Identification of the Yeast Mannoprotein Gene *HZY1* as a Key Genetic Determinant for Yeast-Derived Haze in Beer

Keith Lacy, Rita Mormando, Jeremy R. Smith, Patrick A. Gibney, Lance M. Shaner & Laura T. Burns

View supplementary material

Published online: 11 Jul 2024.

Submit your article to this journal

Article views: 231

View related articles

View Crossmark data

ASBC
THE SCIENCE OF BEER

Taylor & Francis
Taylor & Francis Group

Check for updates

Identification of the Yeast Mannoprotein Gene *HZY1* as a Key Genetic Determinant for Yeast-Derived Haze in Beer

Keith Lacy[a]*, Rita Mormando[a]*, Jeremy R. Smith[b], Patrick A. Gibney[b] (iD), Lance M. Shaner[a], and Laura T. Burns[a] (iD)

[a]Omega Yeast Labs, Chicago, IL 60641, U.S.A.; [b]Department of Food Science, Cornell University, Ithaca, NY 14850, U.S.A.

ABSTRACT

With the sustained popularity of hazy IPAs, brewers have explored multiple approaches to maximizing stable haze that will remain in suspension throughout the shelf life of the beer. Our recent investigations into yeast-dependent haze have uncovered specific brewing yeast strains that promote the formation of haze in heavily dry-hopped beer styles. These brewing strains have been termed "haze-positive" and furthermore, the timing of dry hop additions has been found to be another key factor in producing this stable haze. Classical genetics have identified YIL169C (herein referred to as *HZY1*) as both necessary and sufficient for the haze-positive phenotype in the yeast strain most widely used for Hazy IPAs. *HZY1* encodes a candidate glycoprotein and our recent findings suggest it is localized to the cell wall through a GPI anchor. Surprisingly, using long-read sequencing data we uncovered extensive genetic variation in *HZY1* across brewing strains. The haze-positive phenotype correlates with an expansion in the N-terminal serine-rich region. We propose that the Hzy1 glycoprotein is a critical component to yeast-dependent colloidal haze and the genetic variation in this locus contributes the range of haze phenotypes observed across industrial brewing strains.

Introduction

Hazy IPAs (also referred to as Juicy IPAs, New England IPA, East Coast IPA, NEIPA) have grown rapidly in popularity over the past decade, now representing at least 10% of the craft beer market. With the fast growth in hazy beer styles, brewers have faced challenges in producing stable haze that persists in the beer throughout the product shelf life. Prior beer haze research in lagers and wheat beers indicated that malt and hops were the major drivers of haze, with a mechanism involving protein-polyphenol interactions.[1–3] Many brewers early on had the misconception that non-flocculent yeast was responsible for producing haze, but this approach provided a very short-term haze and became problematic as the yeast settled out after packaging into kegs and cans. Through trial and error, one yeast strain, OYL-011 (trade names include British V, London Ale III, Foggy London, and Juice), became widely popular for Hazy IPAs due to its fruity fermentation profile, lower attenuation, and ability to produce a reliable, stable haze. As this growing category of Hazy IPAs dominated the market, we set out to better understand the genetics behind yeasts' ability to promote stable haze in dry-hopped beer styles.

There has been a strong link between haze and yeast mannoproteins in beer and wine. Yeast mannoproteins are heavily glycosylated cell wall-derived proteins with important functions in flocculation, biofilm and cell sensing.[4,5] In beer, stressed yeast has been proposed to result in increased turbidity.[6–9] Early studies examined the impact of agitation and shear stress and found the release of mannoproteins led to increased haze and difficulties in subsequent beer filtration.[7] A recent investigation into sporadic haze across a brewery's IPA production suggested that yeast mannoproteins and poor yeast management were key factors leading to increased haze.[6] Counter to the haze promoting role for yeast mannoproteins in beer, mannoproteins have been shown to prevent the formation of haze in wine. There is a strong negative correlation of glycoprotein to haze in wine and the overexpression of some mannoproteins results in decreased haze formation.[10–13] Non-*Saccharomyces* strains have been found to release polysaccharides in the form of glycoproteins as well, promoting desirable mouthfeel while also reducing haze.[14] Thus, it is important to note that the influence of specific *Saccharomyces cerevisiae* mannoproteins or mechanisms for mannoproteins in promoting or preventing haze formation in beer and wine may be distinct and largely still unknown. Here, we established

an assay for phenotyping yeast-derived haze and leveraged classic backcrossing and modern whole genome sequencing to identify the genetic component for haze in OYL-011. This approach has uncovered complex genetic variation in the *HZY1* gene and provides the first established role for this novel mannoprotein in promoting stable colloidal haze in beer.

Methods

Assay for yeast-dependent haze

Small scale fermentations of brewer's wort prepared from barley malt (bulk Rahr 2-row malt, Copeland/Synergy blend) were inoculated with each candidate yeast strain to 10 million cells/ml. Separate flask fermentations were dry hopped with 8 g/L of T90 pellet hops (USDA Triple Pearl) for each day following inoculation until day 7. After a total of 14 days of fermentation, the beer samples were centrifuged at 5,000 rpm for 5 min to remove all yeast cells and particulate. The resulting clarified beer samples were measured for haze in an Anton Paar HazeQC turbidity meter.

Yeast strains and genetic backcrossing

Genetic manipulations including crosses, sporulation and tetrad analysis were carried out using standard procedures.[15] Yeast strains used in this study can be found in the strain table (Supplemental Table S1). For genetic backcrossing experiment, the OYL-011, a haze-positive tetraploid heterozygous industrial brewing strain, was backcrossed to a haze-neutral homozygous diploid wine strain. At each backcross, the resulting haze-positive isolates were used for subsequent backcrosses until the seventh backcross in which the resulting isolates were approximately 99.2% identical to the wine strain parent.

Variant calling and variant distribution mapping

Full genome sequences with >1 Gbp of sequencing data were obtained for the parent strains and four isolates from the seventh backcross (BC7-A, BC7-B = haze-positive and BC7-C, BC7-D = haze-neutral). Quality trimming and adapter clipping was performed using Trimmomatic v0.29 with default parameters.[16] The trimmed raw reads were mapped to the *S. cerevisiae* reference genome (S288C) using Minimap2 v2.24.[17] From the resulting alignments, variants were called using Freebayes v0.9.21.[18] A custom bash script was used to parse through the resulting VCF files of each sample. A custom Python script was used to generate variant distribution plots corresponding S288C genome coordinates for each strain using Seaborn v0.12.2 and Matplotlib v3.7.0 Python packages. The position of each variant call within the four isolates was compared to the position of each variant call within the parent wine strain. A genotype ratio of the BC7 isolate to parent wine strain was plotted with a 50 bp sliding window with a step size of 5 along the S288C genome coordinates.

Cloning and comparison of HZY1 alleles

The haze-positive isolate from BC7 (OYR-329) was dissected and tetrads were PCR confirmed for the OYL-011 and wine strain alleles of *HZY1*. Each allele was PCR amplified and subcloned into a shuttling vector with AMP and HYG-B drug-resistance cassettes. The resulting vectors were sequenced with oxford nanopore long-read sequencing. The OYL-011 *HZY1* allele was unstable and exhibited frequent loss of the N-terminal and C-terminal repeat motifs and thus the PCR product was also sequenced with oxford nanopore long-read sequencing. Dotplots were generated identify repetitive regions in the S288C, OYL-011 and Wine Strain alleles of *HZY1* using a Geneious Prime 2023.2.1 (https://www.geneious.com) plugin adapted from the EMBOSS 6.5.7 tool dottup[19] with parameters Low Sensitivity/Fast, Word Size = 10 and Tile Size = 5,000.

Violin plots of HZY1 N-term and C-term lengths

Full long-read genome sequences were obtained for the Omega Yeast Labs collection using Oxford Nanopore sequencing. Three sequence tags were designed in the highest conserved regions surrounding and within the *HZY1* gene (Supplemental Table S2). All reads were mapped to the tag1 sequence using Minimap2 v2.24.[17] The resulting paf file was converted to a bed file resulting in subsequence fragments starting/ending with the sequence tag using a custom python script. The subsequences were then extracted using BEDTools v2.30.0.[20] Using the same methods, the tag1 containing reads were mapped to the tag2 sequence and the resulting sequences trimmed to tag1 and tag2 were combined into one fasta file. The length of all the trimmed reads for each sample were obtained using the sequence-stats v1.1 bash package. Violin plots were generated to show the distribution of N-term lengths using the Ggplot2 v3.4.2 package in R.[21,22] The C-term plots were made following the same methodology as listed above with the exception that the reads were first mapped to tag2 and then the tag3 sequence.

HZY1 disruption and HZY1-011 allele swap

Plasmids and oligonucleotides can be found in Supplemental Table S1. *HZY1* was disrupted in OYL-004, OYL-011, OYL-009 and OYL-077 using CRISPR/Cas9 gene editing. A plasmid containing the CRISPR/cas9 enzyme, *HZY1* targeting sgRNA, G418 drug-resistance cassette (pOY108) and an ssDNA oligo repair template with homology to the 5'UTR and 3'UTR (oligo 256) was transformed using standard Li/Ac transformation protocol.[23] Selection with G418 was used to screen for pOY108 transformants. Subsequently, G418+ transformants were screened to confirm loss of endogenous *HZY1* (primers 229/257) and the resulting scar after *HZY1* deletion (primers 257/258). The corresponding *hzy1Δ* isolates were outgrown and G418- isolates were selected for further experiments. Using a similar approach, a plasmid containing the CRSIPR/cas9 enzyme, *HZY1* targeting sgRNA and G418 drug-resistance cassette (pOY108) was

co-transformed with a plasmid containing the OYL-011 long allele of *HZY1* (*HZY1-011*) including 794 bp of homology upstream and 1350 bp downstream and HYGB drug-resistance cassette (pOY130) were transformed into OYL-106. G418+/ HYG-B + transformants were screened with primers to confirm the loss of endogenous *HZY1* (primers 229/257) and successful swap to the *HZY1-011* allele (primers 260/267). A PCR product of the full region was also verified with direct nanopore sequencing to confirm the presence of the intact ORF.

Brewing trial and tetrad test

An IPA wort was prepared with 85% 2-row base malt and 15% Munich malt to 16.9 plato. Hot side hop additions included 1 g/L Mosaic at 10 min remaining in the boil and 2 g/L of Citra at the beginning of a 15-minute whirlpool. The wort was chilled to 20°C, aerated with oxygen and transferred to two fermentation vessels. OYL-011 or OYL-011 *hzy1Δ* yeast strains were pitched at 10 million cells/ml and fermentations were maintained at 21°C. On day 7 of fermentation, the beers were dry hopped with 16 g/L of Citra. The fermentations were complete on day 14 and were cooled to 4°C for 4 days before transferring to serving vessels to carbonate. Tetrad sensory analysis was performed with a panel of eleven tasters. The beers were served blind in opaque cups and covered to ensure tasters were unable to see differences in the turbidity.

Results

To begin investigating the elusive role for specific yeast strains in promoting haze in beer, we developed an assay for phenotyping haze. We started by testing the popular Hazy IPA strain, OYL-011, and the OYL-004 control strain commonly used for non-hazy American craft styles. Fermentations were performed in 500 ml flasks and dry-hop additions were added during the fermentation to promote the formation of haze. After fermentation was complete (14 days), the samples were collected, centrifuged and measured for haze. The late fermentation dry hop additions between day four and day seven showed the greatest difference in haze measurements in the OYL-011 and OYL-004 strains (Figure 1A,B, respectively). The day seven dry hop was used to phenotype a collection of brewing strains (Figure 1C). A range of haze phenotypes was observed and 200 NTUs was used to define haze-positive (>200 NTUs) and haze-neutral (<200 NTUs) strains. We decided on a 200 NTU threshold for the haze-positive phenotype with this being a very obvious haze and differences in haze measured above 200 NTUs can be difficult to discern by eye. Traditional English and American ale strains were among the most haze-positive strains, whereas Belgian ale strains and German lager strains were the most haze-neutral.

With an established haze phenotyping assay, we set out to determine the genetics of haze. We began by crossing the haze-positive OYL-011 to a haze-neutral wine strain (Figure 2A). We chose the Maxithiol wine strain for these

crossing experiments because it was previously established to be homozygous diploid and therefore could be further used for backcrossing experiments. From the first cross, three hybrid strains were obtained, two haze-positive and one that was haze-neutral indicating the possibility of the haze-positive phenotype linking to a dominant Mendelian trait. We selected one of the haze-positive isolates and continued through seven successive backcrosses (Figure 2B). Viability of OYL-011 spores was <50% and spore viabilities of the resulting crosses with Maxithiol remained compromised in our first three rounds of backcrossing. The zygote viabilities and 2:2 segregation of the haze phenotype was observed after the 4th round of backcrossing once the genome was stabilized to diploid. At backcross 7 (BC7) we obtained two haze-positive (BC7-A, BC7-B) and haze-neutral isolates (BC7-C, BC7-D) that were 99.2% isogenic to the wine strain parent (Figure 2C). We used Illumina sequencing to obtain >1 Gbp of whole genome sequencing data for the parent strains along with the BC7-A, BC7-B, BC7-C and BC7-D isolates. Variant calling was performed for the BC7 isolates and the wine strain against the reference genome S288C. Variant frequencies were plotted by the corresponding genome coordinates as a ratio between the BC7 isolates to the wine strain (Supplemental Figure S1). One region on the left arm of chromosome IX (0-30,000 bp) contained the candidate haze locus, as it was the only region that contained variants unique to the BC7-A and BC7-B haze-positive isolates and not the wine strain parent or BC7-C and BC7-D haze-neutral isolates (Figure 2D).

To further narrow down this region of chr IX to a candidate haze gene, we mapped the short read sequencing data of the parent strains and BC7 isolates to the S288C reference genome. The region between 23,000-26,000 bp showed a several-fold increase in coverage in the OYL-011 parent and BC7-A and BC7-B isolates, suggesting two potential repeat expansions within YIL169C at 25,400 bp and 23,800 bp corresponding to the N-terminus and C-terminus (Figure 3A,B). Since the BC7-A and BC7-B isolates were heterozygous for YIL169C, we sporulated BC7-A to obtain meiotic segregants that were homozygous diploid and either haze-positive or haze-neutral. We then assayed the meiotic segregants for haze. Primers were designed to amplify the N-terminus of YIL169C. Two products of distinct size were identified, with the "long" allele corresponding to the haze-positive phenotype (Figure 3C). Furthermore, PCR products of the full-length OYL-011 "long" and wine strain "short" alleles of YIL169C were directly sequenced using nanopore sequencing and aligned to the YIL169C allele of S288C, confirming the haze-positive OYL-011 allele contains expansions in both the N-terminal and C-terminal regions (Figure 3D). From the amino acid sequence, we noted that the N-terminus is heavily enriched in serine (OYL-011 "long" >55%) and the C-terminus in serine and threonine (OYL-011 "long" >45%). Repeat motifs rich in serine and threonine have been reported in other cell wall proteins.[24] When further examining the YIL169C alleles, we identified two repeat motifs, a 14 aa motif in the N-terminus and a 36 aa motif in the C-terminus (Figure 3E). These repetitive regions can be visualized with self-aligned DNA dotplots,
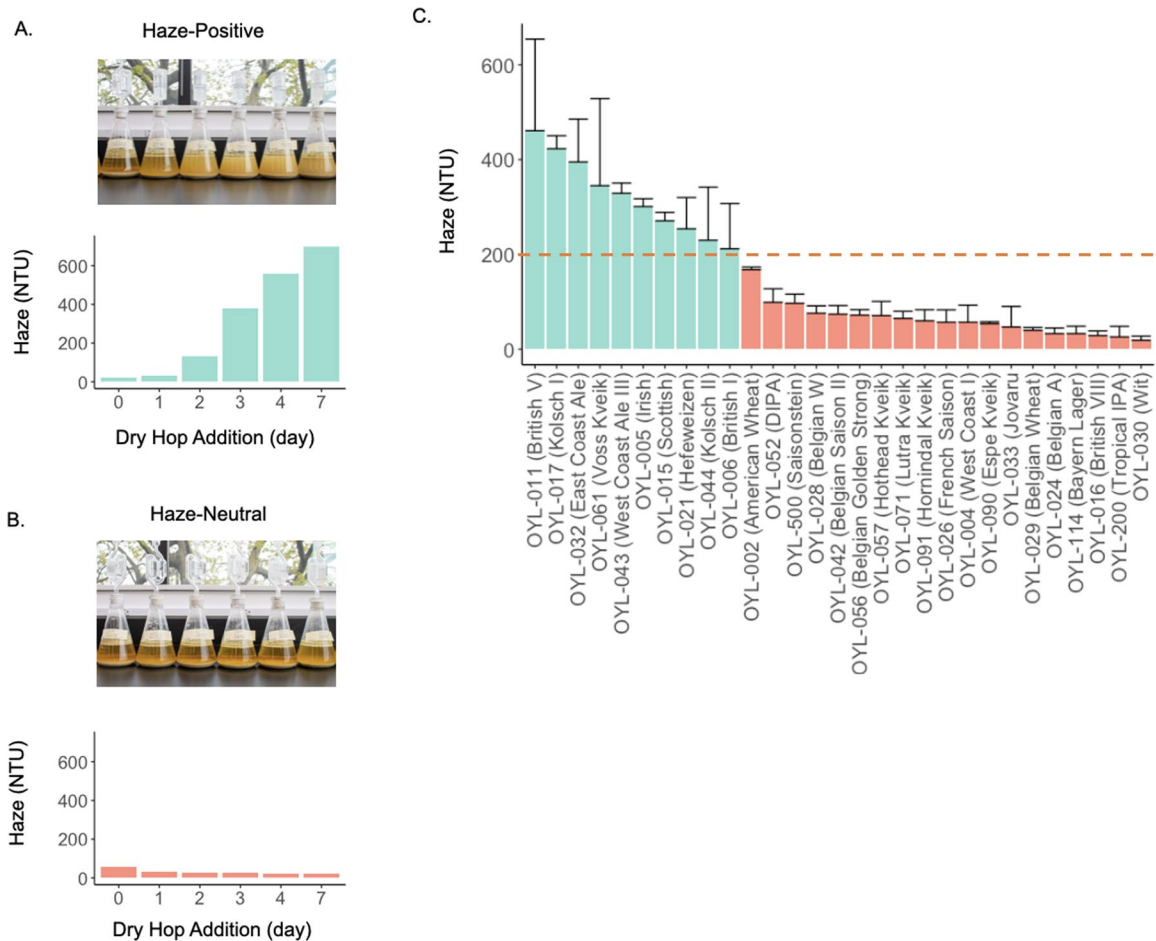
Figure 1. Characterization of the haze phenotype. (A) Image and haze measurements documenting the haze-positive phenotype of OYL-011. The flasks displayed were measured for NTUs reported in the graphs below. From left to right, the flasks were dry hopped on day 0, 1, 2, 3, 4, and 7. The images and haze measurements were taken after fermentation was complete on day 14. Haze increased in OYL-011 fermentations with dry hopping occurring in later fermentation. (B) Image and haze measurements documenting the haze-neutral phenotype of OYL-004. Haze did not form in OYL-004 fermentations regardless of dry-hop timing. (C) Measurements of haze with day seven dry hop addition in a collection of brewing strains. The average of a minimum of three experimental replicates are plotted for each strain with error bars representing standard deviation. Red dashed line at 200 NTUs indicates cutoff to define haze-positive and haze-neutral phenotype.

which can be used to help identify low complexity regions and structural variations in DNA sequences. Two boxes of varying size representing regions of repetitive DNA can be visualized in the N-term and C-term regions of the YIL169C alleles, with the most extensive repeats (multiple diagonal lines) observed in the OYL-011 allele (Supplemental Figure S1). The N-terminal repeat was expanded 53 times in the OYL-011 strain, but was only found 15 times in the S288C lab strain and 7 times in the wine strain. The C-terminal repeat was expanded 17 times in the OYL-011 strain, while only identified once in both the S288C lab strain and wine strain. Another unanticipated finding was a predicted GPI anchor present in the wine parent strain and OYL-011, but not in the S288C lab strain. Through variant calling and SNP analysis we were unable to observe the S288C lab strain mutation resulting in the loss of GPI anchor in any of our brewing strains (data not shown). The loss of the GPI anchor in S288C suggests that the name for YIL169C, *CSS1* for "condition-specific secreted," named as a part of a large-scale localization screen, may not be accurate.

The presence of a GPI anchor may indicate that the YIL169C protein is cell-wall associated, whereas the lab strain Css1 protein without a GPI anchor was observed to be secreted into the media when constitutively expressed. Therefore, we have renamed YIL169C, *HZY1,* and herein provide the first evidence that *HZY1* is necessary for promoting stable colloidal haze in beer.

We hypothesized that if the repeat expansions in OYL-011 *HZY1* allele result in the haze phenotype, then other haze-positive brewing strains may also have repeat expansions in *HZY1*. To look for potential expansions, we obtained long-read nanopore sequencing data for a collection of strains and used three highly conserved sequences upstream (tag1 ChrIX:25,878-26,392), in the central domain (tag2 ChrIX:24,727-25,335) and downstream (tag3 ChrIX:22,324-22,442) of the *HZY1* gene to extract reads containing the full N-term and C-term sequences (Figure 4A). Due to many brewing strains having heterozygous tetraploid genomes and often aneuploidies, we collected the sequence lengths from tag1-tag2 (N-term) and tag2-tag3 (C-term) to
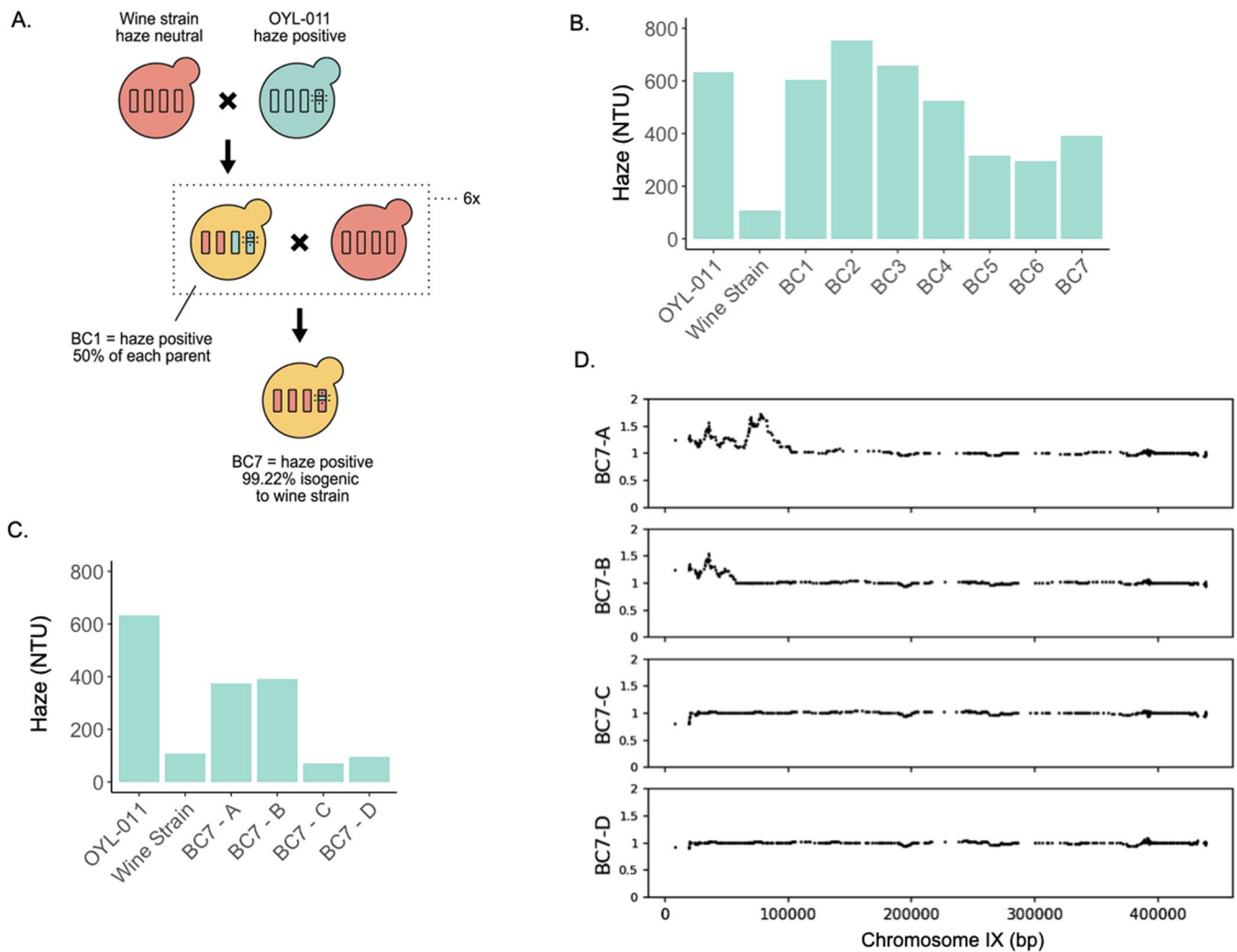
Figure 2. Backcrossing OYL-011 and identification of candidate haze locus in left telomeric region of Chr IX. (A) Schematic representation of the OYL-011 and wine strain backcrossing. (B) Haze measurements of haze-positive isolates from each backcross. (C) Haze measurements of the parent strains and the two haze-positive and two haze-neutral BC7 isolates. (D) Variants specific to the two haze-positive BC7 isolates map to a candidate haze locus on the left arm of Chr IX. The ratio of variants found in the BC7 isolates relative to the parent wine strain (y-axis) are plotted with a sliding window of 50 bp along the Chromosome IX coordinates of the reference genome S288C (x-axis).

capture all potential *HZY1* alleles in each strain. The sequence lengths were then plotted using violin plots to look at both the size and distribution of the N-terminus (Figure 4B) and C-terminus (Figure 4C) in the various brewing strains. Several strains with long N-term regions had previously been identified as haze-positive (OYL-011, OYL-017, OYL-032, OYL-045) in our initial screening. Three haze-neutral strains (OYL-004, OYL-024, OYL-052) were identified to have low frequency alleles with N-term expansions (Figure 4B). While performing a more targeted analysis of these N-term expansions, we observed cases of tandem duplications including the N-term and central domain (data not shown). We could not further resolve these into full coding regions due to the average read lengths, low read accuracy and inherent difficulties resolving the repetitive regions of *HZY1*. It remains possible that these types of tandem duplications influence the haze phenotype of specific brewing strains. We predict that copy number and expression of *HZY1* may be partly responsible

for the degree of haze phenotypes observed. Also, several haze-positive strains did not have long *HZY1* alleles (OYL-061, OYL-043, OYL-015, OYL-021) and therefore it is possible that additional genes are promoting haze in these strains (Figure 4B). The C-term expansion exhibited smaller variation with the majority of expansions falling within a 500 bp distribution (Figure 4C). There was not an observed trend between length of the C-term expansions and haze, though this does not rule out the possibility of a combination of N-term and C-term expansions involved in haze. Due to the length of the reads and the size of the *HZY1* expansions, evaluation of the combinations of N-term and C-term expansions was not possible. Interestingly, OYL-001, OYL-009, OYL-077 were identified to have N-term expansions in *HZY1* and when assayed for haze were found to be among the most haze-positive strains, suggesting the N-term length of *HZY1* is partially predictive of a haze-positive phenotype (Figure 4D). To determine whether N-term length of *HZY1* correlates to the haze phenotype,
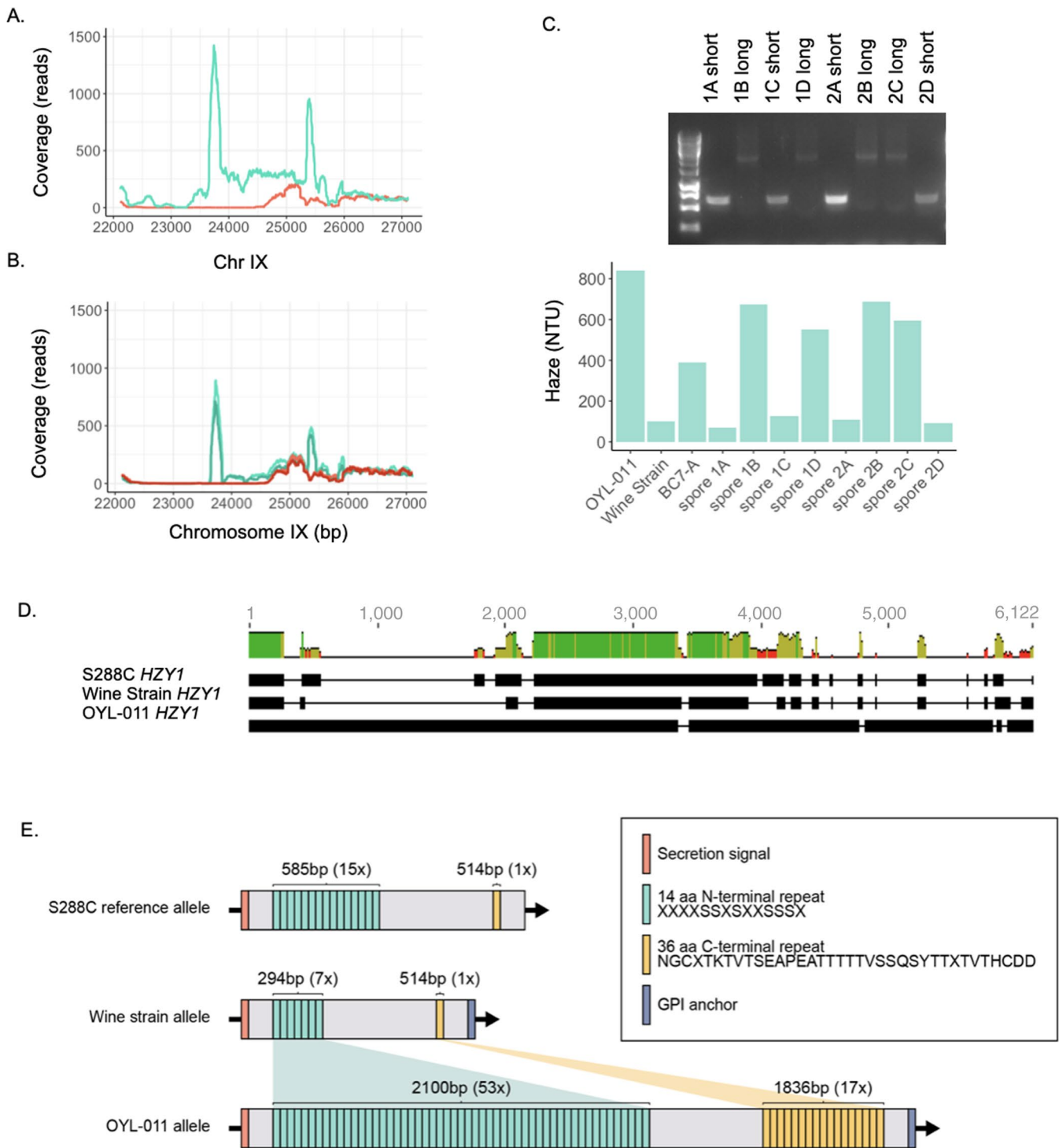
Figure 3. Large intergenic repeat expansions in the OYL-011 *HZY1* allele are associated with haze. (A) Coverage plot using short reads from Illumina whole genome sequencing from the parent wine strain (red line) and the parent OYL-011 strain (blue line) mapped to S288C reference genome. Regions within the N-term and C-term show increased coverage in the OYL-011 strain indicating potential repeat expansions in regions of *HZY1*. (B) Coverage plot using short reads from Illumina whole genome sequencing for the BC7-A and BC7-B isolates (light green, dark green) and BC7-C and BC7-D (light red, dark red). BC7-A and BC7-B isolates also exhibit increased coverage in the N-term and C-term regions. (C) Genotyping of the BC7-A spores for *HZY1* N-terminal expansion, short allele (~580 bp) and long allele (~2415 bp) and the correlation of long allele to the haze-positive phenotype. (D) Alignment of S288C, wine strain and OYL-011 *HZY1* alleles. (E) Schematic representation of the intragenic repeats in *HZY1* for the S288C, wine strain and OYL-011 alleles. Legend indicates identified repeat motifs along with candidate secretory and GPI anchor sequences.

we determined the allele with the longest N-term length (qualified by ≥0.2 allele frequency) in each strain and used the Pearson's correlation test to determine the strength of correlation between the length of the N-terminus of *HZY1* and average NTUs from our haze assay. We found a correlation coefficient of 0.54 which indicates a moderate
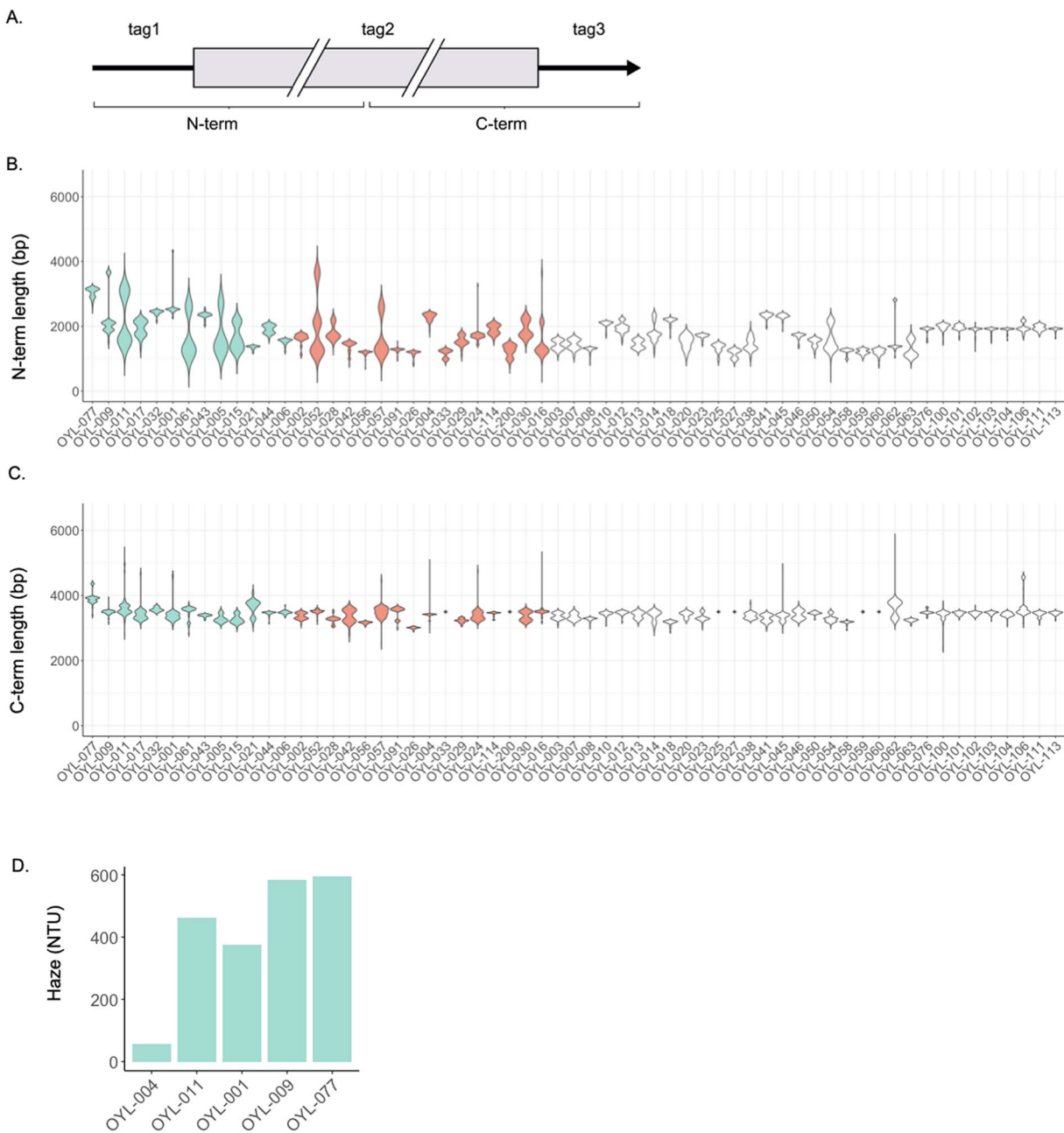
Figure 4. N-terminal and C-terminal expansions in *HZY1* in a collection of brewing strains. (A) Schematic of sequence tags used to extract reads from long read sequencing datasets and determine lengths of N-terminus (tag1 and tag2) and Cterminus (tag2 and tag3) in *HZY1* alleles. (B) Violin plots indicating the size and distribution the N-terminus (region between tag1 and tag2) in long reads obtained from various brewing strains. Haze-positive are indicated as green and haze-neutral as red. Strains uncharacterized for haze phenotype are white. (C) Violin plots indicating the size and distribution the C-terminus (region between tag2 and tag3) in long reads obtained from various brewing strains. Haze-positive are indicated as green and haze-neutral as red. Strains uncharacterized for haze phenotype are white. (D) Haze phenotype of strains identified to have expanded *HZY1* N-terminus.

positive correlation between size of the N-term expansion and haze-phenotype.

With a moderate correlation between repeat expansions in *HZY1* N-term to the haze phenotype, we next disrupted *HZY1* using CRISPR/Cas9 in several of the most haze-positive strains (OYL-011, OYL-009, OYL-077) as well as in one of the haze-neutral strains (OYL-004). The disruption of each allele was confirmed by PCR with complete loss of product (primers 257/229, 213 bp) for *HZY1* N-terminus and gain of product (primers 257/258, 345 bp)

indicating *HZY1* disruption (Figure 5A). Each of the resulting *hzy1Δ* strains showed a substantial decrease in haze (Figure 5B). Even the OYL-004 haze-neutral strain showed reduced haze with *HZY1* disruption. This decrease in haze confirmed *HZY1* was necessary for haze formation in these strains. To determine if the OYL-011 long *HZY1* allele (*HZY1-011*) was sufficient to induce haze in a haze-neutral strain, we used CRISPR/Cas9 to swap the *HZY-011* allele into the endogenous *HZY1* locus in a haze-neutral lager strain, OYL-106. We observed an increase in haze from 48
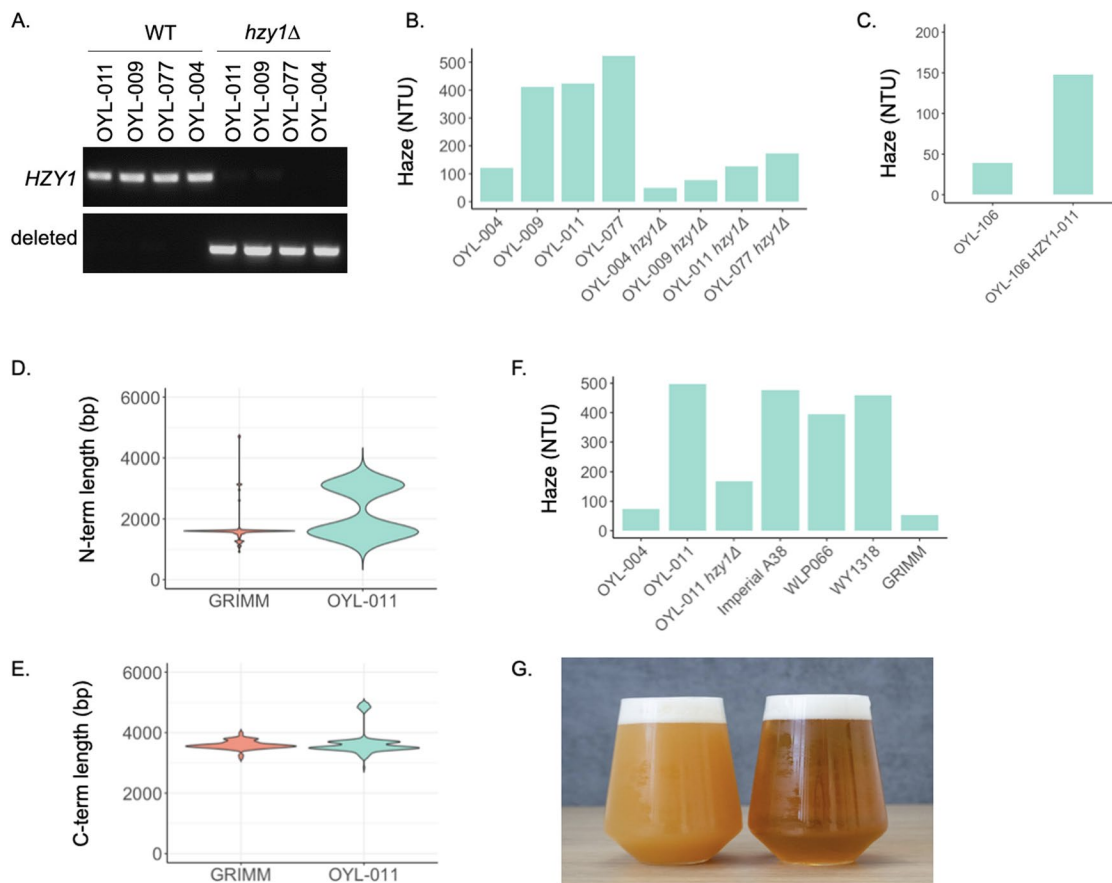
Figure 5. *HZY1* is necessary for the formation of dry hop depending haze. (A) PCR confirmation of the full CRISPR/Cas9 disruption of *HZY1* gene in all *hzy1Δ* strains. (B) The resulting haze phenotype of *hzy1Δ* strains. (C) The resulting haze phenotype of OYL-106 haze-neutral lager strain and OYL-106 with *HZY1-011* allele substitution of the endogenous *HZY1* locus. (D) Distribution of *HZY1* N-term length in GRIMM and OYL-011. (E) Distribution of *HZY1* C-term length in GRIMM and OYL-011. (F) Haze phenotypes of various commercial sources for OYL-011 equivalent strains and a generation 300 isolate from Grimm Artisanal Ales (GRIMM). (G) Typical IPA recipe fermented with OYL-011 (left) and OYL-011 *hzy1Δ* (right).

to 150 NTUs (Figure 5C). This 200% increase in haze is noticeably hazier by eye then the OYL-106 parent strain, but did not reach the level of OYL-011 which is typically >500 NTUs. In the genetic backcrossing experiment, we were able to localize the haze phenotype to the first 30 kb region of the left telomere Chr IX containing *HZY1* and by disrupting only the *HZY1* coding region haze was reduced. These results combined indicate that the *HZY1* gene is necessary and sufficient for the haze phenotype.

While this study was underway, we were in touch with the brewery owners at Grimm Artisanal Ales who had been having difficulty obtaining stable haze using their house culture (GRIMM), which was a derivative of OYL-011 from another commercial yeast supplier. A sample of the GRIMM strain was tested and found to be haze-neutral. The brewery had maintained their house culture for 300 generations by top cropping yeast in active fermentation and utilizing it for successive fermentations. We then assayed several of the other yeast supplier equivalent strains for OYL-011 and found all but the GRIMM isolate to be haze-positive (Figure 5F). We obtained long-read sequencing on the GRIMM isolate to see if *HZY1* had undergone genetic changes that might result in the loss of the haze phenotype. Surprisingly, the GRIMM strain differed from the OYL-011

strain in that it no longer had expansions in the N-term or C-term of *HZY1* (Figure 5D,E). The loss of a long allele in the non-hazy GRIMM isolate suggests that expansion and contraction (or possibly *HZY1* disruption) can result in the gain and loss of the haze phenotype in a brewery setting and provides further evidence for the role of expanded *HZY1* alleles in promoting haze.

The typical beer consumer is heavily influenced by the appearance of beer and haze is often associated with either an increase or decrease in hedonic liking. We set out to determine whether this is strictly based on appearance and associations to beer styles, or if haze also has an impact on the taste and aroma of the beer. A standard IPA recipe was brewed and split into two fermentations, one with OYL-011 and one with OYL-011 *hzy1Δ*. The haze in the resulting beers measured 428 NTUs and 40 NTUs, respectively. This difference was very visually striking and would present very differently to the beer consumer (Figure 5G). A tetrad sensory test was performed where the beer samples were kept covered and in opaque cups with covers to prevent panelists from determining which was hazy or not hazy. Only one out of eleven panelists was able to identify the correct pairing, indicating that the aroma, mouthfeel and taste were not statistically different between the hazy and non-hazy beers.

Discussion

This study investigated the genetics of beer haze in *Saccharomyces cerevisiae*. Much of the haze research in beer to date has focused on malt and hop contributions but a role for yeast has remained largely unknown. From our initial experiments characterizing dry hop-dependent haze in a collection of commercial brewing strains, it became apparent that brewing strains exhibit a wide range of haze phenotypes (referred to as haze-positive and haze-neutral). Employing a classic genetic backcrossing approach, we identified YIL169C, which we have termed *HZY1*, as a novel haze gene. Through *HZY1* knockout experiments and the moderate correlation between the expansion of *HZY1* intragenic repeats and haze-positive phenotype amongst brewing strains, we were able to provide further support a role for *HZY1* in promoting the haze phenotype. While we found that expression of the OYL-011 long *HZY1* allele was sufficient to induce haze when expressed in the haze-neutral strain OYL-106, haze levels did not reach the same as seen with the OYL-011. These results indicate there are other mechanisms that modulate haze levels, perhaps through additional genes that directly contribute to haze or via regulation of *HZY1*. Together our results provide the first evidence that the *S. cerevisiae* gene, *HZY1*, has a critical role in promoting haze in dry-hopped beer styles.

The *HZY1* gene encodes a mannoprotein with a putative GPI anchor and includes intragenic repeats in both the N-terminal serine-rich and C-terminal serine/threonine-rich domains. *HZY1* intragenic repeats possess a high degree of genomic plasticity with massive expansions resulting in a range of haze phenotypes observed amongst brewing strains. The majority of whole genome sequencing studies with brewing strains have employed short read sequencing technologies.[25] Though short read data provides a highly accurate coverage of the genome, complex repetitive sequence cannot be resolved and thus *HZY1* has largely remained uncharacterized. *HZY1* also is located in the subtelomeric region of chromosome IX, and repetitive subtelomeric regions are often genomic hotspots for mitotic recombination and interchromosomal reshuffling,[24,26,27] which partially can help to explain the degree of genomic and phenotypic variability that we have uncovered in this study. In addition, the complex genetics of domesticated *S. cerevisiae* brewing strains including heterozygosity, polyploidy, and chromosomal rearrangements might also suggest that domesticated yeast have a higher degree of evolution in the *HZY1* gene.[26] In future studies, long-read sequencing approaches can help to resolve and phase different *HZY1* alleles across wild and domesticated strains. These studies may also uncover additional yeast genes contribute to the haze in various fermented beverages. A better understanding of *HZY1* allele types and frequencies will help to further elucidate this complex haze trait.

Flocculation and maltose utilization traits exhibit a high degree of biodiversity within brewing strain phenotypes and genotypes.[25] There are clear selective pressures for these two phenotypes as brewers often harvest from the bottom of the fermentation vessel and use wort primarily composed of maltose. However, the selective pressure for haze and any possible fitness advantage to *HZY1* intragenic expansions remains unknown. The expansion of these serine and threonine-rich regions in Hzy1 could lead to heavy glycosylation which could change the cell surface composition, possibly altering cellular adhesion and responses to environmental stress. Among *S. cerevisiae* genes with long intragenic tandem repeats, 75% encode for cell-surface proteins. Many of these genes also exhibit expansions and contractions in the intragenic repeats, including *FLO1, FLO5, FLO9, FLO11* and *HPF1* and impact cell adherence and buoyancy.[24,28] We have not found an association between haze-positive and flocculation phenotypes within brewing strains, but it is possible that the *HZY1* repeat expansions were selected for due to their effect on cell buoyancy. Interestingly, the lager strains analyzed in this study exhibited little diversity in *HZY1* relative to ale strains and some of the hazier ale strains have been historically top cropped. If *HZY1* expansion leads to more buoyant cells, then these cells would be preferentially selected through top cropping, similar to the flocculant cells being selected through bottom cropping.

Our study provides the first evidence for the involvement of *S. cerevisiae* and the gene, *HZY1*, in the promotion of dry hop-dependent haze. The study will likely fuel future research investigating additional genes involved in haze, as well as biochemical studies to determine the composition and the biophysical properties of beer haze. Ultimately, the knowledge shared within will leave brewers with a deeper understanding of yeast-derived haze and the strains and methods that can be used to perfect their hazy (or non-hazy) brands.

Data access

Long-read sequencing data for *HZY1* can be provided upon request.

Acknowledgements

Disclosure statement

A provisional patent application has been filed by Omega Yeast Labs LLC based on the results of this work.

ORCID

Patrick A. Gibney 🆔 http://orcid.org/0000-0003-1392-7553
Laura T. Burns 🆔 http://orcid.org/0000-0002-1774-882X

Literature cited

[1] Bamforth, C. W. Beer Haze. *J. Am. Soc. Brew. Chem.* **1999**, *57*(3), 81–90. DOI: 10.1094/ASBCJ-57-0081.

[2] Asano, K.; Shinagawa, K.; Hashimoto, N. Characterization of Haze-Forming Proteins of Beer and Their Roles in Chill Haze Formation. *J. Am. Soc. Brew. Chem.* **1982**, *40*(4), 147–154. DOI: 10.1094/ASBCJ-40-0147.

[3] Siebert, K. J.; Lynn, P. Y. Effects of Alcohol and pH on Protein–Polyphenol Haze Intensity and Particle Size. *J. Am. Soc. Brew. Chem.* **2003**, *61*(2), 88–98. DOI: 10.1094/ASBCJ-61-0088.

[4] Verstrepen, K. J.; Klis, F. M. Flocculation, Adhesion and Biofilm Formation in Yeasts. *Mol. Microbiol.* **2006**, *60*(1), 5–15. DOI: 10.1111/j.1365-2958.2006.05072.x.

[5] Orlean, P. Architecture and Biosynthesis of the Saccharomyces cerevisiae Cell Wall. *Genetics* **2012**, *192*(3), 775–818. DOI: 10.1534/genetics.112.144485.

[6] Huismann, M.; Gormley, F.; Dzait, D.; et al. Unfilterable Beer Haze Part II: Identifying Suspect Cell Wall Proteins. *J. Am. Soc. Brew. Chem.* **2021**, *80*(1), 26–34. JASBC-2021-0029.

[7] Siebert, K. J.; Stenroos, L. E.; Reid, D. S.; Grablowski, D. Filtration Difficulties Resulting from Damage to Yeast During Centrifugation. *Mbaa TQ* **1987**, *24*(1), 1–8.

[8] Chlup, P. H.; Bernard, D.; Stewart, G. R. The Disc Stack Centrifuge and Its Impact on Yeast and Beer Quality. *J. Am. Soc. Brew. Chem.* **2007**, *65*(1), 29–37.

[9] Chlup, P. H.; Conery, J.; Stewart, G. G. Detection of Mannan from *Saccharomyces cerevisiae* by Flow Cytometry. *J. Am. Soc. Brew. Chem* **2007**, *65*(3), 51–156.

[10] Fusi, M.; Mainente, F.; Rizzi, C.; Zoccatelli, G.; Simonato, B. Wine Hazing: A predictive Assay Based on Protein and Glycoprotein Independent Recovery and Quantification. *Food Control* **2010**, *21*(6), 830–834. DOI: 10.1016/j.foodcont.2009.11.009.

[11] Dupin, I. V.; McKinnon, B. M.; Ryan, C.; Boulay, M.; Markides, A. J.; Jones, G. P.; Williams, P. J.; Waters, E. J. *Saccharomyces cerevisiae* Mannoproteins That Protect Wine from Protein Haze: Their Release During Fermentation and Lees Contact and a Proposal for Their Mechanism of Action. *J. Agric. Food Chem.* **2000**, *48*(8), 3098–3105. DOI: 10.1021/jf0002443.

[12] Schmidt, S. A.; Tan, E. L.; Brown, S.; Nasution, U. J.; Pettolino, F.; Macintyre, O. J.; Lopes, M. d B.; Waters, E. J.; Anderson, P. A. Hpf2 Glycan Structure Is Critical for Protection Against Protein Haze Formation in White Wine. *J. Agric. Food Chem.* **2009**, *57*(8), 3308–3315. DOI: 10.1021/jf803254s.

[13] Brown, S. L.; Stockdale, V. J.; Pettolino, F.; Pocock, K. F.; de Barros Lopes, M.; Williams, P. J.; Bacic, A.; Fincher, G. B.; Høj, P. B.; Waters, E. J.; et al. Reducing Haziness in White Wine by Overexpression of *Saccharomyces cerevisiae* Genes *YOL155c* and *YDR055w*. *Appl. Microbiol. Biotechnol.* **2007**, *73*(6), 1363–1376. DOI: 10.1007/s00253-006-0606-0.

[14] Domizio, P.; Liu, Y.; Bisson, L. F.; Barile, D. Use of Non-*Saccharomyces* Wine Yeasts as Novel Sources of Mannoproteins in Wine. *Food Microbiol.* **2014**, *43*, 5–15. DOI: 10.1016/j.fm.2014.04.005.

[15] Guthrie, C.; Fink, G. R. Guide to Yeast Genetics and Molecular and Cell Biology - Part C. *Methods Enzymol.* **2002**, *351*, 3–735.

[16] Bolger, A. M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*(15), 2114–2120. DOI: 10.1093/bioinformatics/btu170.

[17] Li, H. Minimap2: Pairwise Alignment for Nucleotide Sequences. *Bioinformatics* **2018**, *34*(18), 3094–3100. DOI: 10.1093/bioinformatics/bty191.

[18] Garrison, E.; Marth, G. Haplotype-Based Variant Detection from Short-Read Sequencing. **2012**.

[19] Madeira, F.; Pearce, M.; Tivey, A. R. N.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R. Search and Sequence Analysis Tools Services from EMBL-EBI in 2022. *Nucleic Acids Res.* **2022**, *50*(W1), W276–W279. DOI: 10.1093/nar/gkac240.

[20] Quinlan, A. R.; Hall, I. M. BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics* **2010**, *26*(6), 841–842. DOI: 10.1093/bioinformatics/btq033.

[21] Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer-Verlag: New York, **2016**. https://ggplot2.tidyverse.org.

[22] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria, **2022**. https://www.R-project.org/.

[23] Gietz, R. D.; Schiestl, R. H. High-Efficiency Yeast Transformation Using the LiAc/SS Carrier DNA/PEG Method. *Nat. Protoc.* **2007**, *2*(1), 31–34. DOI: 10.1038/nprot.2007.13.

[24] Verstrepen, K. J.; Jansen, A.; Lewitter, F.; Fink, G. R. Intragenic Tandem Repeats Generate Functional Variability. *Nat. Genet.* **2005**, *37*(9), 986–990. DOI: 10.1038/ng1618.

[25] Gallone, B.; Steensels, J.; Prahl, T.; Soriaga, L.; Saels, V.; Herrera-Malaver, B.; Merlevede, A.; Roncoroni, M.; Voordeckers, K.; Miraglia, L.; et al. Domestication and Divergence of *Saccharomyces cerevisiae* Beer Yeasts. *Cell* **2016**, *166*(6), 1397–1410.e16. DOI: 10.1016/j.cell.2016.08.020.

[26] Yue, J.-X.; Li, J.; Aigrain, L.; Hallin, J.; Persson, K.; Oliver, K.; Bergström, A.; Coupland, P.; Warringer, J.; Lagomarsino, M. C.; et al. Contrasting Evolutionary Genome Dynamics Between Domesticated and Wild Yeasts. *Nat. Genet.* **2017**, *49*(6), 913–924. DOI: 10.1038/ng.3847.

[27] Brown, C. A.; Murray, A. W.; Verstrepen, K. J. Rapid Expansion and Functional Divergence of Subtelomeric Gene Families in Yeasts. *Curr. Biol.* **2010**, *20*(10), 895–903. DOI: 10.1016/j.cub.2010.04.027.

[28] Barré, B. P.; Hallin, J.; Yue, J.-X.; Persson, K.; Mikhalev, E.; Irizar, A.; Holt, S.; Thompson, D.; Molin, M.; Warringer, J.; et al. Intragenic Repeat Expansion in the Cell Wall Protein Gene. *Genome Res.* **2020**, *30*(5), 697–710. DOI: 10.1101/gr.253351.119.