# Systematic profiling of ale yeast protein dynamics across fermentation and repitching

Riddhiman K. Garge [1,2], Renee C. Geck [1], Joseph O. Armstrong [1], Barbara Dunn [1], Daniel R. Boutz [2,3], Anna Battenhouse [2], Mario Leutert [1,4], Vy Dang [2], Pengyao Jiang [1], Dusan Kwiatkowski [5], Thorin Peiser [5], Hoyt McElroy [5], Edward M. Marcotte [2,*], Maitreya J. Dunham [1,*]

[1]Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA
[2]Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX 78712, USA
[3]Antibody Discovery and Accelerated Protein Therapeutics, Department of Pathology and Genomic Medicine, Houston Methodist Research Institute, Houston, TX 77030, USA
[4]Institute of Molecular Systems Biology, ETH Zürich, Zürich 8049, Switzerland
[5]Live Oak Brewing Company, Austin, TX 78617, USA

*Corresponding author: Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX 78712, USA. Email: marcotte@utexas.edu; *Corresponding author: Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA. Email: maitreya@uw.edu

Studying the genetic and molecular characteristics of brewing yeast strains is crucial for understanding their domestication history and adaptations accumulated over time in fermentation environments, and for guiding optimizations to the brewing process itself. *Saccharomyces cerevisiae* (brewing yeast) is among the most profiled organisms on the planet, yet the temporal molecular changes that underlie industrial fermentation and beer brewing remain understudied. Here, we characterized the genomic makeup of a *Saccharomyces cerevisiae* ale yeast widely used in the production of Hefeweizen beers, and applied shotgun mass spectrometry to systematically measure the proteomic changes throughout 2 fermentation cycles which were separated by 14 rounds of serial repitching. The resulting brewing yeast proteomics resource includes 64,740 protein abundance measurements. We found that this strain possesses typical genetic characteristics of *Saccharomyces cerevisiae* ale strains and displayed progressive shifts in molecular processes during fermentation based on protein abundance changes. We observed protein abundance differences between early fermentation batches compared to those separated by 14 rounds of serial repitching. The observed abundance differences occurred mainly in proteins involved in the metabolism of ergosterol and isobutyraldehyde. Our systematic profiling serves as a starting point for deeper characterization of how the yeast proteome changes during commercial fermentations and additionally serves as a resource to guide fermentation protocols, strain handling, and engineering practices in commercial brewing and fermentation environments. Finally, we created a web interface (https://brewing-yeast-proteomics.ccbb.utexas.edu/) to serve as a valuable resource for yeast geneticists, brewers, and biochemists to provide insights into the global trends underlying commercial beer production.

Keywords: proteomics; genomics; systems biology; brewing yeast; mass spectrometry; beer; fermentation

## Introduction

*Saccharomyces cerevisiae*, or brewing yeast, is used in a wide range of commercial processes including beverage fermentation, baking, biofuel generation, and pharmaceutical manufacture. The ease of culture and genetic manipulation of *S. cerevisiae* has also made it one of the most profiled organisms in academic research, including being the first eukaryote to have its genome sequenced (Goffeau *et al.* 1996). Brewing yeast has served as a model organism to understand fundamental cellular and molecular processes and continues to provide valuable insights into human health and disease (Cherry *et al.* 1998; Balakrishnan *et al.* 2012; Amorosi *et al.* 2021; Garge *et al.* 2021; Kachroo *et al.* 2022; Turco *et al.* 2023). Despite its prevalence in laboratory settings, the systematic profiling of yeast in commercial contexts is less common.

These commercial contexts are relevant not just for understanding more about the natural history of yeast, but also due to

their economic and cultural importance. The global beer market was worth $744 billion in 2020 and is expected to grow to $768 billion in 2023 ("Beer Market Size, Industry Share, Growth Rate, Forecast by 2030"). While research on brewing yeasts has been performed over many decades, new techniques—especially in genomics, proteomics, and related high throughput profiling—allow us to better understand the genetic and molecular changes that have occurred in yeast during brewery fermentations.

Throughout brewing and fermentation, yeast must cope with diverse and changing stressors, including fluctuations in nutrient and ethanol levels, oxygen availability, and temperature (Gibson *et al.* 2007). At the start of the brewing process, yeast is inoculated into a fermentation vessel containing aerated wort, a cooled aqueous extract containing the sugars from boiled malted grains along with the aromatic and bittering compounds from hops. Once the yeast cells adapt to the new environment during the lag phase, they begin to grow exponentially and rapidly deplete the available

oxygen creating an anaerobic environment. Along with oxygen, sugars and other essential nutrients are depleted, with concomitant production of ethanol, all of which stress the yeast over the course of fermentation. As the yeast adapt to these successive stressors during the brewing process, cells rapidly shift their gene expression profiles, leading to changes in protein and metabolite levels that help them survive (Gibson *et al.* 2007).

High-throughput approaches for parallel cell-wide measurement of different classes of cellular molecules, such as DNA sequencing (genome), mRNA sequencing (transcriptome), protein profiling (proteome) and metabolite profiling by mass spectrometry (metabolome), offer a path to deeper understanding of how yeast cells respond to the brewing environment. Hundreds of brewing strains, including both *S. cerevisiae* ale yeasts and *Saccharomyces pastorianus* lager yeasts (which are interspecific hybrids between *S. cerevisiae* and *Saccharomyces eubayanus*) have had their full genomes sequenced, yielding insights into domestication history and genetic differences related to flavor and style (Gallone *et al.* 2016, 2018, 2019). Gene expression profiles during growth in wort have been characterized for ale yeasts and lager yeasts in commercial brews and in small wort fermentations (Wang *et al.* 2019; Li *et al.* 2022). Investigators have also characterized ale yeasts' proteomes (Berner *et al.* 2013) and metabolomes (Kang *et al.* 2019; Li *et al.* 2022) during wort fermentation, although high-throughput proteomics based studies are infrequent. The majority of these previous brewing yeast multi-omics studies focused on the analysis of the proteins and metabolites known to be involved in the production of flavor components such as esters and higher alcohols by sampling beer or wort. However, few studies have sampled the actual brewing yeast populations to understand the global aspects of how yeast proteomes temporally change over time during commercial fermentation.

One common commercial brewing condition for which there have been few comprehensive ale yeast studies is "serial repitching", a process in which brewers collect yeast cells at the end of a fermentation cycle and use it to inoculate (or "pitch") a new batch. Serial repitching is mainly done for ease and efficiency of brewing and preserving the sensory or taste profile of the beer. The number of repitches varies across breweries, type of fermentation, and strain of yeast used. However, excessive rounds of repitching can adversely affect yeast fermentation and the taste profiles of the final product (Kalayu 2019). Serial repitching results in repeated exposures to physical, biological, and chemical stresses, which can lead to both reversible and irreversible damage to the yeast cells, with progressive loss in cell viability occurring with increasing pitch number (Kalayu 2019). Plasma membrane damage and stress-activated gene expression programs that result in increased glycogen accumulation and intracellular trehalose levels increase with each subsequent repitch (Smart and Whisker 1996). Changes in flavor and aroma profiles, decreased viability of yeast cells, and the increased likelihood of undesired microbial contamination are among the major reasons that brewers stop repitching after about 8–10 batches and restart brewing with a fresh yeast culture.

Genetic changes in the yeast during repitching can contribute to the altered profiles and viability. Previous studies have characterized mutations and traits that accumulate in ale yeast over repitching (Smart and Whisker 1996; Gibson *et al.* 2007; Bühligen *et al.* 2013; Large *et al.* 2020; Krogerus *et al.* 2021). For instance, 1 study observed repeated changes in copy number of chromosome V and mitotic recombinations that changed allele balance and subtelomeric gene copy number at regions on chromosomes VIII and XV (Large *et al.* 2020). Changes in yeast traits such as flocculation were also observed, though not linked definitively to these

mutations. However, changes in gene expression, protein levels, and metabolite abundance were not measured.

While there have been many ecological studies tracing the origin, evolution, and physiology of fermentation and brewing strains using genomics (Gallone *et al.* 2016, 2018, 2019), few have focused on molecular changes at the protein level and across brewing cycles (Picariello *et al.* 2012; Grochalová *et al.* 2017; Schulz *et al.* 2018; Kerr *et al.* 2019). To gain a more comprehensive understanding of the molecular changes associated with brewing, we characterized the genetic features of a Hefeweizen ale yeast and measured temporal proteomic changes across 2 fermentation cycles separated by serial repitching. Across 64,740 protein abundance measurements, we found many processes altered over fermentation in both time courses: In particular, we observed drastic changes in yeast proteomes across the first 2 days of fermentation largely dominated by ribosome biogenesis and translation. Additionally, we cataloged unique changes between the 2 fermentation batches, observing that lipid and sterol biosynthetic processes were upregulated in the later batch. This dataset serves as a foundational resource to finely characterize the molecular changes underlying commercial ale fermentation and offers a starting point to perturb, modify, or engineer flavor and strain characteristics in commercial and craft brewing settings.

## Methods

### Strain

Wyeast 3068, Weihenstephan Wheat yeast.

### Hefeweizen brewing

Commercial fermentation with Wyeast 3068 was conducted at Live Oak Brewing Company, Austin, Texas, USA. For batch 1, 32 L of yeast cultured on rich media supplied by Wyeast was inoculated into 100 gal of wort; 24 hours later, this was inoculated into 60 bbl of wort. Fermentation proceeded at 20°C in a horizontal tank for 4 to 5 days, after which it was cold crashed via glycol jacket heat exchange to 4.4°C to sediment the yeast. The beer was then separated from the yeast and moved to a conditioning tank to be held at cold temperatures for 2 weeks. Directly after this transfer, the yeast from the fermenter is harvested to be used in the subsequent batch, pitched either that day or the following day. Yeast is harvested manually from the horizontal tank bottom with a large food-grade hoe, gathered from the outlet into a stainless bucket, and poured into collection brinks. The brinks are weighed along with recording the cell counts and viability of the slurry. The same target number of viable cells is pitched into the fermenter inline with the wort stream. After the fermenter is filled with wort and yeast, a final pitch cell count is taken for the beginning of fermentation. Fermentations for batches 2–14 (between batches 1 and 15) occurred over the same 4 to 5-day duration, but now fermenting 180 bbl with 65 L of yeast. All batches were used to produce Hefeweizen ale and both batches sampled for proteomics were subjected to the same wort and fermentation conditions.

### Whole genome sequencing

For genomic DNA isolation, a glycerol stock of Wyeast 3068 strain received from the brewers was streaked onto rich media (YPD) agar plates and incubated for 3–5 days. A single colony was isolated and cultured for DNA isolation. Genomic DNA extraction was performed using Zymo YeaStar Genomic DNA Isolation kit as per manufacturer recommendations. Genomic DNA was sheared to an average length of 410 bp before sequencing on an

Illumina NextSeq 500 platform. Sequences were first analyzed using FastQC to assess overall quality. The 1.6 million read pairs were then mapped to the *Saccharomyces cerevisiae* S288C reference genome R64 (sacCer3) using bowtie2 (v 2.2.6) in local alignment mode. We observed a high overall mapping rate of 92% with low rates of duplication (0.5%) and reads mapping to multiple locations (4.2%). A moderate rate of indel detection (13.4%) and the relatively low 68.8% of reads mapped as proper pairs suggested this commercial yeast differs from the lab strain in some structural ways. Sequencing reads are deposited in the NCBI Sequence Read Archive (SRA) under BioProject PRJNA1011390.

## Copy number analysis

Copy number variation was calculated by read coverage over 1,000-bp sliding windows, as described in a previous study (Large *et al.* 2020). Single-gene deletions were identified using CNVnator (Abyzov *et al.* 2011) and confirmed by viewing bam files in IGV (Robinson *et al.* 2011). The list of affected genes was input into the SGD Gene Ontology Term Finder version 0.86 (https://www.yeastgenome.org/goTermFinder) to identify process ontology aspects.

## Loss of heterozygosity (LOH) analysis

To detect genomic regions of the brewing yeast that experienced the loss of heterozygosity (LOH), we performed variant calling and allele frequency analyses using the CLC Genomics Workbench 11.0 NGS toolkit platform (Qiagen). After importing paired-end Illumina FASTQ read files, Nextera adapters were trimmed, and the trimmed reads were then mapped to the *S. cerevisiae* R64 (sacCer3) reference genome, with no masking. Mapping was random, using the following parameters: mismatch cost 2, insertion and deletion cost 3, length fraction 0.5, and similarity fraction 0.8. We then used Workbench's Basic Variant Detection 2.0 program to identify variants and to calculate the frequency of each variant allele, using the following parameters: ignore broken pairs; ploidy 4; minimum coverage 10; minimum count 2; minimum frequency 0%. Variants were spot-checked on the pile-up viewer of the mapped reads, with all confirmed to be correctly called. After removing the called variants for the mitochondrial genome, the resulting output file was ordered by chromosome number, and whole-genome (concatenated) positions were assigned for each variant. The allele frequencies of the variants were then plotted against the whole-genome position (Supplementary Fig. 1b).

## Sample preparation for proteomics

Yeast proteins for mass spectrometry were isolated using a previously described protocol (Abdullah *et al.* 2023). Briefly, samples were obtained directly from beer fermentation tanks at Live Oak Brewing Company in Austin, Texas, USA. For every time point, 2 replicates were independently sampled from the fermentation tank by collecting 1 L per replicate. Cell pellets were harvested by centrifuging beer at 8,000 g for 5 minutes followed by 2–3 washes in ice cold PBS. Cell pellets were resuspended in Digestion Buffer (50-mM Tris, 2-mM CaCl$_2$) and lysed by bead beating with glass beads for 1 minute cycles repeated 3 times. The whole-cell lysate was separated from the beads and then mixed 1:1 with 2,2,2-trifluoroethanol (TFE). Samples were then reduced by incubation with 5-mM tris(2-carboxyethyl)phosphine (TCEP solution, Pierce) at 60°C for 40 minutes. Reduced samples were alkylated by incubation with 15-mM iodoacetamide at room temperature for 30 minutes. Excess iodoacetamide was quenched by the addition of 7.5-mM dithiothreitol (DTT). Following quenching, samples were diluted 10-fold using digestion buffer and subjected to proteolytic digestion with 2-μg trypsin for 5 hours at 37°C. Tryptic digestion was quenched with 1% formic acid, and samples were concentrated using vacuum centrifugation to reduce the total sample volume to less than 300 μL. Digested samples were cleaned using HyperSep C18 SpinTips (Thermo) according to the manufacturer's protocol. Eluted peptides were briefly dried by vacuum centrifugation, then resuspended in 5% acetonitrile and 0.1% formic acid.

## LC-MS/MS analysis

Tryptic peptides were separated by reverse phase chromatography on a Dionex Ultimate 3000 RSLCnano UHPLC system (Thermo Scientific) with an Acclaim C18 PepMap RSLC column using a 3–42% acetonitrile gradient over 60 minutes. Peptides were eluted directly into a Thermo Orbitrap Lumos mass spectrometer by nano-electrospray. Data-dependent acquisition (DDA) was applied, with precursor ion scans (MS1) collected by FTMS at 120,000 resolution and HCD fragmentation scans (MS2) collected in parallel by ITMS with 3-s cycle times. Monoisotopic precursor selection and charge-state screening were enabled, with ions > +1 charge selected. Dynamic exclusion was applied to selected ions ± 10 ppm for 30 s. Raw mass spectrometry data have been deposited on MassIVE (MSV000092793).

## Proteome database searching and analyses

Raw mass spectrometry data were processed using Proteome Discoverer 2.2, MaxQuant, or converted to mascot generic files (.mgf) using MSConvert to be analyzed by SearchGui and PeptideShaker. Mass spectra were searched against a protein sequence database containing reversed decoy sequences comprising the *S. cerevisiae* reference proteome (UniProt OX: 559292) and a list of common protein contaminants (MaxQuant). All searches were restricted to fully tryptic peptides only, allowing up to 2 missed cleavages. A precursor tolerance of 5 ppm and fragment mass tolerance of 0.5 Da were used. Static modifications of carbamidomethyl cysteine and dynamic modifications of oxidized methionine and protein N-terminal acetylation and/or methionine-loss were considered. High-confidence peptide spectrum matches (PSMs), peptides, and proteins were all filtered at a false discovery rate of <1%.

Protein abundances were calculated using 3 different metrics: (1) estimating parts per million abundance from PSM counts (Weiss *et al.* 2010), (2) intensity-based absolute quantification (iBAQ), and (iii) label-free quantitation (LFQ) (Bantscheff *et al.* 2007; Mueller *et al.* 2008; Tyanova *et al.* 2016). For ppm-based protein abundance, mascot generic files were then searched against MS-GF+, OMSSA, and X!Tandem databases with default settings for each database using SearchGUI version 3.2.20 (Barsnes and Vaudel 2018) and data were analyzed using PeptideShaker version 1.16.12 (Vaudel *et al.* 2015). PeptideShaker report files were parsed to generate a matrix of unique validated PSMs for each protein across the fermentation time course. To normalize the PSM counts, we converted the count matrix for a given protein into parts per million (ppm) using an approach previously described (Weiss *et al.* 2010; McWhite *et al.* 2020). Briefly, unique peptides were trypsin digested in silico and filtered to only those 7–40 amino acids in length. Next, a correction factor was calculated from the sum of the total length of peptides in this range per protein. Detected peptide PSMs were multiplied by the peptide length, and correction factor, multiplied by 1,000,000, and divided by the experiment total to get parts per million. iBAQ and LFQ-based protein abundances were calculated using MaxQuant (Tyanova *et al.* 2016) based on extracted ion-chromatography (XIC) feature intensities.

To compute similarity matrices across time points, pairwise Pearson correlation was calculated between each pair of sampled time points. Since the LFQ values had the highest correlation between replicates and are corrected for technical variation between samples, we proceeded with LFQ abundances for all subsequent analyses. LFQ intensities for a protein were across technical replicates to maximize the number of proteins detected across all samples. To resolve lack of detection vs lack of protein expression, we additionally compared the abundance (Supplementary Fig. 2e) estimated as mean molecules per cell from a previous study (Ho *et al.* 2018) of proteins detected in our dataset to those that were not. We found that in very few cases molecules less than 1,000 copies per cell were detected from fermentation samples.

## Clustering, differential expression, and GO term analysis

Raw LFQ intensity matrices from the MaxQuant output were filtered to ensure that each protein was detected with at least 2 unique peptides. Since LFQ intensity values are log-normally distributed, we log$_2$-transformed the data to obtain a normal distribution of intensities and median centered to ensure equal variances across samples. Only proteins that were detected in both replicates of a particular time point were considered for downstream analyses. Pairwise significance testing across time points was performed using a standard *t*-test with *P*-values adjusted using the Benjamini–Hochberg (BH) procedure with a false discovery rate of 5%. Enriched proteins were further filtered to a log$_2$ fold change of 1 or greater. The above procedures were carried out on the online version of ProVision (Gallant *et al.* 2020).

For hierarchical clustering, log$_2$ protein abundances were normalized to their mean abundance across all time points (Supplementary Table 6) and proteins that did not change more than 2-fold over the mean in any time point were filtered out. Hierarchical row clustering was performed by Morpheus (https://software.broadinstitute.org/morpheus) using an average linkage method. We identified 8 clusters at a node depth of 3 that contained 36–748 proteins. Clusters with over 500 proteins were further subset to a node depth of 5, and resultant subclusters contained 59–387 proteins (Supplementary Table 7). GO enrichment analysis was performed using clusterProfiler (Wu *et al.* 2021) for each cluster and sub-cluster to determine significantly enriched proteins by comparing protein changes in a pairwise manner across all time points. GO term enrichments were filtered to a multiple hypothesis testing FDR of 5% and BH adjusted *P*-value of less than 0.05 (Supplementary Table 8).

## Metabolic pathway analysis

Prediction of metabolic pathways affected by changes in protein levels over time across both batches was performed using the cellular overview tool on SGD Yeast Pathways. Log$_2$ protein abundances normalized to row mean (Supplementary Table 6) with gene name were uploaded onto the Omics Viewer, and output set to a table of 100 top-scoring pathways (Supplementary Table 13). Differential Pathway Perturbation Scores (DPPS) were calculated as described in the tool. Briefly, a differential reaction perturbation score (DRPS) is calculated from the maximum of differences between samples for all entities associated with a given reaction. For a pathway, the square of each DRPS is summed and divided by the total number of reactions in the pathway. The DPPS is the square root of this value, representing the maximal perturbation of the pathway between samples.

## Pathway, protein complex, and subcellular location analysis

Yeast pathways were curated from SGD (Cherry *et al.* 1998, 2012). Yeast protein complexes were obtained from the EBI complex portal (Meldal *et al.* 2019). For every protein complex, the fraction of members detected was computed (Supplementary Table 10). Since the EBI complex portal manually curates protein complexes across the multiple protein databases, there existed multiple versions or variants of a particular complex. In these cases, for our analyses, we considered these to be discrete complexes. We restricted our analysis to proteins that were detected in 2 or more fermentation time points. From this filtered set we next calculated Pearson correlation across all the fermentation time points for all possible protein pairs in the dataset and subset them based on whether they were interacting or not. Protein subcellular localization data were curated from the yeast GFP fusion localization database (Huh *et al.* 2003). In cases where a protein localized to more than 1 location, these proteins were annotated in both subcellular locations or organelles to prevent loss of information (Supplementary Table 9). For each subcellular compartment, we similarly calculated the pairwise Pearson correlation across all proteins localized to a given subcellular location.

# Results
## Hefeweizen yeast shares genetic characteristics with other ale brewing strains

In this study, we used Wyeast 3068, a Weihenstephan Weizen tetraploid S. *cerevisiae* ale strain (Adamczyk *et al.* 2016) frequently used to brew Hefeweizen style beers. This popular German style wheat beer strain ("Strain") has previously been whole genome sequenced and ecologically annotated (Gallone *et al.* 2016, 2019), allowing for comparisons against existing resources. Moreover, the relatively short fermentation cycles in Hefeweizen brewing (4–6 days) allowed us to sample yeast populations across brewing cycles separated by 14 cycles of serial repitching. We picked 14 repitches because it is around this stage that changes in flavor and strain fermentation performance become noticeable. Importantly, since repitched yeast cells are mainly collected from the bottom of the fermentation tank and subsequently propagated, there is a tendency to select flocculating populations.

We confirmed the identity and analyzed the genetic characteristics of our strain using whole genome sequencing and alignment to the S288C *Saccharomyces cerevisiae* genome. Our data demonstrated mostly even coverage across the yeast genome (Supplementary Fig. 1a), with depth differences corresponding to copy number changes (Supplementary Fig. 1b). Several genes found in the reference laboratory strain were deleted in Wyeast 3068 (Supplementary Table 1). Deleted genes were enriched for processes of flocculation ($P \leq 4.59 \times 10^{-5}$, *FLO5*, *FLO9*, *FLO1*, *FLO10*), carbohydrate transmembrane transport ($P \leq 2.99 \times 10^{-3}$, *HXT15*, *HXT16*, *MPH2*, *MPH3*, *AQY3*), asparagine catabolism ($P \leq 3.34 \times 10^{-6}$, *ASP3-3*, *ASP3-2*, *ASP3-1*, *ASP3-4*), and transposition ($P \leq 2.30 \times 10^{-4}$, *YIL082W-A*, *YPL06°C-A*, *YLR157C-A*, *YGR109W-A*, *YGR109W-B*, *YJL113W*, *YJL114W*, *YLR157C-B*), and some of these genes have previously been found to be deleted in ale strains (Gallone *et al.* 2016, 2019). Also like other ale strains, Wyeast 3068 was largely tetraploid with several chromosomal copy gains and losses, most notably possessing an extra copy of chromosome V and losing a copy of chromosome X (Supplementary Fig. 1b). Additionally, a majority of the chromosomes contained genomic regions that had not

changed in copy number, but had experienced a LOH (Supplementary Fig. 1c), as seen in other ale strains (Saada *et al.* 2022).

## A global view of temporal protein changes across brewing

To understand protein dynamics during brewing cycles, we periodically sampled yeast populations directly from the fermentation tank over a brewing time course inoculated with a freshly prepared stock of yeast (hereby termed batch 1). To study the effects of repitching on yeast proteomes, we similarly sampled yeast after 14 repitches (batch 15), with each pitch (fermentation) using a similar Hefeweizen wort composition. As typical for brewery repitching, the yeast cells from the previously finished fermentation were directly inoculated into the subsequent fermentation without any outgrowth. For batch 1 and batch 15, we collected samples representing near matched time points across the 4 days of brewing. To achieve this, we collected beer and isolated yeast populations directly from the fermentation tank. We also collected samples 24 hours post-crash (PC), when the fermentations were chilled or "cold-crashed" to sediment the yeast (Fig. 1). After the cold crash, the majority of the yeast cells were separated from the beer. The beer was moved to a conditioning tank to proceed with maturation where some additional yeast was sedimented (not to be harvested) and the beer flavor matured and turbidity was homogenized. The majority of yeast cells that were harvested from the cold crash were held at 4.4°C for 1–2 days until they were used to inoculate the next batch. For batch 15, we additionally sampled the residual yeast in the conditioning tank over 3 days after the batch 15 fermentation cycle, hypothesizing that the proteomes of yeast in the conditioning tank would represent different physiological states to those in the fermentation tank. In total, we sampled 17 time points: 7 from batch 1 and 10 from batch 15 (Supplementary Table 2). Independent duplicate samples were collected and processed from each time point.

We then performed high-throughput shotgun mass spectrometry on the yeast sampled from each batch (see *Methods*) to generate proteomic snapshots of all assayed yeast proteins and their abundance over rounds of brewing. For every protein, we computed 3 abundance metrics: ppm (Weiss *et al.* 2010), iBAQ, and LFQ values (Bantscheff *et al.* 2007; Mueller *et al.* 2008; Tyanova *et al.* 2016) (see *Methods*). Since the LFQ values had the highest correlation between replicates and were corrected for technical variation between samples, we proceeded with LFQ abundances for all subsequent analyses.

In total, we made 64,740 protein abundance measurements for over 2,600 proteins detected in at least 1 time point, out of a total of 5,610 possible experimentally detectable yeast proteins (Gao *et al.* 2021). We compared the correlation between replicates for each time point by plotting the LFQ intensity values for every protein detected in that time point and found that replicates of a particular time point in 1 batch were highly correlated, with Pearson correlation coefficients ranging from 0.84 to 0.94 (Supplementary Fig. 2a). We generated a similarity matrix by performing all-by-all Pearson correlation calculations across the 34 samples and observed that replicates of a single time point within a particular batch were more correlated to each other than to any other time point (Supplementary Fig. 2b). We observed 3 groups with high correlation: 2 composed of time points within batch 1 and batch 15, respectively, and a third group of matched time points across both batches. Despite high correlation, we found that the replicates of the earliest samples, 6 hours from batch 1 and 3 hours from batch 15, were the least correlated with other time points

assayed, suggesting that the proteome profiles of the earliest fermentation time points were most different compared to the rest of the fermentation.

Since DNA copy number differences can alter protein levels (Dephoure *et al.* 2014), we investigated whether the observed copy number alterations affected initial protein levels. As expected, the protein products of genes that were deleted (Supplementary Table 1) were not detected at the protein level. Among detected proteins, there was no correlation between read coverage and protein abundance at the initiation of batch 1 on a per-gene/protein basis (Supplementary Fig. 1d, Supplementary Table 3, adjusted $R^2 = 0.00008246$, $P \leq 0.2703$), nor by comparing the average read depth per gene to the average abundance of the corresponding protein product for all genes on each chromosome (Supplementary Fig. 1e, adjusted $R^2 = -0.04194$, $P \leq 0.5392$). There was also no correlation at any of the later time points (Supplementary Tables 3 and 4). Given this lack of correlation, we did not normalize the mass spectrometry data to the gene copy number inferred from sequencing read depth.

Given the high correlation between replicates, we summed the intensity values across both replicates to maximize the number of proteins detected in the dataset (Supplementary Fig. 2a and b, see *Methods*). On average, 368 proteins were detected in 1 replicate but not the other (Supplementary Fig. 2c, Supplementary Table 5). We next summarized our dataset to collapse protein measurements by batches. In total, we detected 2518 proteins in batch 1, 2,504 proteins in batch 15, and 2,280 proteins in the conditioning tank (Fig. 2a). We found that 2,274 proteins were common to both batches 1 and 15, while the conditioning tank sampled 231 fewer proteins. We observed 44 proteins exclusive to batch 1, 16 proteins exclusive to batch 15, and 6 proteins specific to the conditioning tank. While we found no significantly enriched gene ontology terms for proteins unique to batch 1, batch 15, and the conditioning tank, among the 231 proteins not detected in the conditioning tank we detected cellular component terms associated with the nucleoplasm and chromatin (Supplementary Fig. 2d). On comparing our dataset to a previous study unifying yeast protein abundance across multiple datasets (Ho *et al.* 2018), we found that distribution of undetected proteins in our dataset tended to have lower average molecules per cell than those that were detected (Supplementary Fig. 2e).

To uncover the trends underlying fermentation, we normalized LFQ intensity values for every protein to the mean abundance across all sampled time points (Supplementary Table 6) and performed hierarchical clustering and GO analysis on all proteins that changed at least 2-fold. We identified clusters by node depth (Supplementary Table 7, see *Methods*) for which we annotated the biological processes, molecular functions, and cellular components by gene ontology analysis using ClusterProfiler (Wu *et al.* 2021) (Fig. 2b, Supplementary Table 8). We found generally consistent expression trends between batches, with protein abundances similarly changing across time in both batches. Specifically, both batches have an initial increase in ribosome biogenesis that decreases after the first day of fermentation (Fig. 2b, cluster 1b), whereas proteins involved in carbohydrate and lipid metabolism and oxidation tended to have low abundance early in fermentation and increased over time (Fig. 2b, cluster 7b). These changes were expected as the yeast cells are rapidly dividing over the first 2 days of fermentation, exhausting the preferred carbon sources. The protein levels in the batch 15 post-fermentation conditioning tank samples were similar to the later fermentation time points in batch 15, with the exception of enrichment for mitochondrial gene expression (Fig. 2b, cluster 7c)
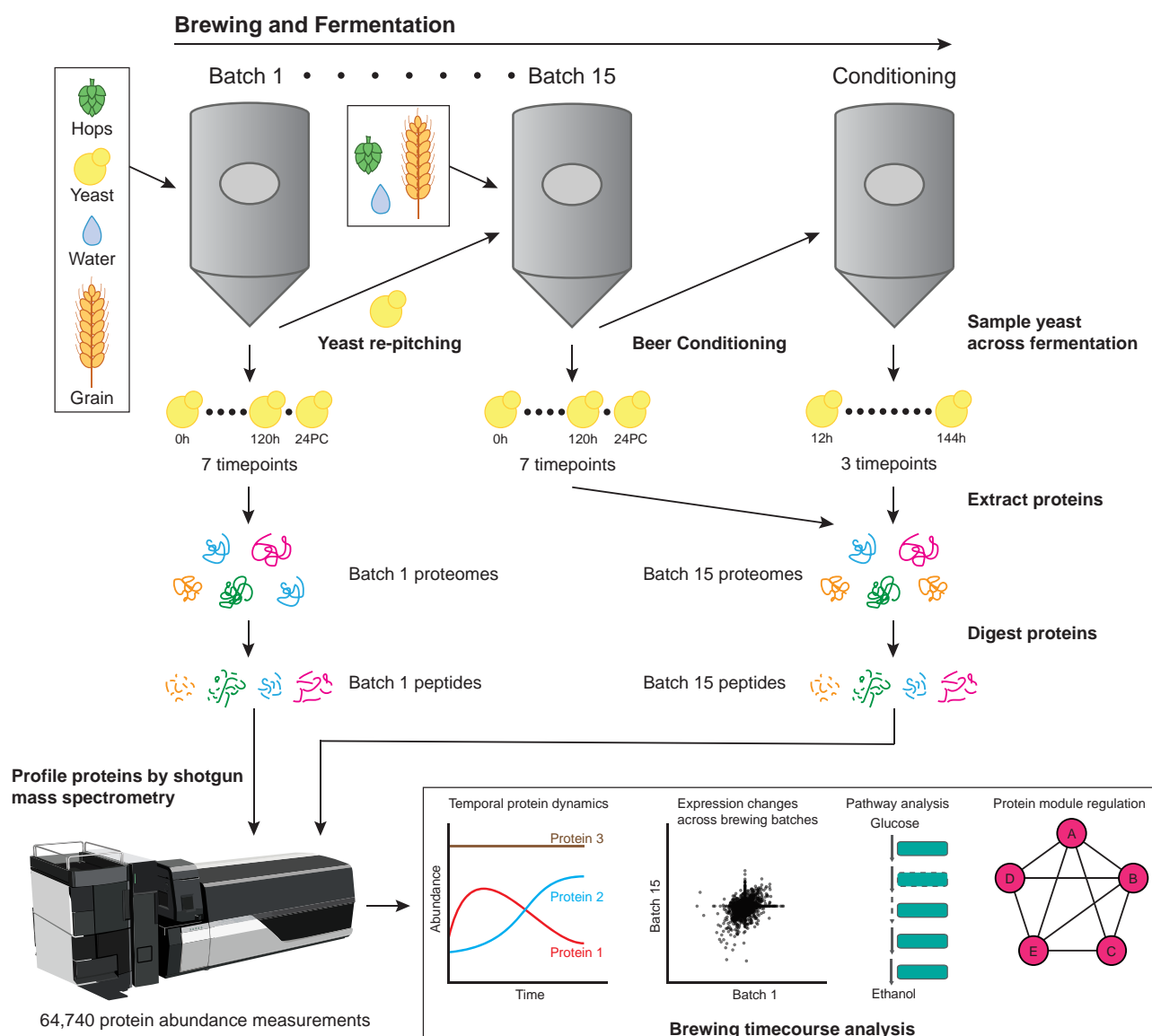
**Fig. 1.** Overview of Wyeast 3068 fermentation analysis via mass spectrometry. To systematically profile the proteomic changes across fermentation, we sampled 2 commercial production scale fermentation batches (batches 1 and 15) consisting of the same parental strain population. Batches 1 and 15 consisted of 7 time points each including a 24 hours post-crash time point (24PC), with each successive time point being no more than a few days apart. At the end of fermentation in batch 15, beer fermentations were subjected to conditioning during which we sampled 3 additional time points. To compare protein dynamics between batches 1 and 15, 6 time points were correspondingly matched to investigate batch-specific protein changes.

and another cluster with proteins strongly downregulated in the final conditioning time point (Fig. 2b, cluster 3), which was not significantly enriched for GO terms.

## Understanding the systems-level regulation of protein modules across brewing

After successfully identifying clusters of globally changing proteins by an unbiased method, we next expanded our analysis to globally analyze all annotated protein complexes regulated across fermentation. On intersecting the proteins detected in our dataset with those annotated in the Yeast GFP Fusion Localization database (Huh *et al.* 2003), we detected 74% of all annotated proteins (Supplementary Fig. 3a). Our approach broadly sampled proteins across yeast cellular compartments. We detected 75% of the cytoplasmic proteome and nearly all the proteins from the smallest classes such as the lipid particles and those that shuttle between the endoplasmic reticulum and Golgi (Supplementary

Fig. 3b, Supplementary Table 9). We next calculated the pairwise correlations between all proteins in a particular subcellular compartment and plotted distributions of the Pearson correlation coefficient values for every given compartment. Most distributions were centered around zero, indicating that entire compartments were generally not shifting in a concerted manner. Of major cellular compartments, the nucleolar proteome, proteins localized to actin fibers, and proteins that traffic between the ER and Golgi contained well-correlated protein pairs (Supplementary Fig. 3c).

Using the Complexome database (Meldal *et al.* 2019), we curated a comprehensive set of protein complexes. On average, annotated yeast protein complexes consisted of 5 members (Supplementary Fig. 3d), and we detected members from 79% (492/620) of annotated complexes (Supplementary Table 10), on average observing 70% of members within a given protein complex (Supplementary Fig. 3e). We generated the pairwise Pearson
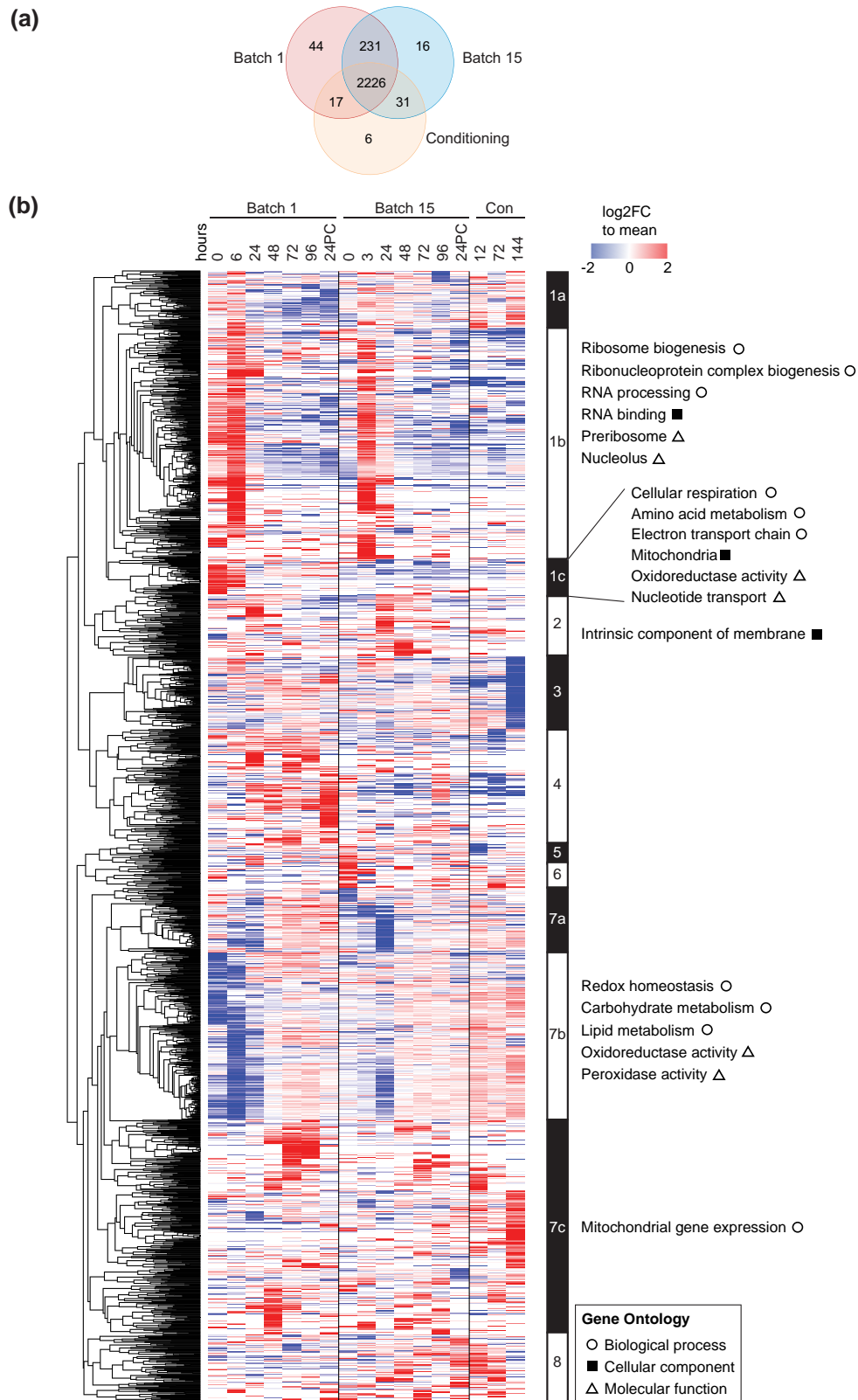
**Fig. 2.** A global view of ale yeast protein changes across brewing cycles. a) Venn diagrams summarizing the overlap of proteins detected across batch 1, batch 15, and the Conditioning tank. b) Heatmap depicting changes in abundances of 1,891 proteins which displayed at least a 2-fold change relative to the mean across all time points (rows) over both brewing batches and final conditioning (columns). Values are colored as $\log_2$ fold change relative to the row mean normalized abundance (see *Methods*). Proteins contained within clusters (defined by node depth; see *Methods*) were annotated from ClusterProfiler (Wu *et al.* 2021) filtering terms to an adjusted *P*-value (Benjamini–Hochberg correction) of <0.05 and a multiple hypothesis testing FDR of 5%. Each cell is the sum of 2 replicates with column time points indicated in hours. 24PC indicates 24 hours post-cold crashing after fermentation.

correlation matrix for all proteins detected across all samples from both batches and identified 4 well-correlated clusters in the similarity matrix (Supplementary Fig. 4a). We hypothesized that interacting protein pairs might share similar expression patterns across fermentation. On plotting the distribution of Pearson correlation across all pairs of well detected proteins in our dataset, we indeed found that the distribution of physically interacting protein pairs was more correlated than noninteracting pairs (Supplementary Fig. 4b). In order to understand the extent of correlation across complexes involved in biological processes, we subsetted our matrix into individual complexes. The 40S and 60S cytosolic ribosomal subunits were highly correlated compared to their mitochondrial counterparts (Supplementary Fig. 4c). Chromatin remodeling complexes like RSC, SWI/SNF, and INO80 showed poor pairwise correlations, with their distributions centered around zero (Supplementary Fig. 4d). Members of RNA polymerase I associated with transcription of rRNAs were highly correlated, unlike mRNA transcribing polymerase II and tRNA transcribing polymerase III (Supplementary Fig. 4e). However, associated transcriptional co-activators/repressors (TFIID, SAGA, and SLIK) exhibited poor correlation. Membrane transport complexes such as the exocyst and vesicle coat complexes COPI and COPII showed varying degrees of correlation among their members (Supplementary Fig. 4f). Finally, protein level regulators (26S proteasome) and chaperone complexes (Prefoldin and T-complex) displayed high correlations among their members (Supplementary Fig. 4g). Therefore, while not all individual protein levels changed over fermentation and repitching, their function could still be impacted by changes in complex partners that show poor correlation.

## Brewing yeast proteomes change drastically in the first 24 hours of brewing

Yeast strains undergo complex molecular changes while they adapt to both physical and chemical changes including nutrient deprivation, loss of oxygen, and increased alcohol production over time. Previous studies have profiled the metabolites, gene transcripts, and proteins present during beer fermentation, but very few studies have performed analyses comparing different times across a complete fermentation cycle (Berner *et al.* 2013; Grochalová *et al.* 2017; Schulz *et al.* 2018; Sun *et al.* 2019; Behr *et al.* 2020; Liu *et al.* 2022). We sought to understand how yeast protein expression patterns change between any 2 time points in the fermentation tank. Since LFQ abundances are log-normally distributed, we $\log_2$ transformed the LFQ abundance values in each time point, subtracted the median abundance across the time point (to equalize variances), and performed differential expression testing across time point pairs. Comparing time points all-by-all, we found that the number of differentially expressed proteins (DEPs) ranged from 2 to 618, with an average of 195 between any 2 time points (Fig. 3a, Supplementary Table 11). When comparing consecutive time points, we observed the largest number of DEPs between early time point pairs in both batch 1 and batch 15, which progressively decreased over time (Fig. 3b, Supplementary Table 11).

To elucidate the processes that are temporally regulated throughout fermentation, we performed GO analysis using the DEPs detected between time points. When comparing the first 2 consecutive time points (batch 1, 0 and 6 hours) we detected 222 DEPs (Supplementary Table 11) and observed an enrichment at the beginning of batch 1 for proteins involved in aerobic respiration, generation of precursor metabolites and energy, and cellular response to stress. On the other hand, proteins detected after

6 hours of fermentation were significantly enriched for processes related to ribosome biogenesis and translation (Fig. 3c, Supplementary Table 12). After the first day of brewing, the proteins involved in these processes along with rRNA regulatory processes were downregulated (Supplementary Table 12).

We also wanted to compare how protein levels differed between batches. Protein levels in matched time points between batches 1 and 15 were generally well correlated with Pearson correlation coefficients (R) ranging from 0.87 to 0.94 (Supplementary Fig. 5a). We compared the initial time points of both batches and found 213 DEPs (Supplementary Fig. 5b). Aerobic respiration and amino acid biosynthesis processes were enriched in batch 1 DEPs, and sterol and lipid biosynthesis processes were enriched in batch 15 DEPs (Supplementary Fig. 5c, Supplementary Table 12). Given the high correlation between protein levels in the early batch 1 (6 hours) and batch 15 (3 hours) time points (Supplementary Fig. 2b), we also compared processes between these time points (Supplementary Fig. 5d) and observed that batch 15 was enriched for sterol biosynthesis and glycogen and cellular alcohol metabolic processes (Supplementary Table 12). Though we observed 116 DEPs between the batches 24 hours post-crash (Supplementary Fig. 5b), no significant GO terms were enriched for this set of differing proteins between the batches at this time point.

## Uncovering the temporal changes in central metabolic proteins across brewing

While we observed global temporal protein changes over fermentation using GO analysis and detected correlation among subcellular locations and protein complexes, we wanted to better understand the degree to which these changes were coordinated by investigating biochemical pathways. We first focused on the 2 main pathways involved in yeast fermentation—glycolysis and the tricarboxylic acid (TCA) cycle—as their regulation is central to alcohol production (Supplementary Fig. 6a). On calculating the pairwise Pearson correlation of LFQ values across all samples for all pairs of proteins in the glycolysis and TCA cycle pathways, we found that sets of proteins in a given pathway tended to be well correlated. Generally, proteins involved in the TCA cycle were more correlated to each other than those involved in glycolysis (Supplementary Fig. 6a). Intriguingly, however, we also found many instances of correlation between the proteins involved in glycolysis and those involved in the TCA cycle. Of the 23 glycolytic proteins detected, Hxk1, Tdh1, Adh4, Pgk1, Gpm1, Eno2, Eno1, and Cdc19 were highly correlated with enzymes in the TCA cycle (Supplementary Fig. 6a).

When considering changes across fermentation, we expected glycolytic proteins to have high abundance early in the fermentation cycle and gradually decrease as sugar was consumed. However, we found that trends in the glycolytic proteins were often noisy (Fig. 4a). The levels of proteins involved in the TCA cycle were in line with the expected profile of a strain in a low-sugar and anaerobic environment (Wiebe *et al.* 2008): TCA cycle enzymes were downregulated over fermentation cycles across both batches and only increased their abundances during the conditioning phase, when yeast were exposed to additional oxygen while being moved to a conditioning tank (Fig. 4b).

We observed similarities between batches for the abundance profiles over time of enzymes catalyzing subsequent steps in the TCA cycle, demonstrated by the high degree of correlation between members of the pathway (Supplementary Fig. 6b). Since enzymes in the TCA cycle interface together to form complexes, we reasoned that enzyme complexes might be coregulated
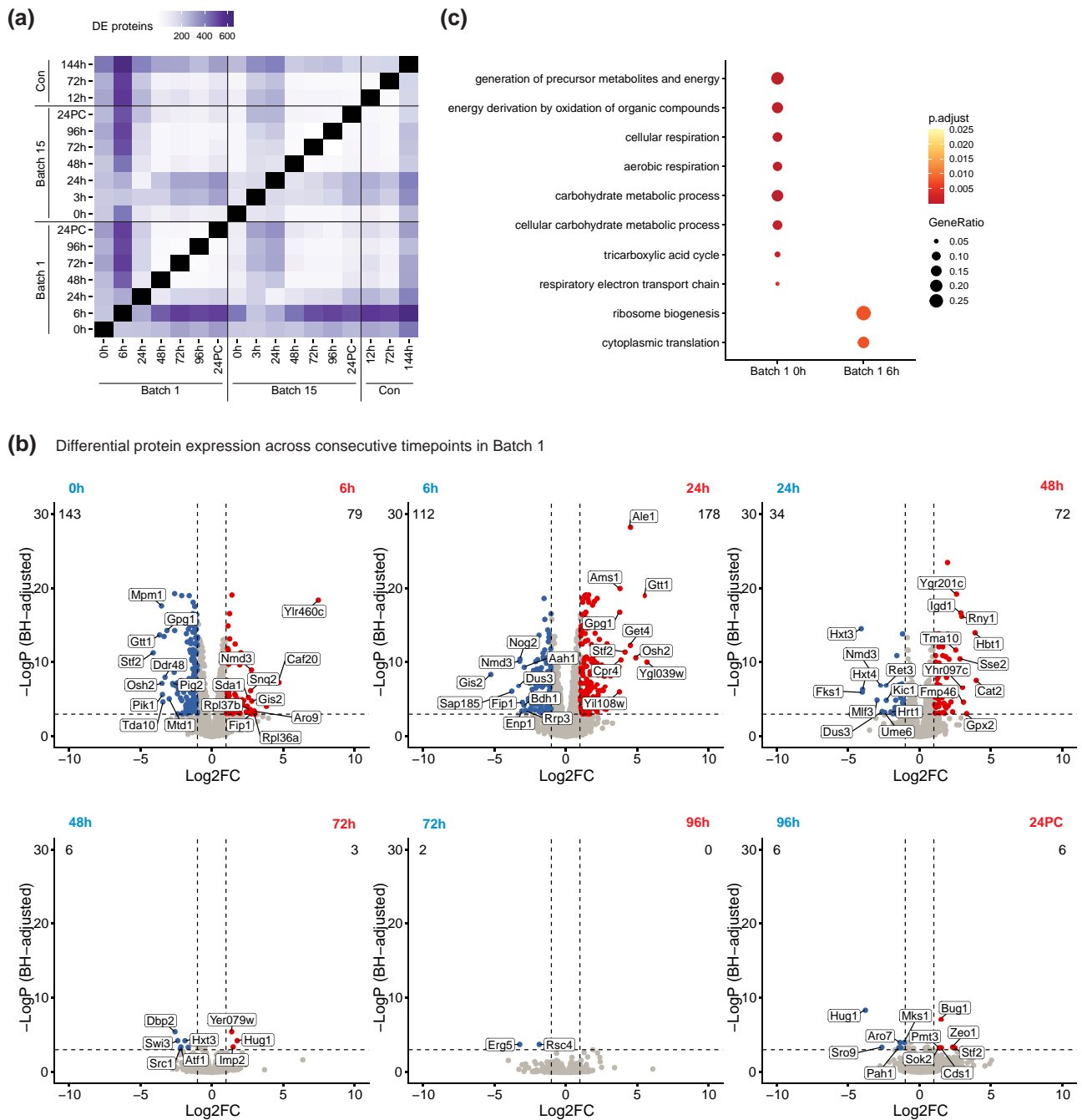
**(a)**



**(c)**



**(b)** Differential protein expression across consecutive timepoints in Batch 1



**Fig. 3.** Differential protein expression dynamics across brewing cycles. a) Heatmap representing the number of differentially expressed proteins (DEPs) across pairs of fermentation time points, indicated in hours; 24PC indicates 24 hours post-cold crashing after fermentation. DEPs were determined by a standard *t*-test. b) Volcano plots determining the protein expression changes across consecutive batch 1 time points. c) Dot plot representing the top enriched Gene Ontology terms between the 0 and 6 hours time point in batch 1.

throughout fermentation. Generally, we found this to be the case, with high correlation among members of the isocitrate dehydrogenase, 2-oxoglutarate dehydrogenase, and succinyl-CoA synthetase complexes (Fig. 4c).

Generally, protein profiles in glycolysis and the TCA cycle were similar between batches (Fig. 4a–b). One striking exception to this was a minor isoform of pyruvate decarboxylase, Pdc5, which converts pyruvate to acetaldehyde. Pdc5 was elevated in batch 1 compared to its mean abundance across the dataset while, in batch 15, we saw the opposite trend (Fig. 4a, Supplementary Fig. 6b). Although peptides unique to major isoform Pdc1 were not

detected, we detected Pdc1 as a part of a protein group along with Thi3. The abundance of this group did not differ between batches or along the time course (Supplementary Fig. 6b). Other proteins that differed between batches included pyruvate carboxylases Pyc1 and Pyc2 and aconitate hydrolases Aco1 and Aco2. Pyc1 and Pyc2 had higher levels for the first 6 hours of fermentation in batch 1, but displayed the opposite trend in batch 15, indicative of strain adaptation to fermentation conditions. Despite acting on the same substrate, Aco1 and Aco2 exhibited unique profiles: Aco1 levels increased across each batch whereas Aco2 decreased, and was higher in batch 1 than batch 15.
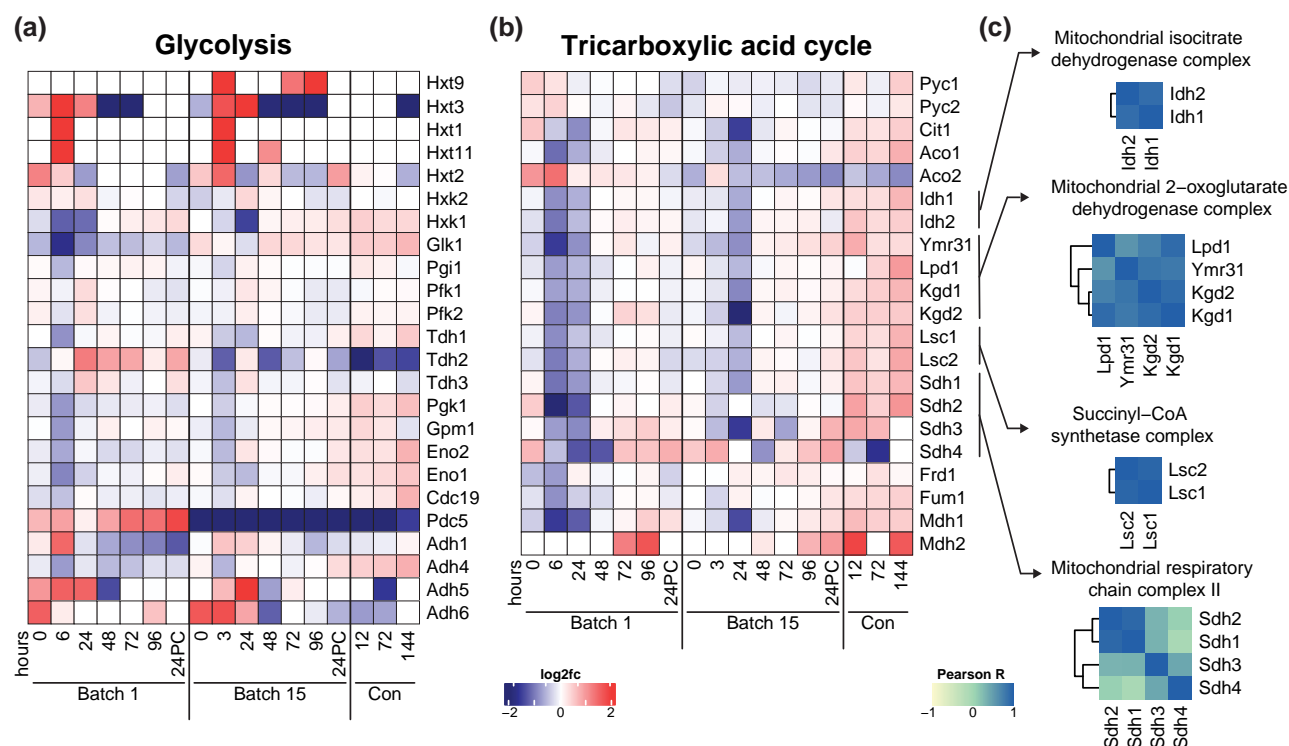
**Fig. 4.** Systems-level regulation within protein modules across brewing. Heatmaps depicting the expression profiles of proteins (rows) across fermentation time (columns) in a) glycolysis and b) the tricarboxylic acid (TCA) cycle. Proteins in heatmaps ordered according to biochemical steps in their respective pathways. c) Pairwise Pearson correlations across proteins within complexes involved in the TCA cycle.

We also examined proteins in other pathways related to pyruvate metabolism (Supplementary Fig. 6c). Pyruvate dehydrogenases Pda1 and Pdb1, which shunt carbons to the TCA cycle, maintained stable levels across fermentation time and between batches (Supplementary Fig. 6b). Alcohol dehydrogenases exhibited varying profiles: The abundances of Adh1 and Adh4 remained relatively unchanged between batches and across fermentation, while Adh5 and Adh6 had lower abundances that further decreased after 1 to 2 days of fermentation (Supplementary Fig. 6b). Bdh1, which converts 2,3-butanediol to acetoin, was markedly lower in abundance in batch 15 and decreased over time.

To take an unbiased approach to identify the metabolic pathways that were most likely to be altered by the observed changes in protein levels, we performed metabolic pathway analysis with Saccharomyces Genome Database (SGD) Yeast Pathways across all samples from both batches. Several pathways predicted to be most altered were those associated with lipid, amino acid, and nucleotide metabolism (Fig. 5a, Supplementary Table 13). We looked more closely at the levels of enzymes with roles in these top pathways and identified key differences between batches. Many of the affected pathways related to sugar and amino acid metabolism contain enzymes Bat1, Bat2, or Pdc5, which are involved in the production of the grainy flavor compound isobutyraldehyde (Fig. 5b). These 3 enzymes are responsible for the high pathway perturbation scores for pyruvate fermentation, acetoin and butanediol biosynthesis, and degradation of amino acids isoleucine, valine, phenylalanine, tryptophan, and tyrosine: All of these pathways drop out of the top 20 most affected metabolic processes when Bat1, Bat2, and Pdc5 are removed from the dataset, with acetoin and butanediol biosynthesis going from second most perturbed pathway to 20 s. Interestingly, both Bat2 and Pdc5 abundances were lower in batch 15, along with a slight decrease in Ald6 which metabolizes

isobutyraldehyde to isobutyric acid (Fig. 5c). Of those pathways that did not involve Bat1, Bat2, or Pdc5, we looked to see if there were other pathways with multiple enzymes that differed between batches. We saw that enzymes involved in ergosterol biosynthesis were present at higher abundances in batch 15 than in batch 1 (Fig. 5d). This was specific to sterol metabolism since enzymes in other lipid metabolism pathways were not significantly different between batches (Supplementary Fig. 6d and e).

## Discussion

Understanding the molecular changes associated with commercial beer production is crucial to inform the brewing process by guiding strain engineering, identifying molecular characteristics of yeast strains that influence the beer flavor, and optimizing workflows for large scale production. Comprehensively mapping protein dynamics during fermentation offers opportunities to globally identify the enzymes and metabolic pathways responsible for generating the diverse range of flavor compounds in beers and delving into the regulatory mechanisms governing the biochemical processes necessary for fermentation. Furthermore, profiling beer yeast proteomes across successive serial repitching cycles provides a unique lens into the evolutionary and physiological dynamics and adaptive responses of yeast populations during brewing.

Here, we characterized the genome of a Hefeweizen ale brewing strain, Wyeast 3068, and the changes in its proteome throughout fermentation in a commercial brewing setting. We characterized a fermentation time course across 2 batches separated by 14 repitches to elucidate the impact of serial repitching on the brewing yeast proteome. While previous work has characterized the genetic changes in brewing yeast genomes (Large *et al.*
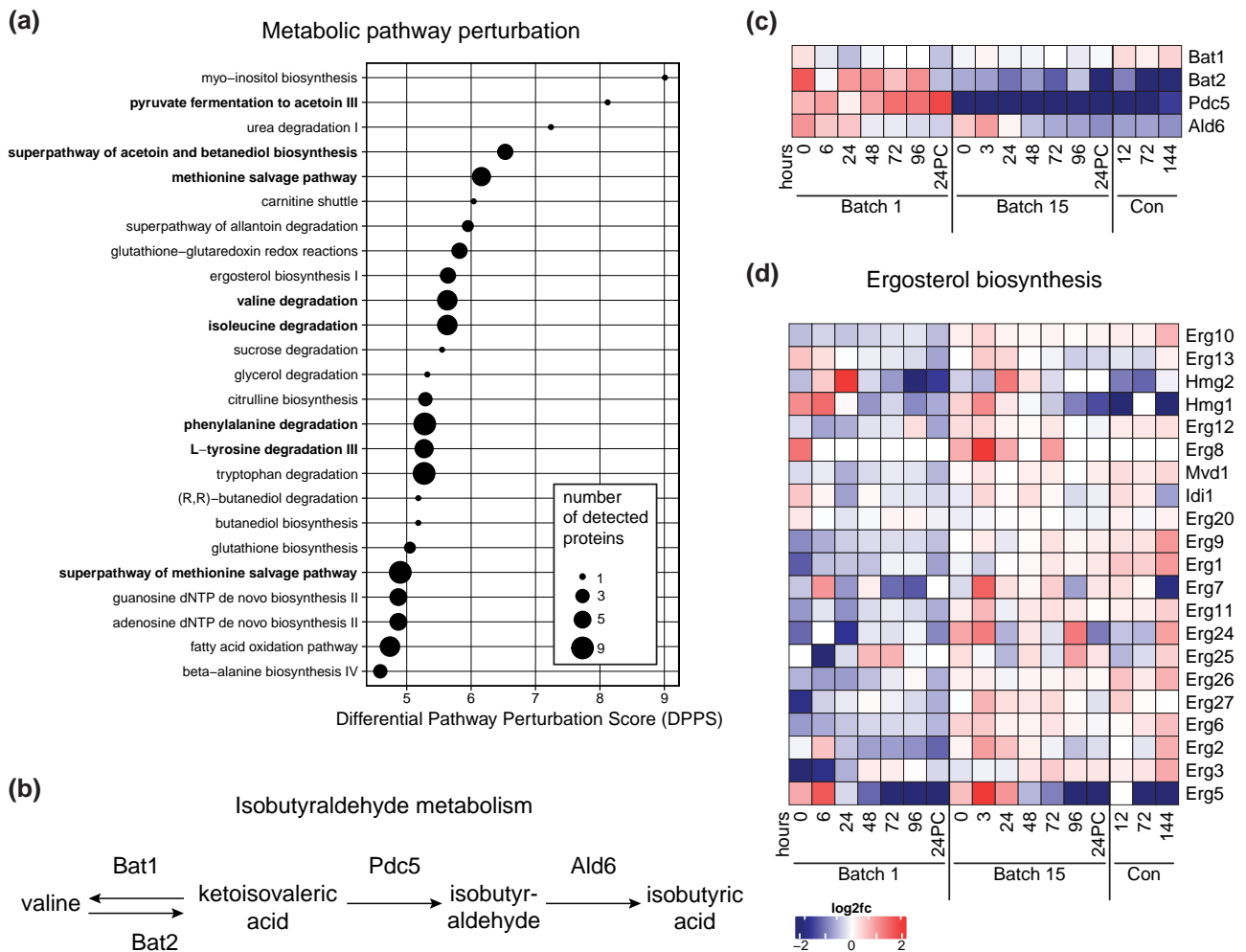
**(a)**



**(b)**



**(c)**



**(d)**



**Fig. 5.** Ergosterol and isobutyraldehyde metabolism differ between early and late batches. a) Top 25 altered metabolic pathways across all samples as determined by SGD Yeast Pathways. Bold pathway names indicate all pathways that contain Bat1, Bat2, or Pdc5. b) Steps in the isobutyraldehyde synthesis and metabolism pathway. c) Isobutyrate synthesis is downregulated in batch 15. Heatmap depicting changes in abundance over time (columns) and enzymes (rows) involved in isobutyrate synthesis. d) Changes in ergosterol synthesis enzyme abundances (rows) over both fermentation and conditioning (columns). Color scale represents log$_2$ fold change vs the mean abundance across all time points.

2020), traced ecological origins of beer brewing strains (Gallone et al. 2016, 2018, 2019), and mapped the residual proteins in beer (Picariello et al. 2012, 2015; Schulz et al. 2018; Kerr et al. 2019; Liu et al. 2022), our work systematically profiled the proteomic changes in an ale yeast strain across entire sets of commercial fermentation cycles using shotgun mass spectrometry. From this study, we have created a comprehensive dataset cataloging 2,572 yeast proteins across 17 time points during industrial beer brewing. Our data reveal global trends in protein expression changes over brewing cycles as well as across serial repitching, a practice widely adopted by many breweries but understudied with respect to molecular changes in the fermenting yeast populations.

Our genetic characterization of Wyeast 3068 showed deletions of many genes as expected for ale yeast: Flocculation genes vary between brewing strains (Van Mulders et al. 2010) and can be lost in aged brewing yeast (Sato et al. 2001), less efficient carbohydrate transporters are lost in some brewing strains (Nakao et al. 2009), ASP3 has been lost in many S. cerevisiae isolates (Coral-Medina et al. 2022), and transposition events and copy number variations of Ty elements are common in industrial yeast strains (Benítez et al. 1996). Previous studies have demonstrated

that genome evolution in brewing yeast strains does occur across the repitching process as selected populations have adapted to the fermentation environment (Large et al. 2020). While it is beyond the scope of this work, the mechanism of strain adaptations warrant further studies such as matched genomic data across fermentation cycles and serial repitching. It remains to be seen to what degree the changes we observed in protein abundances arose from genetic mechanisms (i.e. deletions or duplications) vs gene expression regulation at the RNA and post-translational levels; protein activities might also be changing independently of abundances, such as by post-translational modification or allosteric or feedback regulation, none of which we have attempted to measure here.

By focusing on proteomic changes over the course of fermentation, we found that the largest number of proteins whose abundances significantly change occur within the first 24 hours of fermentation. Sampling more time points in the first day of fermentation would provide finer resolution of the protein expression changes. Importantly, whether these changes are purely associated with growth vs immediate acclimatization to the fermentation tank merits further investigation. Strikingly, protein translation machinery and ribosome biogenesis peaked 3 to

6 hours after the start of fermentation before being downregulated. These trends suggest that the yeast strains are primed for protein synthesis before entering strictly anaerobic conditions later in fermentation. It is interesting to note that while certain protein clusters specific to the conditioning tank do not significantly represent a GO biological process, they tended to consist of mitochondrial ribosomal subunits and those associated with protein translation. Additionally, they also consisted of a handful of proteins encoded by uncharacterized genes (YLR179C, YGR127W, YOR131C, YPR114W, YNL115C in cluster 3 and YKL091C, YLR177W, YDR391C, YER034W, YKL063C, YER152C, YML020W, YJL218W, YNL011C, YGR266W, YBR137W, YBL055C in cluster 7c from Fig. 2, see Supplementary Table 7). Further studies characterizing these sets of proteins will shed light on the molecular processes involved in the conditioning process.

Broadly, we found protein trends across fermentation to be similar before and after serial repitching barring some marked differences. For example, ergosterol synthesis enzymes were elevated in batch 15. Ergosterol can help counter the stress of hypoxic environments, enabling normal growth and flocculation necessary for brewing (Straver *et al.* 1993). This may be caused by the selection for yeast that are better able to survive hypoxic conditions due to increased ergosterol synthesis, suggesting ergosterol biosynthesis as a potential target for beer yeast strain engineering. Alternatively, it may be indicative of cellular stress programs that increase ergosterol production, which could be mitigated by ergosterol supplementation to reduce stress and confer better growth behavior for late-pitch yeast (Casey *et al.* 1984). The other striking difference between batches was in isobutyraldehyde synthesis enzymes Bat2 and Pdc5, which were less abundant in batch 15 than in batch 1. Isobutyraldehyde is associated with a grainy flavor profile, which is desirable in some beers and considered an off-flavor in others. In winemaking, deletions of BAT2 and PDC5 can be used to reduce the production of undesirable fusel alcohols (Yan *et al.* 2023). The observed changes in Bat2 and Pdc5 abundances between early and late batches could contribute to inconsistencies in flavor over many repitches. Why these enzymes change over batches remains to be explored.

While we measured the abundances of thousands of proteins across fermentation cycles, our approach is limited to only measuring protein levels across fermentation. It is important to note that our approach tended to poorly sample proteins with less than 1,000 molecules per cell (Supplementary Fig. 2e). Conversely, it is interesting to note our assays did not detect some proteins that were previously reported to be well detected in other studies (e.g. those involved in cytoplasmic translational machinery and nuclear rRNA export). Future studies incorporating targeted assays and quantitative proteomics geared toward deeper sampling of brewing yeast proteomes will shed light on how these strains have adapted to the fermentation environment. Many metabolic enzymes that regulate important processes in fermentation are not only regulated by absolute levels but also by modifications such as phosphorylation and glycosylation. Many proteins for which we did not observe changes in absolute levels could still have altered activity over fermentation due to post-translational modifications. For example, phosphofructokinase and pyruvate decarboxylase are regulated by phosphorylation (James *et al.* 1995; Dihazi *et al.* 2001). Additionally, it is known that secreted yeast glycoproteins contribute to the proteome of beer, but how they change over time has not been investigated (Kerr *et al.* 2021). A recent study identified changes in the phosphoproteome of yeast during diauxic shift (Dephoure *et al.* 2014), so further studies such as phosphoproteomics over fermentations and before and after repitching would

likely identify other important pathways regulated in brewing. Coupling measurement of yeast enzyme levels to metabolite levels in the fermentation tank could give a comprehensive view of how yeast biology is altered in brewing and how that impacts the fermentation product.

Our study provides a systems biology view of the molecular processes that underlie beer brewing. By analyzing how changes in protein levels alter protein complexes and biochemical pathways, we observed that interacting protein pairs are correlated across samples, suggesting that many yeast cellular modules are co-regulated during brewing. Future studies characterizing the protein-protein interactions using proximity labeling (Li *et al.* 2020; Liu *et al.* 2020) or fractionation (Wan *et al.* 2015) with mass spectrometry across brewing will shed light on the dynamic regulation of protein complexes across fermentation.

Finally, we believe that this resource serves as a comprehensive catalog of fermentation-based protein changes, and we have made it available for exploration on an interactive web interface (https://brewing-yeast-proteomics.ccbb.utexas.edu/). Further analysis of our data and future studies of proteins, post-translational modifications, and metabolite changes across fermentation and repitching will aid beer yeast strain engineering, optimization of brewing workflows, and study of trends that underlie domestication processes.

## Data availability

Weihenstephan Weizen is available from White Labs as Wyeast 3068. Genome sequencing data of the strain in this study have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject PRJNA1011390. Raw mass spectrometry files for the timecourse have been deposited on MassIVE under the MSV000092793. All processed tables related to the manuscript are available in the supplement.

Supplemental material available at G3 online.

## Conflicts of interest

EMM is a co-founder, shareholder, and scientific advisory member of Erisyon, Inc., which played no role in this work. DK is an employee, TP is a former employee, and HME is a founding member of Live Oak Brewing Company. The other authors declare no competing interests.

## Author contributions

RKG (conceptualization, methodology, data curation, formal analysis, investigation, visualization, writing-original draft), RCG (formal analysis, writing-original draft, visualization), JOA (formal analysis, writing-original draft, visualization), BD (formal analysis, writing-original draft), DRB (conceptualization, methodology, analysis, investigation, visualization), AB (investigation, formal analysis), ML (software), VD (software), PJ (formal analysis), DK (resources), TP (resources), HME (resources), EMM (funding acquisition, conceptualization, project administration, supervision, resources, writing-review and editing), MJD (funding acquisition, project administration, supervision, writing-review and editing).

## Literature cited

Abdullah M, Greco BM, Laurent JM, Garge RK, Boutz DR, Vandeloo M, Marcotte EM, Kachroo AH. 2023. Rapid, scalable, combinatorial genome engineering by marker-less enrichment and recombination of genetically engineered loci in yeast. Cell Rep Methods. 3(5):100464 doi:10.1016/j.crmeth.2023.100464.

Abyzov A, Urban AE, Snyder M, Gerstein M. 2011. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. Genome Res. 21(6):974–984. doi:10.1101/gr.114876.110.

Adamczyk J, Deregowska A, Skoneczny M, Skoneczna A, Natkanska U, Kwiatkowska A, Rawska E, Potocki L, Kuna E, Panek A, et al. 2016. Copy number variations of genes involved in stress responses reflect the redox state and DNA damage in brewing yeasts. Cell Stress Chaperones. 21(5):849–864. doi:10.1007/s12192-016-0710-8.

Amorosi CJ, Chiasson MA, McDonald MG, Wong LH, Sitko KA, Boyle G, Kowalski JP, Rettie AE, Fowler DM, Dunham MJ. 2021. Massively parallel characterization of CYP2C9 variant enzyme activity and abundance. Am J Hum Genet. 108(9):1735–1751. doi:10.1016/j.ajhg.2021.07.001.

Balakrishnan R, Park J, Karra K, Hitz BC, Binkley G, Hong EL, Sullivan J, Micklem G, Cherry JM. 2012. YeastMine—an integrated data warehouse for Saccharomyces cerevisiae data as a multipurpose tool-kit. Database (Oxford). 2012:bar062. doi:10.1093/database/bar062.

Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. 2007. Quantitative mass spectrometry in proteomics: a critical review.

Anal Bioanal Chem. 389(4):1017–1031. doi:10.1007/s00216-007-1486-6.

Barsnes H, Vaudel M. 2018. SearchGUI: a highly adaptable common interface for proteomics search and de novo engines. J Proteome Res. 17(7):2552–2555. doi:10.1021/acs.jproteome.8b00175.

Beer Market Size, Industry Share, Growth Rate, Forecast by 2030. https://www.fortunebusinessinsights.com/beer-market-102489.

Behr J, Kliche M, Geißler A, Vogel RF. 2020. Exploring the potential of comparative de novo transcriptomics to classify Saccharomyces brewing yeasts. PLoS One. 15(9):e0238924. doi:10.1371/journal.pone.0238924.

Benítez T, Martínez P, Codón AC. 1996. Genetic constitution of industrial yeast. Microbiologia. 12(3):371–384.

Berner TS, Jacobsen S, Arneborg N. 2013. The impact of different ale brewer's yeast strains on the proteome of immature beer. BMC Microbiol. 13(1):215. doi:10.1186/1471-2180-13-215.

Bühligen F, Rüdinger P, Fetzer I, Stahl F, Scheper T, Harms H, Müller S. 2013. Sustainability of industrial yeast serial repitching practice studied by gene expression and correlation analysis. J Biotechnol. 168(4):718–728. doi:10.1016/j.jbiotec.2013.09.008.

Casey GP, Magnus CA, Ingledew WM. 1984. High-gravity brewing: effects of nutrition on yeast composition, fermentative ability, and alcohol production. Appl Environ Microbiol. 48(3):639–646. doi:10.1128/aem.48.3.639-646.1984.

Cherry JM, Adler C, Ball C, Chervitz SA, Dwight SS, Hester ET, Jia Y, Juvik G, Roe T, Schroeder M, et al. 1998. SGD: Saccharomyces Genome Database. Nucleic Acids Res. 26(1):73–79. doi:10.1093/nar/26.1.73.

Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, Chan ET, Christie KR, Costanzo MC, Dwight SS, Engel SR, et al. 2012. Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res. 40(D1):D700–D705. doi:10.1093/nar/gkr1029.

Coral-Medina A, Fenton DA, Varela J, Baranov PV, Camarasa C, Morrissey JP. 2022. The evolution and role of the periplasmic asparaginase Asp3 in yeast. FEMS Yeast Res. 22(1):foac044. doi:10.1093/femsyr/foac044.

Dephoure N, Hwang S, O'Sullivan C, Dodgson SE, Gygi SP, Amon A, Torres EM. 2014. Quantitative proteomic analysis reveals post-translational responses to aneuploidy in yeast. eLife. 3:e03023. doi:10.7554/eLife.03023.

Dihazi H, Kessler R, Eschrich K. 2001. Phosphorylation and inactivation of yeast 6-phosphofructo-2-kinase contribute to the regulation of glycolysis under hypotonic stress. Biochemistry. 40(48):14669–14678. doi:10.1021/bi0155549.

Gallant JL, Heunis T, Sampson SL, Bitter W. 2020. ProVision: a web-based platform for rapid analysis of proteomics data processed by MaxQuant. Bioinformatics. 36(19):4965–4967. doi:10.1093/bioinformatics/btaa620.

Gallone B, Mertens S, Gordon JL, Maere S, Verstrepen KJ, Steensels J. 2018. Origins, evolution, domestication and diversity of Saccharomyces beer yeasts. Curr Opin Biotechnol. 49:148–155. doi:10.1016/j.copbio.2017.08.005.

Gallone B, Steensels J, Mertens S, Dzialo MC, Gordon JL, Wauters R, Theßeling FA, Bellinazzo F, Saels V, Herrera-Malaver B, et al. 2019. Interspecific hybridization facilitates niche adaptation in beer yeast. Nat Ecol Evol. 3(11):1562–1575. doi:10.1038/s41559-019-0997-9.

Gallone B, Steensels J, Prahl T, Soriaga L, Saels V, Herrera-Malaver B, Merlevede A, Roncoroni M, Voordeckers K, Miraglia L, et al. 2016. Domestication and divergence of Saccharomyces cerevisiae beer yeasts. Cell. 166(6):1397–1410.e16. doi:10.1016/j.cell.2016.08.020.

Gao Y, Ping L, Duong D, Zhang C, Dammer EB, Li Y, Chen P, Chang L, Gao H, Wu J, et al. 2021. Mass-spectrometry-based near-complete

draft of the *Saccharomyces cerevisiae* proteome. J Proteome Res. 20(2):1328–1340. doi:10.1021/acs.jproteome.0c00721.

Garge RK, Cha HJ, Lee C, Gollihar JD, Kachroo AH, Wallingford JB, Marcotte EM. 2021. Discovery of new vascular disrupting agents based on evolutionarily conserved drug action, pesticide resistance mutations, and humanized yeast. Genetics. 219(1):iyab101. doi:10.1093/genetics/iyab101.

Gibson BR, Lawrence SJ, Leclaire JPR, Powell CD, Smart KA. 2007. Yeast responses to stresses associated with industrial brewery handling. FEMS Microbiol Rev. 31(5):535–569. doi:10.1111/j.1574-6976.2007.00076.x.

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, *et al.* 1996. Life with 6000 genes. Science. 274(5287):546–567. doi:10.1126/science.274.5287.546.

Grochalová M, Konečná H, Stejskal K, Potěšil D, Fridrichová D, Srbová E, Ornerová K, Zdráhal Z. 2017. Deep coverage of the beer proteome. J Proteomics. 162:119–124. doi:10.1016/j.jprot.2017.05.001.

Ho B, Baryshnikova A, Brown GW. 2018. Unification of protein abundance datasets yields a quantitative *Saccharomyces cerevisiae* proteome. Cell Syst. 6(2):192–205.e3. doi:10.1016/j.cels.2017.12.004.

Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK. 2003. Global analysis of protein localization in budding yeast. Nature. 425(6959):686–691. doi:10.1038/nature02026.

James AG, Cook RM, West SM, Lindsay JG. 1995. The pyruvate dehydrogenase complex of *Saccharomyces cerevisiae* is regulated by phosphorylation. FEBS Lett. 373(2):111–114. doi:10.1016/0014-5793(95)01020-F.

Kachroo AH, Vandeloo M, Greco BM, Abdullah M. 2022. Humanized yeast to model human biology, disease and evolution. Dis Model Mech. 15(6):dmm049309. doi:10.1242/dmm.049309.

Kalayu G. 2019. Serial re-pitching: its effect on yeast physiology, fermentation performance, and product quality. Ann Microbiol. 69(8):787–796. doi:10.1007/s13213-019-01493-4.

Kang K, Bergdahl B, Machado D, Dato L, Han TL, Li J, Villas-Boas S, Herrgård MJ, Förster J, Panagiotou G. 2019. Linking genetic, metabolic, and phenotypic diversity among *Saccharomyces cerevisiae* strains using multi-omics associations. Gigascience. 8(4):giz015. doi:10.1093/gigascience/giz015.

Kerr ED, Caboche CH, Pegg CL, Phung TK, Gonzalez Viejo C, Fuentes S, Howes MT, Howell K, Schulz BL. 2021. The post-translational modification landscape of commercial beers. Sci Rep. 11(1):15890. doi:10.1038/s41598-021-95036-0.

Kerr ED, Caboche CH, Schulz BL. 2019. Posttranslational modifications drive protein stability to control the dynamic beer brewing proteome. Mol Cell Proteomics. 18(9):1721–1731. doi:10.1074/mcp.RA119.001526.

Krogerus K, Magalhães F, Castillo S, Peddinti G, Vidgren V, De Chiara M, Yue JX, Liti G, Gibson B. 2021. Lager yeast design through meiotic segregation of a *Saccharomyces cerevisiae* × *Saccharomyces eubayanus* hybrid. Front Fungal Biol. 2:733655. doi:10.3389/ffunb.2021.733655.

Large CR, Hanson N, Tsouris A, Saada OA, Koonthongkaew J, Toyokawa Y, Schmidlin T, Moreno-Habel DA, McConnellogue H, Preiss R, *et al.* Genomic stability and adaptation of beer brewing yeasts during serial repitching in the brewery. 166157. https://doi.org/10.1101/2020.06.26.166157, preprint: not peer reviewed.

Li C, Zhang S, Dong G, Bian M, Liu X, Dong X, Xia T. 2022. Multi-omics study revealed the genetic basis of beer flavor quality in yeast. LWT. 168:113932. doi:10.1016/j.lwt.2022.113932.

Li Y, Tian C, Liu K, Zhou Y, Yang J, Zou P. 2020. A clickable APEX probe for proximity-dependent proteomic profiling in yeast. Cell Chem Biol. 27(7):858–865.e8 doi:10.1016/j.chembiol.2020.05.006.

Liu S, Kerr ED, Pegg CL, Schulz BL. 2022. Proteomics and glycoproteomics of beer and wine. Proteomics. 22(15–16):2100329. doi:10.1002/pmic.202100329.

Liu X, Salokas K, Weldatsadik RG, Gawriyski L, Varjosalo M. 2020. Combined proximity labeling and affinity purification–mass spectrometry workflow for mapping and visualizing protein interaction networks. Nat Protoc. 15(10):3182–3211. doi:10.1038/s41596-020-0365-x.

McWhite CD, Papoulas O, Drew K, Cox RM, June V, Dong OX, Kwon T, Wan C, Salmi ML, Roux SJ, *et al.* 2020. A pan-plant protein complex map reveals deep conservation and novel assemblies. Cell. 181(2):460–474.e14. doi:10.1016/j.cell.2020.02.049.

Meldal BHM, Bye-A-Jee H, Gajdoš L, Hammerová Z, Horáčková A, Melicher F, Perfetto L, Pokorný D, Lopez MR, Türková A, *et al.* 2019. Complex Portal 2018: extended content and enhanced visualization tools for macromolecular complexes. Nucleic Acids Res. 47(D1):D550–D558. doi:10.1093/nar/gky1001.

Mueller LN, Brusniak M-Y, Mani DR, Aebersold R. 2008. An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. J Proteome Res. 7(1):51–61. doi:10.1021/pr700758r.

Nakao Y, Kanamori T, Itoh T, Kodama Y, Rainieri S, Nakamura N, Shimonaga T, Hattori M, Ashikari T. 2009. Genome sequence of the lager brewing yeast, an interspecies hybrid. DNA Res. 16(2):115–129. doi:10.1093/dnares/dsp003.

Picariello G, Mamone G, Cutignano A, Fontana A, Zurlo L, Addeo F, Ferranti P. 2015. Proteomics, peptidomics, and immunogenic potential of wheat beer (Weissbier). J Agric Food Chem. 63(13):3579–3586. doi:10.1021/acs.jafc.5b00631.

Picariello G, Mamone G, Nitride C, Addeo F, Camarca A, Vocca I, Gianfrani C, Ferranti P. 2012. Shotgun proteome analysis of beer and the immunogenic potential of beer polypeptides. J Proteomics. 75(18):5872–5882. doi:10.1016/j.jprot.2012.07.038.

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. Nat Biotechnol. 29(1):24–26. doi:10.1038/nbt.1754.

Saada OA, Tsouris A, Large C, Friedrich A, Dunham MJ, Schacherer J. 2022. Phased polyploid genomes provide deeper insight into the multiple origins of domesticated *Saccharomyces cerevisiae* beer yeasts. Curr Biol. 32(6):1350–1361.e3. doi:10.1016/j.cub.2022.01.068.

Sato M, Watari J, Shinotsuka K. 2001. Genetic instability in flocculation of bottom-fermenting yeast. J Am Soc Brewing Chemists. 59(3):130–134. doi:10.1094/ASBCJ-59-0130.

Schulz BL, Phung TK, Bruschi M, Janusz A, Stewart J, Meehan J, Healy P, Nouwens AS, Fox GP, Vickers CE. 2018. Process proteomics of beer reveals a dynamic proteome with extensive modifications. J Proteome Res. 17(4):1647–1653. doi:10.1021/acs.jproteome.7b00907.

Smart KA, Whisker S. 1996. Effect of serial repitching on the fermentation properties and condition of brewing yeast. J Am Soc Brew Chem. 54(1):41–44. doi:10.1094/ASBCJ-54-0041.

Straver MH, vd Aar PC, Smit G, Kijne JW. 1993. Determinants of flocculence of brewer's yeast during fermentation in wort. Yeast. 9(5):527–532. doi:10.1002/yea.320090509.

Sun ZG, Wang MQ, Wang YP, Xing S, Hong KQ, Chen YF, Guo XW, Xiao DG. 2019. Identification by comparative transcriptomics of core regulatory genes for higher alcohol production in a top-fermenting yeast at different temperatures in beer fermentation. Appl Microbiol Biotechnol. 103(12):4917–4929. doi:10.1007/s00253-019-09807-x.

Turco G, Chang C, Wang RY, Kim G, Stoops EH, Richardson B, Sochat V, Rust J, Oughtred R, Thayer N, *et al.* 2023. Global analysis of the

yeast knockout phenome. Sci Adv. 9(21):eadg5702. doi:10.1126/sciadv.adg5702.

Tyanova S, Temu T, Cox J. 2016. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc. 11(12):2301–2319. doi:10.1038/nprot.2016.136.

Van Mulders SE, Ghequire M, Daenen L, Verbelen PJ, Verstrepen KJ, Delvaux FR. 2010. Flocculation gene variability in industrial brewer's yeast strains. Appl Microbiol Biotechnol. 88(6):1321–1331. doi:10.1007/s00253-010-2843-5.

Vaudel M, Burkhart JM, Zahedi RP, Oveland E, Berven FS, Sickmann A, Martens L, Barsnes H. 2015. PeptideShaker enables reanalysis of MS-derived proteomics data sets. Nat Biotechnol. 33(1):22–24. doi:10.1038/nbt.3109.

Wan C, Borgeson B, Phanse S, Tu F, Drew K, Clark G, Xiong X, Kagan O, Kwan J, Bezginov A, et al. 2015. Panorama of ancient metazoan macromolecular complexes. Nature. 525(7569):339–344. doi:10.1038/nature14877.

Wang M, Sun Z, Wang Y, Wei Z, Chen B, Zhang H, Guo X, Xiao D. 2019. The effect of pitching rate on the production of higher alcohols by top-fermenting yeast in wheat beer fermentation. Ann Microbiol. 69(7):713–726. doi:10.1007/s13213-019-01463-w.

Weiss M, Schrimpf S, Hengartner MO, Lercher MJ, von Mering C. 2010. Shotgun proteomics data from multiple organisms reveals remarkable quantitative conservation of the eukaryotic core proteome. Proteomics. 10(6):1297–1306. doi:10.1002/pmic.200900414.

Wiebe MG, Rintala E, Tamminen A, Simolin H, Salusjärvi L, Toivari M, Kokkonen JT, Kiuru J, Ketola RA, Jouhten P, et al. 2008. Central carbon metabolism of Saccharomyces cerevisiae in anaerobic, oxygen-limited and fully aerobic steady-state conditions and following a shift to anaerobic conditions. FEMS Yeast Res. 8(1):140–154. doi:10.1111/j.1567-1364.2007.00234.x.

Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, et al. 2021. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innovation (Camb). 2(3):100141. doi:10.1016/j.xinn.2021.100141.

Yan T, Wang Z, Zhou H, He J, Zhou S. 2023. Effects of four critical gene deletions in Saccharomyces cerevisiae on fusel alcohols during red wine fermentation. Fermentation. 9(4):379. doi:10.3390/fermentation9040379.

*Editor: G. Brown*