Designing and Demonstrating the Use of a Bioluminescent Genetic Tool Set in *Scheffersomyces stipitis*

By

Walter Reichard

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Master of Science in Biology

Middle Tennessee State University
August 2017

Thesis Committee:
Dr. Brian Robertson, Chair
Dr. Matthew Elrod-Erickson
Dr. Elliot Altman

ACKNOWLEDGEMENTS

The amount of support and motivation I have received from others has been incredible, and there are far too many people who have been involved in my studies as a graduate student to name everyone. First I would like to thank my thesis advisor, Dr. Brian Robertson, who has been extremely helpful in guiding me through the program. With his advice and encouragement throughout my graduate studies, I have been able to develop skills necessary to transition from being a student to becoming a scientist. I would also like to thank Dr. Elliot Altman and Dr. Matt Elrod-Erickson for being on my thesis committee and for the guidance and resources for my thesis project. Many of the experiments in this study would not have been possible without them.

The MTSU biology department has been invaluable to my progress through the program by facilitating a helpful and inspiring environment. Thanks to the members of the Robertson lab, Chris, Paul and Virginia for the interesting conversations and keeping the lab a fun place to be. Thank you Dr. Jeff Leblond, Dr. Mary Farone, and Dr. Tony Farone for inspiring my interests in the applications of Microbiology throughout my course work.

I also want to extend my thanks to all of the other faculty and staff in the biology department for being so helpful and friendly, and for truly making the university feel like home.

Lastly I would like to thank my family and friends outside of the academic environment for listening to all my problems and being there with love and understanding when I was getting stressed out. Without the support of my family, none of this would have been possible, thank you.

ABSTRACT

Scheffersomyces stipitis is the most efficient xylose fermenting yeast currently characterized, which could potentially be utilized in a lignocellulosic fermentation consortium. The genetics of *S. stipitis* are not as well studied as other species of yeast because *S. stipitis* is a member of a group of yeast known as the "CUG clade". Yeast species in the "CUG clade" translate the codon CUG differently from most eukaryotic organisms, instead adding a serine to the protein product instead of a leucine. This alternative translation inhibits the modularity of established selectable markers and reporter systems such as *CBG99* in *Saccharomyces cerevisiae*. Foreign genes need to be codon optimized, changing CTG codons into different leucine codons in order to allow the proteins coded from these genes to retain their native amino acid sequence, structure and function. The work presented in this thesis demonstrate a codon optimized bioluminescent tool set used in *S. stipitis*, as well as what it can be used for.

TABLE OF CONTENTS

LIST OF FIGURES vi

LIST OF TABLES vii

CHAPTER I: INTRODUCTION 1

1.1 Cellulosic ethanol 1

1.2 *Scheffersomyces stipitis* 4

1.3 *CBG99* as a reporter gene 5

1.4 Codon optimization 10

1.5 Recoverable selectable marker in gene knockout strategies 12

1.6 Significance of the study 15

1.7 Thesis goals 16

1.8 Hypothesis 16

CHAPTER II: NAVIGATING THE CUG CLADE OBSTACLE TO DESIGN A BIOLUMINESCENT REPORT FOR *S. stipitis* 17

2.1 Introduction 17

2.2 Results and discussion 18

2.2.1 Testing a codon optimized reporter gene in *S. stipitis* 18

2.2.2 Testing suspected inducible promoter regions 20

2.2.3 Creating a rapidly degrading bioluminescent reporter system 22

2.2.4 Measuring yeast competition using two bioluminescent reporter genes 26

2.2.5 Using a codon optimized reporter gene in *Candida albicans* 28

2.3 Materials and methods 30

2.3.1 Naming conventions 30

2.3.2 Plasmid construction 30

2.3.3 Transformation of *S. stipitis* 36

2.3.4 Degron assay experimental design 37

2.3.5 Monitoring yeast competition using two reporter systems experimental design 38

2.3.6 Expressing a codon optimized bioluminescent reporter in *Candida albicans* 39

CHAPTER III: CONCLUSION 40

3.1 Summary 40

3.2 Future directions 40

REFERENCES 42
APPENDICES 47
APPENDIX A: ADDITIONAL FIGURES 48

LIST OF FIGURES

CHAPTER I

Figure 1.1 The net reaction of D-luciferin in the presence of luciferase…….……7

Figure 1.2 Homologous recombination inside the cell…….13

CHAPTER II

Figure 2.1 Testing a codon optimized reporter gene in *S. stipitis*……..…….……19

Figure 2.2 Testing the activity of sugar promoters ……..21

Figure 2.3 *CBG99-CLN2* fusion degron assay……..23

Figure 2.4 *CBG99-CLN2* fusion degron assay normalized…….…….25

Figure 2.5 Monitoring yeast competition in different sugar environments……...27

Figure 2.6 *Candida albicans* transformed with pWR32…….. 29

Figure 2.7 Plasmid map of pAllet compared to other cloning plasmids…….... 31

Figure 2.8 Plasmid map of pAllet (P_{TEF1}-co*HPH*-T_{ACT1})(P_{TDH3}-co*CBG99*-T_{ADH1}) (*ARS2*)……35

LIST OF TABLES

Table 1 Primers used for this study ……..…..…….……..……32

Table 2 Plasmids constructed for this study…….…… 33

CHAPTER I
INTRODUCTION

1.1 Cellulosic ethanol

Plant biomass is one of the most abundant materials on earth. Plants are largely made up of cellulose, and can subsequently be broken down into lignocellulosic biomass which contains an array of sugar monomers including glucose, xylose, and galactose (Kumar *et al*., 2009) (Mosier *et al*., 2005). The carbohydrates found in lignocellulosic biomass are initially found in the form of cellulose, hemicelluloses and lignin (Zaldivar *et al*., 2001). These sugars can be converted into ethanol through fermentation processes that yeast can utilize for their own gain, with one of the byproducts of this reaction being ethanol (Kumar *et al*., 2009). Ethanol is an industrially desired molecule since it is the most easily replenished fuel source, coming from either starch or sugars (Hahn-Hägerdal *et al*., 2006). Lignocellulosic fermentation however has not been industrialized yet in a profitable manner largely due to the low yield and high cost of the hydrolysis process (McAloon *et al*., 2000). This creates a need to develop new technologies that will decrease the cost and yield margins, making lignocellulosic ethanol feasible for commercial use.

Yeast cannot be added directly to ferment the lignocellulosic biomass however, as the biomass requires pretreatment and hydrolysis first before it can become a mixture of fermentable sugars. Cellulose is initially naturally found in the form of a crystalline structure. This structure protects the terminal ends of the cellulose from hydrolysis (Hsu *et*

al. 1980). Aside from cellulose, there are other compounds present called lignin and hemicellulose that provide amorphous character to the crystalline cellulose, strengthening its structure. There are several options for which pretreatment steps are used to dissolve and separate the cellulose and hemicelluloses from the lignin for hydrolysis. Pretreatment of the cellulosic biomass include chemical, physical, and biological pretreatment methods (da Costa Sousa *et al*., 2009.). The diverse options for pretreatment of lignocellulosic biomass are due to the cost burden to the conversion of bioethanol. Pretreatment is one of the most expensive steps in the synthesis of cellulosic ethanol (McAloon *et al*., 2000). These methods of pretreatment all achieve the end goal of breaking up the lignocellulosic biomass at the amorphous regions of the lignin, and freeing the terminal ends of the cellulose molecule for hydrolysis (Mosier *et al*., 2005).

There are three different types of physical pretreatment currently in use. One mechanism for pretreatment under physical conditions includes mashing and milling to reduce the amount of crystalline structures protecting the cellulose. Another method of mechanical pretreatment involves heating the cellulosic material to temperatures of 300˚C. The biomass is then cooled slowly and treated with weak acids to start the hydrolysis of cellulose (Fan *et al*., 1987). There are also chemical based methods for pretreating lignocellulosic biomass; however these have also been proven to be cost inefficient for the production of ethanol (Sun and Cheng, 2002). One of these methods is called ozonolysis, using ozone to degrade the lignin and hemicelluloses present in the biomass (Binder *et al.*, 1980). The last category of pretreatment is biological, the use of enzymes or whole microorganisms to break down the lignocellulosic biomass. These strategies have proved

to be promising over recent years as culturing microbes is a cheap process. Microorganisms used to pretreat lignocelluloses so far include *Ceriporiopsis subvermispor, Irpex lacteus* and *Phanerochaete chrysosporium* (Cianchetta *et al*., 2014) (Du *et al*., 2011) (Potumarthi *et al*., 2013).

Traditionally fermentation would either utilize one robust isolated strain of yeast which could ferment all of the sugars present in this multi sugar environment or use one yeast strain that would solely ferment glucose and ignore the other sugars present in the hydrolysate (Pereira *et al*., 2014) (Zheng *et al*., 2011). Yeast have an adaptive strategy for sugar consumption known as diauxia, where yeast have a preferred sugar carbon source and ignore other sugars present as long as their preferred sugar is present (Young *et* al. 2014). This ultimately slows the process of converting a multi sugar medium into ethanol, and is a contributing factor to the cost effectiveness of the process as a whole. The lack of total sugar fermentation from one single strain of yeast has led to the idea that having a blend of multiple strains of yeast each of which is only capable of recognizing and metabolizing one type of sugar could potentially be a more efficient method for fermenting the sugars of a lignocellulosic biomass. If several strains were present and each could only utilize one monosaccharide in this artificial environment, then we should observe a simultaneous fermentation of all the sugars present, and an overall increase the efficiency of fermentation. Since *Scheffersomyces stipitis* ferments xylose efficiently, there could be a role in this consortium fermentation for *S. stipitis.*

1.2 *Scheffersomyces stipitis*

Scheffersomyces stipitis, formerly known as *Pichia stipitis*, is an ascomycotous yeast species that was discovered and characterized in 1967 (Kurtzman & Suzuki, 2010) (Pignal, 1967). *S. stipitis* is found to be capable of fermenting all of the sugars present in wood (Suh *et al*., 2003) (Melake *et al*., 1996). *S. stipitis* is commonly found in the gut flora of passalid beetles, and is suspected to aid in the digestion of wood materials ingested by these beetles (Lee *et al*., 1986) (Suh *et al*., 2003). *S. stipitis* is one of the strongest fermenters of xylose, making it a valuable yeast species commercially in the production of ethanol (Bruinenberg *et al*., 1984) (Delgenes *et al*., 1996) (Nigam, 2001). Another reason for the interest in *S. stipitis* as a fermenting yeast for cellulosic ethanol is due to *S. stipitis*'s capability to break down cellobiose, which further aids in breaking down cellulose into monosaccharides (Parkeh *et al*., 1998). This could potentially reduce the costs of pretreatment of lignocellulose.

In a pretreated lignocellulosic environment, there are several sugars present such as glucose, xylose, and galactose. This would serve as a challenge for *S. stipitis* to ferment xylose efficiently, since recent studies on *S. stipitis* have shown that this yeast species exhibits a sugar preference with xylose not being the preferred sugar (Kim *et al*., 2012). One potential way around this sugar preference is to remove the genes that code for other preferred sugar transporters, preventing any inhibitory effects from impacting the xylose pathway (Slininger *et al*., 2011). In order to pursue ideas such as removing genes from the chromosome of *S. stipitis*, there is a need for genetic tools to first be developed.

Many xylose fermenting yeast, such as *S. stipitis* are members of a group of yeast known as the "CUG clade" (Ohama *et al*., 1993) (Priest & Lorenz, 2015). Yeast species in this group read the codon CUG during translation differently from most eukaryotic organisms (Jeffries *et al*., 2007) (Santos & Tuite, 1995) (Wohlbach *et al*., 2011). In translation, the CUG codon codes for a leucine in most yeast, however yeast species in the "CUG clade" add a serine instead (Cutfield *et al*., 2000) (Du Preez *et al*., 1985) (Wohlbach *et al*., 2011). This unique translation has been attributed to these yeast species containing a Ser-tRNA which have a Leucine anticodon (Kawaguchi *et al*., 1989) (Massey *et al*., 2003). Due to this altered codon translation, established genetic toolsets such as selectable markers and reporter genes used in model yeast species such as *Saccharomyces cerevisiae* do not retain their function when expressed in yeast belonging to the CUG clade.

1.3 *CBG99* as a reporter gene

The use of a reliable and non-invasive reporter system is a convenient tool to measure and observe gene expression in yeast (Robertson *et al*., 2008). The reporter gene *CBG99*, is found originally in *Pyrophorus plagiophthalamus*, and codes for click beetle luciferase. When the luciferase is introduced to luciferin, a reaction which emits light at 537nm occurs (Mezzanotte *et al*., 2011). The reaction starts off when the gene product of *CBG99*, luciferase is produced. When luciferase comes into contact with D-luciferin, Mg^{+2}, and ATP, the product is D-luciferyl adenylate which becomes oxidized by environmental oxygen and releases light as part of the reaction shown in figure 2 (Baldwin, 1996). The emission of light can be observed and quantified, creating a useful tool in monitoring yeast gene expression and biological rhythms (Brandes *et al*., 1996). A bioluminescent reporter

gene can provide information on transformation success, promoter activity, and protein life spans by quantifying the amount of bioluminescence being emitted from the reaction with the luciferin (Robertson *et al*., 2011). This bioluminescence tool has been used in *Saccharomyces cerevisiae* to better understand yeast gene expression and could serve as a key component in a genetic toolset for better understanding *S. stipitis* (Robertson *et al*., 2008).

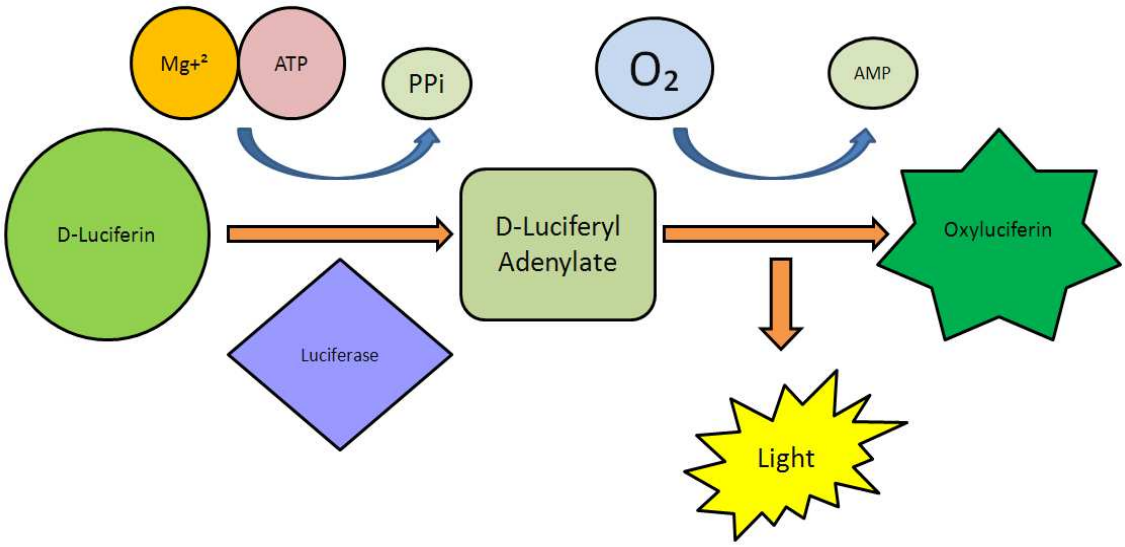


Figure 1.1. The net reaction of D-luciferin in the presence of luciferase. D-Luciferin is converted to D-luciferyl adenylate in the presence of beetle luciferin. In the presence of oxygen, D-luciferyl adenylate is converted into oxyluciferin through a series of reactions and releases light as a byproduct. Using CBG99 as a reporter gene to express D-luciferase in the presence of LH_2, Mg^{+2}, and oxygen, will give off light allowing for detection of yeast cells that are expressing the gene.

CBG99 can be used as an informative tool in regards to plasmid transformation success. This can be done by including *CBG99* in a plasmid that has another gene of interest included. After transformation, cells which emit light when presented with D-luciferin likely contain the entire plasmid. The *CBG99* reporter gene will need to have a constitutive promoter to drive the gene expression in a wide variety of environments. If other types of promoters are used to drive *CBG99* in this situation, false negatives could be observed. By using a constitutive promoter to drive *CBG99*, any yeast cells that take the plasmid and express the genes within the plasmid should be observed to be bioluminescent in the presence of beetle luciferin.

A bioluminescent reporter can be used to monitor gene expression. This can be performed by cloning the upstream promoter region of a gene of interest into an expression vector. Downstream of this promoter region the *CBG99* gene can be cloned into the plasmid. Once transformed into the organism of interest, *CBG99* will then be expressed anytime the targeted gene of interest is expressed due to having the same regulatory elements as the gene of interest. Promoter activity can then be monitored by quantifying the amount of bioluminescence being emitted from the yeast cells. The more luminescence present, the more active the promoter of interest is being expressed. To perform these experiments, there is a need to control for cell density and culture age. This strategy of promoter activity observation is also limited to only one experimental condition, and does not allow for observations on how changes in the environment might affect promoter activity. By observing promoter activity through luminescent expression, comparisons and

inferences can be made across different target genes of interest. This tool can show which genes are being expressed more than others under different conditions.

When a gene is expressed, its mRNA products leave the nucleus to be transcribed into a protein at the ribosome. Once transcribed as a protein, the molecule is no longer dependant on the central dogma of the cell being active. This means that any reactions that occur because of the protein will continue until all of the protein is degraded. Since one of the aims of this study is to use *CBG99* as a tool to measure gene expression, it can be useful to see if the time *CBG99* persists in the cell can be decreased. If the rate of *CBG99* protein regulation is increased, bioluminescence as a tool for measuring gene expression can be improved to show accurate changes in real time.

Yeast G_1 cyclin genes such as *CLN2* have been determined to have a short half-life in yeast (Salama *et al*., 1994). This has been attributed to a C-terminal amino acid region in these genes called a PEST domain, which is composed of proline, glutamic acid, serine and threonine in high amounts (Rogers *et al*., 1986). By creating a protein fusion of *CBG99* to suspected PEST domains, the degradation rate of luciferase proteins expressed should increase. This could potentially create a faster response to promoters being down regulated, allowing for a real time monitoring system for promoter regulation in dynamic environments.

Another application of *CBG99* as a reporter system is when monitoring competition among species of yeast. *Scheffersomyces stipitis* is a potential candidate for being one yeast species in a lignocellulosic fermentation consortium, which will require *S. stipitis* to grow in a co-culture with other species of yeast. *CBG99* glows green in the presence of

beetle luciferin at a wavelength of 537nm (Mezzanotte *et al*., 2011). Another reporter gene known as *CBR* glows red in the presence of beetle luciferin at a wave length of 615nm (Villalobos *et al*., 2010). While both reporter genes use the same substrate, they emit light at different wavelengths. If there are two different yeast species, containing a different bioluminescent reporter such as *CBG99* and *CBR*, the two species could be cultured together and would be distinguishable based on which color the cells luminesce when the substrate is added. This allows studies of competition between yeast species grown in the same environment, as well as allowing experimentation with the variables of the environment in which co-cultures are competing.

1.4 Codon optimization

Amino acids are the subunits that make up proteins, and ultimately give the protein its function. The amino acids used in proteins are determined based on the nucleotides found in the coding sequence of the gene. Every three nucleotides of the coding sequence is transcribed and translated as a codon, which codes for an amino acid. Most amino acids have multiple codons that can code for each amino acid, the exceptions being tryptophan and methionine which only have one codon each (Sharp, *et al*,. 1988). These codons that share a common amino acid often differ by one nucleotide found in the third position. This allows for some tolerance to point mutations, since many of the possible point mutations at the third nucleotide do not change the amino acid coded, which results in no impact on the function of the coded protein. This mutation tolerance could be a way around the CUG alternative translation by editing the third base pair of CUG codons, changing them to different leucine codons.

Despite the equivalent amino acids derived from synonymous codons, the codons do not appear equivalently in the genome (Grantham *et al*., 1980). The genetic code has a tendency to use some codons more than others. This genetic bias is different between species of organisms, meaning that codons found to appear in higher frequencies in *Escherichia coli* are not the same codons with high prevalence in *Scheffersomyces stipitis*. The genetic bias found in these organisms is hypothesized to be attributed to the amount of available isoaccepting tRNAs found within the organism, and has been called the genome hypothesis (Grantham *et al*., 1980). One way to improve gene expression is to change rare occurring codons, and edit them to become more commonly found codons during gene synthesis (Burgess-Brown *et al*., 2008).

Due to the alternative translation phenomenon observed in *Scheffersomyces stipitis*, tools such as reporter genes, selectable markers and negative selectable that have been used in other organisms do not function correctly when directly transformed and expressed in *S. stipitis*. One potential solution to this problem could be through having these genes synthesized based on the sequence of the gene of interest in other organisms, and using codon optimization to change any CUG codons found within the gene to a synonymous codon. The CUG codon codes for a leucine amino acid in all organisms other than those found in the CUG clade which instead translate as serine. The codon CUU similarly codes for a leucine, and appears as 41.8% of the leucine codons in *S. stipitis*, making this the preferred leucine codon. By changing any CUG codons to CUU, synthesized genes should have their structure and function restored when expressed in *S. stipitis*.

1.5 Recoverable selectable marker in gene knockout strategies

One way to remove a gene from the yeast genome is through a multi-step process called homologous recombination (Capecchi, 1989). This gene knockout strategy involves first building a plasmid with three main components which include a selectable marker which is flanked by an upstream portion of DNA and a downstream portion of DNA. These upstream and downstream regions should match the upstream and downstream regions of the gene that is targeted to be knocked out. The plasmid is then digested using restriction enzyme sites that remove genetic components unnecessary for the homologous recombination, such as the origin of replication and bacterial selectable marker. Products from the restriction digest are separated by size using gel electrophoresis. The band that contains the upstream and downstream regions of homology with the selectable marker in between the two regions is then excised from the gel, and purified. This DNA is then introduced to competent cells in a similar method as plasmid transformation. Homologous recombination however is a rare event compared to the uptake and expression of a plasmid. To compensate this infrequently occurring event, typically much more donor DNA is added to the reaction. The linearized DNA is then incorporated into the genome through cell repair mechanisms which replace a portion of the wildtype genome with the donor DNA as shown in figure 1.2. By using a selectable marker, it is possible to have a workable number of colonies on a selective media to PCR screen for successful homologous recombination at the targeted location.

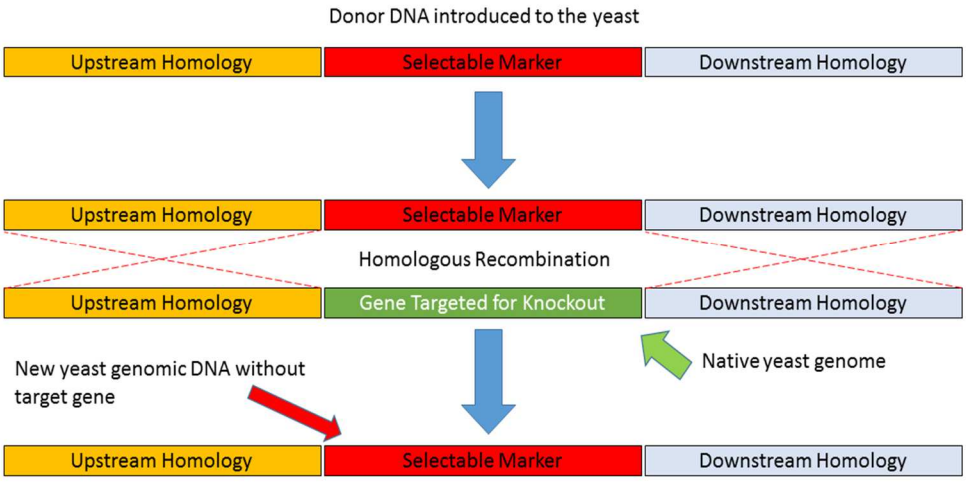


Figure 1.2. Homologous recombination inside the cell. The linear donor DNA is introduced into competent yeast cells. Cell repair mechanisms then use the donor DNA as a template, and can use the regions of homology to target a location on the genome and replace the targeted gene with the selectable marker.

This method of gene deletion is problematic when multiple genes need to be knocked out because the selectable marker used for the first gene knock out now integrated into the yeast cell genome and cannot be easily used for a second knock out. If homologous recombination is used multiple times on the same yeast strain, then multiple selectable markers must be used. This leads to two problems; a multi-drug resistant yeast strain, and fewer genetic markers available for genetic manipulation of the yeast strain. One way around these concerns is by using a *CRE/LoxP* strategy.

The *CRE/LoxP* gene knock out strategy can allow for recovery of the selectable marker after the homologous recombination, allowing the selectable marker to be removed from the genome. The *CRE/LoxP* strategy starts with building a plasmid with upstream and downstream homology and a selectable marker in between these two regions, as the traditional homologous recombination strategy did. However there are also *LoxP* sites that immediately flank the selectable marker. This plasmid is linearized and transformed similarly to how homologous recombination would work. Any successful transformants will be viable on media containing the first selectable marker. These cells are then cultured and transformed with a second plasmid which contains a second selectable marker, as well as a gene called *CRE* which is regulated by an inducible promoter. After the second transformation is successful, the *CRE* promoter is induced so that *CRE* is expressed. The *CRE* recombinase then finds and binds to any *LoxP* sites in the genome. If two *LoxP* sites are close enough, the two Cre recombinase enzymes bind to each other and cleave the DNA, and recombine the two *LoxP* sites. This leaves one *LoxP* site called a *LoxP* scar, and

1.6 Significance of the study

The demand for a profitable, renewable and carbon neutral fuel source is increasing as the current fuel sources are becoming depleted. This creates a need to find alternatives to petroleum. Bio-ethanol is one of the most promising prospects to act as a primary fuel source; however the conversion of ethanol from plant material is not yet efficient enough for commercial use. There are several steps to converting plant mass into ethanol, one of which involves the fermentation of sugars present after hydrolysis. Improving the efficiency of fermentation will help to make lignocellulosic ethanol a more practical fuel source then it is currently.

Scheffersomyces stipitis and other yeast species in the CUG clade are considered to be the strongest fermenters of xylose, making these yeast species valuable in the fermentation of lignocellulosic biomass. These yeast species are not as well studied, and would require several preliminary experiments to be conducted to develop the genetic tools required for more complex studies. By creating a genetic tool set that includes selectable markers, reporter genes, promoters, terminators and negative selectable markers, it will allow researchers to genetically alter and modify yeast in the CUG clade and possibly alter them to better suit a lignocellulosic fermentation.

1.7 Thesis goals

- Design and express a codon optimized version of *CBG99* in *S. stipitis*.
- Identify and synthesize codon optimized selectable markers that function in *S. stipitis*.
- Demonstrate ways that these tools can be used to study *S. stipitis*

1.8 Hypothesis

- Codon optimization will restore protein function of foreign genes expressed in *S. stipitis*.
- Using codon optimized selectable markers and *CBG99*, a toolset for studying the genetics of *S. stipitis* can be constructed.

CHAPTER II
NAVIGATING THE CUG CLADE OBSTACLE TO DESIGN A BIOLUMINESCENT REPORT FOR *S. stipitis*

2.1 Introduction

In order to genetically modify *S. stipitis* in a predictable way, a set of tools must first be developed. The tools required for this strategy include genetic elements such as multiple selectable markers, negative selectable markers, a reporter gene, an origin of replication for *S. stipitis*, and *S. stipitis* promoters and terminators. These tools have been used in other species of yeast, however due to the alternate codon translation for CUG, these tools are not modular. In order to have functional versions of these genes, the sequences were sent to GenScript for gene synthesis. Prior to ordering and synthesizing the selectable markers and reporter genes, the sequences were searched for CUG codons. All CUG codons were edited to CUU which is another leucine codon, as well as the highest frequency of appearance in wild type *S. stipitis* with an occurrence percentage of 41.8%. This change was hypothesized to allow for the genetic tools required for gene knockout to regain their functionality.

Coding sequences mentioned above also required regulatory elements like promoters and terminators that will control expression of the selectable markers, reporter gene and negative selectable markers. A final requirement was a method of plasmid replication by using either an autonomous replication sequence, or through integration of the plasmid into the yeast genome. Neither of the afore mentioned components required codon optimization since they were acquired through polymerase chain reaction of the wild type genome for *S. stipitis*, and as a result already had correct function.

In this chapter, the experiments performed for my thesis aim to demonstrate that the codon optimization of *HPH* and *CBG99* restores function to their protein products, and also demonstrate how this genetic toolset can be used to investigate questions regarding gene regulation found in *S. stipitis*. Several experiments are detailed in this chapter that each build off the results of the previous experiment. The combination of all of these experiments demonstrates the utility of the genetic toolset in *S. stipitis*.

2.2 Results and discussion

2.2.1 Testing a codon optimized reporter gene in *S. stipitis*

The goal of this project was to develop a genetic tool set that could be used to study the regulatory genetics of *S. stipitis*, as well as a tool set that could alter the genome of *S. stipitis* using a recoverable selectable marker to knock out genes. The first investigation in this study was whether or not a codon optimized variation of *CBG99* would express correctly in *S. stipitis*. To demonstrate this, a plasmid was built which contained *CBG99* under a constitutive promoter, as well as *HPH* for hygromycin B resistance and second constitutive promoter, and transformed this plasmid into *S. stipitis*. Colonies grew on the hygromycin B containing media, and were streaked onto a Yeast Peptone Dextrose plate that contained beetle luciferin (LH_2). The plate was then imaged using a ChemiDoc™ MP Imaging System. The image was taken in single accumulation mode, stacking 10 images layered on top of each other as a way to view the accumulation of bioluminescence emitted from the reaction taken across 5 minutes (Figure 2.1).

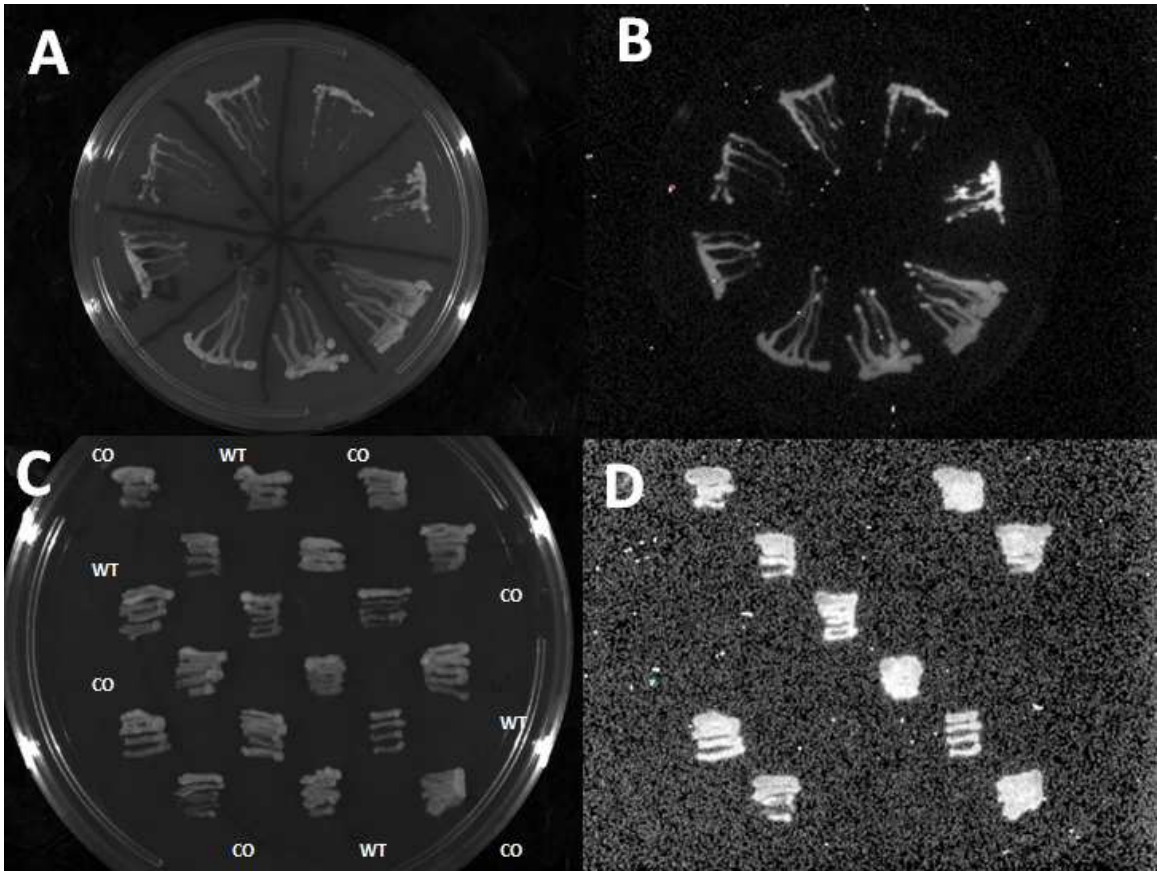


Figure 2.1. Testing a codon optimized reporter gene in *S. stipitis*. (A) Bright field image of streaked suspected transformants of pWR9 on YPD + hygromycin B + LH_2 (B) Bioluminescent of pWR9 transformants using a ChemiDoc™ MP imaging system with a 5 minute exposure time (**C**) Bright field image of patched colonies on YPD + hygromycin B + LH_2. Patches are aligned in alternating diagonals of transformed *S. stipitis* cells containing either wild type *CBG99* or the codon optimized variant of *CBG99* under the control of the constitutive *TDH3* promoter from *S. stipitis*. (**D)** Bioluminescence observed for a 5 minute exposure using the ChemiDoc™ MP imaging system. Colonies of *S. stipitis* containing the codon optimized *CBG99* emit light while the wild type *CBG99* containing colonies do not luminesce. Both patch types had a codon optimized *HPH* gene which allowed for continuing selection on the media.

2.2.2 Testing suspected inducible promoter regions

The next investigation for the genetic toolset was to test different promoters that would regulate *CBG99*. This experiment aimed to detail a few promoters which could be either constitutive or inducible. The promoters for three genes; *GAL1*, *XKS1*, and *POL1*, were placed upstream of *CBG99* using BamHI and KpnI. These transformants were plated on Yeast Peptone Dextrose + hygromycin B + LH_2, as well as similar media which used galactose or xylose as the primary sugar source instead of glucose. The hypothesis of this experiment is that identified xylose (*XYL1* and *XKS1*) and galactose (*GAL1*) import pathway associated genes would be transcriptionally active in the presence of the respective sugar molecule. This could be established by using plasmids with the promoters of these genes controlling the expression of *CBG99*. The results of these experiments are shown in figure 2.2. A promoter for *XYL1* was also tested, however turned out to not be inducible by xylose (See Appendix).

These results show P_{TDH3} and P_{POL1} regulated *CBG99* to have consistent expression in all three sugar environments, with P_{TDH3} showing higher overall expression. When grown in an environment where glucose or xylose is the primary sugar, P_{GAL1} *CBG99* cells were dim. Expression increased a noticeable amount for P_{GAL1} in the presence of galactose. The P_{XYL1} promoter appeared to have high expression in all three sugar environments. P_{XYL1} was not observed to be inducible by having xylose in the media. Another xylose inducible candidate, P_{XKS1} was tested in a similar experiment. The result of P_{XKS1} as a promoter

regulating *CBG99* demonstrated that xylose does induce P_{XKS1} activity. This is shown in figure 2.2.

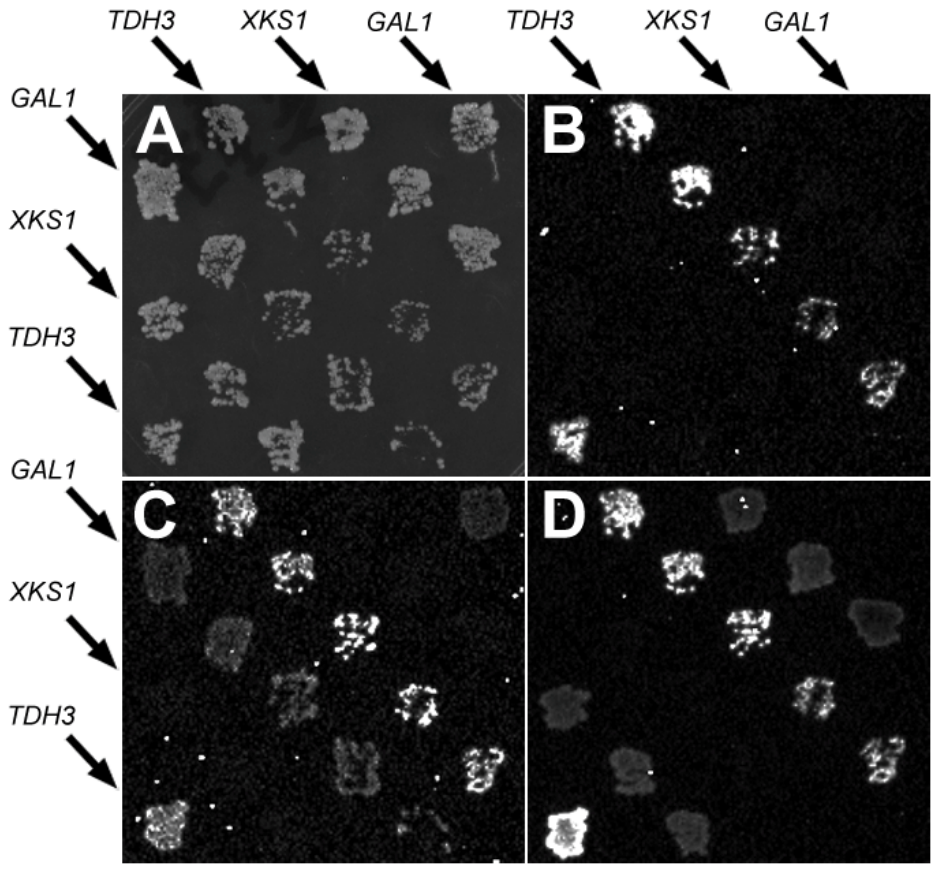


Figure 2.2. Testing the activity of sugar promoters (**A**) Bright field image of patched colonies on YPD + hygromycin B + LH_2 that were replicate plated on onto similar plates containing either galactose or xylose as the primary sugar. (**B)** Patched colonies with media primarily containing glucose. P_{TDH3} is the only promoter active in this environment, (**C**) Patched colonies on galactose based media. In the presence of galactose, the P_{GAL1} is active, producing bioluminescence. P_{TDH3} is constitutive, so therefore also bioluminescent on galactose. (**D)** Patched colonies on xylose based media. In the presence of xylose, the P_{XYL1} promoter is active and produces bioluminescence.

2.2.3 Creating a rapidly degrading bioluminescent reporter system

The use of a reporter gene such as *CBG99* is useful when studying the timing of gene expression, as shown in the previous experiment. Monitoring the expression of *CBG99* can be difficult because when *CBG99* is no longer transcriptionally active, the gene products are still present and can take a long time to degrade. The presence of bioluminescence from yeast can indicate gene expression based on the promoter regulating *CBG99*; however the protein products can remain after the gene is no longer being transcribed. As long as these proteins (and the other bioluminescent reactants) are in the cells, the cells will continue to glow. In order to study when *CBG99* transcription is turned off, the half-life of the protein needs to be reduced.

A solution to the lag time observed between when expression of *CBG99* has stopped, and when light is no longer emitted, could be to create a protein fusion for which *CBG99* would have added amino acids at the C-terminal end that have been associated with short protein half-life such as those found in PEST domains. Cloning an expression vector with *CBG99* and suspected PEST domain sequences could induce protein degradation of *CBG99* and allow gene regulation to be observed with greater temporal fidelity. One such PEST domain has been identified in a gene called *CLN2* (Salama *et al*., 1994). Multiple regions of *CLN2* were PCR amplified from *S. stipitis*. These regions varied in size, sharing the same C-terminus. The sizes of the genes amplified for protein fusion were the full protein (1.4kb), int1 (900bp), int2 (468bp), and int3 (357bp). These protein fusion plasmids were transformed into *S. stipitis*. The experiment for determining an optimum degron region for a *CBG99-CLN2* fusion involved inducing

expression, repressing expression, and quantifying the relative light units (RLU) produced over time (Figure 2.3).

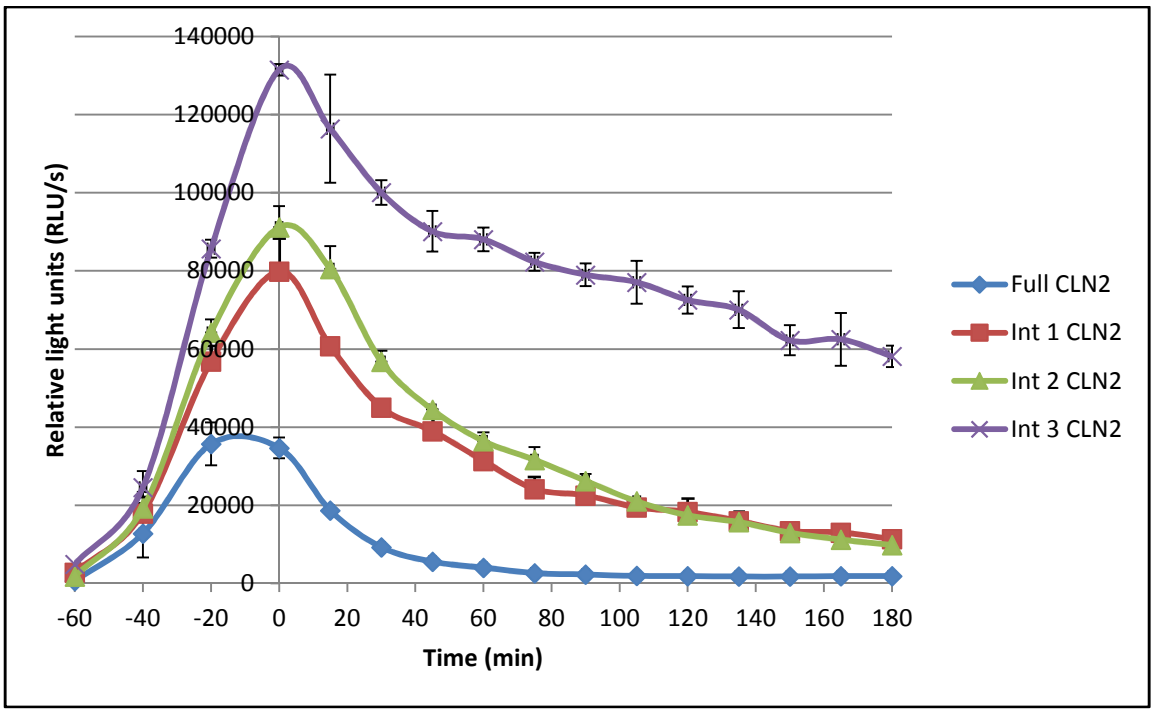


Figure 2.3. *CBG99-CLN2* fusion degron assay. All data sets begin at -60 minutes where the cultures were induced, and grown for one hour while being monitored. Time points were taken every 20 minutes until the zero time point. At zero, the cultures were repressed with galactose and cycloheximide, after which data points were recorded every 15 minutes for up to three hours. Full *CLN2* represents *CBG99* with the full *CLN2* protein fused to it. Int 1, Int 2, and Int 3 represent portions of *CLN2* that are shorter sequences, removing nucleotides from the 5' end in descending order, making Int 3 the shortest of the protein fusions.

CBG99-CLN2 was observed to emit the least amount of light before repression, while *CBG99-Int3* had the highest overall expression. This can be explained by understanding the more amino acid burden that is added to *CBG99*; the less efficient the luciferase can function. Int1 and Int2 showed overlapping levels of expression and repression. Qualitative comparisons can be made regarding expression and degradation between data sets, however quantitative analysis of degradation is skewed due to a lack of a common RLU starting point at the time of repression. A repetition of the experiment was performed, this time monitoring the expression and degradation of *CBG99* unmodified, *CBG99-CLN2*int2, and *CBG99-CLN2*full. The results are shown in figure 2.4. The data from this second experiment when normalized shows that the rate of degradation increases, when more of *CLN2* is fused to *CBG99*. This would not be immediately apparent without normalizing initial bioluminescence before repression and degradation occurs.

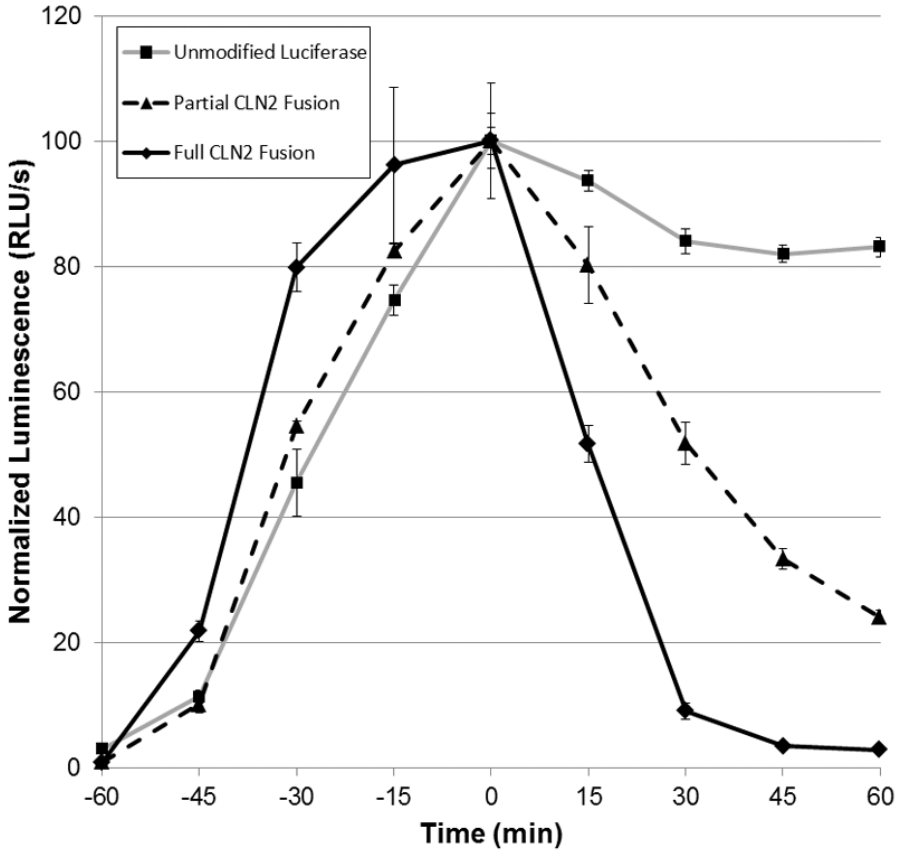


Figure 2.4. *CBG99-CLN2* fusion degron assay normalized. Observing unaltered *CBG99*, *CBG99-Int2* and *CBG99-CLN2*. The data is normalized so that RLU values are at 100% during the peak expression for each culture. The data shows the low rate of degradation of *CBG99*, compared to the partial and full *CLN2* protein fusions.

2.2.4 Measuring yeast competition using two bioluminescent reporter genes

In a theoretical fermentation consortium, there will be several species of yeast present fermenting their own specific sugar. A useful application for a bioluminescent reporter gene could be to observe competition between different yeast species in different sugar environments. The codon optimized *CBG99* for *S. stipitis* allows for the cells to emit a green luminescence. There are other reporter genes such as *CBR* which is a red bioluminescent reporter. If two yeast species have different color reporter systems, colonies can be distinguishable from each other. This can allow for investigating whether or not competition is present among different yeast in a co-culture by plating samples over time as shown in figure 2.5. The results of this study suggest that when *S. stipitis* is in a co-culture with *S. cerevisiae*, *S .stipitis* is at a disadvantage in the presence of only glucose. When xylose is added as half the sugar present in the media however, *S. stipitis* survival is increased.

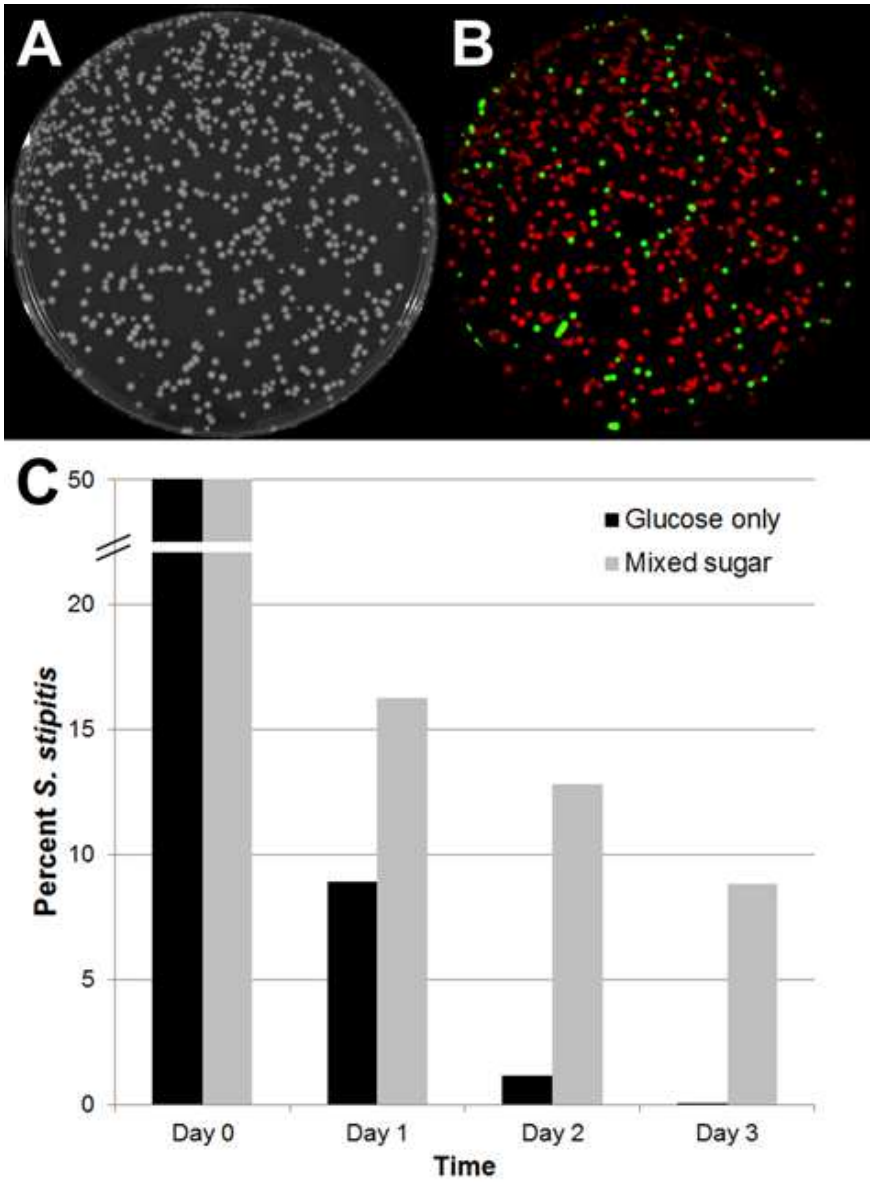


Figure 2.5. Monitoring yeast competition in different sugar environments. (**A**) Bright field example of a plate which has two yeast species growing together. This bright field image is for the plate found in panel B. This plate is a one day co-culture representation grown in the presence of glucose and xylose. (**B**)One day representation of two yeast species in co-culture growing in glucose and xylose producing bioluminescence. Red colonies represent *S. cerevisiae* while green colonies are *S. stipitis*. (**C**) Graph representing samples taken from a co-culture each day for three days. The cultures were started at day 0, and were inoculated with equivalent cells to start off with an initial concentration of 50% *S. stipitis*.

2.2.5 Using a codon optimized reporter gene in *Candida albicans*

Candida albicans is a well-studied species of yeast that has medical importance as one of the most common microorganisms attributed to causing health care- associated infections (Calderone *et al*., 2001). *C. albicans* is also a yeast species that is found in the "CUG clade", similar to *S. stipitis* (Perreau *et al*., 1999) (Santos & Tuite, 1995). This should mean that in theory, genes that maintain function in *S. stipitis* should work in the *C. albicans* species. A plasmid called pYM70 which contains a *C. albicans* autonomously replicating sequence, a synthetic hygromycin B resistance gene and a P_{ACT1} promoter to regulate a gene of interest cloned into the plasmid using the PacI and SacII restriction sites (Basso *et al*. 2010). The codon optimized version of *CBG99* using in *S.* stipitis was cloned into this plasmid and transformed into *C. albicans* to see if bioluminescence was present.

Bioluminescence was not observed on patch plates after transformation. It was hypothesized that since the integrity of the *C. albicans* P_{ACT1} was unknown, and was already contained within the plasmid pYM70, the emitted luminescence might have been below the threshold for the ChemiDoc™ MP imaging system due to low promoter activity. A 10 ml Yeast Peptone Dextrose with 10 µl of LH_2 and 80 µl of hygromycin was inoculated with a colony of *C. albicans* which was resistant to hygromycin but did not appear to be bioluminescent. The culture was monitored as it was growing, using a photomultiplier sensor to compensate for the low emission of bioluminescence. Results show *C. albicans* emitting light as the culture aged, increasing in cell numbers (Figure 2.6).

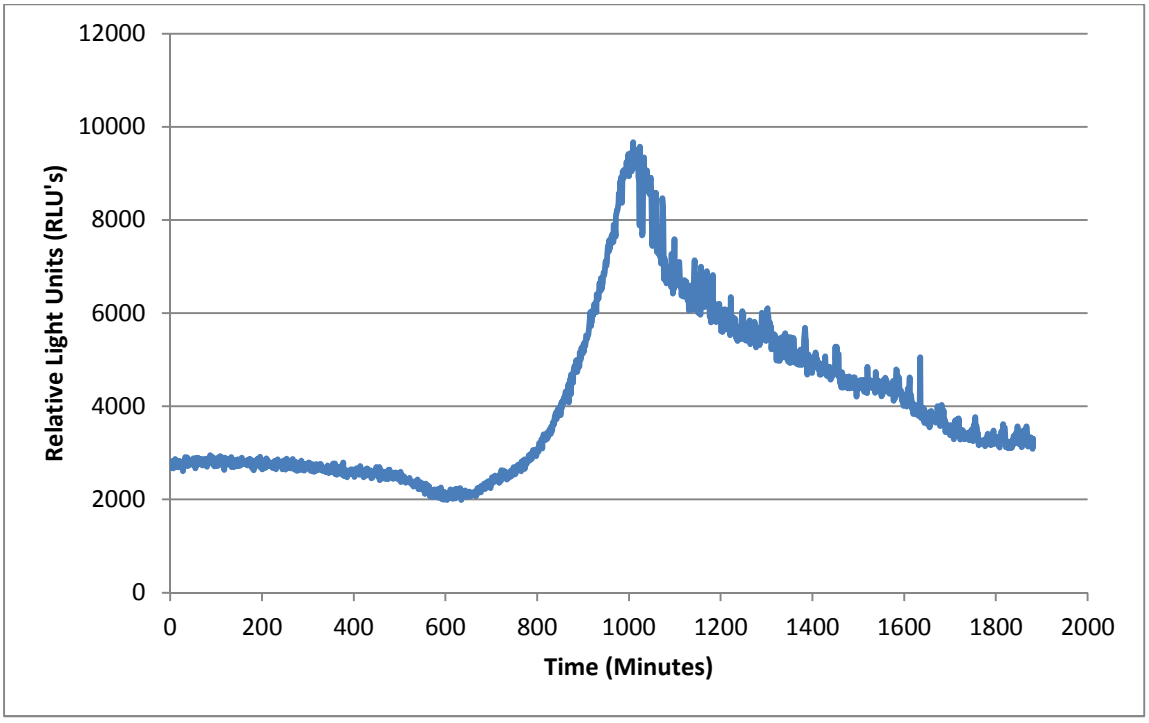


Figure 2.6. *Candida albicans* transformed with pWR32. Cultures were inoculated in YPD hygromycin B LH_2, and measured for bioluminescence by a PMT for 1,811 minutes. The culture had the highest RLU readings from 1,030 until 1,060 minutes. There is a baseline of background noise at 2,800 RLU's.

The data in figure 2.6 shows that a codon optimized *CBG99* does function in *C. albicans* since light was detected. However there was a 2,800 RLU baseline of noise in this experiment, and RLU value that falls below this number is lost in background interference. *C. albicans* does not exhibit the same expression intensities of *S. stipitis* with the same codon optimized *CBG99* in the current pYM70 plasmid.

2.3 Materials and methods

2.3.1 Naming conventions

Plasmids in this study are typically named with a lowercase p, followed by some information of the origin of the plasmid. One of the most commonly used plasmids in this thesis is named pWR9 which stands for plasmid Walter Reichard with a number based chronologically for when the plasmid was constructed. Plasmids are also labeled numerically based on the order they were built. Within the constructed plasmids are promoters and terminators which are indicated by a capital P or T, followed by the gene name in subscript for which the promoter or terminator is naturally associated to such as P_{GAL1} which represents the promoter that regulates the *GAL1 gene*. Synthesized genes in this study have co before the name to indicate that the gene has been codon optimized, such as co*HPH*.

2.3.2 Plasmid construction

The backbone plasmid called pAllet was designed using a large multiple cloning site synthesized by GenScript. The large multiple cloning site was inserted into pUC57-mini. This backbone plasmid has an ampicillin resistance gene, an origin of replication for *Escherichia coli*, as well as 51 unique restriction sites shown in figure 3. Having 51 unique restriction sites allows many pieces of DNA to be ligated into pAllet. This expansive cloning capability allowed a synthesized selectable marker and reporter gene to be ligated into the plasmid, as well as promoters, terminators and an *S. stipitis* autonomous replicating sequence (ARS) acquired through polymerase chain reactions from wild type *S. stipitis* to

create a functional bioluminescence reporter plasmid for the remainder of the project. Throughout the explanation of how the plasmids were constructed, various plasmids and primers will be referred to and are in Table 1 and Table 2.

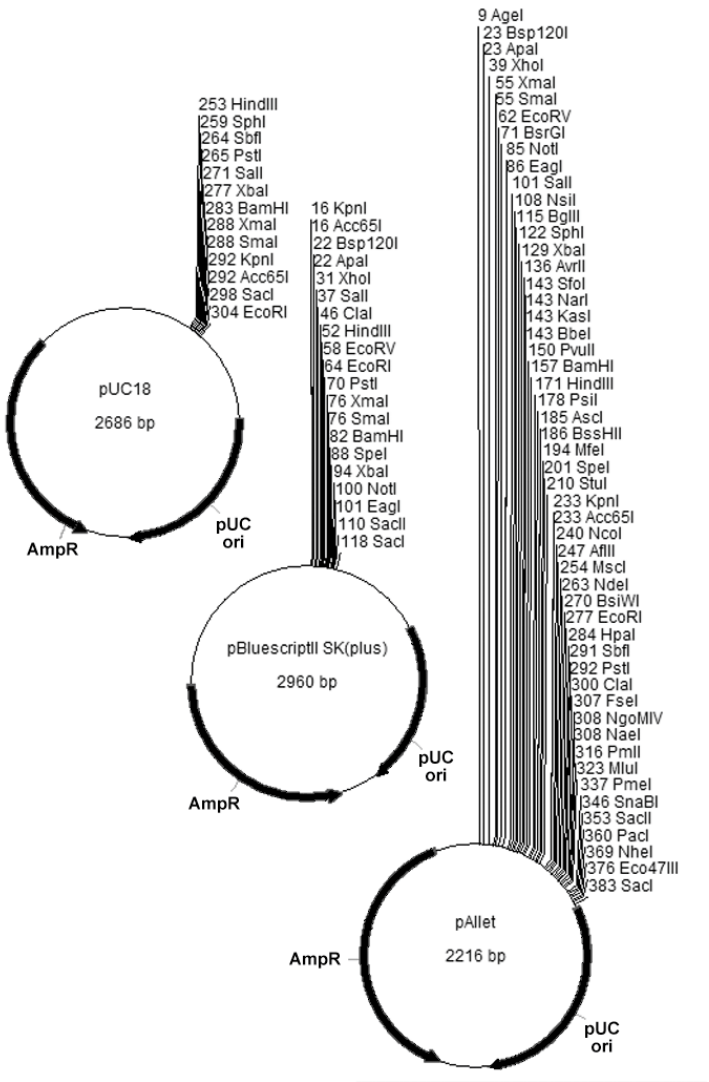


Figure 2.7. Plasmid map of pAllet compared to other cloning plasmids. The pUC18 expression plasmid has 13 unique restriction enzyme sites, pBluescriptII SK(plus) has 20 unique restriction enzyme sites, and pAllet has 51 unique restriction enzyme sites. Having more restriction sites allows for more options in downstream plasmid construction. Similar to pUC18 and pBluescriptII SK (plus), pAllet has an ampicillin resistance gene and an origin of replication for *Escherichia coli* to maintain the plasmid throughout cell divisions.

Table 1: Primers used for this study.

Primer no.	Name	Sequence (5'-3')
1	Ps*ARS*(SacI)5	aacttg**GAGCTC**AGTATAGGATATGGTGATTTAGC
2	Ps*ARS* (SacI)3	caagtt**GAGCTC**TCTGCGGTGTCTACAAGGTC
3	P_{TEF1}(XhoI)5	tgacgt**CTCGAG**CCGTACACTTATTGTTAACTATGAA
4	P_{TEF1}(BrsgI)3	tgccat**TGTACA**TGTAGATAGACTTAGATTGTATGAAA
5	T_{ACT1} (NotI)5	tgacgt**GCGGCCGC**TAACCACTTGCAAAATCCTTTG
6	T_{ACT1} (SalI)3	tgccat**GTCGAC**GAACTTGATTGTAATACAAAATGC
7	P_{TDH3}(BamHI)5	tcatta**GGATCC**TCGTTAGTATTTCCGTGAAG
8	P_{TDH3} (KpnI)3	atgact**GGTACC**GATGAATTGTTTATAGGGAAGA
9	T_{ADH2} (AflII)5	atactg**CTTAAG**CATTTTAGACAAGTGCCTATT
10	T_{ADH2} (EcoRI)3	tgccat**GAATTC**TTCTGCCTTCTGAACGTTTG
11	P_{GAL1}(BamHI)5	attcca**GGATCC**GAAGAACTTGTTCTGAAAGC
12	P_{GAL1} (KpnI)3	acttca**GGTACC**AGTAATAGAAACAACAGAACGT
13	P_{POL1} (Bam)5	atgcat**GGATCC**AAGATTCTACTGTTTTTATTTATAC
14	P_{POL1} (KpnI)3	cagtca**GGTACC**GCTAGATATAGAAGTGGAAAATG
15	*CLN2*stp(EcoRI)3	gcgtca**GAATTC**TCATGGAATGAAAAAACACAGG
16	*CLN2*full(NheI)5	atatat**GCTAGC**ATGACGACCTATTCCCACTC
17	*CLN2*int1(NheI)5	atatct**GCTAGC**GTGAACTCTCCAACATTTGACTC
18	*CLN2*int2(NheI)5	atatct**GCTAGC**TACACACAGTTGATGACAGCTCT
19	*CLN2*int3(NheI)5	atatct**GCTAGC**AAGAGGAACGTCAACATCATC
20	*CLN2*int4(NheI)5	atatat**GCTAGC**GCTATTGCTGCAGTTACAGGT

Table 2: Plasmids constructed for this study.

Plasmid	Description
pUC57-mini	Standard cloning vector from GenScript
pUC57-mini-co*CBG99*	Cloning vector from GenScript containing synthesized codon optimized *CBG99*
pUC57-mini-co*HPH*	Cloning vector from GenScript containing synthesized codon optimized *HPH*
pAllet	Synthesized multiple cloning site inside pUC57-mini
pWR1	pAllet (P_{TEF1}-co*HPH*-T_{ACT1})
pWR5	pAllet (P_{TEF1}-co*HPH*-T_{ACT1}) (P_{TDH3})
pWR6	pAllet (P_{TEF1}-co*HPH*-T_{ACT1})(P_{TDH3}-co*CBG99*)
pWR8	pAllet (P_{TEF1}-co*HPH*-T_{ACT1})(P_{TDH3}-co*CBG99*-T_{ADH1})
pWR9	pAllet (P_{TEF1}-co*HPH* -T_{ACT1})(P_{TDH3}-co*CBG99*-T_{ADH1}) (*ARS2*)
pWR11	pAllet (P_{TEF1}-co*HPH*-T_{ACT1}) (P_{GAL1}-co*CBG99*-T_{ADH1}) (*ARS2*)
pWR12	pAllet (P_{TEF1}-co*HPH*-T_{ACT1}) (P_{XYL1}-co*CBG99*-T_{ADH1}) (*ARS2*)
pWR13	pAllet (P_{TEF1}-co*HPH*-T_{ACT1}) (P_{POL1}-co*CBG99*-T_{ADH1}) (*ARS2*)
pWR16	pAllet (P_{TEF1}-co*HPH*-T_{ACT1}) (P_{TDH3}-co*CBG99*-T_{ADH1}) (*URA*)
pWR22	pAllet (P_{TEF1}-co*HPH*-T_{ACT1}) (P_{XKS1}-co*CBG99*-T_{ADH1}) (*ARS2*)
pWR26	pAllet (P_{TEF1}-co*HPH*-T_{ACT1})(P_{GAL1}-co*CBG99*fuze-T_{ADH1}) (ARS2) (Protein fusion without *CLN2* protein region)
pWR26FULL	pAllet (P_{TEF1}-co*HPH*-T_{ACT1})(P_{GAL1}- co*CBG99*fuze*CLN2* FULL-T_{ADH1}) (*ARS2*)
pWR26INT1	pAllet (P_{TEF1}-co*HPH*-T_{ACT1})(P_{GAL1}-co*CBG99*fuze*CLN2*int1-t T_{ADH1}) (*ARS2*)
pWR26INT2	pAllet(P_{TEF1}-co*HPH*-T_{ACT1})(P_{GAL1}-co*CBG99*fuze*CLN2*int2- T_{ADH1}) (*ARS2*)
pWR26INT3	pAllet(P_{TEF1}-co*HPH*-T_{ACT1})(P_{GAL1}-co*CBG99*fuze*CLN2*int3- T_{ADH1}) (*ARS2*)
pWR26INT4	pAllet(P_{TEF1}-co*HPH*-T_{ACT1})(P_{GAL1}-co*CBG99*fuze*CLN2*int4- T_{ADH1}) (*ARS2*)
pWR27	pAllet (P_{TEF1}-co*HPH*-T_{ACT1})(P_{GAL1}-wt*CBG99*-T_{ADH1}) (*ARS2*)
pWR32	pYM70 (P_{TEF2}-*CaHygB*-T_{ACT1}) (P_{ACT1}-co*CBG99*-T_{TEF2}) (*CaARS2*)
pWR34	pAllet(P_{TEF1}-co*ShBle*-T_{ACT1})(P_{TDH3}-co*CBG99*-T_{ADH1})(P_{GAL1}-*Cre*-T_{ADH2})(*ARS2*)
pWR37	pAllet pHloxer (*URA* US)(P_{TEF1}-co*HPH-*T_{ACT1})(*URA DS*)
pWR38	pAllet (P_{TEF1}-co*HPH*-T_{ACT1})(P_{TDH3}-wt*CBG99*fuze-T_{ADH1}) (*ARS2*)
pWR39	pAllet (P_{TEF1}-co*HPH*-T_{ACT1})(P_{TDH3}-co*CBG99*fuze-T_{ADH1}) (*ARS2*)

The Hygromycin-B 4-O-kinase gene (*HPH*) was synthesized, codon optimized and ligated into pAllet using the restriction enzymes BsrGI and NotI. The promoter for transcription elongation factor (P_{TEF1}) and the terminator for actin were PCR amplified from wildtype *S. stipitis*. P_{TEF1} was PCR amplified using primer sets 3 and 4 from wildtype *S. stipitis* from ATCC (58785) strain CBS 6054, and the amplicon was added to pAllet (co*HPH)* by digesting the vector with XhoI and BsrGI so the promoter is ligated into the upstream region of co*HPH*. Similarly the terminator was PCR amplified using primers 5 and 6, and the amplicon was added to pAllet (P_{TEF1}-co*HPH*) by using the restriction enzyme sites NotI and SalI, and ligating the promoter just downstream of the *HPH* gene. This plasmid then became named pWR1.

The promoter for *TDH3*, the terminator for *ADH2* and an identified autonomous replication sequence were all PCR amplified from wild type *S. stipitis* (Primers 7, 8, 9 and 10). P_{TDH3} was added to pWR1 by using BamHI and KpnI and became pWR5. Codon optimized *CBG99* was cut out of the transport vector from GenScript and ligated into pWR5 using the restriction enzyme sites KpnI and AfiII to become pWR6. The *ADH2* terminator was ligated into pWR6 using EcoRI and AfiII, placing the terminator just downstream the *CBG99*. The last element that was added to this plasmid was *ARS2*, an autonomous replication sequence from chromosome 1 (Yang *et al*. 1994), which was ligated into pWR9 using SacI. This constructed plasmid became the most used plasmid in this study, pWR9 shown in figure 2.2. It was useful as a constitutively active bioluminescent reporter plasmid, and a backbone which could be modified for a variety of experiments.

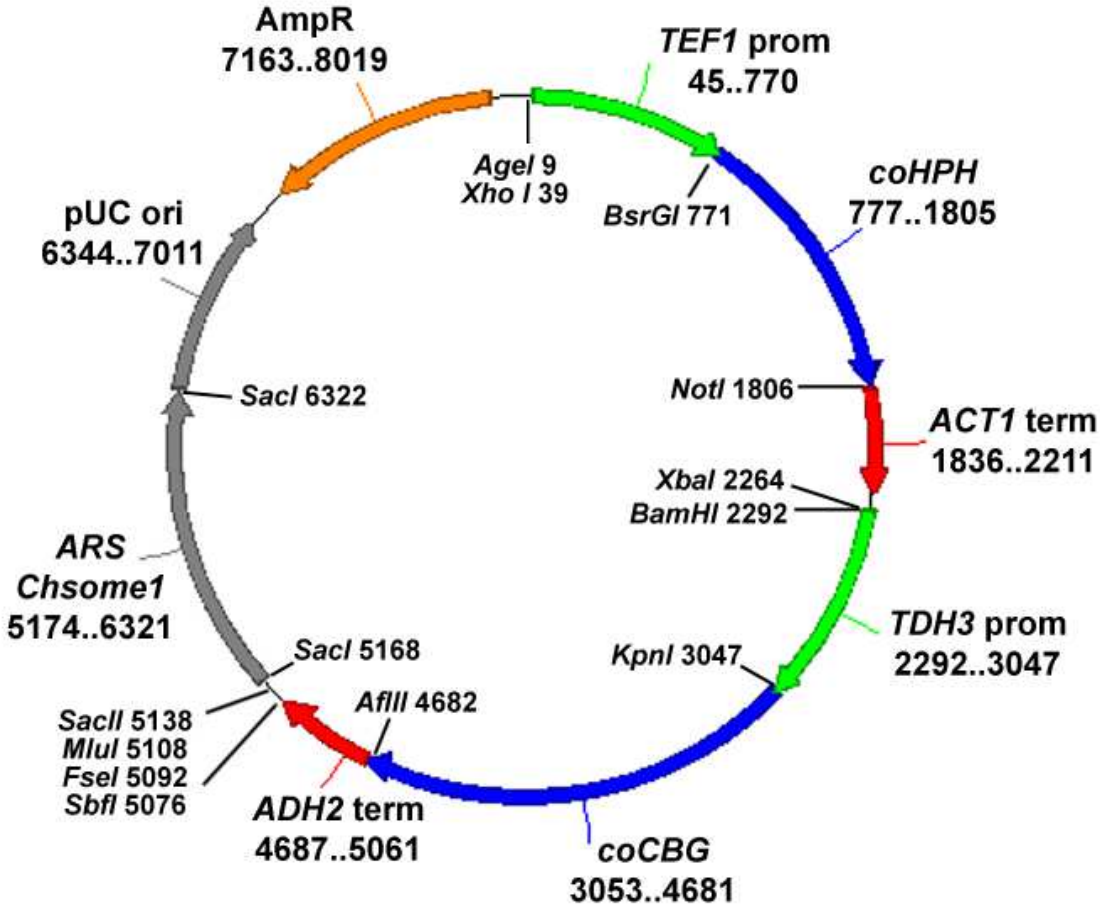


Figure 2.8. Plasmid map of pAllet (PTEF1-coHPH-TACT1)(PTDH3-coCBG99-TADH1) (ARS2). This plasmid supplied *S. stipitis* with a resistance to hygromycin B, providing selectable pressure on cells that do not contain the plasmid. The cells containing the plasmid also express co*CBG99* and provide bioluminescence in the presence of beetle luciferin.

2.3.3 Transformation of *S. stipitis*

S. stipitis transformation protocol was based off of a protocol for *S. cerevisiae* as well as the *S. stipitis* transformation (Gietz *et al*., 1995) (Yang *et al*., 1994). This protocol did not work and required the troubleshooting of different variables until a successful protocol could be established. Transformation efficiency was increased when cell number was increased nine fold, and carrier DNA was heated to become single stranded. The final product of these modifications is described below.

S. stipitis cells were grown to full cell density on the first day. The cells grown to max density were inoculated into a 50 ml flask of YPD at a dilution level of 1:1500 or 1:2000. The cultures then took approximately 10 hours to grow to an OD_{600} of 0.8. The cells were portioned into 9 ml volumes and pelleted at 3000 rpm for 5 minutes. The cells were resuspended in 500 µl of 1X TE and 0.1M lithium acetate to wash the cells and were pelleted again. The cells were then resuspended into 100 µl of 1X TE and 0.1X lithium acetate and allowed to incubate for one hour at 30°C while constantly inverting. Then 5-10 µg of DNA was added in combination with 15µl of single stranded herring sperm DNA, and 700 µl of a mixture of 40% poly ethylene glycol, 1X TE, and 0.1M LiOAc. The cells were left to incubate at 30°C for tumbling for 30 minutes, before being heat shocked at 42°C for 15 minutes. The cells were then pelleted and resuspended in 3 ml of YPD to outgrow for 4 hours. The culture was then pelleted and resuspended in 100 ml of YPD and plated onto YPD containing a selective antibiotic to prevent cells that were not successfully transformed from making colonies.

2.3.4 Degron assay experimental design

CLN2 was PCR amplified from wildtype *S. stipitis*, along with four other PCR products which contained smaller portions of *CLN2*, each product a shorter length with *CLN2*int4 being the shortest. These 5 variations were cloned as protein fusions to *CBG99* and built as pWR26 shown in Table 2. This was done by PCR amplifying co*CBG99* from pWR9 adding an extra NheI restriction site before the *ADH2* terminator, and the modified co*CBG99* was cloned into pWR11. The PCR amplified *CLN2* regions were cloned into this new plasmid using EcoRI and NheI. These plasmids were then transformed and tested for the degradation rate of *CBG99*.

10 ml cultures were grown in Yeast Peptone media containing glycerol (3%) and hygromycin B. These cultures grew for two days, and the optical densities were recorded. Cultures were diluted to equivalent culture densities, by adding a volume of YP glycerol 3% hygromycin B LH_2, depending on culture density to create a 15 ml culture of each series. 3 ml of each culture were aliquoted into scintillation vials allowing for triplicate measurements and optical density measurements. Cultures in the vials were shaken at 180 rpm and 30°C for 3 hours. Cultures were induced with 200 µl of 30% galactose. After one hour of induced *CBG99* expression, cultures were repressed using 120 µl of 50% glucose and 30 µl of 100 mM cycloheximide and monitored for several hours after repression. Throughout the experiment, cultures were measured for bioluminescence using a using a FB 12 luminometer. From galactose inducement, until repression, cultures were measured for bioluminescence every 20 minutes. After repression, cultures were measured every 15 minutes.

2.3.5 Monitoring yeast competition using two reporter systems experimental design

A CBR expressing strain for *S. cerevisiae* had been previously built for an unrelated project in the Robertson lab. It was constructed as follows: The hph cassette from pRS305-*hph*-CBG99 (Krishnamoorthy and Robertson, 2015) was excised and moved to pRS303 (Sikorski & Hieter, 1989) using BamHI and NotI, to create the plasmid pRS303 *hph*. Then the P_{ADH1}-*CBR*- cassette from pRS306-P_{ADH1}-CBR (Krishnamoorthy & Robertson, 2015) was excised and moved to pRS303-hph with XbaI and SalI to create the plasmid pRS303-*hph*-P_{ADH1}-*CBR* (an integrating plasmid conferring hygromycin resistance and constitutively expressed click beetle red luciferase). pRS303-*hph*-P_{ADH1}-*CBR* was linearized with NheI and transformed into *S. cerevisiae* strain CEN.PK113-7D.

Fresh cultures *S. stipitis* (expressing CBG) and *S. cerevisiae* (expressing CBR) were inoculated and grown overnight. The cultures were measured for cell numbers using a hemocytometer and a spectrophotometer collecting OD_{600} values to correlate with cell number. Yeast cultures were then diluted to equivalent cell densities and inoculated into two 4 ml cultures, one broth of Yeast Peptone containing only glucose and the other having 50% of glucose and 50% xylose. At the point of inoculation a sample was plated to ensure that cells were correctly diluted. The following day the cultures were plated and new cultures were started from the previous cultures to maintain the variance in cell numbers for each out grow stage. Plated samples were imaged two days after plating for bioluminescence, and colonies were counted based on the color of light emitted from the colonies.

2.3.6 Expressing a codon optimized bioluminescent reporter in *Candida albicans*

To build and express a codon optimized reporter gene in *C. albicans*, an expression vector called pYM70 was used (Basso *et al*., 2010). The codon optimized *CBG99* gene used in the *S. stipitis* experiments was cloned into this plasmid using the restriction sites PacI and SacII. This put *CBG99* between P_{ACT1} and a T_{TEF2} terminator; both of which were present in pYM70 already and were acquired through PCR of *C. albicans* genome (Basso, *et al*. 2010). Plasmids were transformed into *C. albicans* cells using EZ-YEAST TRANSFORMATION II kit made by Zymo Research. The protocol for transformation was followed by the protocol provided with the kit, however only freshly prepared competent cells were used since frozen competent cells for *C. albicans* were suggested to give low yields by Zymo Research.

The expression plasmid also relied on hygromycin selection using a codon optimized *hph* gene. *C. albicans* demonstrated higher tolerance to hygromycin B in comparison to *S. stipitis* and *S. cerevisiae* which would not grow at a concentration of hygromycin B 200µg/ml in YPD. *C. albicans* was tested at different concentrations of hygromycin B, and was found to be susceptible at a concentration of 400µg/ml in YPD, however after three days small colonies could be observed. Hygromycin B resistant strains were observed to grow much faster, causing a distinct difference in colony size.

Faster growing colonies were tested for bioluminescence using a FB 12 luminometer. One of the colonies was cultured in 10 ml of YPD + 10 µl of LH_2 + 80 µl of hygromycin B. The culture was monitored by a PMT monitor overnight on a stir plate with the temperature stabilized at 32.5˚C-32.9˚C.

CHAPTER III
CONCLUSION

3.1 Summary

Codon optimization is an effective way to solve the barrier to modularity for yeast species in the "CUG clade" and rebuild the genetic tools used in yeast which utilize the traditional method of translation. By codon optimizing a bioluminescent reporter gene and selectable markers, several genetic tools were created which can lead to further studies of *S. stipitis* and potentially other yeast in the "CUG clade". The tool set outlined in this study is capable allowing the transformation of plasmids through either an ARS or integration, the ability to study promoter activity with high degradation for an accurate response to a lack of promoter stimulus, and allows for studies to monitor co-culture competition rates. These genetic tools are a useful stepping stone for better understanding the genetics of *S. stipitis* and potentially other members of the "CUG clade" such as *C. albicans*.

3.2 Future directions

This project concludes with a genetic tool set which allows the researcher to study the genetics of *S. stipitis* and should prove useful to anyone working with this species of yeast. One of the next steps in this project specifically is to demonstrate and create a knockout strain incapable of fermenting glucose by using the tools designed in this study. For *S. stipitis* to have a role in a lignocellulosic consortium, the genes which import glucose into the cell will need to be knocked out of the genome. This aspect of the project has already been started through plasmid construction of a codon optimized *CRE*/*LoxP*

plasmid (see Appendix). Other future directions can include further investigating the use of the codon optimized bioluminescence reporter tool set in other yeast found within the "CUG clade".

REFERENCES

1. Baldwin, Thomas O. "Firefly luciferase: the structure is known, but the mystery remains." *Structure* 4.3 (1996): 223-228.

2. Basso, Luiz R., *et al*. "Transformation of Candida albicans with a synthetic hygromycin B resistance gene." *Yeast* 27.12 (2010): 1039-1048.

3. Binder, Arthur, Luciano Pelloni, and Armin Fiechter. "Delignification of straw with ozone to enhance biodegradability." *Applied Microbiology and Biotechnology* 11.1 (1980): 1-5.

4. Brandes, Christian, *et al*. "Novel features of Drosophila period transcription revealed by real-time luciferase reporting." *Neuron* 16.4 (1996): 687-692.

5. Bruinenberg, Peter M., *et al*. "NADH-linked aldose reductase: the key to anaerobic alcoholic fermentation of Xylose by yeasts." *Applied microbiology and biotechnology* 19.4 (1984): 256-260.

6. Burgess-Brown, Nicola A., *et al*. "Codon optimization can improve expression of human genes in *Escherichia coli*: A multi-gene study." *Protein expression and purification* 59.1 (2008): 94-102.

7. Calderone, Richard A., and William A. Fonzi. "Virulence factors of Candida albicans." Trends in microbiology 9.7 (2001): 327-335.

8. Capecchi, Mario R. "Altering the genome by homologous recombination." Science 244.4910 (1989): 1288-1292.

9. Cianchetta, Stefano, *et al*. "Evaluation of selected white-rot fungal isolates for improving the sugar yield from wheat straw." *Applied biochemistry and biotechnology* 173.2 (2014): 609-623.

10. Cutfield, J. F., P. A. Sullivan, and S. M. Cutfield. "Minor structural consequences of alternative CUG codon usage (Ser for Leu) in Candida albicans exoglucanase." *Protein engineering* 13.10 (2000): 735-738.

11. da Costa Sousa, Leonardo, *et al*. "'Cradle-to-grave'assessment of existing lignocellulose pretreatment technologies." *Current opinion in biotechnology* 20.3 (2009): 339-347.

12. Delgenes, J. P., R. Moletta, and J. M. Navarro. "Effects of lignocellulose degradation products on ethanol fermentations of glucose and Xylose by *Saccharomyces cerevisiae*, *Zymomonas mobilis*, *Pichia stipitis*, and *Candida shehatae*." *Enzyme and microbial technology* 19.3 (1996): 220-225.

13. Du Preez, J. C., and B. A. Prior. "A quantitative screening of some Xylose-fermenting yeast isolates." *Biotechnology letters* 7.4 (1985): 241-246.

14. Du, Wanqing, *et al*. "The promoting effect of byproducts from Irpex lacteus on subsequent enzymatic hydrolysis of bio-pretreated cornstalks." *Biotechnology for biofuels* 4.1 (2011): 37.

15. Fan, L.T., Gharpuray, M.M., Lee, Y.-H., 1987. In: Cellulose Hydrolysis Biotechnology Monographs. Springer, Berlin, p. 57.

16. Grantham, R *et al*. "Codon Catalog Usage and the Genome Hypothesis." *Nucleic Acids Research* 8.1 (1980): r49–r62. Print.

17. Grantham, Richard. "Working of the genetic code." *Trends in Biochemical Sciences* 5.12 (1980): 327-331.

18. Hahn-Hägerdal, Bärbel, *et al*. "Bio-ethanol–the fuel of tomorrow from the residues of today." *Trends in biotechnology* 24.12 (2006): 549-556.

19. Hsu, T., R. Ladisch, and G. Tsao. "Alcohol from cellulose." *Chem. Intermed* 1203.3 (1980): 3.

20. Jeffries, Thomas W., *et al*. "Genome sequence of the lignocellulose-bioconverting and Xylose-fermenting yeast *Pichia stipitis*." *Nature biotechnology* 25.3 (2007): 319-326.

21. Kawaguchi, Yoshiyuki, *et al*. "The codon CUG is read as serine in an asporogenic yeast *Candida cylindracea*." (1989): 164-166.

22. Kim, Soo Rin, *et al*. "Simultaneous co-fermentation of mixed sugars: a promising strategy for producing cellulosic ethanol." *Trends in biotechnology* 30.5 (2012): 274-282.

23. Krishnamoorthy, Archana, and J. Brian Robertson. "Dual-Color Monitoring Overcomes the Limitations of Single Bioluminescent Reporters in Fast-Growing Microbes and Reveals Phase-Dependent Protein Productivity during the Metabolic Rhythms of Saccharomyces cerevisiae." *Applied and environmental microbiology* 81.18 (2015): 6484-6495.

24. Kumar, Parveen, *et al*. "Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production." *Industrial & engineering chemistry research* 48.8 (2009): 3713-3729.

25. Kurtzman, Cletus P., and Motofumi Suzuki. "Phylogenetic analysis of ascomycete yeasts that form coenzyme Q-9 and the proposal of the new genera *Babjeviella*, *Meyerozyma*, *Millerozyma*, *Priceomyces*, and *Scheffersomyces*." *Mycoscience* 51.1 (2010): 2-14.

26. Lee, Hung, et al. "Utilization of xylan by yeasts and its conversion to ethanol by *Pichia stipitis* strains." *Applied and environmental microbiology* 52.2 (1986): 320-324.

27. Massey, Steven E., *et al*. "Comparative evolutionary genomics unveils the molecular mechanism of reassignment of the CTG codon in *Candida spp*." *Genome research* 13.4 (2003): 544-557.

28. McAloon, Andrew, *et al*. "Determining the cost of producing ethanol from corn starch and lignocellulosic feedstocks." *National Renewable Energy Laboratory Report* (2000).

29. Melake, Tsega, Volkmar Passoth, and Ulrich Klinner. "Characterization of the genetic system of the xylose-fermenting yeast *Pichia stipitis*." *Current microbiology* 33.4 (1996): 237-242.

30. Mezzanotte, Laura, *et al*. "Sensitive dual color in vivo bioluminescence imaging using a new red codon optimized firefly luciferase and a green click beetle luciferase." *PloS one* 6.4 (2011): e19277.

31. Millett, Merril A., Andrew J. Baker, and Larry D. Satter. "Physical and chemical pretreatments for enhancing cellulose saccharification." *Biotechnol. Bioeng. Symp.;(United States)*. Vol. 6. Dept. of Agriculture, Madison, WI, 1976.

32. Mosier, Nathan, *et al*. "Features of promising technologies for pretreatment of lignocellulosic biomass." *Bioresource technology* 96.6 (2005): 673-686.

33. Nigam, J. N. "Ethanol production from wheat straw hemicellulose hydrolysate by *Pichia stipitis*." *Journal of biotechnology* 87.1 (2001): 17-27.

34. Ohama, Takeshi, *et al*. "Non-universal decoding of the leucine codon CUG in several *Candida* species." *Nucleic acids research* 21.17 (1993): 4039-4045.

35. Parekh, S. R., R. S. Parekh, and M. Wayman. "Fermentation of Xylose and cellobiose by *Pichia stipitis* and *Brettanomyces clausenii*." *Applied biochemistry and biotechnology* 18.1 (1988): 325-338.

36. Pereira, Francisco B., *et al*. "Industrial robust yeast isolates with great potential for fermentation of lignocellulosic biomass." *Bioresource technology*161 (2014): 192-199.

37. Perreau, Victoria M., *et al*. "The Candida albicans CUG-decoding ser-tRNA has an atypical anticodon stem-loop structure." *Journal of molecular biology* 293.5 (1999): 1039-1053.

38. Pignal, M. C. "Une nouvelle espece de levure isolee de larves d'insectes: *Pichia stipitis*." *Bull Mens Soc Linn Lyon* 36 (1967): 163-168.

39. Potumarthi, Ravichandra, *et al*. "Simultaneous pretreatment and sacchariffication of rice husk by *Phanerochete chrysosporium* for improved production of reducing sugars." *Bioresource technology* 128 (2013): 113-117.

40. Priest, Shelby J., and Michael C. Lorenz. "Characterization of virulence-related phenotypes in *Candida* species of the CUG clade." *Eukaryotic cell* 14.9 (2015): 931-940.

41. Robertson, J. Brian, et al. "Real-time luminescence monitoring of cell-cycle and respiratory oscillations in yeast." *Proceedings of the National Academy of Sciences* 105.46 (2008): 17988-17993.

42. Robertson, J. Brian, and Carl Hirschie Johnson. "Luminescence as a continuous real-time reporter of promoter activity in yeast undergoing respiratory oscillations or cell division rhythms." *Yeast Genetic Networks: Methods and Protocols* (2011): 63-79.

43. Rogers, Scott, Rodney Wells, and Martin Rechsteiner. "Amino acid sequences common to rapidly degrade proteins: the PEST hypothesis." *Science* 234 (1986): 364-369.

44. Salama, Sofie R., Kristin B. Hendricks, and Jeremy Thorner. "G1 cyclin degradation: the PEST motif of yeast *Cln2* is necessary, but not sufficient, for rapid protein turnover." *Molecular and Cellular Biology* 14.12 (1994): 7953-7966.

45. Santos, Manuel AS, and Mick F. Tuite. "The CUG codon is decoded in vivo as serine and not leucine in *Candida albicans*." *Nucleic Acids Research* 23.9 (1995): 1481-1486.

46. Sharp, Paul M., *et al*. "Codon usage patterns in *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster* and *Homo sapiens*; a review of the considerable within-species diversity." *Nucleic acids research* 16.17 (1988): 8207-8211.

47. Sikorski, Robert S., and Philip Hieter. "A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae." *Genetics* 122.1 (1989): 19-27.

48. Slininger, Patricia J., *et al*. "Repression of Xylose-specific enzymes by ethanol in *Scheffersomyces* (*Pichia*) *stipitis* and utility of repitching Xylose-grown populations to eliminate diauxic lag." *Biotechnology and bioengineering* 108.8 (2011): 1801-1815.

49. Suh, Sung-Oui, *et al*. "Wood ingestion by passalid beetles in the presence of Xylose-fermenting gut yeasts." *Molecular Ecology* 12.11 (2003): 3137-3145.

50. Sun, Ye, and Jiayang Cheng. "Hydrolysis of lignocellulosic materials for ethanol production: a review." *Bioresource technology* 83.1 (2002): 1-11.

51. Villalobos, Victor, *et al*. "Dual-color click beetle luciferase heteroprotein fragment complementation assays." *Chemistry & biology* 17.9 (2010): 1018-1029.

52. Wood, Keith V. "Marker proteins for gene expression." *Current opinion in biotechnology* 6.1 (1995): 50-58.

53. Yang, Vina W., *et al*. "High-efficiency transformation of Pichia stipitis based on its URA3 gene and a homologous autonomous replication sequence, ARS2." *Applied and environmental microbiology* 60.12 (1994): 4245-4254.

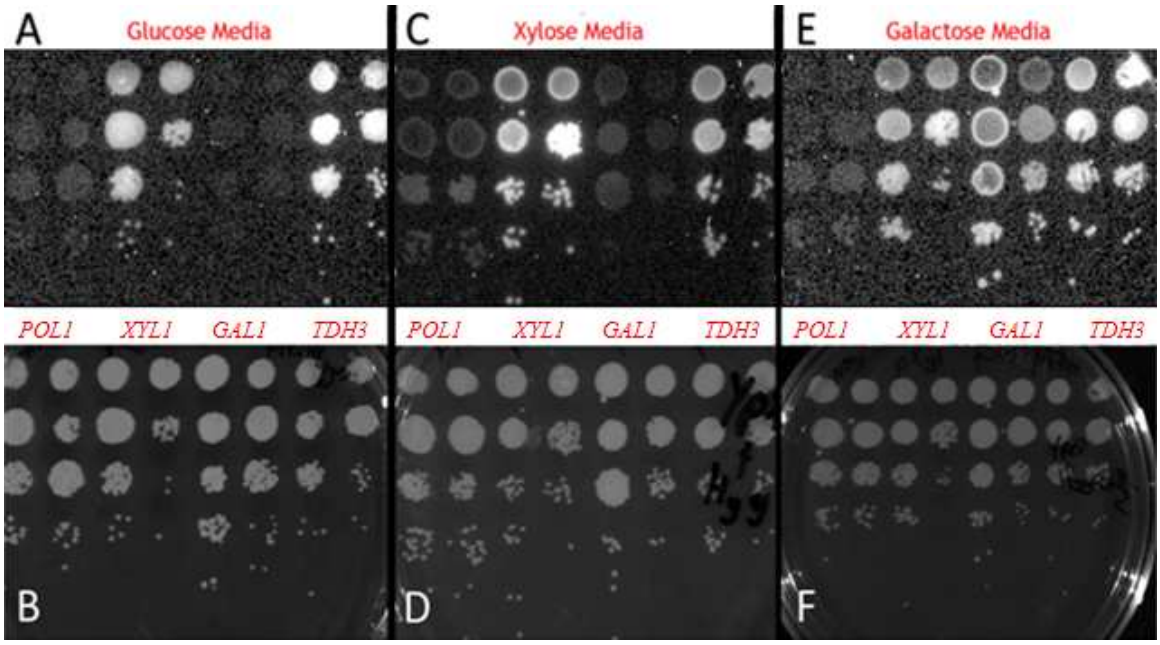
54. Young, Eric M., *et al*. "Rewiring yeast sugar transporter preference through modifying a conserved protein motif." *Proceedings of the National Academy of Sciences* 111.1 (2014): 131-136.

55. Zaldivar, Jesus, Jens Nielsen, and Lisbeth Olsson. "Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process integration." *Applied microbiology and biotechnology* 56.1 (2001): 17-34.

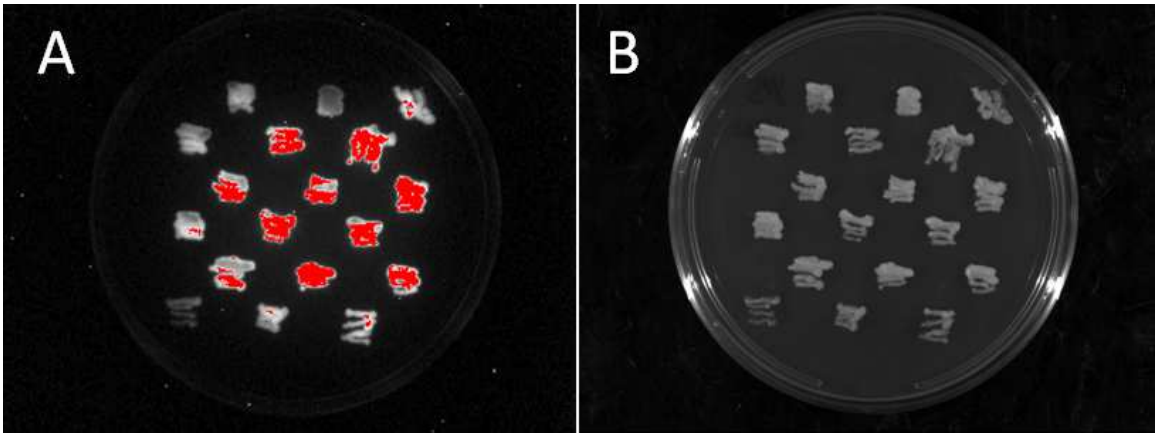
56. Zheng, Dao-Qiong, *et al*. "Screening and construction of *Saccharomyces cerevisiae* strains with improved multi-tolerance and bioethanol fermentation performance." *Bioresource technology* 102.3 (2011): 3020-3027.

APPENDICES

APPENDIX A: ADDITIONAL FIGURES

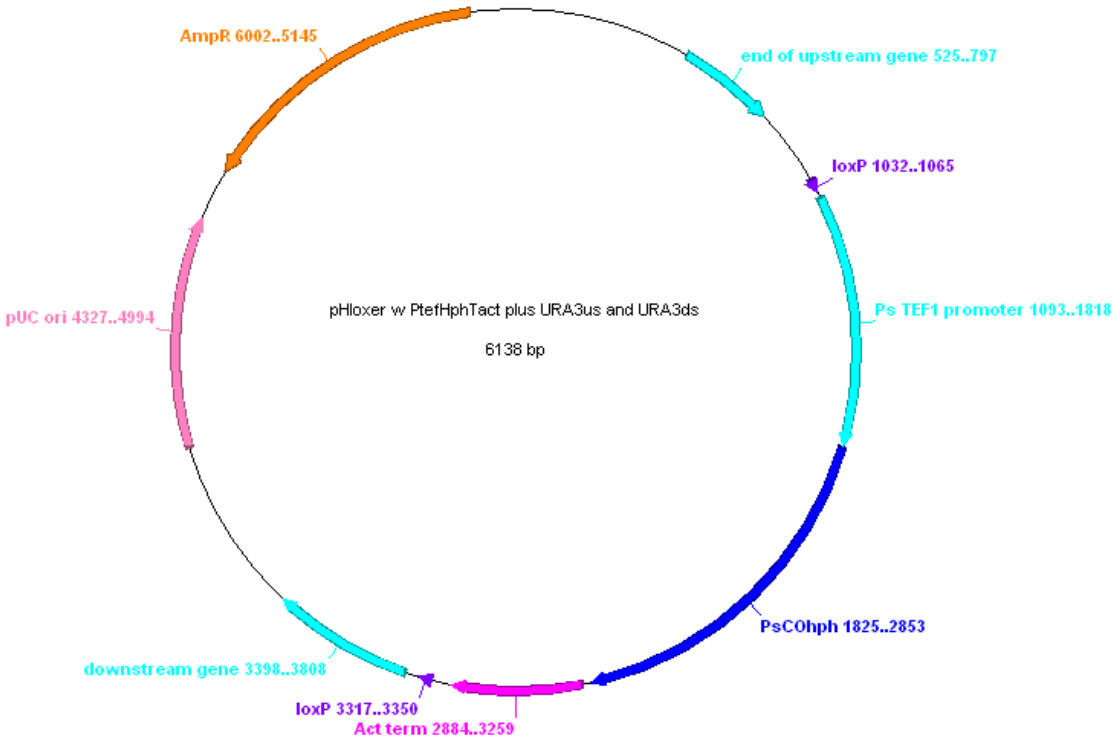


Images of *S. stipitis* with four different promoters controlling the regulation of *CBG99*, and the bioluminescence emitted under each sugar environment (**A**) Bioluminescence controlled P_{POL1}, P_{XYL1}, P_{GAL1}, and P_{TDH3} when glucose is the primary carbon source. (**B**) Bright field image of the glucose plate, showing where colonies were growing on the plate. (**C**) Bioluminescence controlled by all four promoters growing on a plate containing primarily xylose. (**D**) Bright field image of the xylose plate. (**E**) Bioluminescence controlled by all four promoters on a plate with primarily galactose. (**F**) Bright field image of the galactose plate.



Images of S. stipitis transformed with pWR10 (pAllet (PTEF1-coSh Ble -TACT1)(PTDH3-coCBG99-TADH1) (ARS2)) and patched out on Yeast Peptone Dextrose plates containing 25 µl of zeocin and 25 µl of LH2. (A) Bioluminescence monitored using a 5 minute exposure time with the ChemiDoc™ MP system. (B) Bright field image of the plate shown in panel A.

Plasmid map of pWR34. Plasmid contains a codon optimized zeocin resistance marker, a codon optimized CBG99, and the gene CRE recombinase which is inducible by PGAL1. Plasmid was built using pWR11 as a template. PGAL1 was PCR amplified using primers which added SbfI and FseI sites which were using to ligate the amplicon into pWR11. Codon optimized CRE was synthesized by GenScript. This gene was ligated using FseI and MluI in a sequential digest. The TADH2 was PCR amplified from wild type S. stipitis and ligated into the plasmid using MluI and SacII restriction sites in a sequential digest.



Plasmid map of pWR37. GenScript was unable to put the synthesized multiple cloning site with the LoxP sites into pUC57-mini, and instead put it in pUC57 which introduced problematic restriction sites. The pHloxer element was PCR amplified from that plasmid and cloned into pAllet using primers that added AgeI and SacI sites. PTEF1 – coHPH – TACT1 was PCR amplified from pWR5 with BamHI and EcoRI primers and ligated into pHloxer using those sites. The URA3 upstream and downstream elements were PCR amplified from wild type S. stipitis and added into pHloxer using XbaI and XhoI for the upstream region, and for the URA3 downstream region the sites BglII and PstI were used.