4-29-2021

Brewing Meets Biotechnology: A Survey of Genetic Variation in Flocculation Proteins

Sarah A. Dooley

Brewing Meets Biotechnology:

A Survey of Genetic Variation in Flocculation Proteins

Sarah A. Dooley

Department of Biological and Environmental Sciences
Longwood University, 201 High Street, Farmville, VA 23909
Faculty Advisor: Dr. Dale Beach
April 29, 2021

LONGWOOD
SENIOR THESIS

This senior thesis is submitted in partial fulfillment of the requirements for

Honors in Biology

Committee Members: Dr. Consuelo Alvarez (Longwood University), Dr. Erin Shanle (Longwood University), and Dr. Jonathan White (Longwood University).

Brewing Meets Biotechnology:

A Survey of Genetic Variation in Flocculation Proteins

by

Sarah A. Dooley

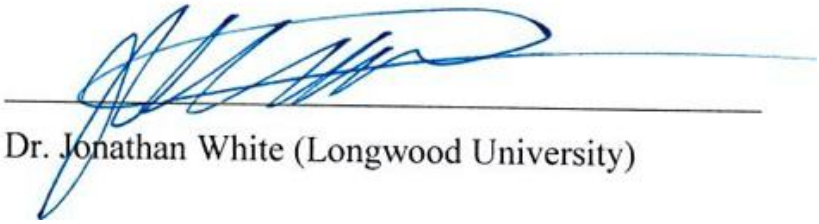
This thesis has been read and approved by the following supervisory committee and submitted in its final draft to the Longwood Senior Thesis Committee in the conferment of Honors towards the degree of Bachelor in Biology.

Committee Members:

Dr. Dale Beach (Longwood University)

Dr. Consuelo Alvarez (Longwood University)

Dr. Erin Shanle (Longwood University)



Dr. Jonathan White (Longwood University)

Table of Contents

List of Figures:

List of Tables:

Abstract

For thousands of years, yeast have been utilized for food and beverage fermentations and more recently have become important model organisms for biomedical research. A catalog of phenotypic characteristics available to brewers can be attributed to domestication and artificial selection as well as chromosomal rearrangements and evolutionary events. Flocculation, the natural capacity for yeast cells aggregate into flocs at the end of the fermentation process, has a direct effect on the final fermentation product.. In order to flocculate, yeast strains must carry genes belonging to the flocculin protein family that will promote cell-cell adhesion. The flocculin genes, *FLO1*, *FLO5*, and *FLO9,* are highly homologous and all share a similar protein structure. The *N*-terminal domain is recognized as the sugar binding domain, allowing flocculin producing cells to selectively bind to mannose residues displayed on the cell wall of adjacent cells. In this study, the genetic variation in the *N*-terminal domain of flocculin genes *FLO1*, *FLO5*, and *FLO9* was investigated. For each strain, the *N*-terminal domain was amplified by PCR and sequenced to assess the genetic variation between strains. Nonsynonymous variation in the flocculin proteins are predicted to correlate with variation in flocculation levels.

Dedication

This Senior Honors Research Thesis is dedicated to the biology and chemistry faculty in the Department of Biological & Environmental Sciences and the Department of Chemistry and Physics at Longwood University, all of which have encouraged me never to stop asking questions. I want to thank my research mentor, Dr. Dale Beach, for supporting me during these past four years of my undergraduate career. I will always think fondly of my time as an undergraduate research student in his lab. I would not be the scientist I am today without his mentorship. Additionally, this thesis is dedicated to my mother, Janette Owens-Dooley, who has forever instilled a strong work ethic in me. I would also like to dedicate this thesis in loving memory to two very special people my Great Aunt, Nancy Moore, and Ernest Smith, who both passed away during the completion of this thesis.

CHAPTER 1
INTRODUCTION

Domestication is a term that refers to breeding and artificially selecting wild-type species in order to pursue genetic variants that enhance desirable features (phenotypes) and specifically thrive in man-made environments. Typical genotypic and phenotypic domestication can be seen and described in pets, livestock, and crops. Interestingly, a similar event can be studied in microorganisms as well. This is seen in the development of industrial bread-making and brewing; wild yeasts from grapes and grains have been domesticated and artificially selected to produce fermented products with ideal phenotypic characteristics over the last thousand years. A powerful yet simple way to exploit biodiversity is by continually selecting a yeast strain that performs best for a particular industrial process (phenotypic outcome). As a result, more than 1,500 strains of *Saccharomyces cerevisiae* have been identified and classified (Bernard 2017). Through artificial selection, evolutionary events including random single nucleotide changes, insertion or deletion of small or large DNA fragments, and chromosomal rearrangements have altered yeast genomes to create a catalog of phenotypic characteristics available.

Flocculation is a phenotypic characteristic of interest for industrial fermentations because it leads to efficient separation of the yeast cells from the fermentation medium (Stewart et al. 1975). Flocculation occurs near the end of the fermentation process, where flocculating cells adhere to one another, forming multicellular aggregates, or ‘flocs’ which then sediment to the bottom of the fermenter to be drained out or rise to the top to be siphoned off (Bauer et al. 2010). Brewmasters and winemakers favor this phenotypic characteristic because it results in a cheap, effective, and environmentally-friendly way to

remove the majority of yeast cells from the final product without solely relying on expensive filtration systems (Vidgren and Londesborough, 2011). Not only is flocculation favorable for minimizing labor, flocculation also has a direct effect on the appearance and flavor of the final product (Speers, 2016). The appearance and flavor are primarily dependent on the timing and variability of flocculation. During fermentation, yeast cells are exposed to different stress conditions such as temperature, acidity, lack of nutrients, and high ethanol content. The formation of flocs responds to the dynamic environmental conditions such as temperature, nitrogen availability, change in pH, glucose depletion, and high ethanol content (Braus et al. 2003; Sampermans et al. 2005). Therefore, flocculation is recognized as a survival response to those harsh conditions, specifically the low sugar and high alcohol content at the end of fermentation (Vestrepen 2003).

Achieving an ideal level and duration of flocculation is challenging because flocculation is a complex process and is dependent on many factors. A low level of flocculation will produce a hazy product (e.g., German Hefeweizen, wheat beers). In comparison, a high level of flocculation leads to a clear and crisp product (e.g., Budweiser, American lagers; Chen 2018). If yeast flocculates too early the product will be under-fermented and sweet from residual sugars and if flocculation occurs too late the product will have an undesirable yeasty, bitter taste. A brewmaster or winemaker must balance the benefits derived in flocculation while considering the effect on the flavor and aroma in the finished product. Brewmasters and winemakers can manipulate environmental conditions, however there are many conditions far beyond reach that contribute to flocculation like the genetic aspects and cell biology of individual yeast

cells. A brewer must balance the benefits derived in flocculation with the effect on flavor and aroma in the finished product.

A mechanism of flocculation can be broadly described as cell-cell adhesion. The flocculation mechanism is attributed to the interaction of two molecules found on cell walls: mannose residues and flocculins (lectins). Flocculent cells have finger-like structures called lectins that protrude from the cell wall and selectively bind to the mannose subunits of adjacent cell walls (Verstrepen et al. 2003; Fig 1). Mannose residues are always present in both flocculent and non-flocculent cells, therefore the presence or absence of flocculin proteins is critical for flocculation. Additionally, flocculation is calcium-ion-dependent. Acting as a cofactor, calcium ions maintain the active conformation of lectins, therefore enhancing the degree of lectin and mannose residue interactions. Additionally, flocculation can be inhibited reversibly by the presence of sugars (Van Mulders et al., 2009). This discovery revealed two distinct flocculation phenotypes, Flo1 phenotype, and NewFlo phenotype. The Flo1 phenotype can be inhibited by mannose but not by glucose, maltose, sucrose, or galactose, whereas the NewFlo phenotype is inhibited by any of these sugar (Stratford and Assinder, 1991).

For flocculation to occur, the yeast strain must produce flocculin proteins that are correctly positioned on the outer cell wall.

The *FLO* protein family consists of five genes where *FLO1*, *FLO5, FLO9,* and *FLO11* encode flocculin proteins, and *FLO8* is a transcriptional activator. *FLO11* is found responsible for being heavily involved in biofilm formation (Verstrepen and Klis, 2006). In commonly used laboratory strains where flocculation is considered undesirable, the

FLO genes are transcriptionally silent because of a nonsense mutation in *FLO8* (Van Mulders et al. 2009). The flocculin genes, *FLO1, FLO5*, and *FLO9,* promote cell-cell adhesion (Teunissen and Steensma, 1995). Members of this protein family all share a similar structure that consists of three entities: *N*-terminal domain, central region, and *C*-terminal transmembrane domain (Fig 2; Gossens 2010). The open reading frame of the *FLO* genes encode a protein length of 1537 amino acids (Engel et al., 2014). The *N*-terminal domain is approximately 240 amino acid residues and is recognized as the sugar-binding domain (Gossens et al., 2011). The central region varies among the genes containing a series of tandem repeats resulting in variable protein lengths (Gemayel et al., 2010), and the *C*-terminal domain serves as an anchor to bind to yeast cell walls (DiGianvito 2017). Though most studies in the literature have focused on *FLO1*, the yeast flocculins, *FLO1, FLO5*, and *FLO9*, are closely related duplications of an original common ancestor.

With next-generation sequencing technologies, scientists can attempt to gain insight into the complex evolution and genetics of the *FLO* protein family in fermentation, lab, and clinical yeast strains (Liti et al., 2011; Strope et al., 2015). However, next-generation sequencing methods fail to adequately sequence FLO genes; the tandem repeat elements in the central region of these genes are too long and repetitive for the small sequence elements (150-300 bps) to assemble reliably. Nevertheless, the demands for increased productivity and changing consumer preferences within the fermentation industry show a greater interest in understanding the phenotypic aspects of industrial strains like flocculation. Using Sanger Sequencing methods, the primary aim of this study is to survey sequence variations in the *N*-terminal region of flocculin genes

FLO1, FLO5, and *FLO9* from different fermentation strains known to flocculate at different strengths.

CHAPTER 2
MATERIALS AND METHODS

Yeast Strains

The brewing yeast strains used in this study are *Saccharomyces cerevisiae* and *Saccharomyces bayanus,* listed in Table 1. All wine strains were purchased as a packet from one manufacturer Red Star (via Amazon.com). All beer strains were provided by 3 Roads Brewery, located in Farmville, Virginia.

Culture

All yeast cells were cultured in Yeast Extract Peptone Dextrose (YPD, 20% B-Peptone, 10% YE, 20% glucose) media. To simulate fermentation yeast cells were cultured in Beer analog media (66g/L CBW Golden Light Dry Malt Extract (Amazon) and 13g/L Glucose (Sigma) for 1.040 specific gravity wort). All cultures were incubated at 30°C 250 RPMs overnight. All cells were plated on solid YPD medium and incubated at 30°C overnight or until colonies were observed.

Temperature Screening

A 1:10 serial dilution of all yeast strains was performed using sterile 96-well plates. Cells were distributed via 48 Pin Microplate Replicator on to solid YPD plates. Plates were incubated for three days at room temperature, 30°C, 37°C, and 4°C. Cell growth was evaluated daily and plates were photographed.

Rapid Isolation of Yeast Genomic DNA

Single colonies were picked and inoculated into 3ml of YPD media and incubated overnight at 30°C shaking 250 RPMs. Cells (1.5 ml) from overnight cultures were pelleted by centrifugation (18,000 RPM microcentrifuge for 1 minute), and the media was discarded. The cell pellets were resuspended in 400 µl yeast lysis buffer (2% Triton

X-100, 1% SDS, 100mM NaCl, 10mM Tris-HCL (pH 8.0), 1 mM EDTA). About 400 µl of 425-600-micron glass beads (Sigma G-8772) and 200 µl Phenol: Chloroform (Fisher Scientific 50:49:1 Isoamyl alcohol) were added to the yeast suspension. Samples were placed into a multi-tube shaker (MP Bio, FastPrep-24 homogenizer) and processed for three cycles of 60 seconds at speed 6.5 M/S. In between cycles, samples were placed on an ice bath. Samples were then centrifuged for 5-10 minutes at max speed. Without disturbing the interphase, the aqueous layer was transferred to a new Eppendorf tube and an equal volume of Isopropanol was added. The samples were allowed to precipitate 2 minutes at room temperature and then centrifuged at maximum speed at room temperature. Pellets were washed with 70% ethanol (20°C) and resuspended with 100 µl of sterile, deionized water. Final genomic DNA was purified following the Monarch Spin Column Cleaning Kit protocol (NEB T1030). Concentration of final genomic DNA was measured using the Thermo Scientific™ NanoDrop 2000 (2 µl sample volume). DNA yield from liquid cultures ranged from 78.5 ng/µl to 304.6 ng/µl.

PCR

Genomic DNA was diluted to ~30ng/µl and then 1 ul was used for each PCR amplification. Genomic DNA was mixed with 2X Master-Mix (AmpliTAQ Gold 360 Master Mix), primers (0.4 µM final concentration) and sterile deionized water according to the manufacturer's recommendations. Primers used to amplify the *N*-terminal region of *FLO1, FLO5,* and *FLO9* can be found in Table 2. Cycling conditions during PCR were as follows: 94°C for 2 min, 30 cycles of 94°C for 30 sec, 50°C for 30 sec, 72°C for 90 sec, followed by 72°C for 4 min. PCR products were purified following the Monarch Spin

Column Cleaning Kit protocol (NEB, T1030) and visualized on a 1% agarose gel (3µl PCR amplicon per lane).

Sequencing

Purified PCR products were submitted for sequencing at Eurofins Genomics (https://www.eurofinsgenomics.com). Samples were submitted according to the service's recommendations, typically each reaction contained 5µl PCR generated DNA (30 ng/ul) and 5µl diluted primer. PCR primers (Table 2) at a final concentration of 2mM were used for DNA sequencing.

Sequencing Analysis

All sequencing analysis was completed with CLC Main Workbench 8.1(Qiagen). Sequence reads were trimmed based on quality score values using the Phred scale. Parameters set for trimming used in this study included a quality control limit set to 0.05.

Helm's Assay

The Helm's flocculation assay was adapted from Bendiak et al (1994). Each yeast strain was cultured in YPD or beer analog media for 2-28 days at 30°C with shaking at 250 RPMs or until yeast enter stationary phase. Cells from the yeast culture (2ml) were harvested by centrifugation (18,000xG, 1 min), washed twice with 1ml of sterile 5mM EDTA solution in water (pH 7.0). After a final wash with 1 ml of deionized water, the cell pellet was suspended into 2ml, 50mM Citrate Buffer (6.702 g Sodium citrate dihydrate) at pH 4.5. The yeast suspension was separated into a control (1 ml) and a treatment (1 ml) group, 10ul calcium chloride solution (0.5M $CaCl_2$) was added to the treatment group and in order to balance volumes, 10 µl of deionized water was added to

the control group. Each group was left to sit for 6 minutes undisturbed before absorbance was measured at 600 nm for 30 sec intervals for 6 minutes.

Percent flocculation values were determined with average absorbance measurements using the following formula:

$$PF = 100\left(\frac{(A-B)}{A}\right)$$

Which percent flocculation (*PF)* is expressed to the tenth of a percent and by the average values of absorbance of control group (A) and the average values of absorbance of treatment group (B). Any values below zero was interpreted as zero, as a strain can be nonflocculent only to zero percent (Helm et al., 1953).

CHAPTER 3
RESULTS

Temperature Screening Assay

The brewing yeast strains employed in this study have specific fermentation temperature ranges that maximize production. Optimal wine fermentation temperatures range from 15 to 30°C, while optimal beer fermentation temperatures range from 7 to 22 °C. To see if rapid cell growth could be achieved at a higher temperature, a temperature screening assay was performed. To observe cell growth, cells were subjected to a 1:10 serial dilution using a 96-well plate and then replica plated onto YPD plates. Plates were grown at room temperature, 30°C, 37°C, and 4°C until cell growth was visible at most temperatures (Fig 3). Yeast strains grew at all subjected temperatures except 4°C. The temperature assay revealed the optimal growth for all yeast strains to be 30°C.

Primer Design and Data Interpretation

The genetic variability of flocculation genes may reveal important identifiers and markers for the molecular evolution of industrial fermentation yeast strains. The *N*-terminal region of targeted genes, *FLO1, 5*, and *9,* were amplified using PCR. Since the *FLO* gene family shares high sequence homology, *FLO1* being 96% homologous to *FLO5* and 94% homologous with *FLO9* (SGD, 2021), it was essential to design primers that were unique to amplify the correct target region. Primers used to amplify the *N*-terminal domain of *FLO1, FLO5,* and *FLO9* can be found in Table 2.

To observe the genetic variability and conserved regions within the *N*-terminal region of *FLO1, FLO5,* and *FLO9,* at the DNA and protein level, a DNA and amino acid (amino acid) multiple sequence alignment were constructed. From the multiple sequence alignments (DNA and amino acid), a DNA and peptide consensus was determined for

each gene. The use of the consensus was an alternative to using any single strain as the baseline. Essentially, none of the yeast strains in this study could be expected to represent an ancestral sequence for all other strains. The consensus sequence provided the best representative of a shared ancestral sequence to provide a baseline to score nucleotide changes in each strain.

Some strains used in this study are not known to be haploid or diploid. For haploid strains, only a single DNA sequence would be expected from each gene. In contrast, for diploid organisms, each gene copy can produce a unique DNA sequence, and at least two PCR amplicons could be expected from a single reaction. The pooled amplicons will produce overlapping nucleotide sequences when sequenced, which will differ at sites harboring a polymorphism or in-del. This results in heterozygous species producing multiple traces at the same nucleotide base pair position (Fig 4, A). Identifying the true sequence for each of the alleles in these base pair locations would require cloning the amplicon into a plasmid and transformation in bacteria such that each bacterial colony would represent a single allelic amplicon. Due to time constraints, plasmid cloning was not used. Instead, IUPAC nucleotide codes were used to indicate sites where two possible nucleotides are indicated from DNA sequence trace files and DNA multiple sequence alignment (Fig 4, B). This allows each gene to be represented as a single DNA sequence for multiple sequence alignments. When translated, the amino acid “X” was inserted at polymorphic sites to maintain a single peptide sequence for multiple sequence alignments (Fig 4, C). Finally, to estimate true amino acid variation, all possible codons were used to predict the true amino acid at each polymeric site (Fig 4, D).

The physicochemical properties of amino acids play an important role in the overall function of the protein. To predict if amino acid changes altered protein function, we characterized the amino acid mutations as “different” or “similar.” If at least one of the physicochemical properties (polarity, charge, and aromaticity) were changed, then the amino acid change was characterized as “different.” Whereas if no physicochemical property was changed, the amino acid change was characterized as “similar”. Lastly, if there were multiple possible amino acid changes and one had different physicochemical properties and one had the same physicochemical properties amino acid mutation was characterized as “similar or different.”

FLO1

PCR products were produced for nine out of the ten strains employed in this study and were subjected to sequence analysis. Strain, Wine E, was subjected to three PCR attempts, all producing little to no DNA as determined by Agarose Gel Electrophoresis and Nanodrop quantification (data not shown); therefore, this strain was omitted from further analysis.

For the fermentation strains that produced PCR data in the amplified *N*-terminal region of *FLO1*, 58 nucleotide mutations were revealed in the DNA multiple sequence alignment (Table 3). The amino acid multiple sequence alignment revealed 9 amino acid changes within fermentation strains (Table 4). Of the 9 amino acid changes within the fermentation strains, 3 had different physicochemical properties, and 1 amino acid change was a nonsense mutation.

FLO5

PCR products were produced for six out of the ten strains employed in this study and were subjected to sequence analysis. Strain, Wine E, was subjected to three PCR attempts, all producing little to no DNA as determined by Agarose Gel Electrophoresis and Nanodrop quantification (data not shown). Two bands were produced as determined by Agarose Gel Electrophoresis for strains, Beer A, Beer B, and Beer C (data not shown). Multiple bands could indicate off-target PCR amplification, and due to lack of time, we did not optimize the PCR reactions these strains were omitted from further analysis.

For the fermentation strains that produced PCR data in the amplified *N*-terminal region of *FLO5*, 4 nucleotide mutations were revealed in the DNA multiple sequence alignment (Table 5). The amino acid multiple sequence alignment revealed a total of 3 amino acid changes within fermentation strains. Of the 3 amino acid changes within the fermentation strains, 2 had different physicochemical properties (Table 6).

FLO9

PCR products were produced for all strains employed in this study and were subjected to sequence analysis. For the fermentation strains that produced PCR data in the amplified *N*-terminal region of *FLO9*, 98 nucleotide mutations were revealed in the DNA multiple sequence alignment (Table 7). The amino acid multiple sequence alignment revealed a total of 21 amino acid changes within fermentation strains. Of the 21 amino acid changes within the fermentation strains, 16 had different physicochemical properties, and 3 had similar different physicochemical properties (Table 8). Of the 21 amino acid changes within the fermentation strains, 20 are found in the functional domain (Fig 3, C).

Peptide Mapping and Lab Strains

The location of the amino acid changes is critical in determining a putative effect on phenotype. For each gene, the amino acids previously described have been mapped to the *N*-terminal domains (Fig 5). In the *N*-terminal domain *FLO1* has a total of 8 amino acid changes (Fig 5, A), *FLO5* has a total of 2 amino acid changes (Fig 3, B), and *FLO9* has a total of 20 (Fig 3, C). The initial goal of this study was primarily focused on the genetic variation in the *N*-terminal domain of flocculin genes *FLO1*, *FLO5*, and *FLO9,* among the fermentation strains, however the lab strains showed a high degree of variation. The amino acid multiple sequence alignment revealed a total of 60 amino acid changes within lab strains for *FLO1*, 80 amino acid changes for *FLO5*, and 46 amino acid changes for *FLO9* (Table 9).

Helm's Assay

One of the goals of this study was to compare genotypic and phenotypic indicators of flocculation. While the flocculation levels of fermentation strains are characterized as high, medium, or low, a more refined and quantitative measurement of flocculation level is needed. To quantitatively determine flocculation capacity for each wine strain, we adapted the Helm's assay. To simulate a fermentation environment, cells were grown under conditions to trigger flocculation, including culturing in a beer analog media with the addition of ethanol to 5% and calcium. However, even with attempts to provide the optimal environment for flocculation, the yeast strains employed in this study did not flocculate in a predictive manner for the Helm's assay, so no data was attainable.

CONCLUSIONS

In this study, we survey the genetic variants of the *N*-terminal domain in the flocculin genes *FLO1, FLO5*, and *FLO9* among *S. cerevisae* and *S. bayanus* fermentation strains. Genetic variants discovered in these flocculin genes can provide novel sequence markers to identify fermentation strains. Whereas gold standard genes like rRNA and cytoskeletal proteins are expressed constitutively and known to be involved in the fundamental processes required for cell maintenance (Hunter et al., 1977). Due to the high sequence homology of *FLO1*, *FLO5*, and *FLO9*, these genes are too similar to identify strains within the same species and instead used to normalize gene expression studies as an internal control (Biederman et al., 2004).

These genetic variants detected can also affect the function of flocculin genes *FLO1, FLO5*, and *FLO9*. In this study we predicted whether an amino acid mutation would have an effective change on phenotype by observing if amino acid physicochemical properties (polarity, charge, and aromaticity) were changed. In addition to observing the amino acid physiochemical properties, the location of the amino acid changes were observed. All mutations were mapped out on the crystal structure created for the N-terminal domain of *FLO1* in complex with calcium and mannose, 4LHN (Gossens et al., 2015; Fig 6). Based on sequence similarity, 4LHN is likely to be similar if not identical to structure *FLO5* and *FLO9.* Even though the amino acid changes detected in this study may illustrate potential correlations with flocculation levels, it is difficult to directly state whether these changes alter flocculation. Some of the brewing strains used in this study are diploid organisms; therefore, only dominant mutations can

alter the phenotype. When two versions of the same allele are present, the allele carrying the recessive carrying mutations may be masked by the dominant allele.

The lab strains surveyed in this study show a large accumulation of mutations in the *N*-terminal regions of *FLO1, FLO5*, and *FLO9*. This is likely since lab strains are naturally non- flocculent due to a nonsense point mutation in the *FLO8* gene, any accumulated mutation will be silent (Liu et al., 1996). Without a transcriptional activator, *FLO8,* the affected genes cannot be expressed. Effectively, flocculin genes *FLO1, FLO5, FLO9* have become generic DNA and are no longer subject to selective pressures. Therefore, these regions have accumulated mutation through many generations. The incredible increase in the number of mutations is surprising, and in further studies, we would like to compare the mutation rate in these genetic elements with other genome regions.

The level of flocculation is often described vaguely as high, medium, or low. Quantitative methods to describe flocculation is often tedious, like total cell counting with a hemocytometer and results are variable. In the late ‘90s recommended methods of analysis by various brewing societies, including the American Society of Brewing Chemists, is the Helm’s Test (Bendiak et al., 1996). Many adaptations to this method have been suggested and attempted in fermentation research (Stratford et al., 1992; D’Hautcourt and Smart, 1999; Speers et al., 2006). To quantitively determine flocculation capacity for each wine strain, we decided to adapt the Helm’s assay. Using the optimal density measures, this method seemed to be the best quantitative measure compared to counting cells in each floc. Unfortunately, when attempting the adapted Helm’s assay in this study using the formula stated previously, all values were negative

inferring no flocculation. A reason no feasible data was accumulated may be because most of the strains in this study are generally low flocculators. In the future, it would be beneficial to perform this modified Helm's assay with strains known to flocculate to a higher degree.

A continuation of this study would benefit from implementing complete protein modeling techniques. Protein modeling can predict the altered three-dimensional structure and function of the protein from the identified amino acid variants. The structural predictions can help to identify functionally important residues. It is important to note that only the *N*-terminal region of *FLO1, FLO5*, and *FLO9* was sequenced and analyzed. For future studies it may be interesting to complete whole-genome sequencing of *FLO1, FLO5*, and *FLO9*, to observe the full genetic variation. While the *N*-terminal region plays a role in the execution of flocculation, the central domain of flocculin genes is also important. The central region varies among these genes containing a series of tandem repeats which result in variable protein lengths. Several studies have shown that the longer the *FLO* protein (carrying more repeats), the stronger is the flocculation ability it confers (Jibiki et al., 2001; Verstrepen et al., 2005; Liu et al., 2009). Additionally, spontaneous point mutations have been shown to more likely occur within tandem repeat and subtelomeric regions (Brown et al., 2010; Gemayel et al., 2010).

Flocculation is a complex and is dependent upon a large number of genes. Beginning to dissect flocculation at the genetic level allows scientists to start mapping out the causative amino acid changes that are a favorable, neutral, or detrimental outcome, these findings illustrate the potential correlations with phenotypic changes and provide evolutionary clues to the history of how these strains were developed over time.

Such knowledge can be used to modify currently used fermentation yeast strains. Modification can improve current phenotypic properties and give rise to novel properties in brewing and wine making but also bread and biofuel industries.

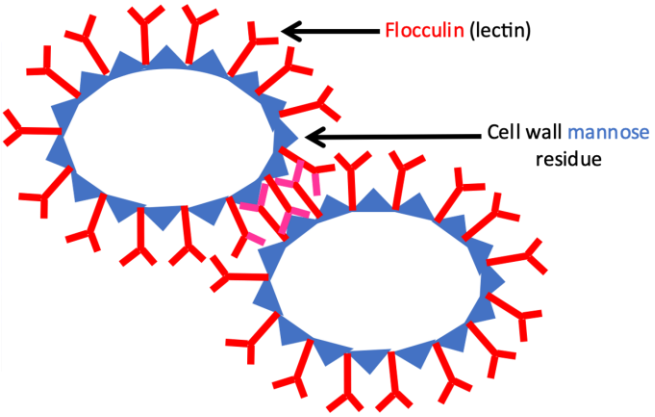


Figure 1. The lectin model for flocculation. Lectin-like proteins, flocculins, (red) stick out of the cell wall of flocculent cells and selectively bind to cell-wall mannose residues (blue triangles) of adjacent cells.

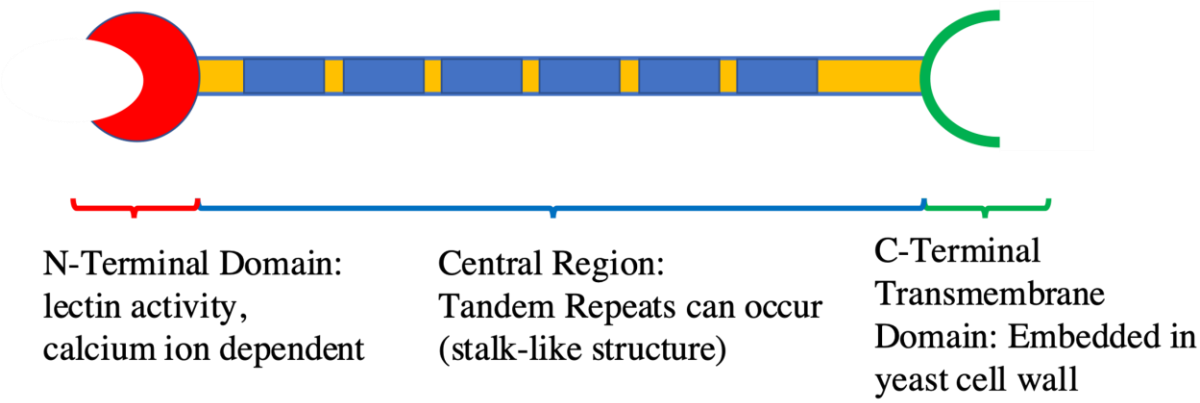


Figure 2. Structure of *FLO* Flocculin Family *FLO1, FLO5,* and *FLO9*. *N*-terminus is the sugar binding domain and needs Calcium ions to function. The *C*-terminal domain serves an anchor to bind to yeast cell walls. The central region contains tandem repeats that are variable in length among *FLO1, FLO5,* and *FLO9*.

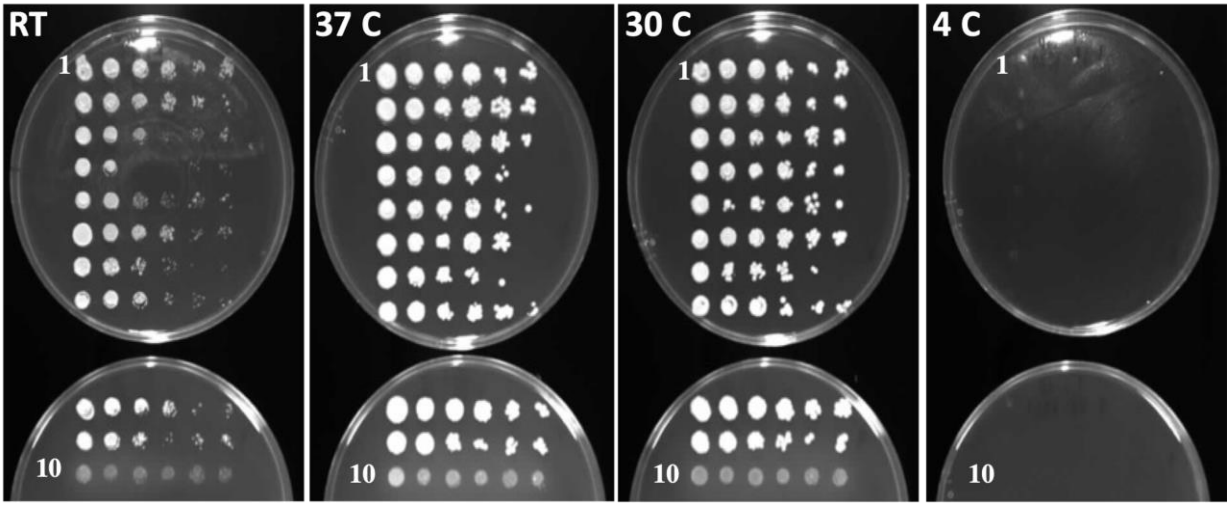


Figure 3. Temperature screening assay. Strain identification (1-10) is listed in Table 1. Photographs were taken after three days of growth at subjected temperatures: room temperature, 30°C, 37°C, and 4°C.

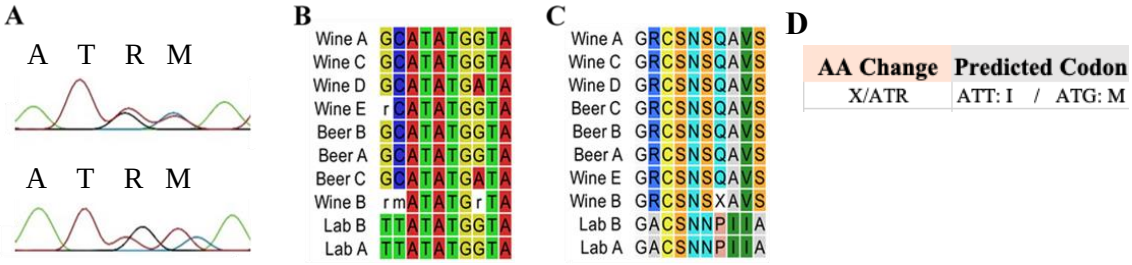


Figure 4. Data Interpretation. **A)** Heterozygous species producing multiple traces at the same nucleotide base pair position. R represents nucleotides thymine (T) and guanine (G) while M represents nucleotides thymine (T) and adenine (A). **B)** DNA multiple sequence alignment showing nucleotide changes. **C)** Amino acid multiple sequence alignment showing an inferred codon with amino acid "X". **D)** Predicted amino acid from inferred codon in panel C.

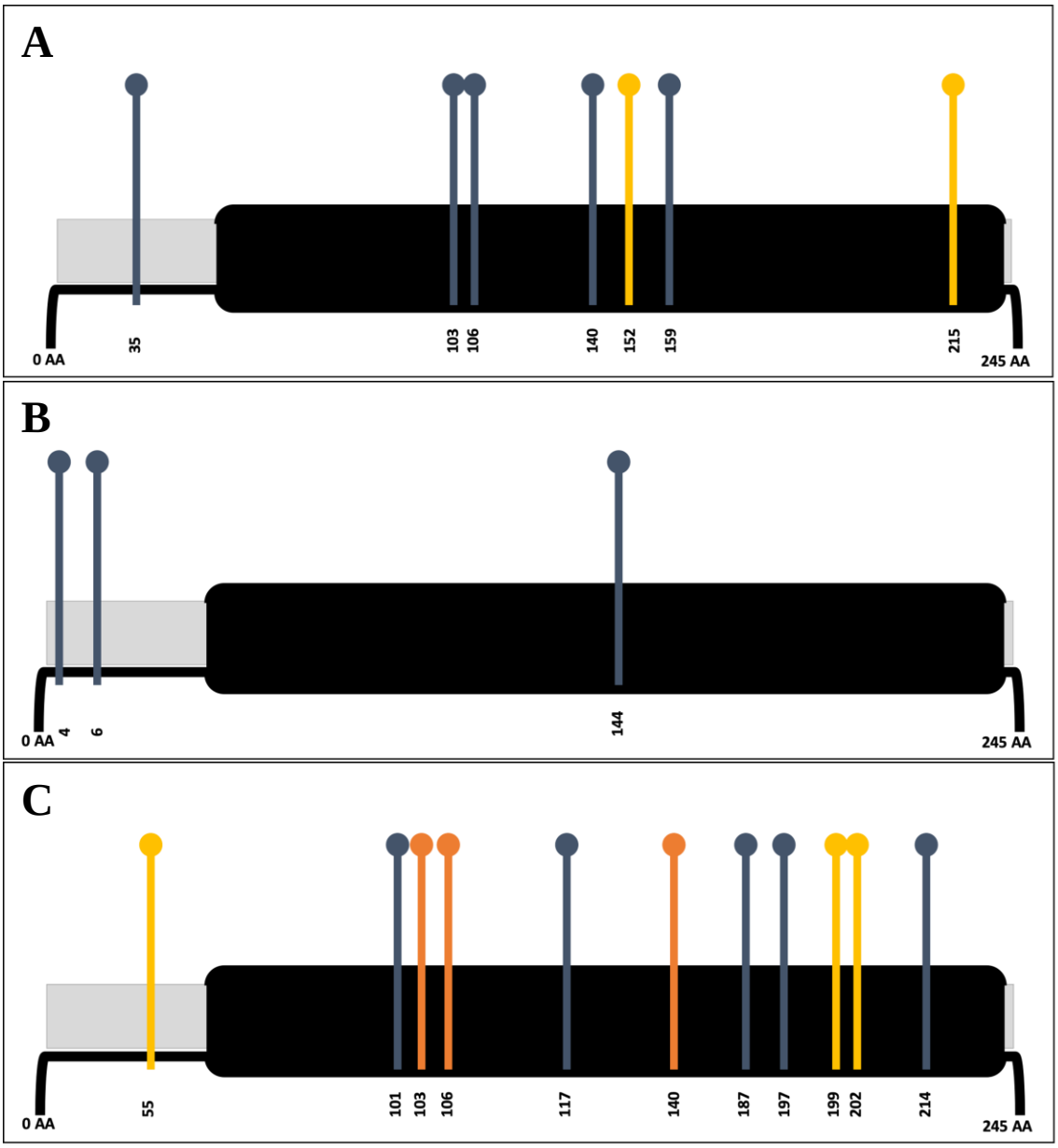


Figure 5. *N*-terminus peptide map. **A)** Amino acid changes in *FLO1*. **B)** Amino acid changes in *FLO5*. **C)** Amino acid changes found in *FLO9*. The hairpins indicate amino acid changes among only in beer strains (yellow hairpin), only in wine strains (blue hairpin), and both beer and wine strains (orange hairpin). The functional domain (black rectangle) is from 74 amino acid to 245 amino acid. Peptide map is not to scale.

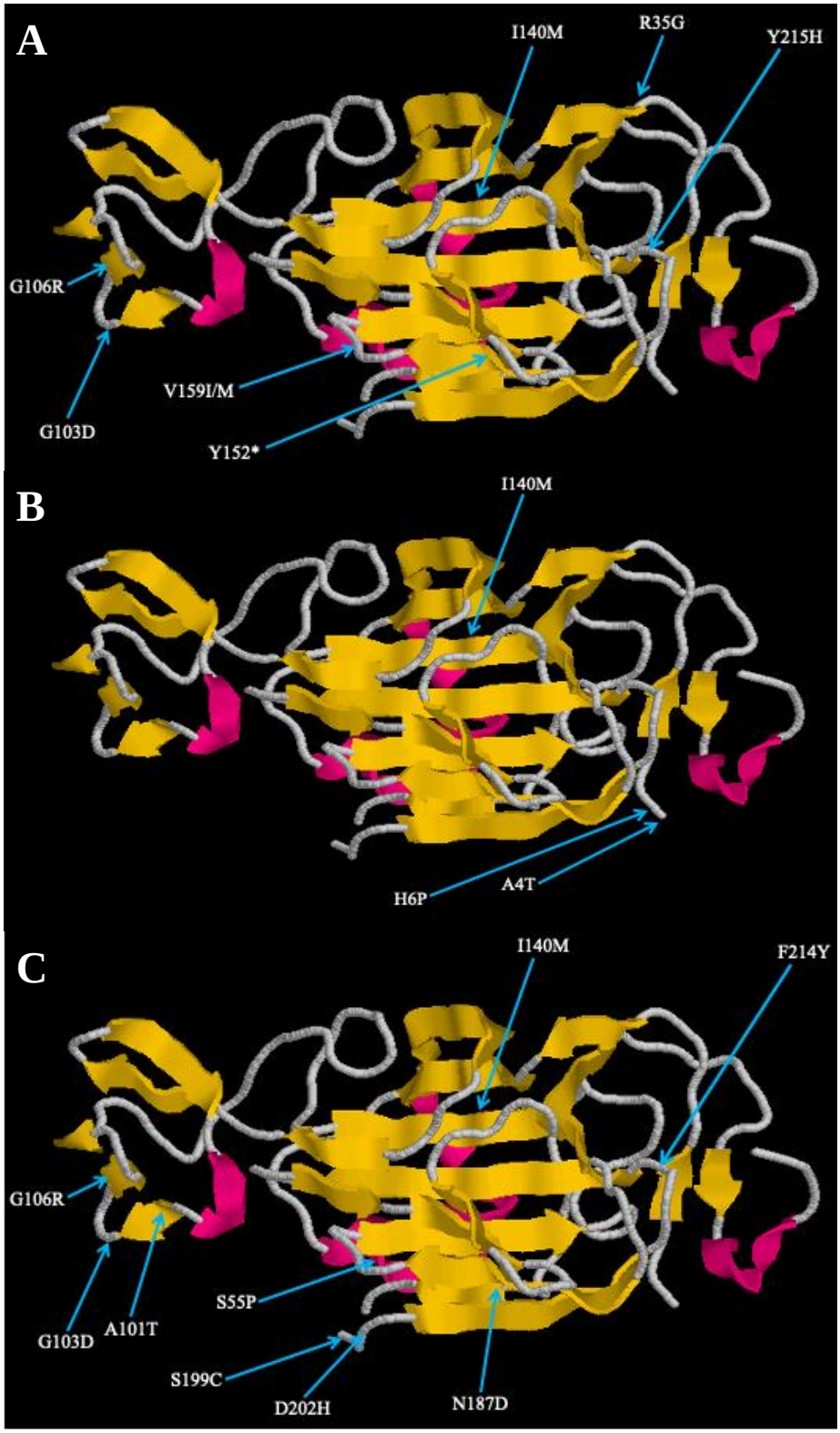


Figure 6. Location of amino acid changes in this study mapped on crystal structure, 4LHN. Amino acid changes occur in ß-sheets (yellow) and α-helices (pink). **A)** Amino acid changes in *FLO1*. **B)** Amino acid changes in *FLO5*. **C)** Amino acid changes found in *FLO9*.

Table 1. *Saccharomyces cerevisiae* and s*accharomyces bayanus* strains employed in this study.

Order	ID	Strain Name	Species	Flocculation Level
1	Wine A	Premier Rouge	*S. cerevisiae*	Low
2	Wine B	Premier Classique	*S. bayanus*	Low
3	Wine C	Premier Cuvée	*S. cerevisiae*	Low
4	Wine D	Premier Côte de Blans	*S. cerevisiae*	Low
5	Wine E	Premier Blanc	*S. bayanus*	Medium/Low
6	Beer A	Safale US-05	*S. cerevisiae*	Medium/Low
7	Beer B	Safale K-97	*S. cerevisiae*	Low
8	Beer C	Southern German Lager	*S. cerevisiae*	Medium/High
9	Lab A	BY4741	*S. cerevisiae*	Non-flocculent
10	Lab B	BY4743	*S. cerevisiae*	Non-flocculent

Table 2. Primers used for PCR to target *N*-terminal domains of *FLO1, FLO5,* and *FLO9*. All primers were designed in this study.

Sequence Name	Sequence 5' to 3'
*FLO1*_-28_F	CTTGTCACCAGTAAACAGAACATCC
FLO5 _-38_F	AAATGATTTTCTTTAAATTGATTAGCACCACTAAA
*FLO9*_-28_F	AGAACAATTGTACAATAAAAGCCCC
*FLO1*_Nterm_Seq_R	CGTACCCTTCGAAGTCATC
*FLO5*_Nterm_Seq_R	CGTACCCTTCAAAGTTATC
*FLO9*_Nterm_Seq_R	CGTACCCTTCAAAGTCATC

Table 3. Total nucleotide changes found in the *N*-terminal Domain *FLO1*. All nucleotide changes were determined when compared to the DNA consensus sequence of the fermentation strains.

FLO1	**Total Changes**
Wine A	NO PCR
Wine B	2
Wine C	4
Wine D	3
Wine E	7
Beer A	16
Beer B	14
Beer C	12
Lab A	117
Lab B	117

Table 4. Total amino acid changes found in the *N*-terminal Domain *FLO1*. All amino acid changes were determined when compared to the peptide consensus sequence of the fermentation strains. All possible codons were used to predict the true amino acid at each polymeric site and is reflected by the multiple amino acids in the "change" column. A star shows when possible codons did not code for any amino acid (stop codon). Chemical changes represent what kind of physicochemical property was changed. Physicochemical properties considered included: polarity, charge, and aromaticity.

FLO1	**Position**	**Consensus**	**Change**	**Chemical Change**
Wine E	35	R	R/G	No Change or Polarity
Wine E	103	G	D	Polarity + Charge
Wine E	106	G	R	Polarity + Charge
Wine C	140	M	I/M	No Change
Wine E	140	M	I/M	No Change
Beer A	152	Y	*	Nonsense
Wine C	159	V	I/M	No Change
Wine B	159	V	I	No Change
Beer B	215	Y	H	Charge

Table 5. Total nucleotide changes found in the *N*-terminal Domain *FLO5*. All nucleotide changes were determined when compared to the DNA consensus sequence of the fermentation strains.

FLO5	**Total Changes**
Wine A	1
Wine B	NO PCR
Wine C	0
Wine D	3
Wine E	No PCR
Beer A	No PCR
Beer B	No PCR
Beer C	No PCR
Lab A	141
Lab B	141

Table 6. Total amino acid changes found in the *N*-terminal Domain *FLO5*. All amino acid changes were determined when compared to the peptide consensus sequence of the fermentation strains. All possible codons were used to predict the true amino acid at each polymeric site and is reflected by the multiple amino acids in the "change" column. Chemical changes represent what kind of physicochemical property was changed. Physicochemical properties considered included: polarity, charge, and aromaticity.

FLO5	**Position**	**Consensus**	**Change**	**Chemical Change**
Wine D	4	A	T	Polarity
Wine D	6	H	P	Polarity + Charge + Aromaticity
Wine A	144	G	A	No Change

Table 7. Total nucleotide changes found in the *N*-terminal Domain *FLO9*. All nucleotide changes were determined when compared to the DNA consensus sequence of the fermentation strains.

FLO9	**Total Changes**
Wine A	2
Wine B	4
Wine C	21
Wine D	5
Wine E	10
Beer A	18
Beer B	19
Beer C	17
Lab A	85
Lab B	85

Table 8. Total amino acid changes found in the N-terminal Domain *FLO9*. All amino acid changes were determined when compared to the peptide consensus sequence of the fermentation strains. All possible codons were used to predict the true amino acid at each polymeric site and is reflected by the multiple amino acids in the "change" column. Chemical changes represent what kind of physicochemical property was changed. Physicochemical properties considered included: polarity, charge, and aromaticity.

FLO9	**Position**	**Consensus**	**Change**	**Chemical Change**
Beer A	199	S	C	Polarity
Beer A	202	D	H	Charge + Aromaticity
Beer B	55	S	P	Polarity
Beer B	199	S	C	Polarity
Beer B	202	D	H	Charge + Aromaticity
Beer C	103	G	D	Polarity + Charge
Beer C	106	G	R	Polarity + Charge
Beer C	140	M	I/M	No Change
Wine B	101	A	K/A/T/E	Polarity + Charge or No Change
Wine B	103	G	D/G	Polarity + Charge
Wine B	106	G	R	Polarity + Charge
Wine B	117	Q	K/Q	Charge or No Change
Wine B	187	N	N/D	No Change or Charge
Wine B	197	H	Q	Charge + Aromaticity
Wine B	214	F	Y	Polarity
Wine D	103	G	D	Polarity + Charge
Wine D	106	G	R	Polarity + Charge
Wine D	140	M	I	No Change
Wine E	101	A	T/A	Charge
Wine E	187	N	D	Charge
Wine E	214	F	Y	Polarity

Table 9. Lab Strains total amino acid changes found in the N-terminal Domain *FLO1, FLO5,* and *FLO9*. All amino acid changes were determined when compared to the peptide consensus sequence of the fermentation strains.

ID	***FLO1* Total Changes**	***FLO5* Total Changes**	***FLO9* Total Changes**
Lab A	30	40	23
Lab B	30	40	23

REFERENCES

Bauer, Florian F., Patrick Govender, and Michael C. Bester. 2010. "Yeast Flocculation and Its Biotechnological Relevance." *Applied Microbiology and Biotechnology* 88 (1): 31–39. https://doi.org/10.1007/s00253-010-2783-0.

Bendiak D, Van der Aar P, Barbero F, Benzing P, Berndt R, Carrick K, Dull C, Dunn S, Eto M, Gonzalez M, Hayashi N, Lawrence D, Miller J, Phare K, Pugh T, Rashel L, Rossmore K, Smart KA, Sobczak J, Speers A, Casey G. 1996. "Yeast flocculation by absorbance." *J Am Soc Brew Chem*

Biederman, Jason, Jerry Yee, and Pedro Cortes. 2004. "Validation of Internal Control Genes for Gene Expression Analysis in Diabetic Glomerulosclerosis." *Kidney International* 66 (6): 2308–14. https://doi.org/10.1111/j.1523-1755.2004.66016.x.

Braus, Gerhard H., Olav Grundmann, Stefan Brückner, and Hans-Ulrich Mösch. 2003. "Amino Acid Starvation and Gcn4p Regulate Adhesive Growth and FLO11 Gene Expression in Saccharomyces Cerevisiae." *Molecular Biology of the Cell* 14 (10): 4272–84. https://doi.org/10.1091/mbc.e03-01-0042.

Brown, Chris A., Andrew W. Murray, and Kevin J. Verstrepen. 2010. "Rapid Expansion and Functional Divergence of Subtelomeric Gene Families in Yeasts." *Current Biology* 20 (10): 895–903. https://doi.org/10.1016/j.cub.2010.04.027.

D'Hautcourt, Olivier, and Katherine A. Smart. 1999. "Measurement of Brewing Yeast Flocculation." *Journal of the American Society of Brewing Chemists* 57 (4): 123–28. https://doi.org/10.1094/ASBCJ-57-0123.

Engel, Stacia R, Fred S Dietrich, Dianna G Fisk, Gail Binkley, Rama Balakrishnan, Maria C Costanzo, Selina S Dwight, et al. 2014. "The Reference Genome Sequence of *Saccharomyces Cerevisiae* : Then and Now." *G3 Genes Genomes Genetics* 4 (3): 389–98. https://doi.org/10.1534/g3.113.008995.

Gallone, Brigida, Stijn Mertens, Jonathan L Gordon, Steven Maere, Kevin J Verstrepen, and Jan Steensels. 2018. "Origins, Evolution, Domestication and Diversity of Saccharomyces Beer Yeasts." *Current Opinion in Biotechnology* 49 (February): 148–55. https://doi.org/10.1016/j.copbio.2017.08.005.

Gemayel, Rita, Marcelo D. Vinces, Matthieu Legendre, and Kevin J. Verstrepen. 2010. "Variable Tandem Repeats Accelerate Evolution of Coding and Regulatory Sequences." *Annual Review of Genetics* 44 (1): 445–77. https://doi.org/10.1146/annurev-genet-072610-155046.

Goossens, Katty V. Y., Catherine Stassen, Ingeborg Stals, Dagmara S. Donohue, Bart Devreese, Henri De Greve, and Ronnie G. Willaert. 2011. "The N-Terminal Domain of the Flo1 Flocculation Protein from Saccharomyces Cerevisiae Binds Specifically to

Mannose Carbohydrates." *Eukaryotic Cell* 10 (1): 110–17. https://doi.org/10.1128/EC.00185-10.

Goossens, Katty V. Y., Francesco S. Ielasi, Intawat Nookaew, Ingeborg Stals, Livan Alonso-Sarduy, Luk Daenen, Sebastiaan E. Van Mulders, et al. 2015. "Molecular Mechanism of Flocculation Self-Recognition in Yeast and Its Role in Mating and Survival." Edited by Sang Yup Lee. *MBio* 6 (2): e00427-15. https://doi.org/10.1128/mBio.00427-15.

Helm E, Nohr B, Thorne RSW. 1953."The measurement of yeast flocculence and its significance in brewing." *Wallerstein Lab Commun* (16):315–325.

Hunter, Tony, and James I. Garrels. 1977. "Characterization of the MRNAs for α-, β- and γ-Actin." *Cell* 12 (3): 767–81. https://doi.org/10.1016/0092-8674(77)90276-8.

Jibiki, Makiko, Tomoko Ishibiki, Toshifumi Yuuki, and Noboru Kagami. 2001. "Application of Polymerase Chain Reaction to Determine the Flocculation Properties of Brewer's Lager Yeast." *Journal of the American Society of Brewing Chemists* 59 (3): 107–10. https://doi.org/10.1094/ASBCJ-59-0107.

Liti, Gianni, David M. Carter, Alan M. Moses, Jonas Warringer, Leopold Parts, Stephen A. James, Robert P. Davey, et al. 2009. "Population Genomics of Domestic and Wild Yeasts." *Nature* 458 (7236): 337–41. https://doi.org/10.1038/nature07743.

Liu, H., C. A. Styles, and G. R. Fink. 1996. "Saccharomyces Cerevisiae S288C Has a Mutation in FLO8, a Gene Required for Filamentous Growth." *Genetics* 144 (3): 967–78.

Liu, Nan, Dong-li Wang, Zhao-yue Wang, Xiu-ping He, and Bo-run Zhang. 2009. "Deletion of Tandem Repeats Causes Flocculation Phenotype Conversion from Flo1 to NewFlo in *Saccharomyces Cerevisiae*." *Journal of Molecular Microbiology and Biotechnology* 16 (3–4): 137–45. https://doi.org/10.1159/000112318.

Miki, B. L., N. H. Poon, A. P. James, and V. L. Seligy. 1982. "Possible Mechanism for Flocculation Interactions Governed by Gene FLO1 in Saccharomyces Cerevisiae." *Journal of Bacteriology* 150 (2): 878–89. https://doi.org/10.1128/JB.150.2.878-889.1982.

Sampermans, S., J. Mortier, and E. V. Soares. 2005. "Flocculation Onset in Saccharomyces Cerevisiae: The Role of Nutrients." *Journal of Applied Microbiology* 98 (2): 525–31. https://doi.org/10.1111/j.1365-2672.2004.02486.x.

Speers, R. Alex, Yong-Quan Wan, Yu-Lai Jin, and Robert J. Stewart. 2006. "Effects of Fermentation Parameters and Cell Wall Properties on Yeast Flocculation [1]." *Journal of the Institute of Brewing* 112 (3): 246–54. https://doi.org/10.1002/j.2050-0416.2006.tb00720.x.

Stewart, G. G., I. Russell, and I. F. Garrison. 1975. “SOME CONSIDERATIONS OF THE FLOCCULATION CHARACTERISTICS OF ALE AND LAGER YEAST STRAINS*.” *Journal of the Institute of Brewing* 81 (3): 248–57. https://doi.org/10.1002/j.2050-0416.1975.tb03686.x.

Stratford, Malcolm, and Stephen Assinder. 1991. “Yeast Flocculation: Flo1 and NewFlo Phenotypes and Receptor Structure.” *Yeast* 7 (6): 559–74. https://doi.org/10.1002/yea.320070604.

Strope, Pooja K., Daniel A. Skelly, Stanislav G. Kozmin, Gayathri Mahadevan, Eric A. Stone, Paul M. Magwene, Fred S. Dietrich, and John H. McCusker. 2015. “The 100-Genomes Strains, an *S. Cerevisiae* Resource That Illuminates Its Natural Phenotypic and Genotypic Variation and Emergence as an Opportunistic Pathogen.” *Genome Research* 25 (5): 762–74. https://doi.org/10.1101/gr.185538.114.

Teunissen, A. W. R. H., and H. Y. Steensma. 1995. “The Dominant Flocculation Genes OfSaccharomyces Cerevisiae Constitute a New Subtelomeric Gene Family.” *Yeast* 11 (11): 1001–13. https://doi.org/10.1002/yea.320111102.

Van Mulders, Sebastiaan E., Els Christianen, Sofie M.G. Saerens, Luk Daenen, Pieter J. Verbelen, Ronnie Willaert, Kevin J. Verstrepen, and Freddy R. Delvaux. 2009. “Phenotypic Diversity of Flo Protein Family-Mediated Adhesion in *Saccharomyces Cerevisiae*.” *FEMS Yeast Research* 9 (2): 178–90. https://doi.org/10.1111/j.1567-1364.2008.00462.x.

Van Mulders, Sebastiaan E., Maarten Ghequire, Luk Daenen, Pieter J. Verbelen, Kevin J. Verstrepen, and Freddy R. Delvaux. 2010. “Flocculation Gene Variability in Industrial Brewer’s Yeast Strains.” *Applied Microbiology and Biotechnology* 88 (6): 1321–31. https://doi.org/10.1007/s00253-010-2843-5.

Verstrepen, Kevin J, An Jansen, Fran Lewitter, and Gerald R Fink. 2005. “Intragenic Tandem Repeats Generate Functional Variability.” *Nature Genetics* 37 (9): 986–90. https://doi.org/10.1038/ng1618.

Verstrepen, Kevin J, and Frans M Kils. 2006. “Flocculation, Adhesion and Biofilm Formation in Yeasts.” *Molecular Microbiology* 60 (1): 15.

Vidgren, Virve, and John Londesborough. 2011. “125th Anniversary Review: Yeast Flocculation and Sedimentation in Brewing.” *Journal of the Institute of Brewing* 117 (4): 475–87. https://doi.org/10.1002/j.2050-0416.2011.tb00495.x.