eight QTLs for the accumulation of five flavor metabolites were identified on chromosomes 2H, 3H, 5H, 6H, and 7H (Table 2.3; Figure 2.1). $-\log_{10}(p\text{-values})$ ranged from 3.32 for one of the oxalic acid dibutyl ester QTLs (*QDBOA.GpFp-3H.2*) to 7.87 for a 2-methoxy-4-vinylphenol QTL (*QMVP.GpFp-5H*; Table 2.3). The percentage of phenotypic variation accounted for ranged from 6.71 % for *QDBOA.GpFp-3H.2* to 17.50 % for *QMVP.GpFp-5H*.

The metabolite QTLs for 2-methoxy-4-vinylphenol (MVP) and oxalic acid dibutyl ester (DBOA) located on 3H overlapped, but with no other QTL for any category of data (Figure 2.1). QTLs for MVP (*QMVP.GpFp-5H*) and acetic acid, 2-phenylethyl ester (PEA) (*QPEA.GpFp-5H*) overlapped on the 5H-end hotspot, where many malt quality and sensory QTLs also colocalized (Figure 2.1).

QTLs contained between 1 (*QDBOA.GpFp-3H.1*) and 4,429 (*QPEA.GpFp-7H*) genes, with an average of 781 genes. *HvMKK3* (Nakamura et al., 2016) and *HvGA20ox1* (Nagel et al., 2019) are candidate genes for the 5H-end hotspot. It should be noted that in the 3H QTL for MVP (*QMVP.GpFp-3H*) there is also a gene (*HORVU.MOREX.r2.3HG0247750*) encoding a cytochrome P450 family cinnamate 4-hydroxylase, which is involved in the synthesis of precursors (cinnamic acids) to 2-methoxy-4-vinylphenol (Harakava et al., 2005; Gómez-López et al., 2019). Candidate genes for singleton metabolite QTLs were not explored further.

Discussion

This research generated multiple data sets on a large biparental mapping population, which allows for the first comprehensive look at the genetic basis of barley contributions to beer flavor, together with metabolomic compounds in beer. The sensory and metabolite data sets are anchored in the malting quality data set: malt precedes beer. Un-malted barleys do not display notable flavor or aroma differences: it is the malting process that leads to these differences. Therefore, an analysis of the contributions of barley genotype to beer flavor is inextricably confounded by the style of malt, and how each genotype responds to the malting protocol that was used to make the malt. This focus on malt, of course, does not account for the significant flavor contributions of hops and yeast to the finished beer. In short, the analysis of the contributions to beer flavor made by barley genotypes (via their malt) is a challenging prospect. In the case of this research, it is important to bear in mind that the malting quality data, while based on the same barley germplasm (the Oregon Promise population) grown at the same location (Corvallis, Oregon and malted at the same facility (the USDA-ARS Cereal Crops Research Unit) using the same methods, traces to samples from a different crop year (2013) than

the samples that were malted and used for nano-brewing (2014). Furthermore, the samples of the parents (Golden Promise and Full Pint) trace to a different location (Lebanon, Oregon). This "imbalance" was an inevitable consequence of the timing and scale of the experiment. Given this caveat these datasets are novel in both nature and scope, as this is a first for nano-brewing of a mapping population, which in turn allows for an unprecedented scale of sensory and beer metabolomics.

Malt modification and beer flavor

Due to the large number of grain samples, under-modification of malts used in QTL studies is an inevitable consequence of the need to use automated, high throughput malting systems. It is impossible and unrealistic to optimize malting regimes for each individual grain sample. It is particularly difficult in a case such as the current research, where neither parent is amenable to current malting protocols, which are traditionally designed for contemporary varieties. Golden Promise is an heirloom variety that continues to persist in the market due to perceived contributions to flavor; Full Pint is a specialty variety that also has perceived contributions to flavor (Mallet, 2014). This leads to the question – are differences in contributions to beer flavor of barley genotypes artifacts of poor modification? Herb et al. (2017a) presented evidence to the contrary. Even when adjusting for modification differences, flavor differences were still present in the subset of the Oregon Promise population they used. Bettenhausen et al. (2020) and Windes et al. (2020) also found differences in flavor in a small sample of contemporary varieties, when dealing with similar degrees of modification which were achieved by tailoring malting to the needs of each variety. Likewise, Craine et al. (2021) used bespoke malting protocols to achieve similar levels of modification in a small set of barley varieties/selections of potential interest to the craft industry. Cumulatively, these results point to

subtle, but definitive contributions of barley genotypes to beers made from pale malts, despite the degree of modification. A key follow-up question for future research remains: what are the contributions of barley genotypes to beer flavor when higher color malts are made from these varieties?

Trained panel sensory analysis of nano beers identifies differences in flavor

The nano-beers prepared at Rahr Malting for this research were produced using a different protocol than that used by Herb et al. (2017b). Furthermore, not all the same sensory descriptors were used in this study as in prior research. In the current study, Golden Promise was rated higher for cereal, malty, and grassy. Full Pint was rated higher for beer color, honey and toasted. In Herb et al. (2017b), Golden Promise was described as significantly higher for floral and fruit, whereas Full Pint was significantly higher for malt, sweet, toasted, and toffee. These commonalities (e.g. toasted) and differences (e.g. malt) between slightly different beer styles brewed from the same two varieties of barley using a different protocol, underscores the challenges of sensory analysis and the importance of beer style and descriptor lexicon in assessing varietal and environmental contributions to beer flavor. In both Herb et al. (2017b) and in the current study, beers brewed from the progeny showed much more variation than those brewed from the parents (Table 2.1). As with malting quality, the positive and negative transgressive segregants for flavor descriptors in the progeny suggests that the parents have different alleles at multiple loci determining these attributes.

Metabolite abundance: barley variety signatures in beer flavor?

The relative abundances of the five metabolites that are the focus of this research varied between Full Pint and Golden Promise. 2-methoxy-4-vinylphenol (a phenol) and phenylethyl

acetate (a benzenoid compound) were more abundant in Full Pint. In Golden Promise, ethyl hexanoate (a lipid ester), linalool (a terpene) and dibutyl oxalate (a carboxylic acid) were more abundant. Bettenhausen et al. (2018) and Windes et al. (2020) also reported that Full Pint beers had higher abundances of benzenoid compounds, phenolics, and lipids and a lower abundance of ethyl hexanoate and many terpenes. Therefore, these relative abundances of metabolic compounds may be useful chemical signatures for specific varieties. Connecting metabolic signatures with sensory attributes, however, can be more challenging. For example, benzenoid compounds, phenolics, and lipids can lead to a fruity/floral/spicy profile and yet Golden Promise beer, rather than Full Pint beer, was described as being higher for fruity and floral attributes by Carpena et al., (2021). Since the fruity and floral descriptors were not significant in the current research, it is not possible to associate them with metabolite abundance. Furthermore, the abundance of certain metabolites, and corresponding flavors, may be due to the interactions of the malt with other components of the finished beer. The higher abundance of 2-methoxy-4-vinylphenol in Full Pint beer could be due to the by enzymatic decarboxylation of the compound ferulic acid by certain strains of *S. cerevisiae* (Coghe et al., 2004). As with malting quality and sensory attributes, the positive and negative transgressive segregants for metabolite abundance in the progeny suggests that the parents have different alleles at multiple loci determining these attributes.

Phenotypic correlations set the stage for QTL analysis

Many of the correlations between malting quality traits conform to expectations based on prior literature: for example, barley grain protein was positively correlated with enzymatic traits and negatively correlated with malt extract (Xue et al., 2008). The phenotypic correlations between malting quality traits and sensory traits and between malting quality traits and metabolic

compounds need to be considered in view of the malting quality data tracing to malts different than those used for brewing. Nonetheless, the negative correlation between wort β-glucan and PEA (-0.55, $p = 0.00$) could merit further investigation. A genetic basis for this correlation is provided by coincident QTLs for these traits at the QTL hotspot located on the 5H-telomeric region, as described below. The positive correlation between malty and toasted (0.46, $p = 0.00$) is also supported by the overlapping QTLs on 3H. The positive correlation of dibutyl oxalate with linalool (0.50, $p = 0.00$) merits further exploration, given the compound is found in both barley and hops. It has a high affinity for calcium and in the context of beer, precipitated oxalate in the beer leads to particulate and haze formation, gushing, and "beer stone," which is particularly a problem in brewing equipment, the latter being responsible for the blocking of beer piping (Oliver, 2012).

Candidate genes for QTL clusters include genes associated with dwarf growth habit and degree of dormancy

Of particular interest, in terms of QTLs and candidate genes, are the clusters (hot spots) of coincident QTLs for multiple traits on chromosomes 3H and 5H. There are intuitively appealing candidate genes for each of these clusters, based on prior literature, and it is also possible that there are multiple physically linked genes that have roles in determining these QTLs.

The 3H QTL cluster is the most diffuse of the three clusters, and further research would be necessary to assign candidate genes to the various malting quality, beer sensory, and beer metabolite QTLs. A candidate on this chromosome is the *Sdw1/Denso* locus, where the determinant gene is *HvGA20ox2* (Xu et al., 2017; Jia et al., 2009) and Full Pint has the recessive (dwarfing allele). The *Denso* locus is within the barley protein QTL, where the higher value

allele was contributed by Golden Promise. The wild type allele, in this case, was associated with higher grain protein, a trait in malting barley that has an upper limit, depending on beer style. For adjunct malts, 12% is the maximum; lower levels are required for all malt brewing. In addition to pleiotropic effects on grain protein, *Denso* alleles are known to affect a range of other agronomic traits (reviewed by Kuczynska et al., 2013). Additional research is required to determine if, in the Oregon Promise population, *Denso* is also the determinant of the QTLs for kernel plumpness, kernel weight, and malty flavor, and toasted flavor. If it is, it would be a positive pleiotropic effect of the wild type allele, as Golden Promise has higher value alleles at these QTLs. Golden Promise also contributes the higher value alleles for malty, honey and toasted flavors, as well as for two volatile metabolites (MVP and DBOA) at QTLs distal to *HvGA20ox2*. In addition to *HvGA20ox2*, there are other genes in this 3H QTL region that could have impacts on malting and flavor traits. These include *HORVU.MOREX.r2.3HG0259410*, which encode ethylene-responsive transcription factors, and *HORVU.MOREX.r2.3HG0256690* and *HORVU.MOREX.r2.3HG0256700*, encoding two amino acid permeases. Ethylene is a plant hormone that stimulates the biosynthesis of gibberellin, a hormone that releases seeds from dormancy (Corbineau et al., 2014). ERFs have also been found to impact kernel size, which would logically also have an impact on kernel weight as well as the protein content of the kernels, as larger kernels tend to have lower protein content (Magliano et al., 2014). Amino acid permeases are involved in nitrogen remobilization (Kohl et al., 2012); nitrogen availability and supply impacts grain protein content, which may affect many malt quality traits (Guo et al., 2019). Assuming that higher levels of all the 3H QTL phenotypes (except perhaps grain protein) are positives, from a breeding standpoint it would seem desirable to maintain the positive relationships by selecting for a large block of this chromosome region with Golden Promise

alleles. Interestingly, the variety “Oregon Promise”, which was top rated for flavor by a consumer panel (Bettenhausen et al., 2020) has a 100% Full Pint haplotype at all alleles for the chromosome 3H QTLs (Supplementary Table 2.3). Assuming linkage, rather than pleiotropy, additional research will be required to determine if the agronomic advantages of the *Denso* allele from Full Pint can be combined, via recombination, with the potentially favorable alleles for other traits from Golden Promise.

Candidate genes for the mid-5H QTL cluster are *HvDep1* and *HvAlaAT1*. The former is a dwarfing gene – the *Ari-e* locus. Golden Promise has the loss of function dwarfing allele (*ari-e.GP*). Full Pint has the wild type (functional) allele. The Golden Promise allele, the result of an induced mutation, was a breakthrough in reducing plant height and lodging. *HvAlaAT1* is the determinant of *SD1*, a major dormancy gene (Sato et al., 2016). Allele resequencing shows that Full Pint and Golden Promise are identical at the causal SNP in *HvAlaAT1* (Sweeney et al., submitted). While it is possible that regulation of the structural gene could account for differences in dormancy, with pleiotropic effects on malting and sensory traits, this leaves *HvDep1* as the most obvious candidate. The *ari-e.GP* allele has negative pleiotropic effects on thousand grain weight and grain length (Wendt et al., 2016). This supports our detection of a QTL for kernel plumpness, with Full Pint contributing the positive (favorable) allele, and barley protein, with Golden Promise contributing the higher value (generally unfavorable) allele. In this same QTL cluster, Full Pint has the higher value and positive allele for grain color; brighter grain has a higher Agtron score. There are no reports of pleiotropic effects of *ari-e.GP* on grain color. Further research is warranted, perhaps following the lead in rice, protein and seed color are positively correlated (Tan et al., 2001). Golden Promise contributes the positive allele for honey flavor. Further research is necessary to determine the basis of this QTL, which may relate to

grain protein level and sensory panel perceptions. Storage proteins are important in all cereals for the embryo once germination occurs, and these proteins typically have high amounts of the amino acid, proline (Fox, 2010). Although not directly associated with honey flavor, proline has a sweet flavor (Sorensen and Sammis, 2004), which may be chemically altered during the malting process to be similar enough (along with other metabolic factors) to be perceived as a honey flavor by a sensory panel, but further exploration is needed. From a breeding standpoint, moderate grain protein and plump kernels are desirable. Therefore, in this population, the *denso* dwarfing allele on 3H (tracing to Full Pint) would be more favorable than the *arie*-dwarfing allele on 5H. However, selection for the wild type allele at 5H would compromise selection for the coincident honey flavor QTL, where Golden Promise contributes the favorable allele. Interestingly, the Oregon Promise variety has Full Pint alleles at all markers in this this QTL region (Supplementary Table 2.3). Perhaps, if honey flavor is desirable in beers, the 3H QTL allele (tracing to Golden Promise) is sufficient.

The most obvious candidate gene for the end-5H region QTLs is *HvMKK3*, the most cited determinant gene for *SD2* (Nakamura et al., 2016). Full Pint has the most non-dormant *HvMKK3* allele (*MKK3_N**) (Sweeney et al., submitted). However, *HvGA20ox1* may also have role in these QTL, as proposed by Nagel et al. (2019). Full Pint and Golden Promise have contrasting alleles at both *HvMKK3* and *HvGA20ox1*, but according to Sweeney et al., (submitted) *HvGA20ox* is not a determinant of dormancy and malting quality in North American spring barley germplasm. Precedent for malting quality QTL coincident with the *SD2* locus are provided by Castro et al. (2010), who used a biparental population with Full Pint as a parent and reported that the most QTL, and QTL with the largest effects, were found on the long of chromosome 5H at a location coincident with as *SD2*. In terms of validation, Oregon Promise

has Full Pint alleles at all markers in end-5H QTL region (Supplementary Table 2.3). Castro et al. (2010) also reported QTLs for dormancy and water sensitivity in the *SD2* region – with Full Pint contributing the non-dormant and non-water sensitive alleles. In the current research, absolute dormancy was not encountered, because grain was malted one year after harvest. Furthermore, pre-harvest sprouting and water sensitivity were not observed. As argued by Vetch et al. (2019) and Sweeney et al. (submitted), the effects of *SD2* on malting quality traits can be in terms of degree of dormancy. If a lower degree of dormancy is equated with higher germination rate and metabolic activity, then we would expect Full Pint to achieve a greater degree of modification. Indeed, Full Pint has higher value alleles for malt extract, wort protein, S/T, AA, FAN, quality score, and the lower value allele for BG.

Conclusions

This work represents an important first step towards integrating malting quality, beer sensory, and metabolomics via an understanding of the determinant genes. The data presented herein support that morphological traits (e.g. semidwarf growth habit) and seed physiology traits (e.g. dormancy) may have profound downstream effect of malting quality, beer flavor, and metabolite abundance. QTL data indicate potential causal relationships between beer flavor outcomes and the genes determining malting quality and volatile metabolites. Our results lay the groundwork for future genetics and breeding research, including (i) editing of candidate genes to determine flavor outcomes and (ii) marker assisted selection for key QTL haplotypes in other genetic backgrounds. Further research is also warranted in malting and brewing sciences involving the same genotypes, or subsets thereof. These could include (i) different malt styles (ii) different beer styles, and/or different growing environments.

TABLES AND FIGURES

Table 2.1 Parental and DH population mean, standard error, range, and skewness values. Malt quality abbreviations: S/T = soluble/total protein, DP = diastatic power, AA = α-amylase, BG = β-glucan, FAN = free amino nitrogen. Unit abbreviations: mg = milligrams, % = percent, DU = dextrinizing units, ppm = parts per million, a.u. = arbitrary units.

Category	Trait	Golden Promise	Full Pint	DH population			
				Mean	SE	Range	Skewness
Malt quality	Kernel weight (mg)	44.00	49.70	42.58	0.33	32.75 – 52.47	0.12
	Kernel plumpness (%)	41.00	100.00	91.77	0.62	67 – 100	-1.16
	Barley color (Agtron)	96.00	43.00	46.53	0.56	31 – 65	0.00
	Malt extract (%)	77.80	78.10	77.51	0.13	73.67 – 80.88	-0.03
	Wort color	2.00	2.60	2.13	0.03	1.48 – 4	1.03
	Barley protein (%)	11.90	13.90	12.78	0.09	10.58 – 15.98	0.38
	Wort protein (%)	3.64	5.07	4.51	0.06	3.29 – 6.29	0.42
	S/T (%)	32.20	36.40	36.78	0.45	26.29 – 49.64	0.33
	DP (°ASBC)	98.00	204.00	137.02	2.08	87.70 – 221	0.68
	AA (20°DU)	52.10	122.40	77.09	1.77	43.23 – 126.53	0.32
	BG (ppm)	677.00	421.00	361.25	13.15	48.13 – 743.72	0.36
	FAN (ppm)	172.00	245.00	179.29	3.58	109.97 – 284.84	0.37
	Quality score	29.00	42.00	38.84	0.94	13 – 67	0.39
	Overall rank	199.00	53.00	75.95	3.70	1 – 156	0.03
Sensory	Beer color (-4 – +4 scale)	-1.13*	-0.50*	-0.59	0.06	-2.36 – 1.08	-0.11
	Cereal flavor (-4 – +4 scale)	0.88*	0.75*	0.49	0.03	-0.4 – 1.25	0.00
	Malty flavor (-4 – +4 scale)	0.38*	-0.13*	0.21	0.02	-0.45 – 1	0.09
	Honey flavor (-4 – +4 scale)	0.38*	0.57*	0.41	0.02	-0.38 – 1.13	0.13
	Grassy flavor (-4 – +4 scale)	0.75*	0.25*	0.80	0.03	0 – 1.69	0.27
	Toasted flavor (-4 – +4 scale)	-0.25*	0.00*	-0.04	0.02	-0.71 – 0.80	0.48
Metabolites	2-methoxy-4-vinylphenol (a.u.)	3,745,686.15*	4,826,833.96*	3,513,127.46	62,018.71	1,917,265 – 6,797,483	0.51
	Acetic acid, 2-phenylethyl ester (a.u.)	279,516,231.50*	643,562,567.60*	790,741,728.85	19,694,988.91	151,122,999 – 1,425,190,984	-0.06
	Linalool (a.u.)	6,103,479.69*	5,183,004.04*	4,152,971.83	94,007.80	1,919,898 – 7,788,532	0.60
	Ethyl hexonate-like (a.u.)	76,978.45*	21,639.55*	35,891.96	2,214.40	0.005 – 104,310	0.34
	Oxalic acid dibutyl ester (a.u.)	23,780,778.51*	21,752,774.06*	23,260,970.58	448,241.33	11,862,902 – 39,044,261	0.54

*Data from samples grown in Lebanon, OR, USA (2016).

Table 2.2 Distribution of SNPs in the Oregon Promise genetic map

Chr	1H	2H	3H	4H	5H	6H	7H	Total
Markers	1,512	2,284	1,487	1,310	2,145	1,492	2,223	12,453
Bins	143	174	151	128	194	102	181	1,073
cM	170.63	190.12	193.94	134.77	212.60	135.92	183.79	1,221.76

Table 2.3 QTLs identified for sensory, malt quality, and metabolite traits

Category	Trait	QTL	Peak SNP	Chr	Position (kb)*	-Log$_{10}$(P)	QTL region (cM)	QTL region (kb)*	% Phenotypic variation	Effect	Known gene(s) in QTL region
Malt quality	Kernel weight	*QKw.GpFp-3H*	JHI-Hv50k-2016-207525	3H	582,098	4.51	132.30 - 154.37	577,460 - 594,378	9.91	1.28	
	Kernel plumpness	*QKp.GpFp-3H*	JHI-Hv50k-2016-205406	3H	572,529	3.40	126.15 - 127.03	572,529 - 573,139	6.64	0.11	
		QKp.GpFp-5H	JHI-Hv50k-2016-307371	5H	435,709	4.48	46.61 - 53.63	374,134 - 446,936	11.06	-2.77	*HvAlaAT1, HvDep1*
	Barley color	*QBc.GpFp-2H*	JHI-Hv50k-2016-98501	2H	492,803	3.77	66.52 - 67.40	489,367 - 520,441	13.71	2.92	
		QBc.GpFp-5H	JHI-Hv50k-2016-301330	5H	349,008	6.95	43.98 - 49.68	35,704 - 437,198	25.61	-3.82	*HvDep1*
	Malt extract	*QMe.GpFp-2H*	JHI-Hv50k-2016-103558	2H	563,603	3.52	76.69	563,603 - 565,975	10.46	0.59	
		QMe.GpFp-5H	JHI-Hv50k-2016-365534	5H	594,137	12.82	200.76 - 212.60	588,682 - 598,994	28.92	-0.86	*HvMKK3, HvGAox1*
	Wort color	*QWc.GpFp-5H*	JHI-Hv50k-2016-362729	5H	590,865	16.00	195.05 - 212.60	586,795 - 598,994	32.73	-0.25	*HvMKK3, HvGAox1*
	Barley protein	*QBp.GpFp-3H*	SCRI_RS_103215	3H	572,324	5.16	125.27 - 127.91	571,521 - 573,139	10.01	0.36	*HvGA20ox2*
		QBp.GpFp-5H	JHI-Hv50k-2016-301330	5H	349,008	4.09	46.17 - 48.36	349,008 - 431,457	15.00	0.51	
	Wort protein	*QWp.GpFp-5H*	JHI-Hv50k-2016-362943	5H	591,069	16.00	196.81 - 212.60	587,561 - 598,994	37.19	-0.42	*HvMKK3, HvGAox1*
	S/T	*QSt.GpFp-5H*	JHI-Hv50k-2016-363828	5H	592,162	16.00	199.44 - 212.60	588,466 - 598,994	46.00	-3.57	*HvMKK3, HvGAox1*
	DP	*QDp.GpFp-1H*	JHI-Hv50k-2016-4906	1H	4,865	3.60	5.97 - 8.60	4,798 - 5,592	6.25	6.65	
		QDp.GpFp-7H	SCRI_RS_161101	7H	1,401	3.92	0 - 1.32	227 - 2,619	7.43	-7.42	
	AA	*QAa.GpFp-5H*	JHI-Hv50k-2016-363791	5H	592,087	15.00	201.66 - 212.60	590,717 - 598,994	42.14	-12.92	*HvMKK3, HvGAox1*
	BG	*QBg.GpFp-2H*	BOPA1_3608-2133	2H	648,278	5.01	151.81 - 162.78	644,545 - 655,253	9.56	-52.41	
		QBg.GpFp-5H	JHI-Hv50k-2016-360298	5H	586,834	3.39	195.49 - 202.51	586,834 - 590,865	7.16	45.03	
		QBg.GpFp-7H	JHI-Hv50k-2016-438742	7H	3,362	3.83	4.83 - 11.41	3,362 - 7,668	7.66	45.63	
	FAN	*QFa.GpFp-5H*	JHI-Hv50k-2016-362943	5H	591,069	10.00	200.76 - 212.60	589,596 - 598,994	42.69	-28.00	*HvMKK3, HvGAox1*
	Quality score	*QQs.GpFp-5H*	JHI-Hv50k-2016-363828	5H	592,162	16.00	198.12 - 212.60	587,972 - 598,994	36.87	-6.78	*HvMKK3, HvGAox1*
	Overall rank	*QOr.GpFp-5H*	JHI-Hv50k-2016-364126	5H	592,490	16.00	200.76 - 212.60	588,682 - 598,994	48.27	31.21	*HvMKK3, HvGAox1*
Sensory	Beer color	*QCo.GpFp-5H*	JHI-Hv50k-2016-361935	5H	588,466	14.00	195.05 - 212.16	586,795 - 598,994	21.36	-0.332	*HvMKK3, HvGAox1*
	Cereal flavor	*QCe.GpFp-7H*	JHI-Hv50k-2016-511500	7H	619,240	3.52	157.32	619,240	6.88	0.089	
	Malty flavor	*QMa.GpFp-3H*	JHI-Hv50k-2016-207283	3H	581,633	3.71	138.02 - 138.90	581,663 - 582,615	7.04	0.038	
	Honey flavor	*QHo.GpFp-3H*	JHI-Hv50k-2016-225245	3H	625,551	3.85	192.19 - 193.94	623,023 - 625,680	7.11	0.035	
		QHo.GpFp-5H	JHI-Hv50k-2016-284934	5H	19,967	5.20	32.57 - 73.05	10,652 - 491,116	7.73	0.037	*HvAlaAT1, HvDep1*
	Grassy flavor	*QGr.GpFp-2H*	BOPA1_816-265	2H	34,276	3.82	47.05	34,276 - 34,355	7.81	0.05	
	Toasted flavor	*QTo.GpFp-3H*	JHI-Hv50k-2016-207283	3H	581,633	4.96	132.30 - 138.90	577,460 - 582,615	10.35	0.09	
		QTo.GpFp-5H	BOPA1_6873-531	5H	592,173	3.65	202.07 - 205.14	590,798 - 592,247	7.35	-0.08	
Metabolites	2-methoxy-4-vinylphenol	*QMVP.GpFp-3H*	SCRI_RS_146347	3H	528,679	5.54	86.54 - 99.28	526,926 - 548,613	11.68	270460.71	
		QMVP.GpFp-5H	JHI-Hv50k-2016-367564	5H	597,237	7.87	204.70 - 212.60	592,087 - 598,994	17.50	-334270.87	*HvMKK3, HvGAox1*
	Acetic acid, 2-phenylethyl ester	*QPEA.GpFp-5H*	JHI-Hv50k-2016-367061	5H	596,380	7.08	199.44 - 212.60	588,466 - 598,994	14.04	-94588275.18	*HvMKK3, HvGAox1*
		QPEA.GpFp-7H	JHI-Hv50k-2016-470701	7H	82,589	6.00	87.87 - 97.52	66,454 - 503,740	10.70	-84295784.80	
	Linalool	*QLOO.GpFp-2H*	JHI-Hv50k-2016-113871	2H	611,171	7.12	106.13 - 119.29	606,789 - 624,863	10.95	-395681.89	
	Ethyl hexonate-like	*QEHEXL.GpFp-6H*	JHI-Hv50k-2016-380526	6H	27,476	3.89	44.89 - 47.97	25,117 - 28,720	7.09	7591.54	
	Oxalic acid dibutyl ester	*QDBOA.GpFp-3H.1*	JHI-Hv50k-2016-195050	3H	527,696	3.38	87.42	527,696	7.37	1600930.34	
		QDBOA.GpFp-3H.2	JHI-Hv50k-2016-215710	3H	602,711	3.32	168.87 - 169.31	602,711 - 603,732	6.71	1523937.55	

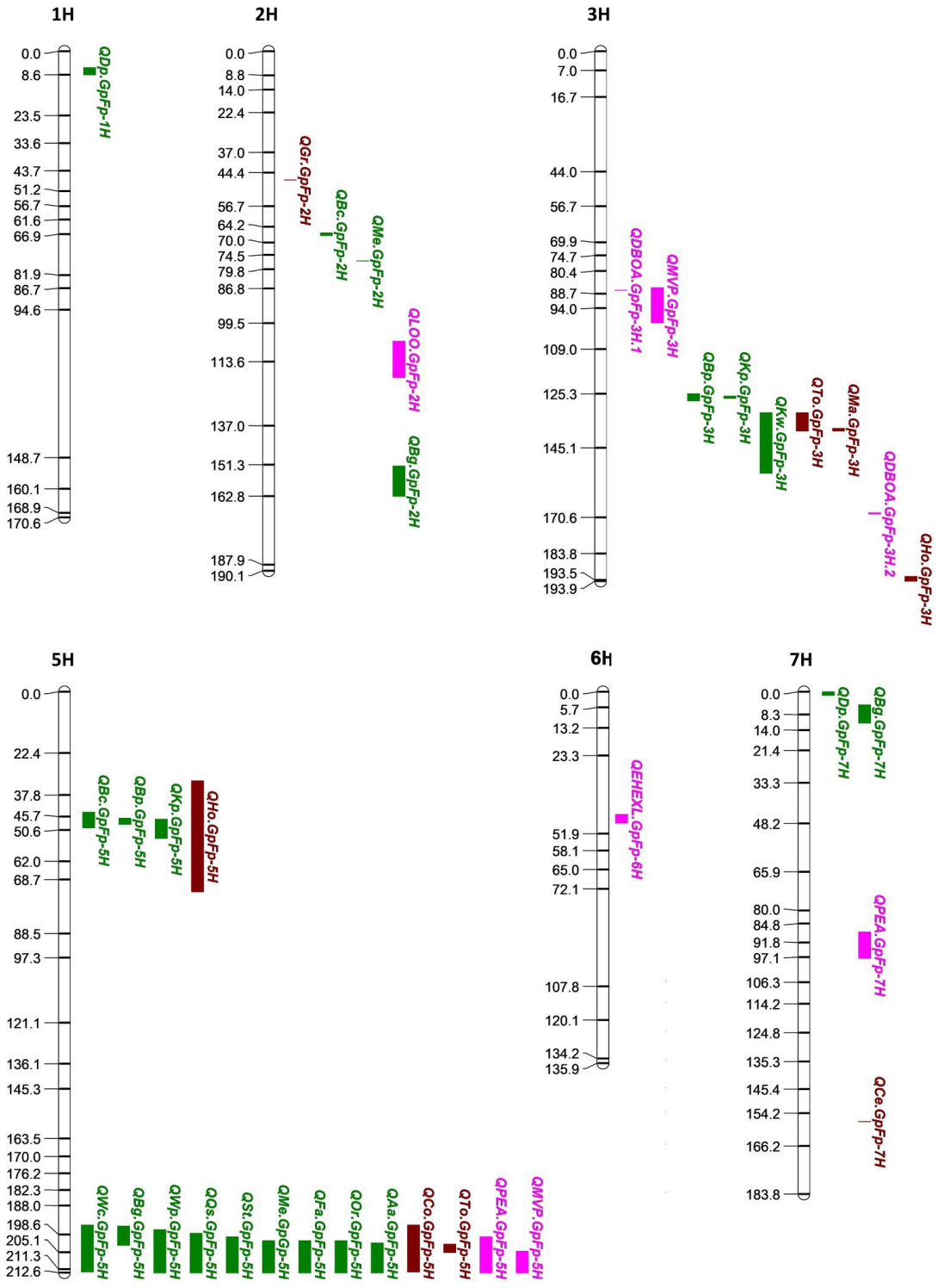

Figure 2.1 QTL regions shown by chromosome. cM positions are on the left, with only one every tenth position being shown. Malt quality QTLs are represented in green, sensory QTLs are represented in brown, and metabolite QTLs are shown in pink.

REFERENCES

Bamforth, C.W., 2009. Current perspectives on the role of enzymes in brewing. J. Cereal Sci. 50, 353–357.

Bayer, M.M., Rapazote-Flores, P., Ganal, M., Hedley, P.E., Macaulay, M., Plieske, J., Ramsay, L., Russell, J., Shaw, P.D., Thomas, W., Waugh, R., 2017. Development and evaluation of a barley 50k iSelect SNP array. Front. Plant Sci. 8, 1–10.

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. 57, 289–300.

Bettenhausen, H.M., Barr, L., Broeckling, C.D., Chaparro, J.M., Holbrook, C., Sedin, D., Heuberger, A.L., 2018. Influence of malt source on beer chemistry, flavor, and flavor stability. Food Res. Int. 113, 487–504.

Bettenhausen, H.M., Benson, A., Fisk, S., Herb, D., Hernandez, J., Lim, J., Queisser, S.H., Shellhammer, T.H., Vega, V., Yao, L., Heuberger, A.L., Hayes, P.M., 2020. Variation in Sensory Attributes and Volatile Compounds in Beers Brewed from Genetically Distinct Malts: An Integrated Sensory and Non-Targeted Metabolomics Approach. J. Am. Soc. Brew. Chem. 78, 136–152.

Briggs, D.E., 1978. Grain quality and germination. In: Barley. Springer Netherlands, Dordrecht, pp. 174–221.

Broeckling, C.D., Afsar, F.A., Neumann, S., Ben-Hur, A., Prenni, J.E., 2014. RAMClust: A Novel Feature Clustering Method Enables Spectral-Matching-Based Annotation for Metabolomics Data. Anal. Chem. 86, 6812–6817.

Burton, R.A., Zhang, X.Q., Hrmova, M., Fincher, G.B., 1999. A single limit dextrinase gene is expressed both in the developing endosperm and in germinated grains of Barley. Plant Physiol. 119, 859–871.

Carpena, M., Fraga-Corral, M., Otero, P., Nogueira, R.A., Garcia-Oliveira, P., Prieto, M.A., Simal-Gandara, J., 2021. Secondary aroma: Influence of wine microorganisms in their aroma profile. Foods 10, 1–26.

Castro, A.J., Benitez, A., Hayes, P.M., Viega, L., Wright, L., 2010. Coincident qualitative trait loci effects for dormancy, water sensitivity and malting quality traits in the BCD47Baronesse barley mapping population. Crop Pasture Sci. 61, 691–699.

Colmsee, C., Beier, S., Himmelbach, A., Schmutzer, T., Stein, N., Scholz, U., Mascher, M., 2015. BARLEX - the Barley Draft Genome Explorer. Mol. Pant 8, 964–966.

Cistué, L., Vallés, M.P., Echávarri, B., Sanz, J.M., Castillo, A., 2003. Barley anther culture. Doubled Haploid Prod. Crop Plants 29–34.

Clark, S.E., Hayes, P.M., Henson, C.A., 2003. Effects of single nucleotide polymorphisms in β-amylasel alleles from barley on functional properties of the enzymes. Plant Physiol. Biochem. 41, 798–804.

Coghe, S., Benoot, K., Delvaux, F., Varderhaegen, B., Delvaux, F.R., 2004. Ferulic Acid Release and 4-Vinylguaiacol Formation during Brewing and Fermentation: Indications for Feruloyl Esterase Activity in Saccharomyces cerevisiae. J. Agric. Food Chem. 52, 602–608.

Corbineau, F., Xia, Q., Bailly, C., El-Maarouf-Bouteau, H., 2014. Ethylene, a key factor in the regulation of seed dormancy. Front. Plant Sci. 5, 1–13.

Craine, E.B., Murphy, K.M., Bramwell, S., Ross, C.F., Fisk, S., 2021. Strategic malting barley improvement for craft brewers through consumer sensory evaluation of malt and beer. J. Food Sci. 1–17.

Edney, M.J., Mather, D.E., 2004. Quantitative trait loci affecting germination traits and malt friability in a two-rowed by six-rowed barley cross. J. Cereal Sci. 39, 283–290.

Erkkilä, M.J., Leah, R., Ahokas, H., Cameron-Mills, V., 1998. Allele-dependent barley grain β-amylase activity. Plant Physiol. 117, 679–685.

Evans, D.E., Li, C., Eglinton, J.K., 2010. The Properties and Genetics of Barley Malt Starch Degrading Enzymes. In: Zhang, G., Li, Chengdao (Eds.), Genetics and Improvement of Barley Malt Quality. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 143–189.

Fang, Y., Zhang, X., Xue, D., 2019. Genetic analysis and molecular breeding applications of malting quality QTLs in barley. Front. Genet. 10, 1–11.

Fox, G.P., 2010. Chemical Composition in Barley Grains and Malt Quality. In: Zhang, G., Li, C. (Eds.), Genetics and Improvement of Barley Malt Quality. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 63–98.

Gómez-López, P., Lázaro, N., Alvarado-Beltrán, C.G., Pineda, A., Balu, A.M., Luque, R., 2019. One-pot Cu/TiO2 nanoparticles synthesis for trans-ferulic acid conversion into vanillin. Molecules 24, 1–9.

Guo, B., Li, D., Lin, S., Li, Y., Wang, S., Lv, C., Xu, R., 2019. Regulation of nitrogen availability results in changes in grain protein content and grain storage subproteomes in barley (Hordeum vulgare L.). PLoS One 14, 1–13.

Han, F., Ullrich, S.E., Chirat, S., Menteur, S., Jestin, L., Sarrafi, A., Hayes, P.M., Jones, B.L., Blake, T.K., Wesenberg, D.M., Kleinhofs, A., Kilian, A., 1995. Mapping of β-glucan content and β-glucanase activity loci in barley grain and malt. Theor. Appl. Genet. 91, 921–927.

Han, F., Ullrich, S.E., Kleinhofs, A., Jones, B.L., Hayes, P.M., Wesenberg, D.M., 1997. Fine structure mapping of the barley chromosome-1 centromere region containing malting-quality QTLs. Theor. Appl. Genet. 95, 903–910.

Harakava, R., 2005. Genes encoding enzymes of the lignin biosynthesis pathway in Eucalyptus. Genet. Mol. Biol. 28, 601–607.

Herb, D., Filichkin, T., Fisk, S., Helgerson, L., Hayes, P., Meints, B., Jennings, R., Monsour, R., Tynan, S., Vinkemeier, K., Romagosa, I., Moscou, M., Carey, D., Thiel, R., Cistue, L., Martens, C., Thomas, W., 2017a. Effects of barley (Hordeum vulgare L.) variety and growing environment on beer flavor. J. Am. Soc. Brew. Chem. 75, 345–353.

Herb, D., Filichkin, T., Fisk, S., Helgerson, L., Hayes, P., Benson, A., Vega, V., Carey, D., Thiel, R., Cistue, L., Jennings, R., Monsour, R., Tynan, S., Vinkemeier, K., Li, Y., Nguygen, A., Onio, A., Meints, B., Moscou, M., Romagosa, I., Thomas, W., 2017b. Malt modification and its effects on the contributions of barley genotype to beer flavor. J. Am. Soc. Brew. Chem. 75, 354–362.

Hori, K., Sato, K., Takeda, K., 2007. Detection of seed dormancy QTL in multiple mapping populations derived from crosses involving novel barley germplasm. Theor. Appl. Genet. 115, 869–876.

Hummel, J., Strehmel, N., Bölling, C., Schmidt, S., Walther, D., Kopka, J., 2013. Mass Spectral Search and Analysis Using the Golm Metabolome Database. In: The Handbook of Plant Metabolomics. pp. 321–343.

Igartua, E., Edney, M., Rossnagel, B.G., Spaner, D., Legge, W.G., Scoles, G.J., Eckstein, P.E., Penner, G.A., Tinker, N.A., Briggs, K.G., Falk, D.E., Mather, D.E., 2000. Marker-based selection of QTL affecting grain and malt quality in two-row barley. Crop Sci. 40, 1426–1433.

Ishibashi, Y., Aoki, N., Kasa, S., Sakamoto, M., Kai, K., Tomokiyo, R., Watabe, G., Yuasa, T., Iwaya-Inoue, M., 2017. The interrelationship between abscisic acid and reactive oxygen species plays a key role in barley seed dormancy and germination. Front. Plant Sci. 8, 1–10.

Jia, Q.J., Zhang, J.J., Westcott, S., Zhang, X.Q., Bellgard, M., Lance, R., Li, C.D., 2009. GA-20 oxidase as a candidate for the semidwarf gene sdw1/denso in barley. Funct. Integr. Genomics 9, 255–262.

Knox, C.A.P., Sonthayanon, B., Chandra, G.R., Muthukrishnan, S., 1987. Structure and organization of two divergent α-amylase genes from barley. Plant Mol. Biol. 9, 3–17.

Kohl, S., Hollmann, J., Blattner, F.R., Radchuk, V., Andersch, F., Steuernagel, B., Schmutzer, T., Scholz, U., Krupinska, K., Weber, H., Weschke, W., 2012. A putative role for amino acid permeases in sink-source communication of barley tissues uncovered by RNA-seq. BMC Plant Biol. 12, 1–18.

Kreis, M., Williamson, M.S., Shewry, P.R., Sharp, P., Gale, M., 1988. Identification of a second locus encoding β-amylase on chromosome 2 of barley. Genet. Res. 51, 13–16.

Kuczyńska, A., Surma, M., Adamski, T., Mikołajczak, K., Krystkowiak, K., Ogrodowicz, P., 2013. Effects of the semi-dwarfing sdw1/denso gene in barley. J. Appl. Genet. 54, 381–390.

Kyraleou, M., Herb, D., Reilly, G.O., Conway, N., Bryan, T., Kilcawley, K.N., 2021. New Make Spirit. Foods 1–23.

Li, C., Ni, P., Francki, M., Hunter, A., Zhang, Y., Schibeci, D., Li, H., Tarr, A., Wang, J., Cakir, M., Yu, J., Bellgard, M., Lance, R., Appels, R., 2004. Genes controlling seed dormancy and pre-harvest sprouting in a rice-wheat-barley comparison. Funct. Integr. Genomics 4, 84–93.

Litts, J.C., Simmons, C.R., Karrer, E.E., Huang, N., Rodriguez, R.L., 1990. The isolation and characterization of a barley 1,3-1,4-β-glucanase gene. Eur. J. Biochem. 194, 831–838.

Liu, X., Zhang, H., Zhao, Y., Feng, Z., Li, Q., Yang, H.Q., Luan, S., Li, J., He, Z.H., 2013.

Auxin controls seed dormancy through stimulation of abscisic acid signaling by inducing ARF-mediated ABI3 activation in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. 110, 15485–15490.

Lo, S., Muñoz-Amatriaín, M., Boukar, O., Herniter, I., Cisse, N., Guo, Y.N., Roberts, P.A., Xu, S., Fatokun, C., Close, T.J., 2018. Identification of QTL controlling domestication-related traits in cowpea (Vigna unguiculata L. Walp). Sci. Rep. 8, 1–9.

Magliano, P.N., Prystupa, P., Herman Gutierrez Boem, F., 2014. Protein content of grains of different size fractions in malting barley. J. Inst. Brew. 120.

Mahalingam, R., Graham, D., Walling, J.G., 2021. The Barley (Hordeum vulgare ssp. vulgare) Respiratory Burst Oxidase Homolog (HvRBOH) Gene Family and Their Plausible Role on Malting Quality. Front. Plant Sci. 12, 1–16.

Mallet, J., 2014. Malt: A Practical Guide from Field to Brewhouse. Brewers Publications, Boulder, Colorado.

Mao, X., Zhang, J., Liu, W., Yan, S., Liu, Q., Fu, H., Zhao, J., Huang, W., Dong, J., Zhang, S., Yang, T., Yang, W., Liu, B., Wang, F., 2019. The MKKK62-MKK3-MAPK7/14 module negatively regulates seed dormancy in rice. Rice 12.

Mather, D.E., Tinker, N.A., LaBerge, D.E., Edney, M., Jones, B.L., Rossnagel, B.G., Legge, W.G., Briggs, K.G., Irvine, R.B., Falk, D.E., Kasha, K.J., 1997. Regions of the genome that affect grain and malt quality in a North American two-row Barley Cross. Crop Sci. 37, 544–554.

Mohammadi, M., Blake, T.K., Budde, A.D., Chao, S., Hayes, P.M., Horsley, R.D., Obert, D.E., Ullrich, S.E., Smith, K.P., 2015. A genome-wide association study of malting quality across eight U.S. barley breeding programs. Theor. Appl. Genet. 128, 705–721.

Monat, C., Padmarasu, S., Lux, T., Wicker, T., Gundlach, H., Himmelbach, A., Ens, J., Li, C., Muehlbauer, G.J., Schulman, A.H., Waugh, R., Braumann, I., Pozniak, C., Scholz, U., Mayer, K.F.X., Spannagl, M., Stein, N., Mascher, M., 2019. TRITEX: Chromosome-scale sequence assembly of Triticeae genomes with open-source tools. bioRxiv 1–18.

Morrissy, C., Bettenhausen, H.M., Simaeys, K.R. Van, Fisk, S., Hernandez, J., Mathias, K., Benson, A., Shellhammer, T.H., Hayes, P.M., 2021. Continued exploration of barley genotype contribution to base malt and beer flavor through the evaluation of lines sharing Maris Otter ® parentage. J. Am. Soc. Brew. Chem.

Muñoz-Amatriaín, M., Xiong, Y., Schmitt, M.R., Bilgic, H., Budde, A.D., Chao, S., Smith, K.P., Muehlbauer, G.J., 2010. Transcriptome analysis of a barley breeding program examines gene expression diversity and reveals target genes for malting quality improvement. BMC Genomics 11.

Muthukrishnan, S., Gill, B.S., Swegle, M., Chandra, G.R., 1984. Structural genes for alpha-amylases are located on barley chromosomes 1 and 6. J. Biol. Chem. 259, 13637–13639.

Nagel, M., Alqudah, A.M., Bailly, M., Rajjou, L., Pistrick, S., Matzig, G., Börner, A., Kranner, I., 2019. Novel loci and a role for nitric oxide for seed dormancy and preharvest sprouting in barley. Plant Cell Environ. 42, 1318–1327.

Nakamura, S., Pourkheirandish, M., Morishige, H., Kubo, Y., Nakamura, M., Ichimura, K., Seo, S., Kanamori, H., Wu, J., Ando, T., Hensel, G., Sameri, M., Stein, N., Sato, K., Matsumoto, T., Yano, M., Komatsuda, T., 2016. Mitogen-activated protein kinase kinase 3 regulates seed dormancy in barley. Curr. Biol. 26, 775–781.

Nakamura, S., Pourkheirandish, M., Morishige, H., Sameri, M., Sato, K., Komatsuda, T., 2017. Quantitative trait loci and maternal effects affecting the strong grain dormancy of wild barley (Hordeum vulgare ssp. spontaneum). Front. Plant Sci. 8, 1–11.

Oliver, G. 2012. The Oxford Companion to Beer. Oxford University Press. Oxford, United Kingdom.

Paynter, B., Young, K., 1996. Managing grain protein in malting barley.

Potokina, E., Caspers, M., Prasad, M., Kota, R., Zhang, H., Sreenivasulu, N., Wang, M., Graner, A., 2004. Functional association between malting quality trait components and cDNA array based expression patterns in barley (Hordeum vulgare L.). Mol. Breed. 14, 153–170.

Prada, D., Ullrich, S.E., Molina-Cano, J.L., Cistué, L., Clancy, J.A., Romagosa, I., 2004. Genetic control of dormancy in a Triumph/Morex cross in barley. Theor. Appl. Genet. 109, 62–70.

Sato, K., Yamane, M., Yamaji, N., Kanamori, H., Tagiri, A., Schwerdt, J.G., Fincher, G.B., Matsumoto, T., Takeda, K., Komatsuda, T., 2016. Alanine aminotransferase controls seed dormancy in barley. Nat. Commun. 7, 1–9.

Sharp, D.C., Steensels, J., Shellhammer, T.H., 2017. The effect of hopping regime, cultivar and β-glucosidase activity on monoterpene alcohol concentrations in wort and beer. J. Inst. Brew. 123, 185–191.

Shewry, P.R., Darlington, H., 2002. The proteins of the mature barley grain and their role in determining malting performance. In: Barley Science: Recent Advances from Molecular Biology to Agronomy of Yield and Quality. pp. 503–521.

Shorinola, O., Balcárková, B., Hyles, J., Tibbits, J.F.G., Hayden, M.J., Holušova, K., Valárik, M., Distelfeld, A., Torada, A., Barrero, J.M., Uauy, C., 2017. Haplotype analysis of the pre-harvest sprouting resistance locus Phs-A1 reveals a causal role of TAMKK3-A in global germplasm. Front. Plant Sci. 8, 1–14.

Sorensen, E.J., Sammis, G.M., 2004. A Dash of Proline Makes Things Sweet. Science (80-.). 305, 1725 LP – 1726.

Sweeney, D.W., Rooney, T.E., Walling, J.G., Sorrells, M.E., 2021. Pleiotropic effects of *HvMKK3* alleles on preharvest sprouting, seed dormancy, and malting quality. Crop Science. In press.

Szűcs, P., Blake, V.C., Bhat, P.R., Chao, S., Close, T.J., Cuesta-Marcos, A., Muehlbauer, G.J., Ramsay, L., Waugh, R., Hayes, P.M., 2009. An Integrated Resource for Barley Linkage Map and Malting Quality QTL Alignment. Plant Genome 2.

Tan, Y.F., Sun, M., Xing, Y.Z., Hua, J.P., Sun, X.L., Zhang, Q.F., Corke, H., 2001. Mapping quantitative trait loci for milling quality, protein content and color characteristics of rice using a recombinant inbred line population derived from an elite rice hybrid. Theor. Appl.

Genet. 103, 1037–1045.

Tautenhahn, R., Cho, K., Uritboonthai, W., Zhu, Z., Patti, G.J., Siuzdak, G., 2012. An accelerated workflow for untargeted metabolomics using the METLIN database. Nat. Biotechnol. 30, 826–828.

Tibbot, B.K., Skadsen, R.W., 1996. Molecular cloning and characterization of a gibberellin-inducible, putative alpha-glucosidase gene from barley. Plant Mol. Biol. 30, 229–241.

Vetch, J.M., Stougaard, R.N., Martin, J.M., Giroux, M., 2019. Allelic impacts of TaPHS1, TaMKK3, and Vp1B3 on preharvest sprouting of northern great plains winter wheats. Crop Sci. 59, 140–149.

Vetch, J.M., Walling, J.G., Sherman, J., Martin, J.M., Giroux, M.J., 2020. Mutations in the HvMKK3 and HvAlaAT1 genes affect barley preharvest sprouting and after-ripened seed dormancy. Crop Sci. 60, 1897–1906.

Wendt, T., Holme, I., Dockter, C., Preu, A., Thomas, W., Druka, A., Waugh, R., Hansson, M., Braumann, I., 2016. HvDep1 Is a Positive regulator of culm elongation and grain size in barley and impacts yield in an environment-dependent manner. PLoS One 11, 1–21.

Windes, S., Bettenhausen, H.M., Simaeys, K.R.V., Clawson, J., Fisk, S., Heuberger, A.L., Lim, J., Queisser, S.H., Shellhammer, T.H., Hayes, P.M., 2020. Comprehensive Analysis of Different Contemporary Barley Genotypes Enhances and Expands the Scope of Barley Contributions to Beer Flavor. J. Am. Soc. Brew. Chem. 0, 1–25.

Wishart, D.S., Jewison, T., Guo, A.C., Wilson, M., Knox, C., Liu, Y., Djoumbou, Y., Mandal, R., Aziat, F., Dong, E., Bouatra, S., Sinelnikov, I., Arndt, D., Xia, J., Liu, P., Yallou, F., Bjorndahl, T., Perez-Pineiro, R., Eisner, R., Allen, F., Neveu, V., Greiner, R., Scalbert, A., 2013. HMDB 3.0-The Human Metabolome Database in 2013. Nucleic Acids Res. 41, 801–807.

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

Xu, S., 2013. Mapping quantitative trait loci by controlling polygenic background effects. Genetics 195, 1209–1222.

Xu, Y., Jia, Q., Zhou, G., Zhang, X.Q., Angessa, T., Broughton, S., Yan, G., Zhang, W., Li, C., 2017. Characterization of the sdw1 semi-dwarf gene in barley. BMC Plant Biol. 17, 1–10.

Xue, D.W., Dong, J., Wu, F.B., Zhang, G.P., 2008. Genotypic and Environmental Variation in Grain Protein Components and their Relations to Beta-amylase and Beta-glucanase Activity in Malting Barley. Cereal Res. Commun. 36, 125–134.

Zhang, L., Liu, P., Wu, J., Qiao, L., Zhao, G., Jia, J., Gao, L., Wang, J., 2020. Identification of a novel ERF gene, TaERF8, associated with plant height and yield in wheat. BMC Plant Biol. 20, 1–12.

Zhu, Z.J., Schultz, A.W., Wang, J., Johnson, C.H., Yannone, S.M., Patti, G.J., Siuzdak, G., 2013. Liquid chromatography quadrupole time-ofcas-flight mass spectrometry characterization of metabolites guided by the METLIN database. Nat. Protoc. 8, 451–460.

CHAPTER 3 – NON-TARGETED METABOLOMICS OF COOKED COWPEAS (VIGNA UNGUICULATA) AND PIGEON PEA (CAJANUS CAJAN) FROM GHANA[2]

Overview

Legumes, a global staple food with human health properties, merit detailed composition analysis in cooked forms. This study analyzed cowpea [*Vigna unguiculata*] (three varieties: Dagbantuya, Sangyi, and Tukara), pigeon pea [*Cajanus cajan*], and common bean [*Phaseolus vulgaris*] using two distinct ultra-performance liquid chromatography mass spectrometry (UPLC-MS) workflows. Comparisons between cowpea and pigeon pea locally consumed in Ghana, and common bean, revealed 75 metabolites that differentiated cowpeas from the other legumes. Cowpea and pigeon pea metabolite fold-change comparisons revealed 142 metabolites with significantly higher abundance in cowpea than pigeon pea, and 154 with significantly higher abundance in pigeon pea than cowpea. There were 479 metabolites that remained similar between legume varieties. Legume-type specific markers were identified by cowpea variety, namely tonkinelin (Dagbantuya), pheophytin A (Sangyi), and linoleoyl ethanolamide (Tukara). Identification of novel exposure biomarkers for cowpea varieties that are distinct from other legumes merit attention for evaluation following consumption in people.

---

[2] This chapter is a modified form of a manuscript submitted to Food Chemistry: Molecular Sciences:

Brooke Sayre-Chavez, Bridget Baxter, Corey D. Broeckling, María Muñoz-Amatriaín, Mark Manary, and Elizabeth P. Ryan

Introduction

Cowpea (*Vigna unguiculata*), native to Africa, is a warm-season and nitrogen-fixing legume well adapted to the sandy soils and low-input farming practices of sub-Saharan Africa (Ji et al., 2019; Rawal, 1975; Vaillancourt and Weeden, 1992). Cowpea is a nutritious food due to its high protein content, along with its abundance of other various vitamins, trace minerals, antioxidants, amino acids, fibers, lipids, and phytochemicals which contribute to an assortment of health benefits (Abizari, Pilime, Armar-Klemesu, and Brouwer, 2013; Awika and Duodu, 2017; da Silvia et al., 2018; Jayathilake et al., 2018). In addition to containing many nutrients that can combat malnutrition, cowpeas are also known to lower cholesterol and blood pressure, reduce inflammation, and even help prevent diseases such as diabetes and cancer (Awika and Duodu, 2017; da Silvia et al., 2018; Jayathilake et al., 2018). Cowpea is also highly valued as a cash crop and used for animal feed, since livestock production makes up a large part of farmers income (Gómez, 2004). In Ghana, consumers pay a premium for certain cowpea varieties so despite the many known benefits of consuming cowpea, the incentive to sell the crop often outweighs the incentive to consume it (Langyintuo et al., 2003).

Pigeon pea (*Cajanus cajan*) is another important nitrogen-fixing, warm season legume grown in sub-Saharan Africa (Adjei-Nsiah, 2012). Like cowpea, it is drought tolerant and well adapted to low-input farming practices (Saxena, 2008). Pigeon pea is often used as a border crop or grown in an intercropping system to help improve soil fertility, which is a major factor impacting food security or lack thereof (Abunyewa and Karbo, 2005). Pigeon pea is also valued for human and animal consumption due to its high protein content (Abunyewa and Karbo, 2005; Adjei-Nsiah, 2012; Nwokolo, 1987; Pal, Mishra, Sachan, and Ghosh, 2011; Saxena, 2008). Although less researched than cowpea, pigeon pea is a nutritious food to help alleviate

malnutrition (Nwokolo, 1987; Pal et al., 2011). Based on anecdotal evidence of a local focus group in Ghana (described in Materials and Methods), pigeon peas are also part of their diet.

Dietary biomarkers exist for soybeans, green peas, chickpeas, lentils, and various dry beans (Borresen et al., 2017; Lu et al., 2010; Madrid-Gambin et al., 2017; Perera et al., 2015; Sri Harsha et al., 2018; Tsopmo and Muir, 2010; Zarei et al., 2021). For example, pipecolic acid and s-methylcysteine have been proposed as biomarkers for dry bean consumption including pinto, navy, kidney, lima, and black beans (Perera et al., 2015). There is a need for a more detailed look into the metabolomes of cowpea and pigeon pea to be better equipped for future nutritional studies in undernourished populations including those in Ghana. An understanding of the metabolomic profiles may also be useful in breeding for improved nutritional quality.

In this study, we analyzed the metabolite profiles of three different cowpea varieties commonly consumed in Ghana. We also included a pigeon pea that was also identified as consumed in combination with cowpeas in Ghana. A common bean of the Navy market class was used as control legume that has been characterized (Perera et al., 2015; Zarei et al., 2021). The main objective was to examine and compare the metabolomes of cowpeas and pigeon pea types consumed in Ghana using a non-targeted metabolomics approach.

Materials and Methods

Legume flours

Four varieties of local "cowpea" flours (Dagbantuya, Sangyi, Tukara, and Adua) were collected from a local market in Tamale (northern Ghana). These varieties were identified by a local community focus group in the region. All samples were made into pre-cooked flours for

metabolite analysis. Notably, seed morphology together with metabolomics data (see Results) revealed that one of the varieties (Adua) was a pigeon pea. The common bean flour (Navy market class) was collected purchased from ADM Edible Bean Specialties, Inc. (Archer Daniels Midland Company, Decatur, Illinois).

All legume flours used were cooked prior to extraction for metabolite profiling analysis. Flours were prepared by boiling the legumes for 45 min., draining them, and then drying them on a flat sheet in an oven at 40°C. The dried, cooked legumes were then ground to a fine powder with a mortar and pestle. Flours were stored in sealed conical tubes until the time of analyses.

Metabolomics Platform 1: CSU Analytical Resources Core – Bioanalysis and Omics laboratory (ARC-BIO) (Fort Collins, CO).

Sample preparation

For each legume sample, 50 (+/-1) mg of each cooked legume flour was weighed into a 2.0 mL eppendorf tube with 1.5 mL of absolute methanol. Three process blanks were prepared alongside the legume samples, where solvent was used to extract from empty tubes. Samples were vortex mixed and extracted with shaking for one hour at 4ºC. After centrifugation at 4ºC, 13,000 xg,1.0 mL of supernatant was collected and transferred to an autosampler vial. 100 uL of supernatant was collected from each sample to generate a pooled QC. Sample processing order was randomized.

Ultra performance liquid chromatography-time of flight mass spectroscopy (UPLC-TOF-MS)

Three microliters of legume flour sample extract were injected onto a Waters Acquity UPLC system in randomized order with a pooled quality control (QC) injection after every 5

samples. Separation was achieved using a Waters Acquity UPLC CSH Phenyl Hexyl column (1.7 μM, 1.0 x 100 mm), using a gradient from solvent A (Water, 2mM ammonium formate) to solvent B (Acetonitrile, 0.1% formic acid). Injections were made in 99% A, held at 99% A for 1 min, ramped to 98% B over 12 minutes, held at 98% B for 3 minutes, and then returned to starting conditions over 0.05 minutes and allowed to re-equilibrate for 3.95 minutes, with a 200 μL/min constant flow rate. The column and samples were held at 65 °C and 6 °C, respectively. The column eluent was infused into a Waters Xevo G2-XS Q-TOF-MS with an electrospray source in positive mode, scanning 50-1200 m/z at 0.1 seconds per scan, alternating between MS (6 V collision energy) and MSE mode (15-30 V ramp). Calibration was performed using sodium formate with 1 ppm mass accuracy. The capillary voltage was held at 700 V, source temperature at 140 °C, and nitrogen desolvation temperature at 600°C with a desolvation gas flow rate of 1000 L/hr.

Data normalization, filtration, and grouping

RAMClustR version 1.1.0 in R version 3.6.2 (2019-12-12)) was used to normalize, filter, and group features into spectra from XCMS output data (Smith, Want, O'maille, Abagyan, and Siuzdak, 2006; Tautenhahn, Bottcher, and Neumann, 2008). Features which failed to demonstrate signal intensity of at least 3-fold greater in QC samples than in blanks were removed from the feature dataset. 18561 of 52141 features were removed. Features with missing values were replaced with small values to simulate noise and then the minimum detected or simulated value was multiplied by 0.1. The filled value was the absolute value of this value. Features were normalized by linearly regressing run order versus qc feature intensities to account for instrument signal intensity drift. Only features with a regression *p*-value less than 0.05 and an r-squared greater than 0.1 were corrected. Features were filtered based on their qc sample CV

values. Only features with CV values less than or equal to 0.3 in MS or MSMSdata sets were retained. 22091 of 33580 features were removed. Features were additionally normalized to total extracted ion signal to account for differences in total solute concentration. Features were clustered using the ramclustR algorithm. Parameter settings were as follows: st = 2.22, sr = 0.7, maxt = 222, deepSplit = FALSE, hmax = 0.3, minModuleSize = 2, and cor.method = pearson. Charge state detection was performed using the assign.z function using parameters: chargestate = 3, mzError = 0.005, nEvents = 2, minPercentSignal = 10, and assume1 = TRUE. Molecular weight was inferred from in-source spectra (Broeckling et al., 2016) using the do.findmain function, which calls the interpretMSSpectrum package (Jaeger, Hoffman, Schmitt, and Lisec, 2016). Parameters for do.findmain were set to: mode = positive, mzabs.error = 0.002, ppm.error = 10, ads = default, scoring = auto, and use.z = TRUE.

MSFinder (Tsugawa et al., 2016) was used for spectral matching, formula inference, and tentative structure assignment, and results were imported into the RAMClustR object. Annotations were assigned using the RAMClustR annotate function. Annotation priority was assigned from highest priority to lowest: MSFinder structure, MSFinder formula, interpretMSSpectrum M. Database priority was set to HMDB, PubChem, UNPD,ChEBI, PlantCyc, KNApSAcK, FooDB, DrugBank, LipidMAPS, and Urine. Compounds were assigned to chemical ontogenies using the ClassyFire API (Djoumbou, 2016).

MetaboAnalyst and Statistical Analyses

The normalized spectral abundance data was grouped by legume type and comparison grouping prior to input into MetaboAnalyst Version 5.0 (https://www.metaboanalyst.ca/), where the following statistical functions were performed; one-way ANOVA, dendrogram, heatmap, and fold-change analysis. Data was not additionally filtered, normalized, or transformed. Relative

abundance data was auto-scaled for visualization using the formula: $\bar{x}_{ij} = (x_{ij} - \bar{x}_{ij})/S_i$. ANOVA *p*-value cutoff was set to 0.05 and Fisher's LSD post-hoc analysis was used. Dendrograms used Euclidean distances and Ward clustering. Correlation heatmaps mapped the features with Pearson r distances. Fold change threshold was set to 2.

Metabolomics Platform 2: Metabolon Inc. (Durham, NC)

The samples of cowpea and pigeon pea flours were also sent to Metabolon, Inc. (Durham, NC, USA) for a more comprehensive varietal analysis of the cowpea flours.

Sample preparation

Samples were inventoried and accessioned into the Metabolon LIMS system where they were assigned a unique identifier and then stored at -80°C until processed. Samples were prepared using the automated MicroLab STAR® system from Hamilton Company. Several recovery standards were added prior to the first step in the extraction process for QC purposes. To remove protein, dissociate small molecules bound to protein or trapped in the precipitated protein matrix, and to recover chemically diverse metabolites, proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000) followed by centrifugation. The resulting extract was divided into four fractions: two for analysis by two separate reverse phase (RP)/UPLC-MS/MS methods with positive ion mode electrospray ionization (ESI), one for analysis by RP/UPLC-MS/MS with negative ion mode ESI, one for analysis by HILIC/UPLC-MS/MS with negative ion mode ESI. Samples were placed briefly on a TurboVap® (Zymark) to remove the organic solvent. The sample extracts were stored overnight under nitrogen before preparation for analysis.

Ultrahigh performance liquid chromatography-tandem mass spectroscopy (UPLC-MS/MS)

All methods utilized a Waters ACQUITY ultra-performance liquid chromatography (UPLC) and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer operated at 35,000 mass resolution. The sample extract was dried then reconstituted in solvents compatible to each of the four methods. Each reconstitution solvent contained a series of standards at fixed concentrations to ensure injection and chromatographic consistency. One aliquot was analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophilic compounds. In this method, the extract was gradient eluted from a C18 column (Waters UPLC BEH C18-2.1x100 mm, 1.7 µm) using water and methanol, containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA). Another aliquot was also analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophobic compounds. In this method, the extract was gradient eluted from the same afore mentioned C18 column using methanol, acetonitrile, water, 0.05% PFPA and 0.01% FA and was operated at an overall higher organic content. Another aliquot was analyzed using basic negative ion optimized conditions using a separate dedicated C18 column. The basic extracts were gradient eluted from the column using methanol and water, however with 6.5mM Ammonium Bicarbonate at pH 8. The fourth aliquot was analyzed via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1x150 mm, 1.7 µm) using a gradient consisting of water and acetonitrile with 10mM Ammonium Formate, pH 10.8. The MS analysis alternated between MS and data-dependent MSn scans using dynamic exclusion. The scan range varied slighted between methods but covered 70-1000 m/z.

Data extraction and compound identification

Raw data was extracted, peak-identified and QC processed using Metabolon's hardware and software. These systems are built on a web-service platform utilizing Microsoft's .NET technologies, which run on high-performance application servers and fiber-channel storage arrays in clusters to provide active failover and load-balancing. Compounds were identified by comparison to library entries of purified standards or recurrent unknown entities. Metabolon maintains a library based on authenticated standards that contains the retention time/index (RI), mass to charge ratio (m/z), and chromatographic data (including MS/MS spectral data) on all molecules present in the library, which include more than 3300 commercially available purified standards. Furthermore, biochemical identifications are based on three criteria: retention index within a narrow RI window of the proposed identification, accurate mass match to the library +/- 10 ppm, and the MS/MS forward and reverse scores between the experimental data and authentic standards. Standard statistical analyses are performed in ArrayStudio on log transformed data.

Results

Non-targeted metabolomics of legume flours

The CSU ARC-BIO metabolomics (Platform 1) yielded 775 metabolites across diverse classes. By superclass, there were 410 lipids and lipid-like molecules, 260 unclassified metabolites, 26 organic acids and derivatives, 18 organoheterocyclic compounds, 17 phenylpropanoids and polyketides, 14 organic oxygen compounds, 8 benzenoids, 8 hydrocarbons, 8 organic nitrogen compounds, 2 alkaloids and derivatives, 2 lignans/neolignans/related compounds, 1 nucleoside/nucleotide/analogue, and 1 organosulfur

compound in the data (data available at: www.ebi.ac.uk/metabolights/MTBLS3619) (Haug et al., 2020).

The Metabolon analysis (Platform 2) yielded 441 metabolites in the cowpea and pigeon pea. There were 400 metabolites with known identifications categorized by super pathways. This included 134 lipids, 130 amino acids, 43 carbohydrates, 40 nucleotides, 24 cofactors/prosthetic groups/electron carriers, 21 secondary metabolites, 5 peptides, 2 xenobiotics, and 1 hormone (Supplementary Table 3.1).

Metabolite profile comparison across legume types

Clear differences were observed between the legumes. Figure 3.1 shows all legumes clustered by type. Principal component (PC) 1 explained 52.7% of the variation and mainly differentiated the pigeon pea (i.e., Adua) samples from cowpea and common bean, while PC2 explained 22.3% of the variation that separated the common bean samples. A dendrogram constructed with metabolite abundance profiles showed similar relationships between the three legumes (Figure 3.1). One-way ANOVA was applied to reveal 551 metabolites with significant differences in relative abundances between legume types. The most significant metabolites were TG(48:7), PG(44:2), UNPD93557, C11H23N2O24PS10, and pipecolic acid, which all have higher abundance in common bean compared to cowpea and pigeon pea (Supplementary Table 3.2). There were 75 metabolites that differentiated cowpea, 121 that differentiated pigeon pea, and 185 that differentiated common bean after post-hoc analysis (Figure 3.1).

The 75 metabolites that differentiated cowpeas from the other legumes include 45 lipids and lipid-like molecules, 22 unclassified metabolites, 3 organoheterocyclic compounds, 2 organic oxygen compounds, 2 phenylpropanoids and polyketides, and 1 organic acid or

derivative (Table 3.1, Supplementary Table 3.2). Most of the metabolites that differentiated cowpeas from the other legumes, were significantly higher in cowpeas than the other two legumes. Some of the most significant ones include TG(50:3), C41H102N10O3S2, TG(52:3), 1-[(9Z)-octadecenyl]-3-[(9Z)-octadecenoyl]-sn-glycerol, and TG(56:5). Apart from C41H102N10O3S2, the listed metabolites belong to the glycerolipid class. Based on comparisons of the metabolites that differentiate cowpea in the legume analysis and the cowpea varietal analysis (described in Results), 3-(all-trans-nonaprenyl)benzene-1,2-diol, N-tetracosanoylphytosphingosine, and sitoindoside II warrant further investigation as metabolites indicative of cowpea consumption. These metabolites were all higher in cowpea than the other two legumes.

There were 121 metabolites that differentiated pigeon pea from the other legumes. This list included 63 lipid and lipid-like molecules, 46 unclassified metabolites, 6 phenylpropanoids and polyketides, 2 hydrocarbons, 1 benzenoid, 1 organic acid or derivative, 1 organic oxygen compound, and 1 organoheterocyclic compound (Table 3.1, Supplementary Table 3.2). Most of the differentiating metabolites were significantly higher in pigeon pea than the other legumes and some of the most significant ones include proline betaine, PE(38:6), C14H9NS12, 22:0-Glc-Sitosterol, and DG(36:5). These are from the carboxylic acid, glycerophospholipid, unclassified, steroid, and fatty acyl classes, respectively.

Fold change comparisons between legume types

Fold change analysis of the metabolites identified on Platform 1 revealed notable differences between cowpea and common bean. 173 metabolites were significantly higher in abundance for cowpea than in common bean, 183 were significantly higher in common bean than in cowpea, and 419 did not show a significant log2(fold change) (Figure 3.2). The most

extreme log2(fold change) values observed in either direction were -17.235 and 16.594 for pipecolic acid and C33H79N13S, respectively (Supplementary Table 3.3). Pipecolic acid was one of the most significant metabolites found by ANOVA analysis (see Results). Post-hoc and fold change analyses both confirm that it is higher in abundance in common bean than the other two legumes.

Comparing pigeon pea and common bean metabolite abundance by fold changes on Platform 1, revealed 233 significantly higher metabolites in pigeon pea than in common bean, 216 were significantly higher in common bean than in pigeon pea, and 326 did not show a significant fold change (Figure 3.2). The most extreme log2(fold change) values observed in either direction were -20.07 and 15.416 for soyasaponin V and piptamine, respectively (Supplementary Table 3.3). Soyasaponin V was significantly higher in common bean and did not have a significant difference in means between cowpea and pigeon pea. Piptamine was significantly higher in pigeon pea than both cowpea and common bean and may warrant further investigation as a metabolite unique to pigeon pea.

Cowpea and pigeon pea fold change comparisons from Platform 1 revealed that 142 metabolites were significantly higher in cowpea than in pigeon pea, 154 were significantly higher in pigeon pea than in cowpea, and 479 did not show a significant fold change (Figure 3.2). The most extreme log2(fold change) values observed were -12.909 and 16.783 for pipecolic acid and TG(59:6) [iso6], respectively (Supplementary Table 3.3). Pipecolic acid was one of the most significant metabolites found by ANOVA analysis (see Results), being higher in common bean than the other two legumes. Post-hoc and fold change analysis also establish a difference between cowpea and pigeon pea.

Comparison of three cowpea varieties from Ghana

In comparing the metabolic profiles of three cowpea varieties (Dagbantuya, Sangyi, and Tukara), we observed noticeable differences according to principal component analysis and the metabolite-based dendrogram showing similar results (Figure 3.3). Both plots also indicate that Dagbantuya and Sangyi are more similar to each other than they are to Tukara.

The one-way ANOVA testing supported 320 metabolites with significant differences in means, out of the 775 detected from Platform 1. Of those, 101 metabolites differentiated Dagbantuya, 27 differentiated Sangyi, and 82 metabolites differentiated Tukara after post-hoc analysis (Figure 3.3). Cowpea varietal comparisons bring an intentional highlight on the following compounds C18H34N6O6S16, C45H106N12O3S2, C14H47N4O32PS12, 1,2-Di-(9Z,12Z,15Z-octadecatrienoyl)-3-(Galactosyl-alpha-1-6-Galactosyl-beta-1)-glycerol, and uvarigrin;(+)-uvarigrin based on ANOVA *p*-values. Apart from 1,2-Di-(9Z,12Z,15Z-octadecatrienoyl)-3-(Galactosyl-alpha-1-6-Galactosyl-beta-1)-glycerol, these were all higher in Dagbantuya than the other varieties (Supplementary Table 3.2).

The 101 metabolites that differentiated Dagbantuya from the other cowpeas on Platform 1 include 56 lipids and lipid-like molecules, 40 unclassified metabolites, 3 organic oxygen compounds, 1 organic acid or derivative, and 1 phenylpropanoid and polyketide (Table 3.2). Most of the metabolites that differentiated Dagbantuya from the other varieties were higher in abundance in Dagbantuya than the other varieties. The 27 metabolites that differentiated Sangyi from the other cowpeas on Platform 1 include 18 lipids and lipid-like molecules, 4 unclassified metabolites, 2 benzenoids, 1 nucleoside/nucleotide/analogue, 1 organic acid or derivative, and 1 phenylpropanoid and polyketide (Table 3.2). Most of the metabolites differentiating Sangyi from the other varieties were lower in abundance in Sangyi than the other varieties. The 82

metabolites that differentiated Tukara from the other cowpeas on Platform 1 include 48 lipids and lipid-like molecules, 25 unclassified metabolites, 4 organic oxygen compounds, 3 organic acids and derivatives, and 2 organoheterocyclic compounds (Table 3.2). Most of the metabolites differentiating Tukara from the other varieties were higher in abundance in Tukara than the other varieties. Details can be found in Supplementary Table 3.2.

Cowpea and pigeon pea metabolites

We found notable differences in the types of compounds detected and identified from Platform 1 and 2. The results from Platform 2 are shown as a supplement to the metabolite lists that differentiate cowpea varieties from pigeon pea. There were 49 metabolites common to all cowpea varieties, and 337 common to both cowpea and pigeon pea on Platform 2. By cowpea variety, Platform 2 did not detect metabolites unique to Dagbantuya or Tukara, but Pheophytin A was uniquely detected for Sangyi. There were eight metabolites detected uniquely to pigeon pea (Adua), although none of these were detected on Platform 1.

Both Platform 1 and 2 were used to create a list of potentially unique metabolites. Both platforms have a unique set of benefits and complement each other's data sets, since no one platform can detect every metabolite present. Although the number of detected metabolites and annotations differ, we took into account that the platforms are different so the sensitivity and annotation software will differ as well.

Platform 1 consistently detected pipecolic acid (p-value = 1.22E-16) as differentiating between legume types, whether by ANOVA or fold change analysis. Although it has a higher abundance in common bean, pipecolic acid may still be a common metabolite for all legume types analyzed in this study. Relative abundance is visualized in Figure 3.4.

Since Platform 1 compared cowpea to both pigeon pea and common bean, the focus was on metabolites from this data. Based on comparisons of the metabolites that differentiated cowpea from the other two legumes in the legume analysis (described in Results), along with the commonality in the cowpea varietal analysis (described in Results), 3-(all-trans-nonaprenyl)benzene-1,2-diol (p-value = 0.0057), N-tetracosanoylphytosphingosine (p-value = 0.0002), and sitoindoside II (p-value = 0.0039) may be identifiers of cowpea consumption. These metabolites were all higher in cowpea than the other two legumes (Figure 3.4).

Platform 2 did not detect any metabolites unique to Dagbantuya so the focus was turned to Platform 1. Based on the ANOVA significance, post-hoc differentiation, and presence in the literature, tonkinelin (p-value = 3.25E-05) is proposed as a metabolite that may be unique to the cowpea variety Dagbantuya.

Platform 2 reported one metabolite unique to the cowpea variety Sangyi, named pheophytin A. Due to its detection in one variety only and because Platform 2 has a higher level of annotation confidence, pheophytin A warrants further investigation as a metabolite identifying of the cowpea variety Sangyi.

Both Platform 1 and 2 detected the compound linoleoyl ethanolamide. Platform 1 showed that linoleoyl ethanolamide had an abundance significantly lower in Tukara than in Dagbantuya and Sangyi (p-value = 0.0045), and Platform 2 only detected it in Dagbantuya and Sangyi. Because important metabolites can also be low in abundance, or absent, linoleoyl ethanolamide warrants further investigation as a metabolite indicative of consumption for Tukara.

Platform 1 detected piptamine (p-value = 0.0031) as significant and differentiating in pigeon pea (Figure 3.4). Although Platform 2 detected eight compounds unique to pigeon pea,

based on the ANOVA significance, post-hoc differentiation, and fold change of piptamine on Platform 1, this warrants further investigation as a metabolite that may be indicative of pigeon pea consumption.

Discussion

Legumes contain a diverse and beneficial range of chemical compounds with many established health and nutritional benefits. A non-targeted approach was utilized for investigating the metabolic profiles of different legume types in cooked forms consumed by people, and we applied an intentional focus to compare local cowpea varieties commonly consumed by households in Ghana. Two separate metabolite detection and analysis platform workflows were applied that provided novel compound lists for cooked cowpea and pigeon pea. This analysis deciphered metabolite relationships between legumes and between cowpea varieties to support that the cowpea metabolite profile is more similar to that of common bean than to the pigeon pea. Notably, these findings mirror the phylogenetic relationships (Ji et al., 2019). Our results support that cowpea Dagbantuya and Sangyi metabolite profiles are more similar to each other than to Tukara.

Metabolomics is an incredibly informative tool for food composition profiling and aids plant breeding as well as nutrition. The identification of food biomarkers may allow for exposure assessments from consumption but can also be used to assess compliance in feeding intervention studies without the bias of self-reported data (Hedrick et al., 2012). Knowing which metabolites are present in staple foods such as legumes and what benefits or absorption patterns that unique food metabolites may have, can also guide crop breeding efforts aimed at either increasing,

maintaining, or reducing the amounts of certain metabolites. Metabolomics-assisted breeding would also allow for screening to select desired phenotypes early in the breeding process (Fernie and Schauer, 2009). As presented in the results, cowpeas and pigeon pea contain metabolites with many established health benefits that could be targeted in breeding programs within West Africa and across geographically distinct regions.

Pipecolic acid was identified as a common metabolite for all legumes in general despite differences to the relative abundances. This compound has been previously postulated as a biomarker for dry bean (including pinto, navy, kidney, lima, and black beans) consumption (Perera et al., 2015). Pipecolic acid is a conjugate acid of pipecolate, and may also be a candidate dietary biomarker of navy bean (Zarei et al., 2021). Pipecolic acid helps regulate immunity in both plants and humans and is an important precursor to secondary metabolites with antitumor, antibiotic, anthelmintic, and anti-inflammatory properties (Natarajan, Muthukrishnan, Khalimonchuk, Mott, and Becker, 2017; Wang et al., 2018). Although pipecolic acid had a higher abundance in common bean than cowpea and pigeon pea, additional attention is needed to establish differences in bioavailability from feeding studies.

Currently, food metabolite profile analysis of cowpea varieties is limited even though there is promising agronomic traits, nutritional value, and preferential consumption by local communities (Abizari et al., 2013; Gómez, 2004; Jayathilake et al., 2018). We highlighted three cowpea metabolites; 3-(all-trans-nonaprenyl)benzene-1,2-diol, N-tetracosanoylphytosphingosine, and sitoindoside II that are novel to this study, and for which information of health benefits exists (Bentinger, Tekle, and Dallner, 2010; Dahlén and Pascher, 1972; Poon et al., 1999; Satmbekova et al., 2018). The 3-(all-trans-nonaprenyl)benzene-1,2-diol is a prenol lipid that plays a role in *E. coli* for Coenzyme Q biosynthesis (Poon et al., 1999). Coenzyme Q has well

established anti-inflammatory properties (Bentinger et al., 2010). N-tetracosanoylphytosphingosine is a sphingolipid, which exhibits immunological activity (Dahlén and Pascher, 1972). Sitoindoside II is a steroid/steroid derivative that is found in the plant *Cichorium intybus L.*, often used in traditional medicine for its diuretic, anti-inflammatory, cardiotonic, liver tonic, and digestive benefits (Satmbekova et al., 2018).

Other cowpea metabolites of varietal distinction include tonkinelin, pheophytin A, and linoleoyl ethanolamide, for Dagbantuya, Sangyi, and Tukara, respectively. These components have not been previously reported from cowpea, but information is known from other systems. Tonkinelin is a fatty acyl that has been identified in *Uvaria tonkinensis* and has established acetogenic effects (Chen and Yu, 1996). Pheophytin A is involved in chlorophyll metabolism and contributes to dark pigment colors (Yilmaz and Gökmen, 2015). Sangyi has the darkest pigmentation of the cowpea varieties analyzed. Linoleoyl ethanolamide is a carboximidic acid/derivative that has anti-inflammatory effects (Ishida et al., 2013).

For pigeon pea, the benezenoid metabolite piptamine was found, and this is a known antibiotic, first isolated from *Piptoporus betulinus* (Schlegel, Luhmann, Hartl, and Grafe, 2000). The impacts for this compound from food remains unclear and merits follow up attention for impact by post-harvest and processing conditions.

Legumes and particularly cowpeas can help alleviate malnutrition and support healthy growth of children in low-income countries where cowpeas are prevalent and well adapted for local climates. Future analysis of studies completed with these cooked legume flours in children and pregnant women with increasing doses over a period of 20-day period will guide more specific candidate dietary marker identifications and assessments in blood and urine.

Study limitations for this non-targeted metabolic profiling is that metabolite identification and annotation can vary slightly across platforms due to differences in instrument sensitivity and software. The majority of metabolites identified using both platforms, were annotated and classified though RAMClustR, MSFinder and in-house libraries, supporting the premise that computational annotation tools can provide valuable insight in the absence of spectral libraries. Further, quantifying metabolite levels using internal standards would give useful information on the absolute quantities available during consumption. Comparing data sets from this study with additional cowpea varieties is warranted alongside the blood and urine from individuals following consumption for an integrated identification of nutritional biomarkers.

Conclusions

This study analyzed the metabolic profiles of the cooked flours of three different cowpea varieties commonly consumed in Ghana, along with a pigeon pea, and a control legume (common bean) on two distinct metabolite profiling platforms. Relationships between cowpea, pigeon pea, and common bean, based on metabolic profiles helped to establish differences between the legumes that mirrored genetic relationships. Metabolic profiles of the cowpea varieties also presented differences between them, with Dagbantuya and Sangyi being more similar to each other than Tukara. This study supports identification of novel metabolites associated with each legume species and variety, including Dagbantuya (tonkinelin), Sangyi (pheophytin A), Tukara (linoleoyl ethanolamide), and pigeon pea (piptamine). Future work to quantify and validate the metabolites that differentiated the respective legumes is needed alongside an integrated biomarker analysis from feeding studies and also in breeding strategies geared towards improving nutrition and food function from cowpea and pigeon pea intake.

Table 3.1 Classification of metabolites that differentiate between legume types

		Number of differentiating metabolites		
Superclass	Class	Cowpea	Pigeon Pea	Common Bean
Benzenoids	Benzene and substituted derivatives	0	1	2
Hydrocarbons	Unsaturated hydrocarbons	0	2	0
Lipids and lipid-like molecules	Fatty Acyls	2	7	14
	Glycerolipids	12	19	15
	Glycerophospholipids	24	29	36
	Prenol lipids	2	2	15
	Saccharolipids	0	0	2
	Sphingolipids	1	3	2
	Steroids and steroid derivatives	4	3	12
Organic acids and derivatives	Carboximidic acids and derivatives	0	0	7
	Carboxylic acids and derivatives	0	1	0
	Organic phosphoric acids and derivatives	0	0	1
	Peptidomimetics	1	0	0
Organic nitrogen compounds	Organonitrogen compounds	0	0	3
Organic oxygen compounds	Organooxygen compounds	2	1	4
Organoheterocyclic compounds	Azoles	0	0	1
	Benzopyrans	1	0	0
	Indolizidines	1	0	1
	Lactones	0	0	1
	Pyrrolidines	0	1	0
	Quinolizines	0	0	1
	Tetrapyrroles and derivatives	1	0	0
Phenylpropanoids and polyketides	Cinnamic acids and derivatives	1	5	0
	Linear 1,3-diarylpropanoids	1	0	0
	Macrolactams	0	0	1
	Macrolides and analogues	0	0	1
	Tannins	0	1	0
NA	NA	22	46	66
Total		75	121	185

Table 3.2 Classification of metabolites that differentiate between cowpea varieties

		Number of differentiating metabolites		
Superclass	Class	Dagbantuya	Sangyi	Tukara
Benzenoids	Benzene and substituted derivatives	0	2	0
Lipids and lipid-like molecules	Fatty Acyls	11	3	6
	Glycerolipids	13	6	13
	Glycerophospholipids	16	5	23
	Prenol lipids	4	1	3
	Sphingolipids	2	0	1
	Steroids and steroid derivatives	10	3	2
Nucleosides, nucleotides, and analogues	Purine nucleosides	0	1	0
Organic acids and derivatives	Carboximidic acids and derivatives	0	0	1
	Carboxylic acids and derivatives	1	1	1
	Organic phosphoric acids and derivatives	0	0	1
Organic oxygen compounds	Organooxygen compounds	3	0	4
Organoheterocyclic compounds	Quinolines and derivatives	0	0	1
	Tetrapyrroles and derivatives	0	0	1
Phenylpropanoids and polyketides	Cinnamic acids and derivatives	1	0	0
	Linear 1,3-diarylpropanoids	0	1	0
NA	NA	40	4	25
Total		101	27	82

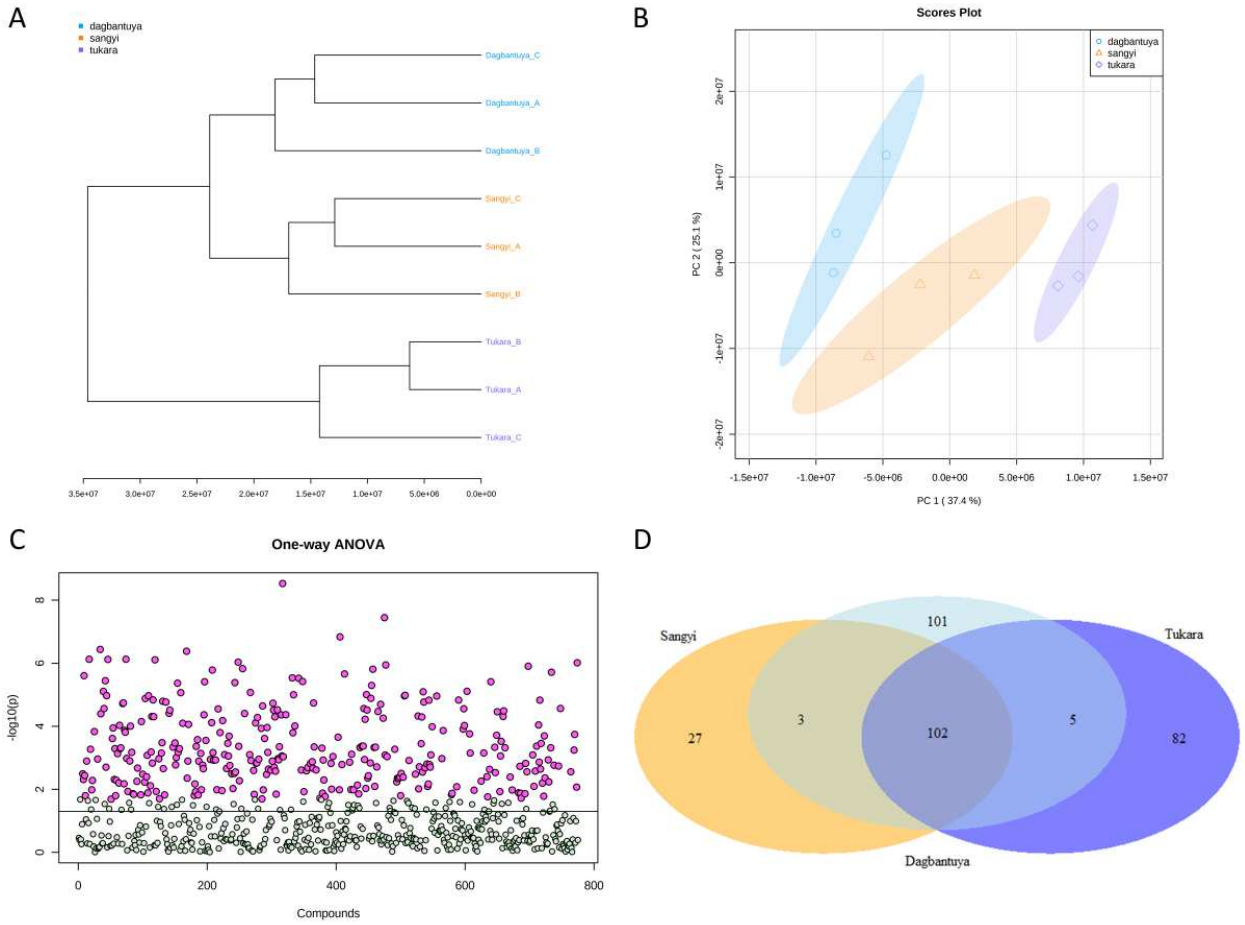

Figure 3.1 (A) Hierarchal cluster dendrogram of legume sample data based on Euclidean distance and Ward clustering. (B) PCA scores plot of legume sample data. (C) one-way ANOVA plotting $-\log_{10}$ (*p*-values) of all detected metabolites. Metabolites with significant differences in means across legume type are plotted in red (n=551), metabolites with no significant difference in mean across legume types are plotted in green (n=224). (D) Venn diagram indicating the number of metabolites that differentiate legume type or types based on Fisher's LSD post-hoc analysis.

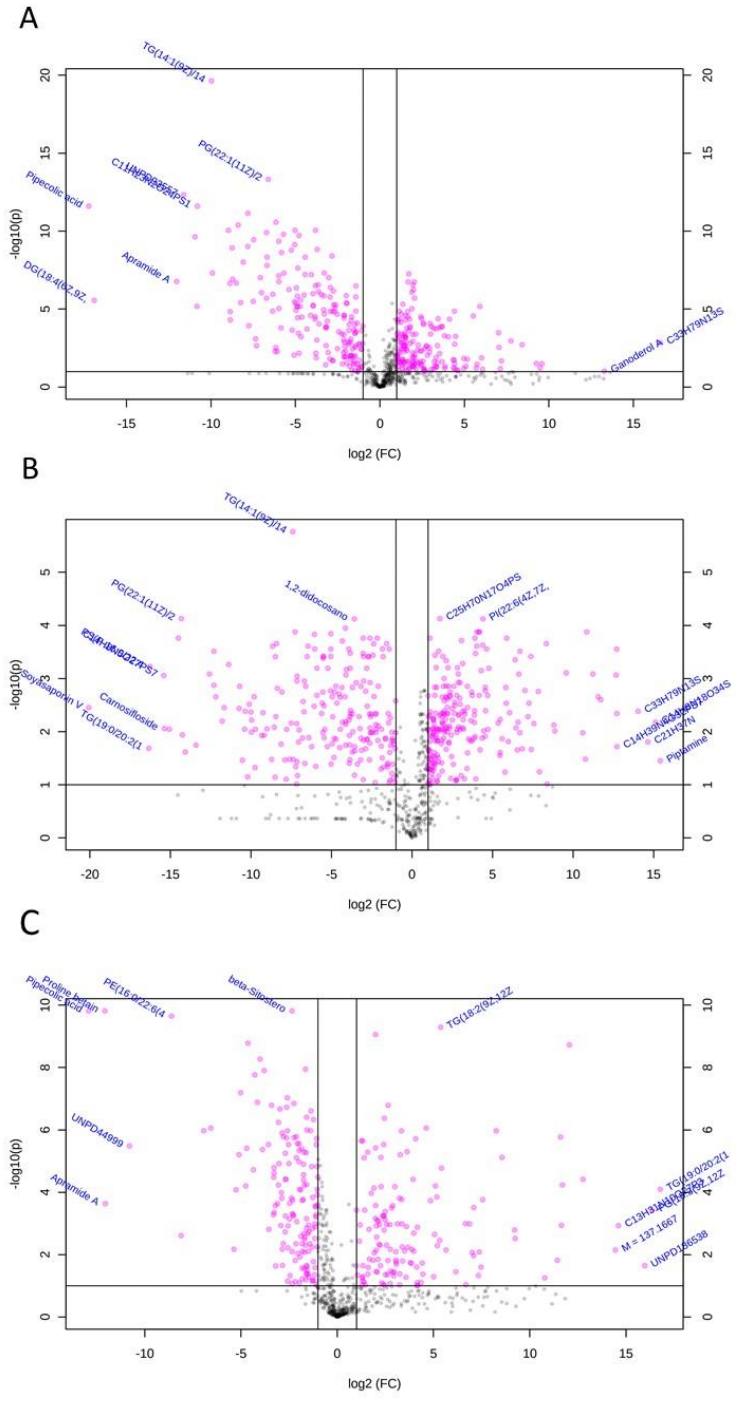

Figure 3.2 Volcano plots illustrating compound differences between two legume types. (A) cowpea versus common bean fold change volcano plot. (B) pigeon pea versus common bean fold change volcano plot. (C) cowpea versus pigeon pea fold change volcano plot. First legume listed indicates right side of the plot, second legume listed indicates left side of the plot.

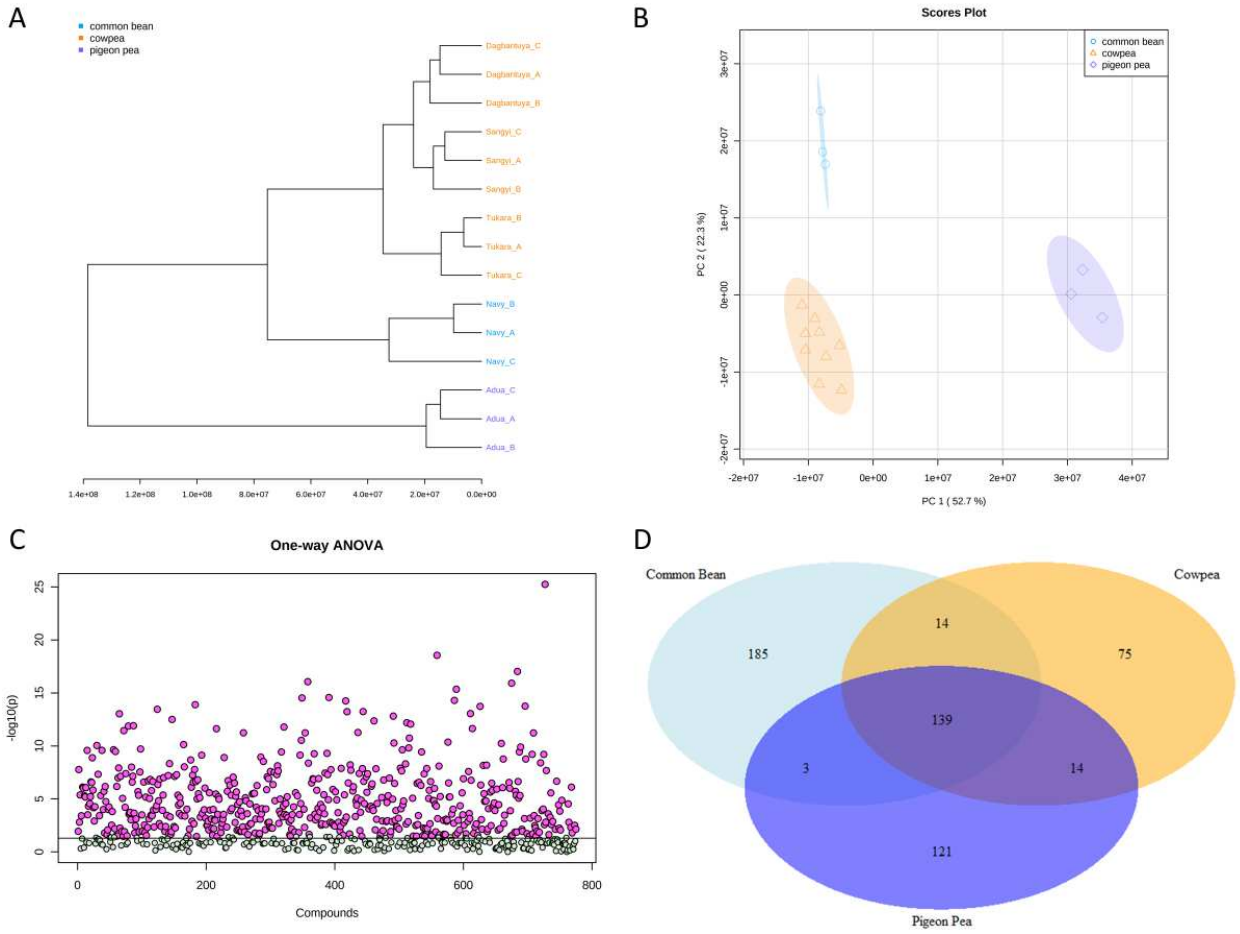

Figure 3.3 (A) Hierarchal cluster dendrogram of cowpea variety data based on Euclidean distance and Ward clustering. (B) PCA scores plot of cowpea variety data. (C) one-way ANOVA plotting $-\log_{10}$ (p-values) of all detected metabolites. Metabolites with significant differences in means across cowpea variety are plotted in pink (n=320), metabolites with no significant difference in mean across cowpea variety are plotted in grey (n=455). (D) Venn diagram indicating the number of metabolites that differentiate cowpea variety or varieties based on Fisher's LSD post-hoc analysis.

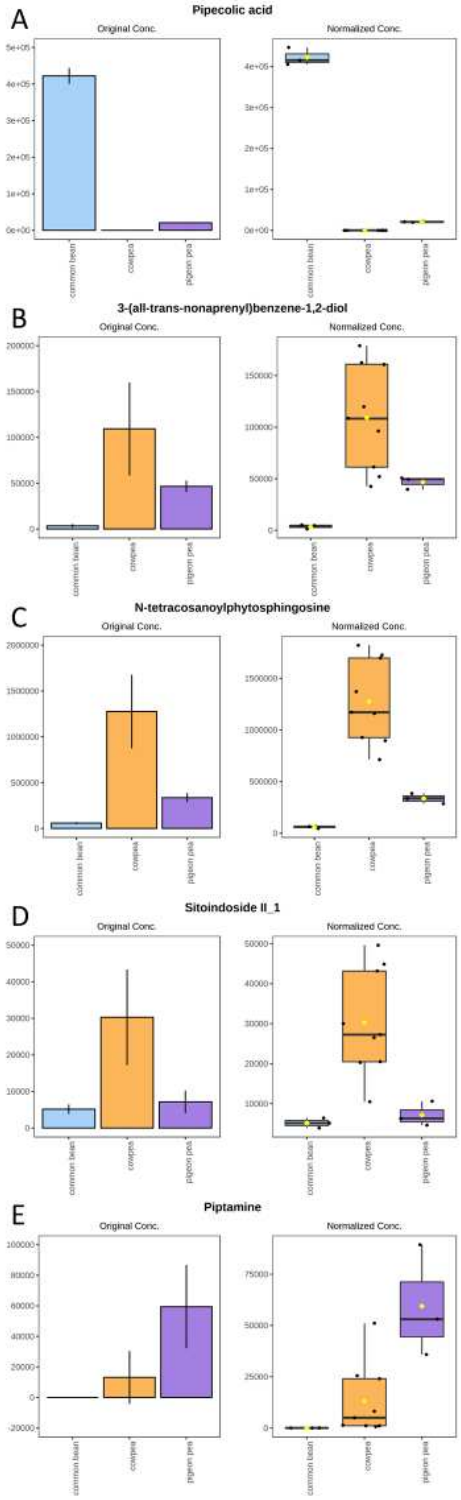

Figure 3.4 Median scaled relative abundance for metabolites distinguishing cowpeas, common bean and pigeon pea using Platform 1. Metabolites for (A) all legume types, (B-D) cowpea, (E) pigeon pea.

REFERENCES

Abizari, A.R., Pilime, N., Armar-Klemesu, M., Brouwer, I.D., 2013. Cowpeas in Northern Ghana and the Factors that Predict Caregivers' Intention to Give Them to Schoolchildren. PLoS One 8, 1–8.

Abunyewa, A.A., Karbo, K.N., 2005. Improved fallow with pigeon pea for soil fertility improvement and to increase maize production in a smallholder crop-livestock farming system in the subhumid zone of Ghana. L. Degrad. Dev. 16, 447–454.

Adjei-Nsiah, S., 2012. Role of Pigeonpea Cultivation on Soil Fertility and Farming System Sustainability in Ghana. Int. J. Agron. 2012, 1–8.

Awika, J.M., Duodu, K.G., 2017. Bioactive polyphenols and peptides in cowpea (Vigna unguiculata) and their health promoting properties: A review. J. Funct. Foods 38, 686–697.

Bentinger, M., Tekle, M., Dallner, G., 2010. Coenzyme Q - Biosynthesis and functions. Biochem. Biophys. Res. Commun. 396, 74–79.

Borresen, E.C., Zhang, L., Trehan, I., Nealon, N.J., Maleta, K.M., Manary, M.J., Ryan, E.P., 2017. The nutrient and metabolite profile of 3 complementary legume foods with potential to improve gut health in rural Malawian children. Curr. Dev. Nutr. 1, 1–11.

Broeckling, C.D., Ganna, A., Layer, M., Brown, K., Sutton, B., Ingelsson, E., Peers, G., Prenni, J.E., 2016. Enabling Efficient and Confident Annotation of LC-MS Metabolomics Data through MS1 Spectrum and Time Prediction. Anal. Chem. 88, 9226–9234.

Chen, Y., Yu, D.Q., 1996. Tonkinelin: A Novel Annonaceous Acetogenin from Uvaria tonkinensis. Planta Med. 62, 512–514.

da Silvia, A.C., da Costa Santos, D., Teixeira Jr., D.L., da Silvia, P.B., dos Santos, R.C., Siviero, A., 2018. Cowpea: A Strategic Legume Species for Food Security and Health. In: Legume Seed Nutraceutical Research. pp. 47–65.

Dahlén, B., Pascher, I., 1972. Molecular arrangements in sphingolipids. Crystal structure of N-tetracosanoylphytosphingosine. Acta Crystallogr. B, 2396–2404.

Djoumbou, Y., 2016. ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. J. Cheminform. 8.

Fernie, A.R., Schauer, N., 2009. Metabolomics-assisted breeding: a viable option for crop improvement? Trends Genet. 25, 39–48.

Gómez, C., 2004. Introduction: Economic and Social Impact, COWPEA: Post-Harvest Operations.

Haug, K., Cochrane, K., Nainala, V.C., Williams, M., Chang, J., Jayaseelan, K.V., O'Donovan, C., 2020. MetaboLights: A resource evolving in response to the needs of its scientific community. Nucleic Acids Res. 48, D440–D444.

Hedrick, V.E., Dietrich, A.M., Estabrooks, P.A., Savla, J., Serrano, E., Davy, B.M., 2012. Dietary biomarkers: advances, limitations and future directions. Nutr. J. 11.

Ishida, T., Nishiumi, S., Tanahashi, T., Yamasaki, A., Yamazaki, A., Akashi, T., Miki, I., Kondo, Y., Inoue, J., Kawauchi, S., Azuma, T., Yoshida, M., Mizuno, S., 2013. Linoleoyl ethanolamide reduces lipopolysaccharide-induced inflammation in macrophages and ameliorates 2,4-dinitrofluorobenzene-induced contact dermatitis in mice. Eur. J. Pharmacol. 699, 6–13.

Jaeger, C., Hoffman, F., Schmitt, C.A., Lisec, J., 2016. Automated Annotation and Evaluation of In-Source Mass Spectra in GC/Atmospheric Pressure Chemical Ionization-MS-Based Metabolomics. Anal. Chem. 88, 9386–9390.

Jayathilake, C., Visvanathan, R., Deen, A., Bangamuwage, R., Jayawardana, B.C., Nammi, S., Liyanage, R., 2018. Cowpea: an overview on its nutritional facts and health benefits. J. Sci. Food Agric. 98, 4793–4806.

Ji, J., Zhang, C., Sun, Z., Wang, L., Duanmu, D., Fan, Q., 2019. Genome Editing in Cowpea Vigna unguiculata Using CRISPR-Cas9. Int. J. Mol. Sci. 20, 1–13.

Langyintuo, A.S., Lowenberg-DeBoer, J., Faye, M., Lambert, D., Ibro, G., Moussa, B., Kergna, A., Kushwaha, S., Musa, S., Ntoukam, G., 2003. Cowpea supply and demand in West and Central Africa. F. Crop. Res. 82, 215–231.

Lu, C.T., Tang, H.F., Sun, X.L., Wen, A.D., Zhang, W., Ma, N., 2010. Indole alkaloids from chickpea seeds (Cicer arietinum L.). Biochem. Syst. Ecol. 38, 441–443.

Madrid-Gambin, F., Llorach, R., Vazquez-Fresno, R., Urpi-Sarda, M., Almanza-Aguilera, E., Garcia-Aloy, M., Estruch, R., Corella, D., Andres-Lacueva, C., 2017. Urinary 1H Nuclear Magnetic Resonance Metabolomic Fingerprinting Reveals Biomarkers of Pulse Consumption Related to Energy-Metabolism Modulation in a Subcohort from the PREDIMED study. J. Proteome Res. 16, 1493–1491.

Natarajan, S.K., Muthukrishnan, E., Khalimonchuk, O., Mott, J.L., Becker, D.F., 2017. Evidence for Pipecolate Oxidase in Mediating Protection Against Hydrogen Peroxide Stress. J. Cell. Biochem. 118, 1678–1688.

Nwokolo, E., 1987. Nutritional evaluation of pigeon pea meal. Plant Foods Hum. Nutr. 37, 283–290.

Pal, D., Mishra, P., Sachan, N., Ghosh, A.K., 2011. Biological activities and medicinal properties of Cajanus cajan (L) Millsp. J. Adv. Pharm. Technol. Res. 2, 207–214.

Perera, T., Young, M.R., Zhang, Z., Murphy, G., Colburn, N.H., Lanza, E., Hartman, T.J., Cross, A.J., Bobe, G., 2015. Identification and monitoring of metabolite markers of dry bean consumption in parallel human and mouse studies. Mol. Nutr. Food Res. 59, 795–806.

Poon, W.W., Barkovich, R.J., Hsu, A.Y., Frankel, A., Lee, P.T., Shepherd, J.N., Myles, D.C., Clarke, C.F., 1999. Yeast and rat Coq3 and Escherichia coli UbiG polypeptides catalyze both O-methyltransferase steps in coenzyme Q biosynthesis. J. Biol. Chem. 274, 21665–21672.

Rawal, K.M., 1975. Natural Hybridization Among Wild, Weedy and Cultivated Vigna unguiculata (L.) Walp. Euphytica 24, 699–707.

Satmbekova, D., Srivedavyasasri, R., Orazbekov, Y., Omarova, R., Datkhayev, U., Ross, S.A., 2018. Chemical and biological studies on Cichorium intybus L. Nat. Prod. Res. 32, 1343–1347.

Saxena, K.B., 2008. Genetic Improvement of Pigeon Pea — A Review. Trop. Plant Biol. 1, 159–178.

Schlegel, B., Luhmann, U., Hartl, A., Grafe, U., 2000. Piptamine, a new antibiotic produced by Piptoporus betulinus Lu 9-1. J. Antibiot. (Tokyo). 53, 973–974.

Smith, C.A., Want, E.J., O'maille, G., Abagyan, R., Siuzdak, G., 2006. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. Anal. Chem. 78, 779–787.

Sri Harsha, P.S.C., Wahab, R.A., Garcia-Aloy, M., Madrid-Gambin, F., Estruel-Amades, S., Watzl, B., Andrés-Lacueva, C., Brennan, L., 2018. Biomarkers of legume intake in human intervention and observational studies: A systematic review. Genes Nutr. 13, 1–16.

Tautenhahn, R., Bottcher, C., Neumann, S., 2008. Highly sensitive feature detection for high resolution LC/MS. BMC Bioinformatics 9, 1–16.

Tsopmo, A., Muir, A.D., 2010. Chemical profiling of lentil (Lens culinaris Medik.) cultivars and isolation of compounds. J. Agric. Food Chem. 58, 8715–8721.

Tsugawa, H., Kind, T., Nakabayashi, R., Yukihira, D., Tanaka, W., Cajka, T., Saito, K., Fiehn, O., Arita, M., States, U., 2016. Hydrogen Rearrangement Rules: Computational MS/MS Fragmentation and Structure Elucidation Using MS-FINDER Software. Anal. Chem. 88, 7946–7958.

Vaillancourt, R.E., Weeden, N.F., 1992. Chloroplast DNA Polymorphism Suggests Nigerian Center of Domestication for the Cowpea, Vigna Unguiculata (Leguminosae). Am. J. Bot. 79, 1194–1199.

Wang, C., Liu, R., Lim, G.H., De Lorenzo, L., Yu, K., Zhang, K., Hunt, A.G., Kachroo, A., Kachroo, P., 2018. Pipecolic acid confers systemic immunity by regulating free radicals. Sci. Adv. 4, 1–12.

Yilmaz, C., Gökmen, V., 2015. Chlorophyll. Encycl. Food Heal. 37–41.

Zarei, I., Baxter, B.A., Oppel, R.C., Borresen, E.C., Brown, R.J., Ryan, E.P., 2021. Plasma and urine metabolite profiles impacted by increased dietary navy bean intake in colorectal cancer survivors: A randomized-controlled trial. Cancer Prev. Res. 14, 497–508.

CHAPTER 4 – CONCLUSIONS AND FUTURE DIRECTIONS

Chapters 2 and 3 of this thesis explored quality and nutrition aspects of barley and cowpea as finished foods using genomic and metabolomic approaches. We found that barley genotype does contribute to beer flavor and that cowpea metabolite profiles have distinct and characterizable differences from pigeon pea and common bean. The results of these projects set a foundation for future genetic and breeding, as well as metabolomic work involving both foods, explored in more detail below.

Barley genotype contributes to beer flavor

Chapter 2, Genetic Basis of Barley Contributions to Beer Flavor, aimed to 1) test the hypothesis that barley genotype contributes to beer flavor, 2) identify regions of the genome that control traits associated with flavor, and 3) identify candidate genes that control traits associated with flavor. The work integrated malt quality, beer sensory, and metabolomics data to map QTLs and gain a deeper understanding of the genetic control of barley beer flavor. QTLs were identified for many traits and overlapping QTLs (hotspots) were found that included traits from all categories (malt quality, beer sensory, and metabolomics). Candidate genes we identified in those regions included genes related to seed dormancy and plant height, which seem to have downstream effects on malt quality, beer flavor, and metabolite content.

The results of this study lay the groundwork for future genetics and breeding research on the connection between barley genotype and beer flavor. Gene-editing technologies could be used to validate candidate genes, marker-assisted breeding could help select desired haplotypes

in different genetic backgrounds, and the effects of genotype by environment (GxE) interaction could be explored in depth on a subset of this population.

The malt quality, beer sensory, and metabolite datasets we collected can stand independently, but are more complete when integrated, as malt is an intermediate between barley and beer, and the chemistry of malting and brewing impacts metabolic changes of the barley grain and resulting metabolites. Craft brewing especially, finds that flavor is an attribute consumer are most interested in, so the economic impact of high-quality barleys with improved flavor profiles could be significant.

Cowpea metabolite profiles are distinctly different from pigeon pea and common bean

Chapter 3, Non-targeted Metabolomics of Cooked Cowpeas (*Vigna unguiculata*) and Pigeon Pea (*Cajanus cajan*) From Ghana, aimed to 1) characterize the metabolic profiles of three cowpea flours, and 2) test the hypothesis that there are metabolites unique to cowpea (and cowpea variety) that can be potential biomarkers for nutritional studies. Two non-targeted UPLC-MS approaches were used to compare cowpea to pigeon pea and common bean. Metabolite profiles of the legumes analyzed established differences between them, and metabolites that are unique to each species were identified. Metabolites common to all cowpea varieties were identified, as well as those unique to each cowpea type. Metabolomic profiles of cowpea were largely composed of lipid and lipid-like molecules, along with other compounds with established health benefits.

The results of this study will be integrated with an analysis of blood and urine from women and children who consumed those same cowpeas at increasing concentrations over a set

period of time. Validation of the metabolites as nutritional biomarkers will allow for measurement of consumption and compliance in feeding studies involving cowpea. In addition to having an impact in nutritional studies, this work could be used in future breeding efforts aiming to increase nutrient content. Metabolomics-assisted breeding would allow for early selection of desired phenotypes, and multi-trait genomic selection could further current breeding efforts for traits related to agronomic performance or quality with nutritional components as well. The scope of this study could also be expanded to other cowpea varieties, as the three we studied are not the only ones consumed in undernourished parts of the world.

Malnutrition is prevalent in many developing parts of the world, and cowpea is an important crop in many of those areas. As a food security crop with well-known health benefits, detailed cowpea metabolite profiles are necessary to determine biomarkers of this important legume. Biomarkers can take nutritional studies to the next level, removing the requirement of participant self-reported data. With the current lack of metabolite profiling in cowpea, this work brings us one step closer to validating biomarkers for cowpea, which will make nutritional studies of this important crop more effective.