Comparing the Meta-Proteomic Composition of Fermented Yeast Plant Proteome

By: Sidharth Sirdeshmukh

Lab of: Dr. Laura Herring, UNC Department of Pharmacology

Senior Honors Thesis
Department of Biology
University of North Carolina at Chapel Hill

May 5th, 2021

Approved:

_____
Dr. Laura Herring, Thesis Advisor
Dr. Jason Reed, Reader
Dr. William Marzluff, Reader

Abstract:

Beer commonly contains proteins such as hordeins, non-specific lipid-transfer protein 1 (LTP-1), and protein Z. The normal physiological functions of these proteins include seed storage and lipid transfer, but in beer they influence desirable beverage characteristics like head foam, aroma, and the formation of chill haze. Prior beer proteomics studies have detected variations in barley protein content across different brewing stages while others have observed changes due to known behavior of yeasts (ie. flocculation). Previous studies have mainly assessed samples brewed under lab conditions. Our study, in a departure from earlier studies, addresses the gap in research examining the proteome profiles of commercial beers. We compare the abundance of specific proteins in multiple species across different types of beer and cider to determine the proteomic impact of brewing conditions and ingredients by comparing the proteins in our beer and cider samples to data from previously published studies. We isolated proteins from our samples by performing a sample preparation protocol developed by the Protein Research Group (PRG), a subgroup of the Association of Biomolecular Resources Facilities (ABRF). Once peptides were cleaned, quantified, and recovered we analyzed samples by liquid chromatography coupled to mass spectrometry (LC-MS) analysis at the UNC Proteomics Core Facility. We assessed our proteomic workflow by comparing the number of proteins we identified with other beer proteomics datasets. Protocol validity and reproducibility was assessed through PCA plots and Pearson correlations among sample replicates. Overrepresentation analysis of yeast proteins highlights functional biological pathways that contribute to desirable protein expression and correspond to yeast genes. We studied other ingredient-related effects by comparing protein abundances of barley and yeast proteins previously

explored by beer proteomic literature to make predictions about beer quality. This allowed us to evaluate differences in sample preparation and LC-MS methods, as well as to compare beer protein results with the other 50 participants of the ABRF PRG beer study from around the world. Our research will contribute to the growing body of research in the field of beer proteomics.

Introduction:

Beer is a popular beverage throughout the world. What began as an ancient practice of combining sugary broths and yeast has become a mechanized, industrial process that serves many millions of people. Given the wide scale of the production and consumption of beer, it is surprising that there is limited research currently available on the biochemical composition of the drink. Previous studies have identified proteins in beer that are responsible for particular desirable qualities like chilled haze and foam.[1] Other studies have documented changes in protein representation based on changes in malting and the use of various yeast strains.[2]

Brewing can be conceptually understood as the process of feeding sugar-rich solutions to organisms that release alcohol into solution that can be separated and consumed. The process differs regionally and depending on the size/ capacity of the brewery, as some brewers may carry out particular ingredient modifications on-site, others importing ingredients from third-party suppliers. From beginning to end, beer brewing traditionally begins with the malting of milled barley grain. Wheat and rye are commonly malted for the same purpose. Dried, cleaned barley is malted by soaking and drying the grain, and then drying again after germination has taken place. Germination

produces proteins and enzymes that are required to break down barley starch in subsequent mashing stages, and malted barley proteins conserved by the end of the brewing process influence the taste and texture of beer.

Mashing follows malting and is responsible for producing the final “sugary broth,” known as wort, that yeast is added to during the fermentation process consume to generate alcohol. During mashing, crushed malt is mixed with hot water between 62.2-70C, which activates malt enzymes. These enzymes convert sugars in malted barley to maltose and dextrin, then the wort is separated from the grain. The wort is subsequently boiled for pasteurization, and then boiled with additional ingredients to impart particular flavors that are expected to be maintained for the remainder of the brew stages.[3] Hops, and refined sugars are traditional additives, and fruits and spices are commonly added at this stage as well. Once the boiled and flavored wort has been cooled depending on the kind of beverage desired, specific strains of yeast are added to generate alcohol via the consumption of sugar and the excretion of alcohol and carbon dioxide in a process known as fermentation. The product that we now have is called “bright beer” and can be filtered and carbonated before being packed for human consumption. Heating during mashing and fermentation steps during the brewing process are expected to influence the kinds of proteins that are present in the final product.[4] Appendix Figure 2 is a schematic of the beer brewing process**.**

There are many varieties of beers and ciders, and I will discuss changes in the brewing process that account for their differences. The main distinction between the brewing process for beers and ciders is that is that in ciders apple juice is substituted for wort. Barley and other grains are not ingredients in cider, whereas they are in beer. This

also means that additional flavor can be imparted after fermentation has concluded via the addition of sweeteners, juices, and concentrates. Hops are not used in the brewing process because ciders are not meant to be bitter. Most commercial beers fit into one of two categories, which are lagers and ales. Ales and lager differ based on whether they top or bottom-ferment, respectively. Flocculation behavior in yeast is what this difference is attributed to, and top-fermenting strains are less flocculent. Ales typically ferment between 18-22C and lagers ferment between 7-15C.[5] The temperature and length of aging that follows fermentation is also different between the two, with ales aging between 4.4-13C for a short time and lagers aging between 0-7C for a relatively long time. While it is known that the brewing process contributes to flavor and foam characteristics of beer, little is known about the specific effects that proteome profiles have these attributes.

Mass spectrometry-based proteomics is an invaluable tool for identify, quantifying and characterizing proteins in an unbiased manner. In the 'bottom-up' proteomics approach, a protease digests a protein sample, converting the sample into a peptide mixture that is then analyzed by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). During the MS/MS analysis, peptides are fragmented inside the mass spectrometer, and the harmonic oscillations of the particles within the orbitrap are converted via a Fourier transformation to generate mass to charge (m/z) ratios. The MS and MS/MS data are then searched against a public protein database and measured peptide sequence information is mapped back to respective theoretical peptide sequences determined by molecular weight. These databases are built from genome sequence data, since protein sequences can be derived from the genomic

data. With this technique not only can peptide/protein sequences be identified, but proteins can also be quantified based on the relative abundance of the protein across samples.[6]

Determining the abundance of particular proteins can help us predict modifications to known ingredients, however the broader benefits of better and more frequent biochemical characterization of beer is analogous to those of testing drinking water or other widely consumed products; These analyses are important to perform because they can inform the public of potential public health dangers or dietary concerns. Additionally, quantitative analysis of beer sample proteomic data can provide insights that can enable brewers and food chemists to optimize recipes to improve the expression of desirable proteins. Phylogenetic trees can be produced based on proteins identified and genetic improvement of grain and yeast strains are important goals in the field of beer proteomics.

There are multiple reasons why proteomic analysis of beers and ciders is interestingly challenging: Firstly, beer is made with multiple ingredients from different species, so special attention has to be paid when selecting databases to search raw data against. Secondly, many of the species where ingredients specific to beer originate are not well-studied, so genomic sequence information/protein databases for those species may be lacking. Thirdly, we are measuring fragments of proteins produced by peptide-level cleavage that takes place during brewing. As mentioned previously, the various stages in the brewing process contribute to the proteins that are present in the final beverage. Prior research has recognized non-tryptic cleavage events in peptides that have been attributed to yeast or barley enzyme-driven proteolysis that occurs

during brewing. At a physical level, the hydrolytic activity that takes place following malting and during the mash phase limits the molecular weights of proteins present in beer to a range of 5-100kDa.[7] Figure 1 of the appendix contains a diagrams of barley germination and hydrolytic cleavage.

While challenges exist in this analysis, current publications in beer proteomics have identified proteins that make significant contributions to beer quality which is determined by characteristics like foam and haze formation. The field has already established relationships between the enrichment of particular proteins and physically observed qualities in beer. Proteins that influence foam formation and stability include low molecular weight barley proteins like hordeins, lipid-transfer proteins, and a serine-like protease (serpin) known as protein Z. LTP-1 is a common lipid-transfer protein in beer, and is known to have strong foam forming properties, and improved foam stabilizing properties in the presence of hordeins or protein Z.[8] Protein Z has multiple isoforms that have been studied and has known foam stabilization properties. Previous publications have also demonstrated relationships between relative proportions of hordeins, lipid-transfer proteins, and protein Z, and malting level, or Kolbach (KI) index of barley used to make the wort.[7] Additionally, proteinase A from yeast is known to degrade LTP-1 during fermentation.[9]

In our study, we aim to employ an adapted version of the PRG sample preparation protocol for proteomic analysis of commercial beer samples to determine the accuracy and wider application of the protocol, and to assess the proteomic profiles of a diverse group of commercial beer and cider samples. Our beers primarily come from microbreweries around North Carolina, so we will be able to better understand the

quantitative impact of ingredient choices and brewing techniques by modifying database search parameters and performing additional comparisons. Ingredient species were divided into *flavorants*, or spice and fruit proteins, and *grains and yeasts*.

To determine the protein composition of our beverages, we took aliquots of commercial beers and ciders, precipitated proteins with acetone, digested with trypsin enzyme, and cleaned the peptides with C18 spin-columns. According to peptide concentration data from BCA quantitation, we normalized peptide amounts and analyzed the same amount of each peptide sample using a LC-MS/MS on a Thermo Easy nLC 1200 coupled to a QExactive mass spectrometer (Thermo). Raw data were analyzed in Proteome Discoverer 2.4 (Thermo), by searching against 14 UniprotKB protein databases. Within Proteome Discoverer, label-free quantitative values were extracted using area under the curve, and protein abundances could be compared across samples to assess protein abundance differences among different species, beverages, and other quantitative categories. All results were filtered using a 1% false discovery rate (FDR) to filter out low scoring proteins. Perseus software and R were used for Pearson correlation and analysis of genetic pathway enrichment respectively.

Overall, we identified 547 number of total proteins derived from 20 of different species across 10 different commercial beers/ciders. We demonstrate that beer of similar type shares similar proteomics profiles, according to principal component analysis, with the exception of the two replicates of the same cider were very compositionally different. We selected a group of commonly expressed proteins previously identified in beer, for example, the barley proteins that show varying abundances depending on malting and foam presence, to make additional predictions

about beverage quality. This work will contribute to the larger ABRF PRG project; the raw data collected for locally sourced commercial beer samples at UNC has been uploaded to a public repository and will be compared with data collected from international commercial beer samples. From this, we will be able to compare the proteomic results of the ABRF pale ale 'standard' across multiple labs to determine how different preparations and mass spec analyses can influence the results. We will also be able to compare beers from around the world to one another to shed insight into how location, specific brewing processes, and ingredient choices can affect the proteomic profiles. This study is among the first known of its kind and will add to the growing body of beer proteomics research.

Analysis/ Methods:

Wet-lab sample preparation:

The method used for preparing the beer proteomics samples was presented by Brett Phinney (UC Davis) and Ben Neely (NIST) in the PRG's Beer Proteomics project proposal. The described method was designed so that any proteomics lab across the world could apply it to their beer samples; regardless of participant's location, they should have all supplies listed in the method already in their lab. Meta-proteomic analysis will eventually be conducted using the large pool of submitted raw data to draw conclusions about geography-based protein representation, and other factors that contribute to variable protein abundance across international commercial beer samples. Tables and figures corresponding to wet-lab sample preparation are presented in the supplementary methods section (Tables 1-4, Figure 1).

Protein Precipitation:

We began by collecting our beer samples and cider samples (Supplemental methods, Table 1). Including the ABRF standard, we had 8 beers and 2 ciders to digest, clean, and analyze. We aliquoted 2 mL of each drink into conical tubes and stored them at 4C. Two replicates of each beverage were prepared by aliquoting 50 uL of each sample into Eppendorf tubes before adding 200 uL of acetone to each tube to precipitate the proteins from solution. After vortexing for 15 seconds, tubes were moved to the -20C freezer to incubate overnight to ensure complete precipitate formation. The next day, samples were centrifuged at maximum speed to pellet the protein, washed with acetone, and decanted twice before air drying the protein pellet.

Trypsin Digestion:

Once the protein was dried, we performed in-solution digestion of each sample with trypsin according to the PRG sample preparation guidelines. We calculated the volumes of our reagents, and proceeded with reconstituting, reducing, alkylating, and adding trypsin to each of our samples (Supplemental methods, Table 2). Adding trypsin cleaves proteins at the c-termini of arginine and lysine. To improve the digestion efficiency of trypsin, dithiothreitol (DTT) is used to linearize folded proteins, and iodoacetamide (IAA) is used to alkylate the linearized proteins to prevent cysteine bonds from reforming. The samples were incubated at 37C overnight, then the next day they were acidified with trifluoracetic acid (TFA) to stop the digestion.

Peptide Desalting for Mass Spectrometry:

Peptide cleaning with desalting spin columns (Pierce) is performed after digestion to remove any salts or small molecules that could interfere with downstream LC/MS/MS analysis. We began by opening refrigerated columns and spinning at 5000g

for 1 minute, before discarding the storage solution. We added 200 uL ACN (pH=10) to the column, and spun at 5000 x g for 1 minute, adding 200 uL of ACN again before repeating. We performed 2 wash steps in the same manner as the ACN wash, but with 0.1% TFA (pH=2). Then the samples, reconstituted in 0.1% TFA, are loaded onto the spin columns and centrifuged at lower speed (3000 x g) for subsequent spins to avoid sample loss. The samples are washed with 0.1% TFA two times, then eluted with 70% ACN, 0.1% TFA into a new Eppendorf tube. Eluted, cleaned peptides are dried down via vacuum centrifugation.

BCA Colorimetric Assay:

The BCA peptide colorimetric assay (Pierce) is performed to determine the concentration of cleaned peptides in each of our samples. From this method we created two calibration curves in case sample concentrations are particularly high or low. We first prepared a serial dilution by adding 100 uL of peptide standard to tube A, and then transferred 50 uL into tubes B through G. We added a diluent, which is 0.1% formic acid, to each tube B-G before transferring the standard. Finally, we prepared the working reagent, adding enzymatically active reagent C last. After the plate was incubated adequately, we analyzed it on the plate reader (make, model) we assumed that there were about 100 ug of lyophilized peptides in each tube, so a 2x dilution was performed by adding 200 uL 0.1% formic acid to get to a concentration of 0.5 ug/ uL. Then, we transferred 25 uL of each of the 2x diluted solutions to new tubes and added 25 uL of formic acid to get to bring the new solution to a 4x dilution.

Peptide Resuspension and Vialing:

All of the samples had a concentration of 0.3 ug/ul or higher, which is ideal for proteomics analysis, with the exception of the cider samples which had lower concentration. We ultimately decided to lyophilize and resuspend the cider samples then reconstituted them in a lower volume of 0.1% formic acid to increase concentration. This was done by placing samples 11-14 in the lyophilizer for 1.5 hours before adding the appropriate volume of diluent to normalize the sample concentration to 0.3 ug/uL, which is what we also normalized all other samples to (Supplemental methods, Table 3). For all of our samples we pipetted thawed peptides and necessary amounts of diluent to get to 20 uL of total volume of a 0.3 ug/uL solution. Samples were ready for vialing after vortexing for 15 seconds to ensure all volumes were thoroughly mixed. Finally, mass spectrometer vials were labeled, and the total 20 uL sample volume was transferred to the 22 respective vials. 4 additional 'pooled sample' vials were created, containing 1 uL of each sample to run for quality assessment purposes.

Data analysis:

We performed database searches against the raw data generated via mass spectrometry. Based on the ingredients reported by manufacturers, we determined potential organisms whose proteins would be represented in our various samples. In total we obtained 14 database files from UniProtKB, including various yeasts and fruits (Supplemental methods, Table 4). We analyzed the raw data to Proteome Discoverer 2.4 (PD2.4) software and applied standard search parameters to generate lists of high-confidence proteins that were present in our samples. From these lists, we could draw several conclusions regarding the presence/absence of proteins derived from the

different ingredients in each beer, the relative abundance of proteins across the samples, and assess the molecular weight range of the proteins identified.

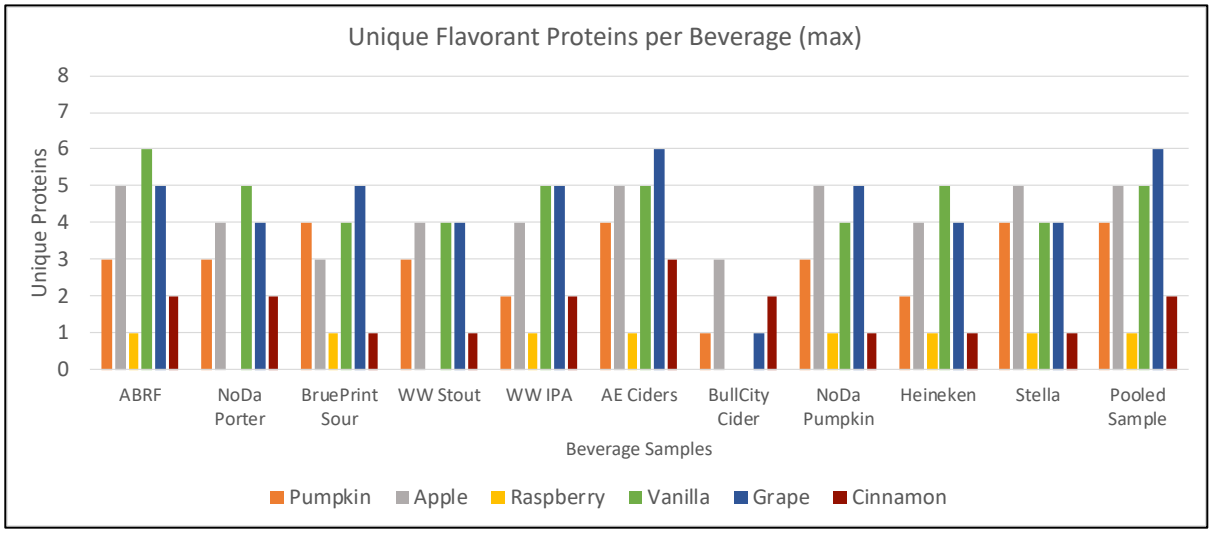
We used Perseus software to perform Pearson correlation tests among replicate samples. Normalized scaled abundance data was uploaded in tab-delimited form (.txt) and non-values were converted to 0 to avoid errors. The 'column correlation' function was selected in Perseus because sample replicates were arranged in columns in the Excel exported from PD2.4 following database comparison searches. A new matrix was produced with values of Pearson coefficient R between -1 and 1, and visualized in a heatmap with clustered terms to assess replicate similarity (Supplemental methods, Figure 1).

To determine the functional biological groups of proteins enriched, particularly for the strains of yeast identified in our database searches, we used DAVIDBioinformatics and Morpheus web tools to produce heatmaps. Yeast proteins were selected for this analysis even though they have low abundances across our samples, because they are essential to fermentation, and yeast already has many known targets for optimization via genetic engineering (ie. Mutations, knockouts etc.). Additionally, yeast is the most well-annotated species of all species used in this analysis; therefore, there is a good amount of information known about the yeast protein's biological functions. We selected the top-15 high-confidence proteins from all yeast species represented in our analysis for every beverage replicate. Replicate protein lists for single beverages were combined and protein accessions were converted to gene entrez IDs. Gene entrez IDs were loaded into DAVIDBioinformatics to perform over-enrichment analysis, and data

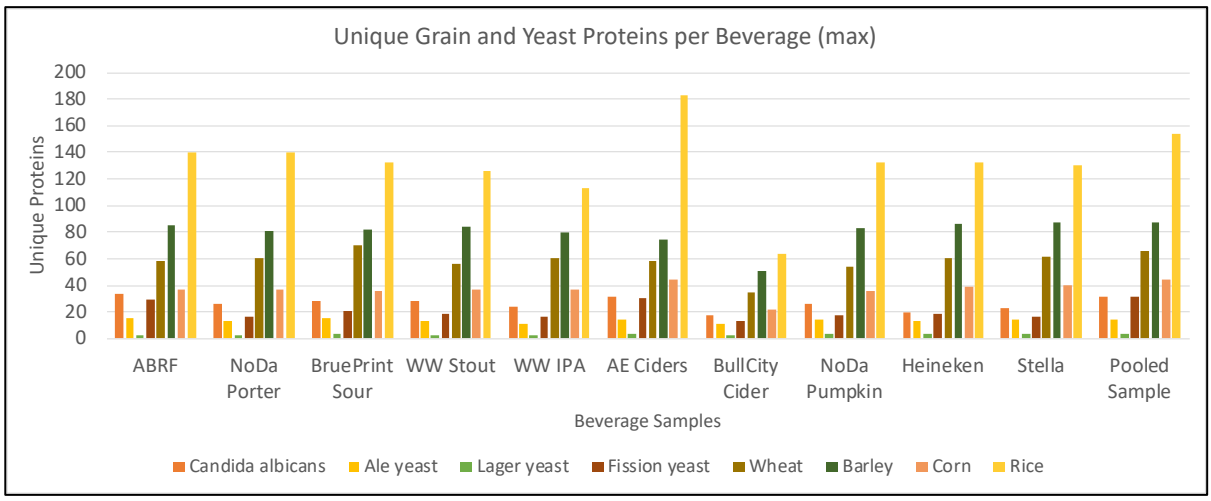
was downloaded, organized, and visualized using Morpheus (Supplemental methods, Figure 2).

Results:

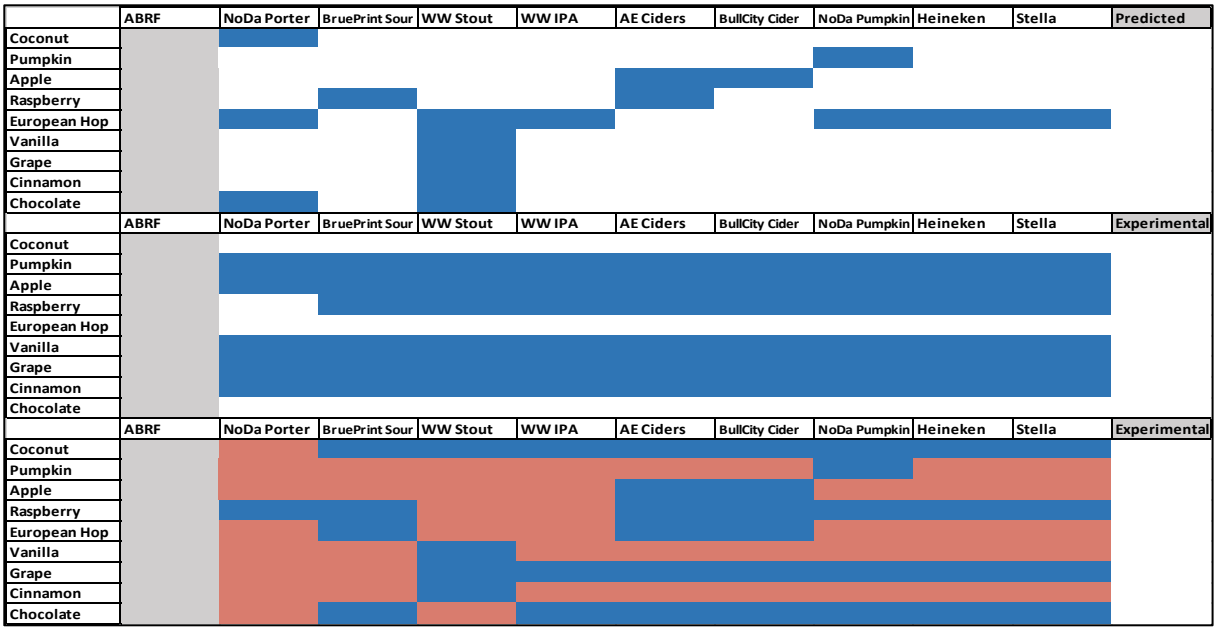
a.



b.



C.



	ABRF	NoDa Porter	BruePrint Sour	WW Stout	WW IPA	AE Ciders	BullCity Cider	NoDa Pumpkin	Heineken	Stella	Predicted
Coconut											
Pumpkin											
Apple											
Raspberry											
European Hop											
Vanilla											
Grape											
Cinnamon											
Chocolate											

	ABRF	NoDa Porter	BruePrint Sour	WW Stout	WW IPA	AE Ciders	BullCity Cider	NoDa Pumpkin	Heineken	Stella	Experimental
Coconut											
Pumpkin											
Apple											
Raspberry											
European Hop											
Vanilla											
Grape											
Cinnamon											
Chocolate											

	ABRF	NoDa Porter	BruePrint Sour	WW Stout	WW IPA	AE Ciders	BullCity Cider	NoDa Pumpkin	Heineken	Stella	Experimental
Coconut											
Pumpkin											
Apple											
Raspberry											
European Hop											
Vanilla											
Grape											
Cinnamon											
Chocolate											

Figure 1: *Identification of protein species in beverages.* ***A.*** *Unique flavorant proteins per beverage for maximum counts detected among replicate pairs.* ***B.*** *Unique grain & yeast proteins per beverage for maximum counts detected among replicate pairs.* ***C.*** *Color-coded table of protein species identifications (flavorants, grain & yeast) per beverage analyzed. Grey cells indicate no prediction was recorded, blue cells in Predicted and Experimental tables indicate 1 or greater high-confidence hits for proteins of particular ingredients (species). In Overall table, blue cells denote correct predictions and red cells denote incorrect predictions.*

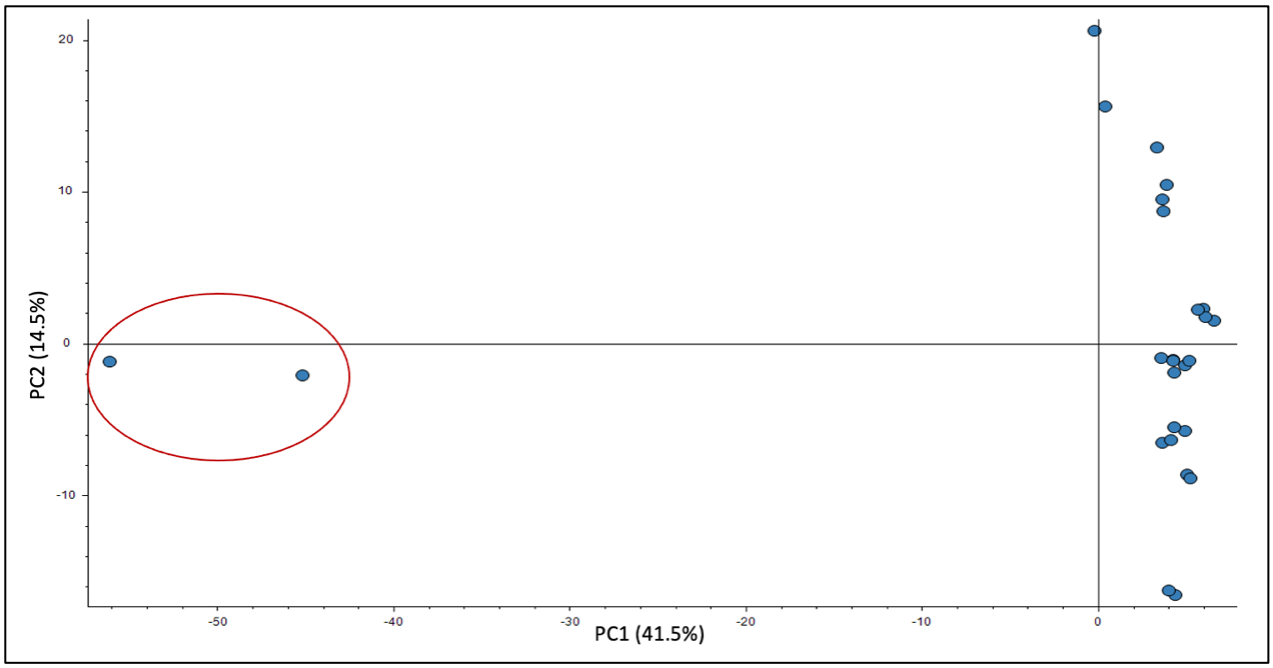
One of the central goals of this project was to measure and identify proteins that contribute to the flavors and structure of commercial beers and ciders because of the limited research that is currently available. We noted that ingredient contributions are two-sided; flavorant proteins include fruits and spices and all other ingredients included in the comparison databases were grain and yeast proteins. To measure the unique flavorant protein compositions of the beverages, we searched the raw mass spectrometry data against publicly available UniProtKB protein databases. The majority of the database files are ‘reviewed’, denoting protein lists were submitted from literature and also manually annotated. The flavorant protein species included coconut, pumpkin, apple, raspberry, tomato, and vanilla, grape, and cinnamon. Non-flavorant, or grain and

yeast-derived proteins, included ale, lager, and fission yeast, barley, yeast, wheat, corn, and rice.

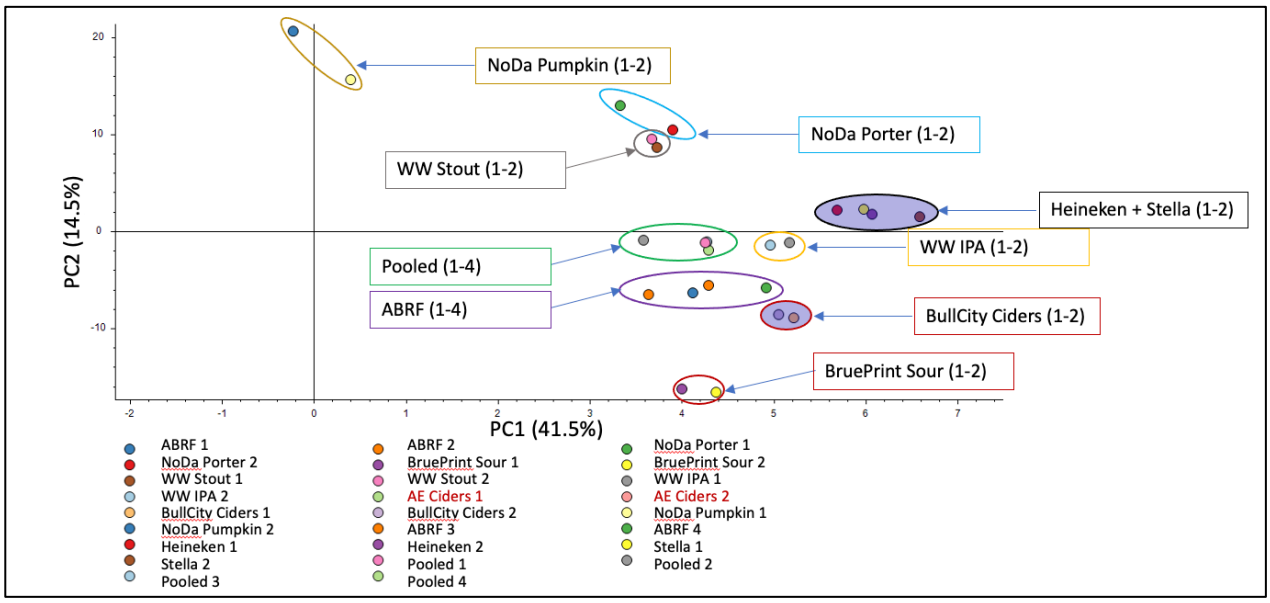
Overall, we identified 547 proteins altogether; of these, roughly 80 proteins were identified in all beverage samples analyzed belonging to barley, yeast, wheat, and corn species. Of all the flavorant protein species considered in our database search, vanilla had the highest number of unique (only vanilla-specific) high-confidence proteins identified. AE Ciders contained the most unique flavorant proteins of any beverages, and the ABRF standard had the most unique flavorant proteins among beers specifically (Figure 1A). From the comparisons against our grain and yeast protein databases, we found that Schizosaccharomyces pombe (fission yeast) was the best represented among proteins from all yeast species, and rice had the most high-confidence proteins recorded in the analysis (Figure 1B).

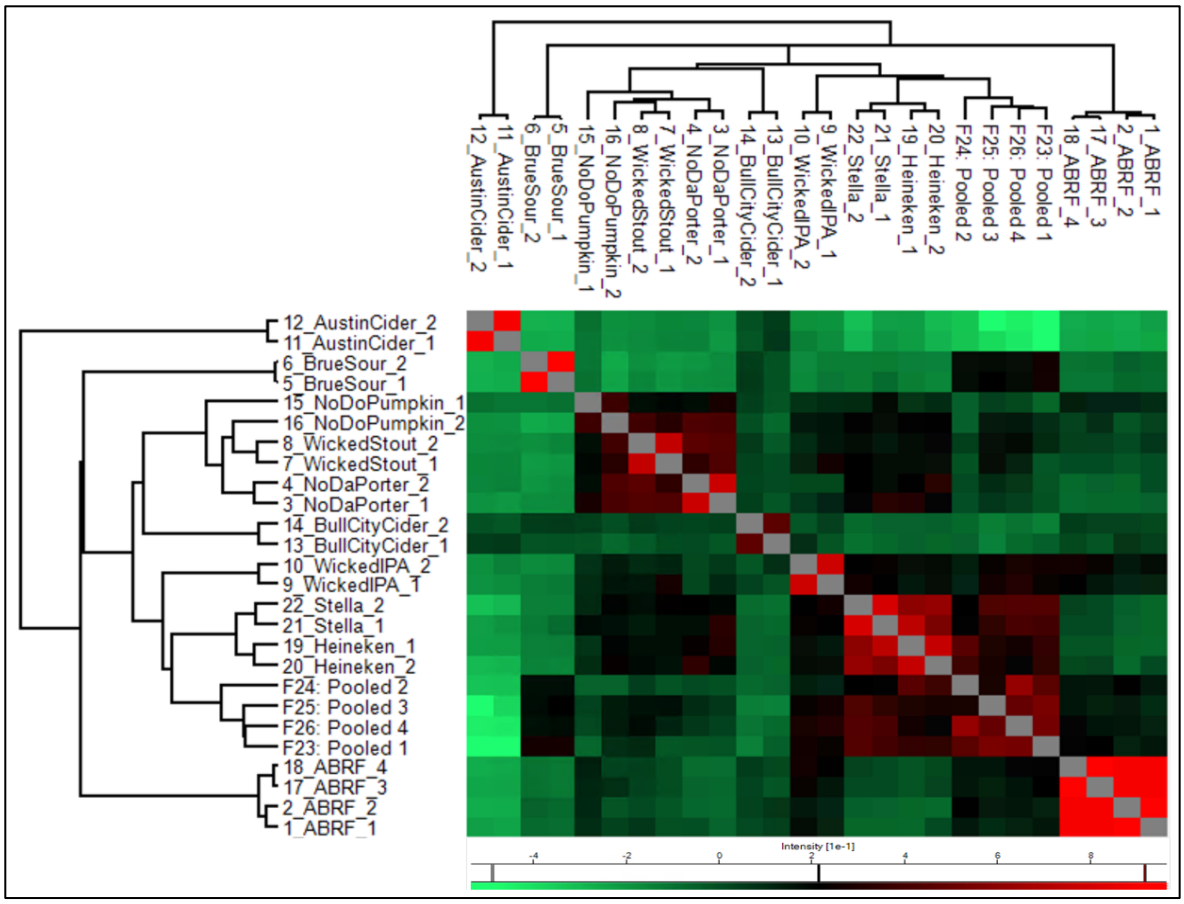
Another goal of our analysis was to make species-level identifications as a means of determining the accuracy and validity of our own analysis. We sorted our high-confidence protein abundance data generated by the database searches by species and noted which species had proteins of *any* abundance per beverage. Figure 1C provides an overview of flavorant protein species that were predicted and identified prior to and following the flavorant database searches. We normalized unique flavorant protein totals to remove obviously errant species identifications (ie. Tomato), and determined that 44.4% of our initial predictions of ingredient presence/ absence were correct. We also did not detect any proteins belonging to hops.

a.


PC2 (14.5%)
PC1 (41.5%)
20
10
0
-10
-50
-40
-30
-20
-10
0

b.


NoDa Pumpkin (1-2)
NoDa Porter (1-2)
WW Stout (1-2)
Heineken + Stella (1-2)
WW IPA (1-2)
Pooled (1-4)
ABRF (1-4)
BullCity Ciders (1-2)
BruePrint Sour (1-2)
PC2 (14.5%)
PC1 (41.5%)
ABRF 1
NoDa Porter 2
WW Stout 1
WW IPA 2
BullCity Ciders 1
NoDa Pumpkin 2
Heineken 1
Stella 2
Pooled 3
ABRF 2
BruePrint Sour 1
WW Stout 2
AE Ciders 1
BullCity Ciders 2
ABRF 3
Heineken 2
Pooled 1
Pooled 4
NoDa Porter 1
BruePrint Sour 2
WW IPA 1
AE Ciders 2
NoDa Pumpkin 1
ABRF 4
Stella 1
Pooled 2



C.

Figure 2: Assessment of compositional similarities among replicates *via PCA plots and Pearson correlation.* ***A.*** *Replicates circled in red are AE Ciders (1-2)* ***B.*** *(Zoom) Highlighted samples denote cider (BullCity) and lager (Stella + Heineken) samples. AE Ciders replicates are omitted in this plot.* ***C.*** *Heatmap and clustered samples to visualize Pearson correlation across replicates. Grey cells indicate no value (same replicate comparison), Green, Black, and Red cells indicate low, medium, and high correlations between compared samples, respectively.*

Beyond reviewing the general, relative abundance and overall presence of species and unique proteins, we sought to assess the clustering, or the compositional similarities among samples based on identified peptides. From the database comparison run with all flavorant, yeast, and grain protein lists, ProteomeDiscover2.4 generated a principal component analysis (PCA) plot from which we could visually assess similarities. An initial takeaway is that the AE Ciders replicates cluster on the far left of the first PCA plot, while all other samples cluster towards the right (Figure 2A). This indicates that the protein composition of the AE Ciders replicates is very different

relative to all the other samples analyzed. Annotation of the zoomed in plot (excluding the AE Ciders) reveals a clustering pattern that is potentially ingredient-related: The highlighted, grouped samples that cluster furthest to the right of the zoomed-in PCA plot correspond to the replicates of the cider and lager samples (Figure 2B). Furthermore, the beers brewed at the largest commercial volume, and also our only lager samples, Stella and Heineken, cluster very well together.

Another goal of this project was to determine the validity and accuracy of our wet-lab sample preparation protocol. We did this by reviewing the PCA plots, as well as by running a Pearson correlation between replicates in Perseus (MaxQuant) using protein abundance data. The grouping of replicates in the PCA plot illustrate the similarity in replicate composition. The majority of replicates were grouped along the horizontal axis (PC1- 41.5%) (Figure 2B). To better represent replicate similarity, we conducted a Pearson correlation and visualized data using conditionally formatted cells. We can see red cells for all replicate comparisons, and a larger proportion of black cells for pooled samples (samples 23-26) and non-pooled sample comparison, which was expected given pooled samples contained elements of every sample analyzed (Figure 2C). Overall these results suggest high reproducibility in the wet-lab sample preparation method, and ability to use a common proteomic wet-lab method to study beer samples.

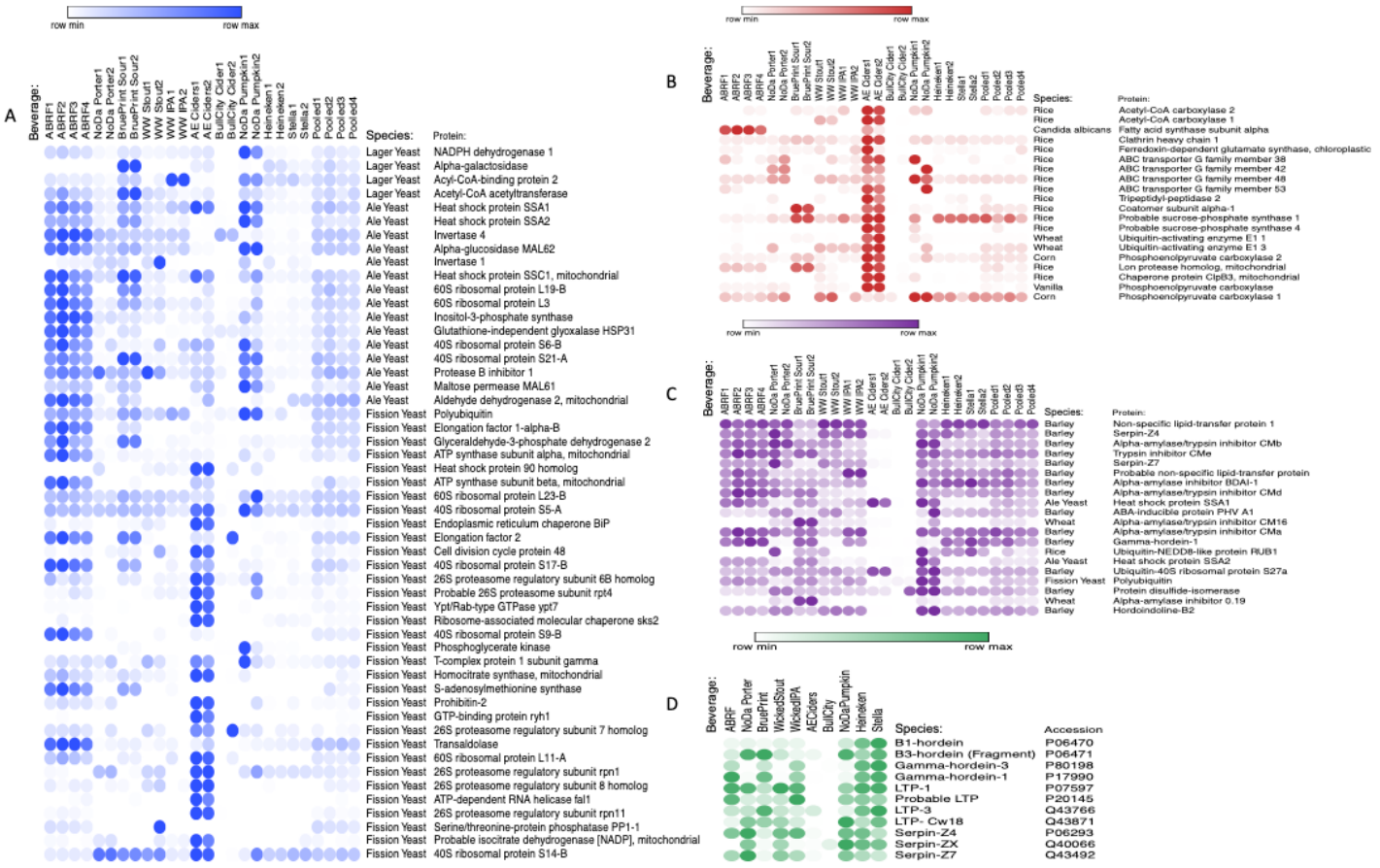


Figure 3: *Protein abundance assessment for estimation of beer quality (barley, yeast)* ***A.*** *All yeast high-confidence proteins, organized by species.* ***B.*** *Highest molecular weight proteins represented among all samples, decreasing in weight going down.* ***C.*** *Highest PSM number proteins represented among all samples, decreasing in PSM number going down.* ***D.*** *Barley proteins previously identified in beer proteomics literature which relate to beer and ingredient quality.*

Once we were able to assess beverage similarity via the PCA plots, we sought to understand the protein composition of beverages using parameters that do not necessarily categorize proteins by species. We sorted all proteins from highest to lowest molecular weight, discarding keratin and trypsin proteins. Keratin proteins are common contaminants that can enter samples from human hair, skin, nails during wet-lab preparation stages, while trypsin proteins correspond to inactivated trypsin enzyme that was not removed during desalting or separation on the UPLC.

All of our beer and cider samples contained yeast, because fermentation conducted by yeast is responsible for the production of alcohol. Our database search yielded high-confidence proteins for 3 species of yeast, which were *Saccharomyces pastorianus* (Lager yeast), *Saccharomyces cerevisiae* (Ale yeast), and *Schizosaccharomyces pombe* (Fission yeast) that are known to be used for beer and wine brewing.[10] It is evident that ale yeast proteins are relatively higher in abundance in ABRF (ale), BruePrint Briarberry (sour ale), and NoDa Gordgeous (pumpkin ale). Pooled samples showed consistent, intermediate protein abundance values across proteins from all species, which was expected. Lager yeast proteins made up 4 out of 52 of the high-confidence yeast proteins identified, and they were better represented in lager samples (Heineken and Stella Artois) as expected. Austin Eastciders replicates showed relatively high abundances of fission yeast, while Bull City Ciders showed almost none (3A).

Two major takeaways from the high molecular weight proteins chart are that rice proteins make up the vast majority of high molecular weight proteins in our beverages, and AE Ciders replicates contain the highest abundance of high molecular weight proteins (Figure 3B). We then sorted all proteins from highest to lowest number of peptide spectrum matches (PSMs), which is a measure of relative abundance, to assess abundances of the top-20 highest-PSM proteins in our samples. The protein in this list with the highest number of PSMs across all beverages is a barley protein called ‘non-specific lipid-transfer protein 1’ (highlighted in purple). This means that this protein had the most corresponding peptides detected relative to any other protein recorded by the mass spectrometer. This common protein was measured at a similar abundance

among all beverages except the ciders. Barley proteins generally made up the majority of the top-20 highest-PSM proteins (Figure 3C). Barley proteins are highlighted in orange.

Previous literature has determined the relationship between the relative abundances of certain barley proteins and corresponding measures of beer quality.[7] Supplementary table 4 lists the 11 barley proteins, their functions, and known contributions to beer. We took the given abundance data for the selected proteins and calculated the relative proportions of protein abundance using maximum values among sample replicates.

It is evident that Heineken and Stella beers (lagers) had the highest proportions of B1-hordein (P06470) and Gamma-hordein-3 (P80198), while NoDa Porter, NoDa Pumpkin, BruePrint Sour, and Stella show higher proportions of B3-hordein (P06471) and Gamma-hordein-1 (P17990) relative to all other beverages. Both cider samples had the lowest proportions of all barley proteins, which was expected given barley is not an ingredient commonly used in cider. Barley lipid-transfer proteins (LTPs) were generally expressed in higher proportions in lagers and ales, relative to all other beverages. Including lagers, NoDa Porter and Pumpkin beverages also had relatively higher proportions of serpins (Figure 3D).

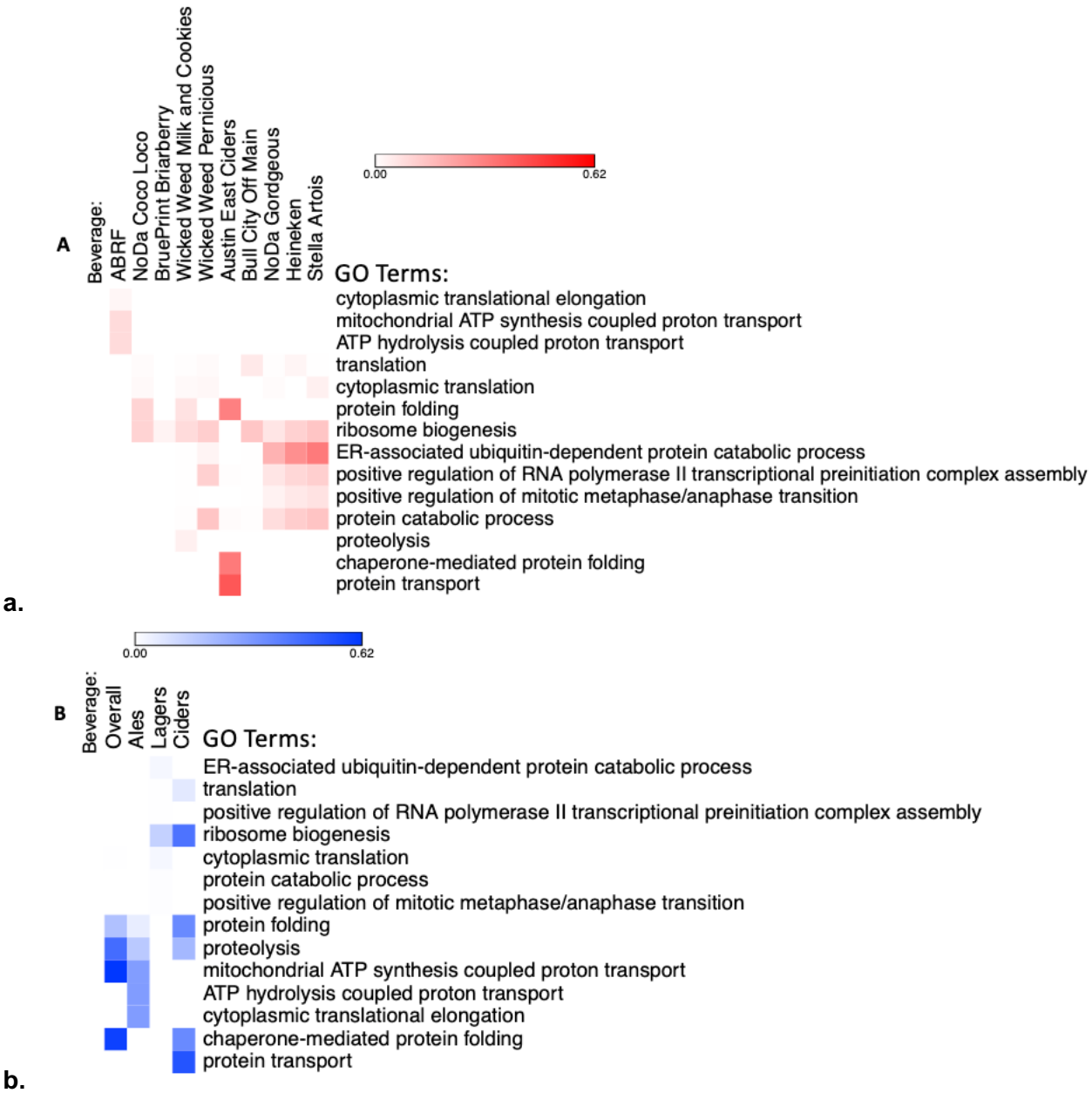


Figure 4: *Yeast Protein GO Term Over-enrichment Analysis* ***A.*** *GO Term over-enrichment by beverage.* ***B.*** *GO Term over-enrichment, organized by beverage style***.**

Yeast is a model organism that is well understood genetically, and a critical ingredient in beer brewing. Yeast is responsible for producing alcohol and carbon dioxide from sugars present in wort, under oxygen free conditions. Yeast proteins remain in beer at the time of packaging and influence beer quality broadly. Previous studies sought to assess the proteomic effects of recipe-level modifications. This

analysis may serve as the first steps of using proteomic data to generate research hypotheses for the genetic optimization of yeast in beer and cider production. Gene Ontology (GO) Biological Process (BP) terms were selected for review because categories pertain to organism-specific gene functionality of varying specificity. Genes that correspond to GO terms can be found on the Amigo database and then mapped to yeast vectors and modified for desired effects (ie. mutated, upregulated etc.).

Among all beverages analyzed, the ribosome functionality term was commonly moderately over-enriched. We expected that beverages brewed by the same brewery and of the same style (ale or lager) would show the over-enrichment of similar terms due to the use of the same strain of yeast. What we found however was that different NoDa and Wicked Weed beers did not show similar patterns of over-enrichment. This is likely due to the proteolytic or homology-level effects of the various flavorant and grain species used, or perhaps effects of modified gene expression that is a result of different yeast diets during fermentation. In a comparison of beverage categories, it is apparent that GO term over-enrichment is quite different which may be partially due to the species-level differences in yeast used to brew ales, lagers, and ciders. While this analysis is not comprehensive enough to make direct connections between beverage-specific GO term over-enrichment and protein abundance, it can provide clues about what genes to perform initial shotgun analysis on from which more detailed and extensive experimentation and proteomic characterization may proceed.

Discussion

Our overall goals for this project were to use an adapted wet-lab sample preparation protocol to first efficiently prepare proteomic samples from beers and ciders,

We ultimately compared differences in ingredient-protein abundances against known values from previously published literature. We began by providing a higher-level view of unique proteins counts for individual protein species (flavorants vs. non-flavorants), then assessed replicate and beverage sample similarity with PCA plots and visualized Pearson correlations. Abundance data was then further assessed according to overall protein molecular weight, and also among proteins identified in literature that vary in expression according to the malt-level (KI index) of barley used, and contribute to beer foam generation and stability. Finally, highly abundant yeast protein accessions were converted into gene lists from which heatmaps to visualize the enrichment of certain functional gene groups (GO-terms). Our findings validate the PRG methodology for LC/MS analysis of alcoholic beverages, and open the door for further analysis that may be useful for beverage optimization for commercial brewers.

From this research we were able to generate the following conclusions: Firstly, as demonstrated by the peptide BCA results, as well as the mass spectrometry proteomics results, the proteomic sample preparation method was effective and highly reproducible between replicates. Secondly, we can use this method to identify ingredients present in beverages, measure differences in protein abundance, and make further assessments regarding beer and ingredient quality, and identify relevant biological categories for additional research. Thirdly, peptide-level fragmentation and sequence homology among species may contribute to overlapping and errant protein identifications. Future studies should account for high levels of certain proteins and discrepancies in database sizes to improve the accuracy of the database comparison portion of this analysis.

We expected to identify flavorant proteins listed in the ingredients of the respective beverages in our database search against the same ingredients. For example, NoDa Pumpkin is a pumpkin ale, so we expected pumpkin and ale yeast proteins to be present in our samples of the beverage. We identified both proteins in the NoDa Gordgeous (pumpkin ale) sample, however these proteins were also identified at similar levels (number of unique proteins identified) for all other samples. We also reviewed the unique proteins present in ciders and found similar overlap. Ciders are not brewed with barley, rather apple juice is used instead, so we expected to see apple proteins present in cider samples. Instead, we found apple proteins in all beverages, and despite having the lowest numbers among all samples the ciders contained unique barley proteins as well (Figures 1A,1B).

There are two main reasons why we generally observe erroneous protein identifications in our beverages: One reason is that there may be protein sequence homology among these yeast strains, and with so many additional non-lager yeast proteins it is possible that the lager yeast proteins recorded are simply those that are very similar in sequence to some of the non-lager yeast proteins that were identified. Additionally, other unknown or non-sequenced/ un-characterized flavorants could be in these beers and their sequences could be very similar to those of a database we did include. This could be a reason why we identified, for example, pumpkin in several beers not known to contain pumpkin, and tomato and vanilla proteins in all beverages.

The accuracy of our wet-lab sample preparation methodology was confirmed by assessing the compositional similarities of replicates. It is evident that replicates of the same samples cluster together along the horizontal axis of the PCA plot, representing

overall compositional similarity (Figure 2B). Pearson correlation performed in Perseus yielded high correlation values (red) between replicates, and low correlation scores (green) for non-replicates. Pooled samples showed intermediate correlation levels (black), which was expected given those samples were produced by combining all aliquots together after peptide isolation and cleaning. Additionally, the PCA plots revealed that Austin East Ciders (AE Ciders) was the most compositionally different beverage (Figure 2A), which may be due in part to higher abundances of rice and fission yeast proteins relative to all other beverages.

While it is difficult to assess which proteins account for most of the compositional difference in Austin Eastciders, it is worth noting that this beverage showed high abundances of rice proteins (Figure 3B). Rice may be added to improve the volume of wort produced by mashing, but is excluded from some recipes because it is not as desirable of a brewing grain as barley. The main reason that this beverage shows higher abundances of rice proteins is likely to improve cider yield. Research from Kirin Labs also notes that high molecular weight proteins added to beverages brewed with low levels of malt show improvements in texture. Though rice protein molecular weights are much higher than those listed by Kirin (10-20 kDa), it is possible that hydrolytic cleavage driven by yeast proteinases may cleave rice proteins to desirable weights by the time fermentation has concluded.[11] Cider does not contain malt, so though rice is viewed as an inferior ingredient in the brewing process, it is worth exploring the foam-generation properties of rice and other high molecular weight proteins in traditional low-malt or no-malt beverages.

The majority of beer proteomics literature seeks to better understand the relationships between the proteomic profiles of alcoholic beverages and their observed physical characteristics. We reviewed protein abundance proportion differences between specific barley proteins across all beverages to predict physical beverage and ingredient characteristics such as the formation and stabilization of beer foam and KI index, which pertains to the level of malting that occurs in barley used for brewing (Appendix, Table 1). According to Schulz, Phung et al. beers with higher proportions of particular hordeins (accessions: P06470, P06471, and P80198) are those that use high KI index malt.[4]

We noted earlier that Heineken and Stella beers (lagers) had the highest proportions of B1-hordein (P06470) and Gamma-hordein-3 (P80198), so this comparison suggests that our lager samples use the most extensively malted barley during brewing. One beer that may also be brewed with high KI index malt is the NoDa Coco Loco (porter), which shows the highest proportions of both B3-hordein (P06471) and Gamma-hordein-1 (non-literature). The literature suggests other barley proteins that are expressed in high proportions when malting is at a low KI index, however they were not present in our final high-confidence protein list suggesting that the majority of known barley-containing beverages were using malted barley in the higher end of the KI index.

We compared proportions of additional barley proteins with functional effects in beer to make more predictions about beer quality. All LTPs were more highly expressed in lagers and ales, and BruePrint Sour was the only beer that had a low LTP-1 proportion, but a very high LTP-3 proportion. While the LTP-1 reaction with yeast proteinases is a known process in foam degradation in alcoholic beverages, the impact

of LTP-3 on foam formation should be studied further given sour beers are known to have foam of poor stability. Though some common explanations attribute the poor stability to the lack of hops used in the sour brewing process, barley LTP-3 abundance may have an inverse relationship with foam stability.

The final literature barley proteins that we compared the proportions of among our beverages were serpin-like Z proteins. The Z proteins from barley that were identified among our beverages Z4, ZX, Z7. Isoforms Z4 and Z7 are known to be commonly found in beer and were predicted to be identified in our samples. These proteins have also been noted to show a linear relationship between their contribution to the stabilization of beer foam and the amount of malting in the barley used for brewing.[8] So, for beverages with highly malted barley, we would expect a larger contribution to beer foam stability from Z proteins. From our previous analysis, we noted that our lager samples likely used the most malted barley (high KI). We would add to our earlier prediction, that somewhat higher proportions of Z proteins in lagers generally may be due to this linear interaction previously noted. Additionally, given NoDa Porter has the highest respective proportions of Z4 and Z7, and may use higher KI malt given the enrichment of B3-hordein (P06471), NoDa Porter may have the most stable foam among all beverages analyzed in this study.

Yeast proteins have been minimally explored in beer proteomics literature, with regards to connections between abundance and ingredient-level quality. Proteinase A and other literature-characterized yeast proteins found in previous studies were not present among the roughly 550 high-confidence proteins identified in all beverages. It is possible that higher levels of other identified species (ie. rice proteins in Austin

Eastciders) negatively biased the calculated yeast protein abundances by decreasing them. We performed GO term over-enrichment analysis primarily to make up for this inconvenience. Additionally, because many elements of this research is exploratory in nature, GO term over-enrichment analysis serves to outline future steps to better understand and control the proteome of alcoholic beverages. However, because yeast is a popularly used organism in synthetic biology research and for wide applications of cloning and genetic engineering, the field may seek to upregulate or knockout genes that correspond to high-enriched terms in the lab and prepare aliquots of beer samples for analysis via LC/MS.

As mentioned earlier, protein representation is responsible for the taste of beer. While taste itself is difficult to gauge objectively, beer proteomics offers alternative quantitative tools to make such assessments. A brewer or researcher may seek to improve their brew by reinforcing the abundances of specific yeast proteins. The proteins selected may be those that are highly represented already, so for Austin Eastciders we might continually measure prohibitin-2 and heat shock proteins after running experimental treatments. While GO terms and proteins are not correlated in this analysis, brewers may consult over-enriched categories to pick genes to modulate with the hopes of improving or controlling the yield of yeast proteins of interest. While yeast is not the only contributor to flavor, genetic improvement of yeasts paired with similar LC/MS approaches may spurn novel research outcomes in the field of beer proteomics, and perhaps influence paradigm changes in brewing globally. Previous beer proteomics research has established methods for preparation of lab-brewed samples.[5] Such methods may be modified to incorporate upstream genetic manipulation of brewing

yeast of choice for the subsequent characterization of protein content. In the same way that commercial breweries maintain lines of active yeast that were evolutionarily optimized via natural selection and cross breeding, breweries may leverage lab resources to improve yeast strains and ultimately beverage quality in a rapid, targeted fashion.

Though we were successful in making species-level protein identifications and drawing conclusions from protein abundance data for beverages, it is apparent that some of the aforementioned difficulties for this analysis introduced some error to our overall analysis. The initial concern related to the homology-related overlap of peptide identifications during our database search conducted in PD2.4. It is evident that proteolytic cleavage can be driven by barley proteins during mashing and yeast proteinase A during fermentation, so our starting samples contained many more polypeptides of shorter length due to hydrolysis. This means that after trypsin digestion, we would have reduced peptides again to units of smaller size and more similar sequence identity. The mass spectrometer would have read peptides of similar weight as simpler units of proteins belonging to species that were not used as ingredients (ie. tomato, pumpkin, vanilla). We believe that this issue skewed our analysis because proteins belonging to unrepresented species had more of a chance of being included in our final list of high-confidence proteins.

Another source of error may have come from the relative sizes of our UniProtKB species databases. High-confidence proteins in our analysis are determined by how many times particular peptide sequences are recognized by calculations based on the harmonic motion of peptides. In our analysis we noticed that lager yeasts only made up

only 7.7% of all yeast proteins (Figure 3A), and UniProtKB lists 6,721 proteins for ale yeast, while lager yeast only has 15. There were likely many lager yeast proteins included in Heineken and Stella Artois beers, however this public database does not include them. Future yeast proteomics efforts may be able to improve the characterization of lager yeast and subsequently improve assessments of yeast protein abundance.

As mentioned previously, future beer proteomics research would be aided by the improvement of protein database mapping. It is unclear how the ambiguity of peptide and ultimately protein identifications due to evolutionary and proteolytic homology can be resolved, so perhaps future studies should also characterize the beer *peptidome* to improve the resolution of this analysis. LTP-3 has been characterized in work studying the mash stage of brewing, however our conclusion that high levels of LTP-3 in sour ales may degrade beer foam motivates research to better understand its molecular interactions given cleavage of LTP-1 in mashing influences foam formation in bright beers. Future beer meta-proteomics studies of this kind should also include steps to measure beer quality parameters like density of gas and foam depth because of the importance of measuring these qualities to the beer industry, consumers, and previous beer proteomics research.

Just as our evaluation of commercial alcoholic beverages using LC/MS-based proteomic analysis is a relatively new research approach, so too would be the characterization of yeast protein abundance from beers brewed with genetically engineered yeast strains. We were able to highlight over-enriched biological groups that are mapped to known yeast genes, and future studies may seek to reinforce the

expression of desirable yeast proteins by modulating gene expression in living yeasts. Beer proteomics has established methods to brew and analyze small quantities of beer in the lab, and optimization via cloning may be coupled to optimize beer via proteome-driven hypothesis making. Synthetic biology is interested in optimizing the fermenting abilities of brewers yeast specifically to produce materials like plastics and biofuels, so perhaps similar approaches can be adopted in commercial brewing.

Finally, the broad characterization of flavorant and non-flavorant proteins in commercial beers is important in the realm of citizen science given the popularity of beer and homebrewing globally. Just as labs commonly sequence DNA, and people choose to submit saliva for genetic analysis by companies like 23andMe, improving the accessibility of beer proteomics research and identifying parameters with obvious connections to elements of beer and ingredient quality may lead to demand for proteomic characterization of the beers and ciders they brew and consume. The opportunities to expand on this research are endless, and collaboration across a host of technological disciplines will improve efforts to characterize the proteome of beer.

Supplement: *Supplemental Methods and Appendix*

Supplemental Methods:

Table 1: *Beverage samples, ingredients and ABV%, and measured concentrations*

Sample	Sample Name	Beverage Type	Fermentation	Location	Flavorants	Grains and Yeast	ABV%	Conc. (ug/uL)
MD13-1	ABRF PRG Standard Rep. 1	Ale	Top	CA, USA				0.42
MD13-2	ABRF PRG Standard Rep. 2	Ale	Top	CA, USA				0.42
MD13-3	NoDa Coco Loco Rep. 1	Ale	Top	NC, USA	Coconut, chocolate	Brown malt, chocolate malt	6.2	0.93
MD13-4	NoDa Coco Loco Rep. 2	Ale	Top	NC, USA	Coconut, chocolate	Brown malt, chocolate malt	6.2	0.88
MD13-5	BruePrint Briarberry Rep. 1	Ale	Top	NC, USA	Blackberry, raspberry, orange, lemon, bacteria		4.5	0.42
MD13-6	BruePrint Briarberry Rep. 2	Ale	Top	NC, USA	Blackberry, raspberry, orange, lemon, bacteria		4.5	0.41
MD13-7	Wicked Weed Milk & Cookies Rep. 1	Ale	Top	NC, USA	Cinnamon, golden raisins, vanilla	California ale yeast	8.5	1.1
MD13-8	Wicked Weed Milk & Cookies Rep. 2	Ale	Top	NC, USA	Cinnamon, golden raisins, vanilla	California ale yeast	8.5	1.06
MD13-9	Wicked Weed Pernicious IPA Rep. 1	Ale	Top	NC, USA		California ale yeast	7.3	0.46
MD13-10	Wicked Weed Pernicious IPA Rep. 2	Ale	Top	NC, USA		California ale yeast	7.3	0.47
MD13-11	Austin East Ciders Blood Orange Rep. 1	Cider		TX, USA	Orange, apple, raspberry	White wine yeast	5	0.1
MD13-12	Austin East Ciders Blood Orange Rep. 2	Cider		TX, USA	Orange, apple, raspberry	White wine yeast	5	0.09
MD13-13	Bull City Off Main Rep. 1	Cider		NC, USA	Apple	Gluten-free yeast	6	0.1
MD13-14	Bull City Off Main Rep. 2	Cider		NC, USA	Apple	Gluten-free yeast	6	0.1
MD13-15	NoDa Gordgeous Rep. 1	Ale	Top	NC, USA	Pumpkin		6.4	0.73
MD13-16	NoDa Gordgeous Rep. 2	Ale	Top	NC, USA	Pumpkin		6.4	0.78
MD13-17	ABRF PRG Standard Rep. 3	Ale	Top	CA, USA				0.47
MD13-18	ABRF PRG Standard Rep. 4	Ale	Top	CA, USA				0.43
MD13-19	Heineken (standard) Rep. 1	Lager	Bottom	Amsterdam		Heineken A-yeast, Irish malt, hops	5	0.37
MD13-20	Heineken (standard) Rep. 2	Lager	Bottom	Amsterdam		Heineken A-yeast, Irish malt, hops	5	0.37
MD13-21	Stell Artois Rep. 1	Lager	Bottom	Belgium		Stella Artois yeast	5	0.4
MD13-22	Stell Artois Rep. 2	Lager	Bottom	Belgium		Stella Artois yeast	5	0.41

Yeast and grain proteins are referred to as non-flavorant proteins in this analysis, as they are not explicitly listed in beverage ingredients by manufacturers. Ingredients and ABV% were unavailable for ABRF standard.

Table 2: *Reagent volumes for digestion of precipitated proteins*

Sample #	Sample	Sample Nar	Digested vo	Sample Typ	Acetone (uL)		8M Urea 50 mM ABC (uL)	100 mM DTT (uL)	375 mM IAA (uL)	ABC to add to dilute to <1M Urea (uL)	Trypsin 0.5ug/uL to add (uL)
1	MD13-17	ABRF PRG s	50	In-Solution	200		25	1.25	1.05	200	6.5
2	MD13-18	ABRF PRG s	50	In-Solution	200		25	1.25	1.05	200	6.5
3	MD13-19	Heineken st	50	In-Solution	200		25	1.25	1.05	200	6.5
4	MD13-20	Heineken st	50	In-Solution	200		25	1.25	1.05	200	6.5
5	MD13-21	Stella Artois	50	In-Solution	200		25	1.25	1.05	200	6.5
6	MD13-22	Stella Artois	50	In-Solution	200		25	1.25	1.05	200	6.5
										total volume (uL)	39

Table 3: *Volumes of diluent and sample to normalize peptide concentrations*

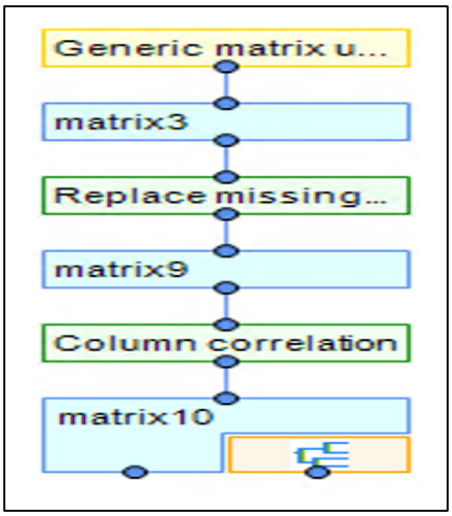
Sample	Beverage	Conc ug/uL	Normalize to 0.3 ug/ul - sample volume (ul)	Volume of diluent to add to get 20ul total volume
MD13-1	ABRF PRG standard Rep 1	0.42	14.2	5.8
MD13-2	ABRF PRG standard Rep 2	0.42	14.4	5.6
MD13-3	NoDa Coco Loco Porter Rep 1	0.93	6.5	13.5
MD13-4	NoDa Coco Loco Porter Rep 2	0.88	6.8	13.2
MD13-5	BruePrint Briarberry Rep 1	0.42	14.1	5.9
MD13-6	BruePrint Briarberry Rep 2	0.41	14.7	5.3
MD13-7	Wicked Weed Milk & Cookies Rep 1	1.1	5.4	14.6
MD13-8	Wicked Weed Milk & Cookies Rep 2	1.06	5.6	14.4
MD13-9	Wicked Weed Pernicious IPA Rep 1	0.46	13	7
MD13-10	Wicked Weed Pernicious IPA Rep 2	0.47	12.9	7.1
MD13-11	Austin East Ciders Blood Orange Rep 1	0.1	13	43.5
MD13-12	Austin East Ciders Blood Orange Rep 2	0.09	11.8	39.3
MD13-13	Bull City off Main Rep 1	0.1	12.6	42.1
MD13-14	Bull City off Main Rep 2	0.1	12.4	41.2
MD13-15	NoDa Gordgeous Rep 1	0.73	8.2	11.8
MD13-16	NoDa Gordgeous Rep 2	0.78	7.7	12.3
MD13-17	ABRF PRG standard Rep 1	0.47	12.8	7.2
MD13-18	ABRF PRG standard Rep 2	0.43	13.8	6.2
MD13-19	Heineken standard Rep 1	0.37	16.1	3.9
MD13-20	Heineken standard Rep 2	0.37	16.4	3.6
MD13-21	Stella Artois Rep 1	0.4	15	5
MD13-22	Stella Artois Rep 2	0.41	14.7	5.3

Table 4: *Protein species database names and sources*

Common Name	Species Name	Category	Reviewed?	Source
Ale Yeast	*Saccharomyces cerevisiae*	Yeast	Yes	UniProtKB
Lager Yeast	*Saccharomyces pastorianus*	Yeast	Yes	UniProtKB
Wine Yeast	*Saccharomyces bayanus*	Yeast	Yes	UniProtKB
Fission Yeast	*Schizosaccharomyces pombe*	Yeast	Yes	UniProtKB
Fungal Yeast	*Candida albicans*	Yeast	Yes	UniProtKB
Wheat	*Triticum aestivum*	Grain	Yes	UniProtKB
Corn	*Zea mays*	Grain	Yes	UniProtKB
Barley	*Hordeum vulgare*	Grain	Yes	UniProtKB
Rice	*Oryza sativa subsp. japonica*	Grain	Yes	UniProtKB
Apple	*Malus domestica*	Flavorant	Yes	UniProtKB
Chocolate	*Theobroma cacao*	Flavorant	Yes	UniProtKB
Cinnamon	*Cinnamomum verum*	Flavorant	Yes	UniProtKB
Coconut	*Cocos nucifera*	Flavorant	Yes	UniProtKB
European Hop	*Humulus lupulus*	Flavorant	Yes	UniProtKB
Grape	*Vitis vinifera*	Flavorant	Yes	UniProtKB
Pumpkin	*Cucurbita maxima*	Flavorant	Yes	UniProtKB
Raspberry	*Rubus idaeus*	Flavorant	Yes	UniProtKB
Vanilla	*Vanilla planifolia*	Flavorant	No	UniProtKB

Protein species databases organized by yeast, grain or flavorant.

Figure 1: *Perseus data import instructions*



Nodes were selected to replace missing values in original protein abundance matrix, and then to calculate column correlations according to replicate protein abundances. Heatmap of conditionally formatted R values with hierarchical clustering along axes.

Figure 2: *GO term over-enrichment analysis instructions*

1. Select protein lists of interest and upload accessions on Uniprot Retrieve ID/Mapping website. Convert UniprotKB AC/ID protein accessions to Entrez Gene (GeneID). Duplicate values will not influence gene list output.
    a. https://www.uniprot.org/uploadlists/
2. Save GeneID terms for each beverage or beverage category (ie. ales, lagers, ciders).
3. Upload each GeneID term list separately on DAVIDBioinformatics website. Ale yeast (*Saccharomyces cerevisiae)* and fission yeast will automatically be selected as background species. Run analysis.
    a. https://david.ncifcrf.gov/tools.jsp
4. Select the General Ontology tab on the right when the analysis concludes and then select GO BP Direct terms. Download the data for this term group.
    a. http://amigo.geneontology.org/amigo/search/ontology (*For gene information corresponding to over-enriched GO BP terms*)
5. Organize log-fold enrichment (ie. log-fold, p-value, Benjamini) data as desired. For our analysis we chose Benjamini correct p-values to visualize.
6. Import excel into Morpheus website and export heatmap.
    a. https://software.broadinstitute.org/morpheus/

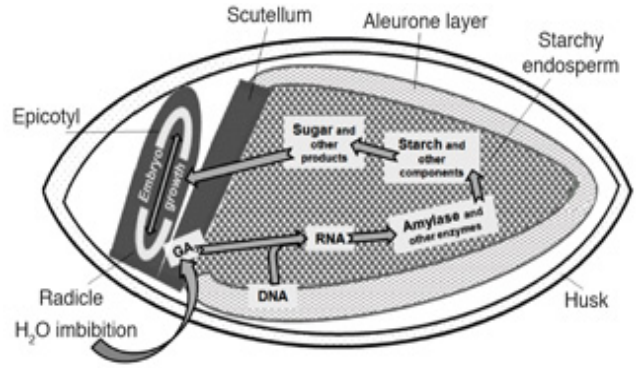
Appendix:

Table 1: *Literature barley proteins and roles in beer foam formation and stabilization*

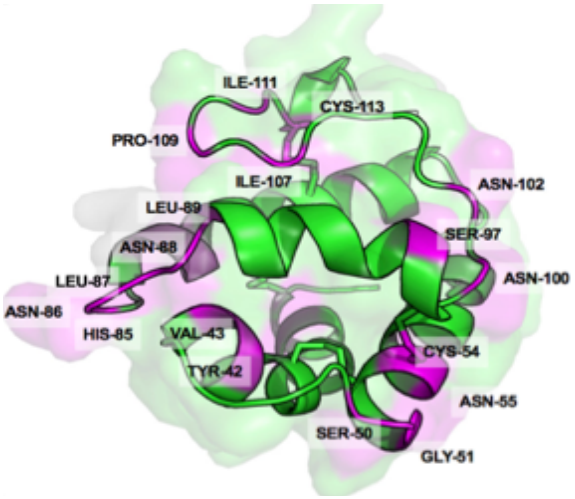
Protein	Accession	Role in beer
B1-hordein OS	P06470	Storage proteins that are soluble when hydrolyzed, foam production
B3-hordein (Fragment) OS	P06471	Storage proteins that are soluble when hydrolyzed, foam production
Gamma-hordein-3 OS	P80198	Storage proteins that are soluble when hydrolyzed, foam production
Gamma-hordein-1 OS	P17990	Storage proteins that are soluble when hydrolyzed, foam production
Non-specific lipid-transfer protein 1 OS	P07597	Degraded by yeast proteinase A, foam stabilization
Probable non-specific lipid-transfer protein OS	P20145	Degraded by yeast proteinase A, foam stabilization
Non-specific lipid-transfer protein 3 OS	Q43766	Degraded by yeast proteinase A, foam stabilization
Non-specific lipid-transfer protein Cw18 OS	Q43871	Degraded by yeast proteinase A, foam stabilization
Serpin-Z4 OS	P06293	Serine-like protease, foam stabilization
Serpin-ZX OS	Q40066	Serine-like protease, foam stabilization
Serpin-Z7 OS	Q43492	Serine-like protease, foam stabilization

Barley proteins identified in previous literature with known contributions to beer quality. Proteins are grouped into three categories: hordeins, LTPs, and protein Zs. Hordein proportions were compared to previous beer proteomics trends to make predictions about ingredient quality. (Blasco et al.)

Figure 1: *Diagrams of barley germination (malting) and proteolytic cleavage sites*



a.



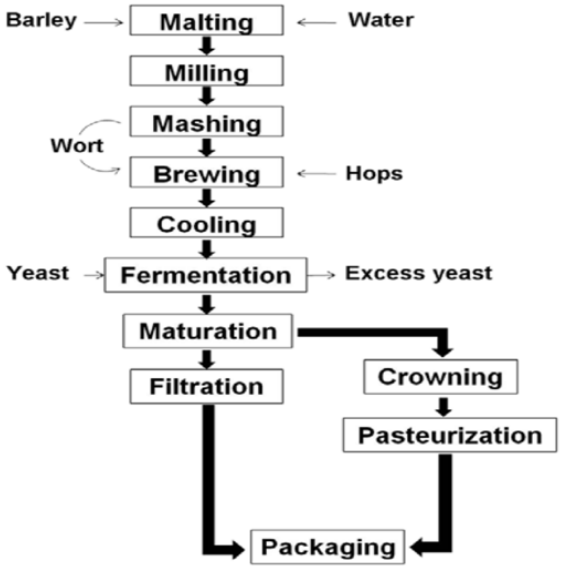
b.

c.

```
                                                    30
MARAQVLLMA AALVLMLTAA PRAAVALNCG

                                                    60
QVDSKMKPCL TYVQGGPGPS GECCNGVRDL

                                                    90
HNQAQSSGDR QTVCNCLKGI ARGIHNLNLN

NAASIPSKCN VNVPYTISPD IDCSRIY
```

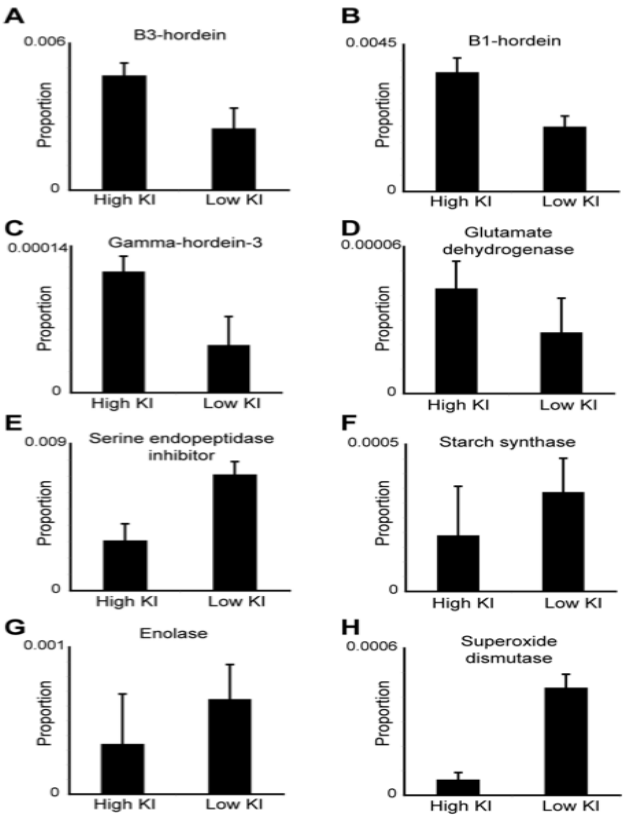
A) Diagram of enzyme and starch production proceeding during barley germination. Source
B) Cartoon X-ray crystal structure of commonly cleaved LTP-1 protein. (Kerr et al.) Cleavage site sequences (underlined not included in B) (Kerr et al.)

Figure 2: *Schematic of beer brewing process*



Hydrolytic activity among barley proteins during mashing, hydrolytic activity among barley and yeast proteins during fermentation, and further denaturation effects driven by temperature influence protein expression in commercial beers and ciders. (Blasco et al.)

Figure 3: *Barley protein abundance proportions corresponding to malting (KI) index*



Selected proteins with significant differences in relative abundance in bright beer made from high or low KI malt. (*Schulz, Phung et al.*)

Acknowledgements: *Thank you to Dr. Laura Herring, Alex Prevatte, and Dr. Allie Mills from my lab in the Michael J. Hooker Proteomics Research Center for mentorship and guidance, my family and friends for support, my thesis group, and Dr. Maddox for suggestions and feedback.*

Works Cited:

1. Asano, K., Shinagawa, K., & Hashimoto, N. (1982). Characterization of haze-forming proteins of beer and their roles in chill haze formation. *Journal of the American Society of Brewing Chemists*, *40*(4), 147–154. https://doi.org/10.1094/asbcj-40-0147
2. Berner, T., Jacobsen, S., & Arneborg, N. (2013). The impact of different ale brewer's yeast strains on the proteome of immature beer. *BMC Microbiology*, *13*(1), 215. https://doi.org/10.1186/1471-2180-13-215
3. *The Brewing Process.* Aslan Brewing Co. (n.d.). https://aslanbrewing.com/thebrewingprocess.
4. Kerr, E. D., Caboche, C. H., & Schulz, B. L. (2019). Posttranslational Modifications Drive Protein Stability to Control the Dynamic Beer Brewing Proteome. *Molecular & Cellular Proteomics*, *18*(9), 1721–1731. https://doi.org/10.1074/mcp.ra119.001526
5. Kerr, E., McDiarmid, D., Fraser, J., & Schulz, B. (2018). Cell Wall and Whole Cell Proteomes Define Flocculation and Fermentation Behavior of Yeast. *Fermentation*, *4*(3), 80. https://doi.org/10.3390/fermentation4030080
6. Steen, H., & Mann, M. (2004). The abc's (and xyz's) of peptide sequencing. *Nature Reviews Molecular Cell Biology*, *5*(9), 699–711. https://doi.org/10.1038/nrm1468
7. Schulz, B. L., Phung, T. K., Bruschi, M., Janusz, A., Stewart, J., Meehan, J., … Vickers, C. E. (2018). Process Proteomics of Beer Reveals a Dynamic Proteome with Extensive Modifications. *Journal of Proteome Research*, *17*(4), 1647–1653. https://doi.org/10.1021/acs.jproteome.7b00907
8. Blasco, L., Viñas, M., & Villa, T. G. (2011). Proteins influencing foam formation in wine and beer: the role of yeast. *International microbiology : the official journal of the Spanish Society for Microbiology*, *14*(2), 61–71. https://doi.org/10.2436/20.1501.01.136
9. Leisegang, R., & Stahl, U. (2005). Degradation of a Foam-Promoting Barley Protein by a Proteinase from Brewing Yeast. *Journal of the Institute of Brewing*, *111*(2), 112–117. https://doi.org/10.1002/j.2050-0416.2005.tb00656.x
10. Gonzalez Viejo, C., Caboche, C. H., Kerr, E. D., Pegg, C. L., Schulz, B. L., Howell, K., & Fuentes, S. (2020). Development of a Rapid Method to Assess Beer Foamability Based on Relative Protein Content Using RoboBEER and Machine Learning Modeling. *Beverages*, *6*(2), 28. https://doi.org/10.3390/beverages6020028
11. Kato, M., Kamada, T., Mochizuki, M., Sasaki, T., Fukushima, Y., Sugiyama, T., … Imai, T. (2020). Influence of high molecular weight polypeptides on the mouthfeel of commercial beer. *Journal of the Institute of Brewing*, *127*(1), 27–40. https://doi.org/10.1002/jib.630