Optimizing Tyrosine Production in Yeast via *in vivo* Continuous Evolution

Presented by Su Min Cho

In partial fulfillment of the requirements for graduation with Dean's Scholars Honor's Degree in Biochemistry.

Dr. Hal Alper Supervising Professor	**Date**
Dr. Jeffrey E. Barrick Honor's Advisor for Biochemistry	**Date**

Optimizing Tyrosine Production in Yeast via *in vivo* Continuous Evolution

Department of Biochemistry

Department of Chemical Engineering

Su Min Cho **Date**

Dr. Hal Alper **Date**
Supervising Professor

Abstract

Tyrosine is an aromatic amino acid with a wide range of uses. Tyrosine is a precursor for numerous biological compounds ranging from neurotransmitters to hormones. It therefore plays a big role in the pharmaceutical industry. Tyrosine is also used by the chemical industry to make non-biological products such as polymers, adhesives, and semiconductors. Yeasts are unicellular eukaryotic organisms capable of producing tyrosine through the shikimate pathway. We seek to evolve specific genes in the Shikimate pathway – ARO2, ARO3, ARO4 – to increase yeast tyrosine production and improve efficiency. These genes are evolved through a method called *in vivo* continuous evolution, or ICE, which is a novel way to generate mutations using a retrotransposon. ICE allows for mutations to target genes instead of random mutations throughout the genome. We used ICE to induce mutations to the shikimate genes and produced a library of mutant yeast cells. We screened the cells for high tyrosine production and found four possible mutants. Three of the four strains could not be sequenced due to failure to isolate the mutant gene through PCR. The mutant gene from the AROp-G strain was successfully isolated and amplified. Sequencing showed that AROp-G strain was a mutant with a truncated C-terminus at the ARO4 gene. We inserted this mutant gene into a plasmid which was then transformed into wild-type yeast. There was a significant increase in tyrosine production, confirming that increased tyrosine production was due to the mutant gene. We will continue to make additional mutant libraries and find more mutant sequences.

Background

Cells construct proteins from various amino acids. Tyrosine is an aromatic amino acid with a hydroxyl functional group attached to the phenyl ring, as shown in figure 1. Tyrosine plays an important role in cellular and physiological processes. For example, tyrosine is a component of receptor kinases, which aid in signal transduction pathways and hence cell growth and differentiation (Schlessinger and Ullrich, 1992). There are reports that tyrosine supplement improve physical performance and increase depression and anxiety control (Deijen and Orlebeke, 1994). Tyrosine supplements are especially critical to phenylketonuria patients who cannot produce tyrosine naturally (Lutke-Eversloh, Santos, and Stephanopoulos, 2007). Tyrosine is a precursor not only to biological compounds such as neurotransmitters and hormones but also to non-biological compounds such as polymers and adhesives (Rasmussen et al., 1983; Qi et al. 2007; Sariaslani 2007; Vannelli et al. 2007). It can be used to synthesize melanin which can be used as drug carriers, UV absorbers, semiconductors, and even cation exchangers (Lutke-Eversloh, Santos, and Stephanopoulos, 2007; Bell and Wheeler 1986; della-Cioppa et al. 1990; Cabrera-Valladares et al. 2006). The wide scope and range of use for this compound therefore makes it ideal for mass production. Many pharmaceutical, plastic, and even electronics-based industries that use this compound could benefit from more effective tyrosine production.

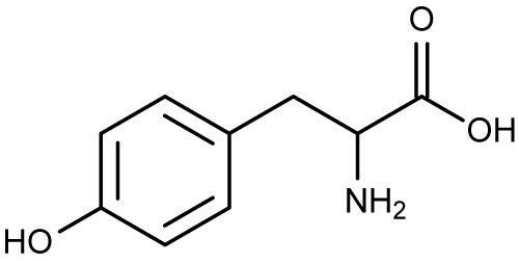

Figure 1. Molecular structure of tyrosine

Currently, two main approaches to mass tyrosine production are utilized: chemical synthesis and biological synthesis. The chemical synthesis of tyrosine utilizes substrates at

specific steps to build tyrosine. Reaction conditions need to be highly controlled to chemically synthesize pure tyrosine. On the other hand, biological synthesis of tyrosine uses utilizes microbes such as yeast to produce tyrosine. This method is less costly and more environmentally friendly compared to the traditional chemical approach (Lutke-Eversloh, Santos, and Stephanopoulos, 2007).

Microbes such as yeast can manufacture tyrosine. However, the host organisms' genes must first be engineered to not only synthesize tyrosine but to also do so in a manner that is both efficient and quick. We seek to increase tyrosine production by altering genes responsible for tyrosine production in yeasts. This is achieved through a method known as *in vivo* continuous evolution, or ICE, which mutates specific genes by utilizing native mobile retrotransposon elements. ICE utilizes yeast as the host organism. Yeast is a good model for ICE because yeasts have multiple well-characterized retroelements. It is also a single-celled eukaryotic organism, which is advantageous because of its low maintenance in cultivation, fast growth, and possession of complex RNA modifications such as intron splicing. Other advantages of using yeast include greater tolerance to low temperatures, easier separation, and lower contamination susceptibility (Curran et al., 2013).

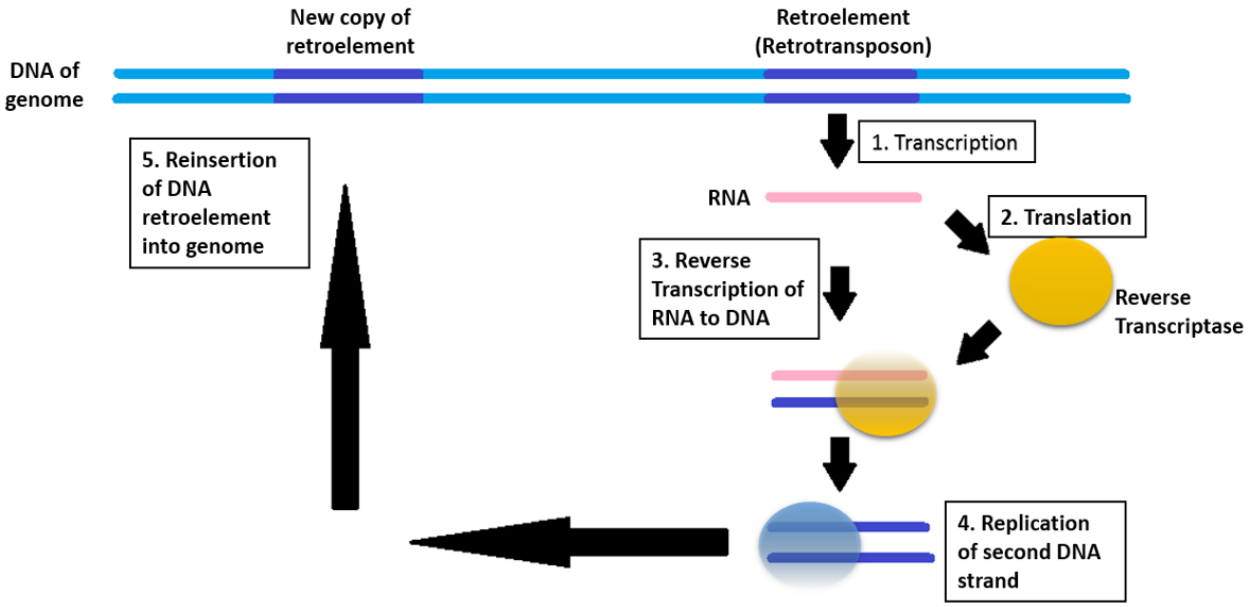

Figure 2. Operation of retroelement (retrotransposon). Su Min Cho, 2016. Adapted from Pearson Education, Inc.

In ICE, the target gene is inserted inside the ICE retroelement. A retroelement is a gene segment capable of copying itself and inserting the copy into new genomic locations (Craig et al., 2002). In the presence of galactose, the repressor that blocks transcription of the ICE retroelement is removed. The repressor removal allows for a series of transcriptions and reverse transcriptions in an error-prone process. The error inherent to reverse transcription is due to lack of proofreading by reverse transcriptase (Ricchetti and Buc, 1990). This is responsible for the mutations to the target gene within the retroelement. Reverse transcriptase in the retroelement is transcribed and translated along with the target gene. Reverse transcriptase then reverse-transcribes the RNA retroelement back into DNA to be re-inserted back into the genome, as indicated in figure 2. Ultimately, ICE produces a library of mutants in the presence of galactose.

The mutants with the desired phenotype is screened for, and the the mutated DNA is isolated and sequenced to confirm target gene mutation.

We seek to utilize ICE to evolve specific genes in the shikimate pathway, a biosynthesis pathway to increase tyrosine production in yeasts. These genes are ARO2, ARO3, and ARO4, which we collectively referred to as AROp. ICE is more advantageous than whole-genome random mutation techniques because in ICE a mutated gene be subsequently isolated and transferred to another yeast strain to exhibit the same phenotype. This is possible because only a single DNA region is mutated via ICE. It is difficult to determine which mutation is responsible for a desired phenotype if there are random mutations all over the genome. It is difficult to test whether the mutation exhibits the same phenotype when transferred to another yeast strain. ICE also enables much larger library sizes than traditional methods of single-gene mutagenesis. Through ICE, AROp in the tyrosine synthesis pathway will be evolved and inserted into a yeast strain rationally engineered to produce high tyrosine levels to further increase tyrosine production.

Materials and Methods

ICE production of mutant library

The AROp genes - ARO2, ARO3, and ARO4 - were inserted into the ICE retroelement in a plasmid. Each of the genes contained an intron which would later be used for additional screening. First, restriction enzymes were used to insert the AROp into the ICE retroelement ligated with a T4 ligase. The entire retroelement, along with an ampicillin marker gene upstream of the retroelement, was transformed into electrocompetent *E. coli* cells. Subsequent incubation

in LB + ampicillin agar plates screened for *E. coli* cells that took in the plasmid and subsequently the ICE retroelement and the marker gene.

After 24 hours of incubation at 37 °C, colonies were selected for further growth in LB + ampicillin liquid media. After cells reached an OD_{600} of 2, DNA from the cells was isolated and purified using the Promega miniprep kit (A1125). The DNA samples were sequenced to confirm successful AROp insertion. A confirmed sample was then transformed into wild-type yeast strain (By4741) to produce a new yeast strain AROp-WT. This was plated onto a CSM – URA plate (114511214 MP BIOMEDICALS) to screen to successful transformants. After colony growth, yeast colonies were picked and incubated in CSM – URA (114511212 MP BIOMEDICALS) liquid media at 30 °C for 24 hours. The yeast DNA was isolated and purified using the Zymoprep yeast miniprep kit (D2004). The purified DNA was then sequenced to ensure the presence of the ICE retroelement in the AROp-WT strains. AROp-WT was pipetted into a 2 mL CSM + galactose media for a final OD_{600} of 0.5. These cells were incubated at 22 °C for 48 hours. During this step, the retroelement was induced with galactose, initiating a cascade of transcription and reverse transcription to create a mutant library. The mutant library was transferred into a minimal media lacking aromatic amino acids to isolate the cells that exhibited at least one cycle of transcription and reverse transcription. The cells that transcribed and reverse-transcribed the retroelement at least once would have a fully functional gene coding for production of aromatic amino acids needed to grow in the media.

Mutant library screen via aptamer fluorescence

The mutant library was shipped to collaborators at the University of California at San Francisco to select mutant yeast cells that produced relatively high amounts of tyrosine. The

mutant library was exposed to aptamers that fluoresce when bound with tyrosine. Tyrosine-yield was measured by first encapsulating each individual yeast cells in an aqueous droplet surrounded by oil. As tyrosine was produced and secreted into the droplet media, the surrounding aptamers interacted with the produced tyrosine to fluoresce. Droplets with fluorescence above a certain threshold were individually sorted into media. These were the Sort 1 mutants. The entire library was sorted again to produce Sort 2 mutants.

Tyrosine Quantification Assay

Tyrosine quantification assays were performed on four yeast strains demonstrating highest fluorescence. The methodology for this assay was derived from a research study by Lütke-Eversloh and Stephanopoulos (2007). The four strains were AROp-D, AROp-G, AROp-I, and AROp-J. In addition to these strains, the assay was performed on By4741, B44, ΔZWF1, By4741-low OD, and B44-low OD. 150 uL of each culture was removed in a sterile environment. Supernatant was collected after spinning down the 150 uL samples. The remaining cultures were returned to a 30 °C shaker. The following assay reagents were first prepared: 250 mL of 0.1% (w/v) 1-nitroso-2-naphthol in ethanol and 250 mL of 20% (v/v) nitric acid containing 0.5 g/l (w/v) $NaNO_2$. The reagents were mixed in 1:1 (v/v) ratio. 25 uL of each supernatant was mixed with 75 uL of water and 100 uL of mixed assay reagents. The 100 uL of mixed assay reagents were also mixed with 100 uL of 1 mM tyrosine, 0.5 mM tyrosine, 0.25 mM tyrosine, 0.125 mM tyrosine, 0.0625 mM tyrosine, and distilled water to generate a standard curve. The mixtures were incubated at 55°C for 45-60 minutes then moved to room temperature for 2-3 hours. Each sample was transferred to a 96-well plate to measure fluorescence at 485/590 nm. The supernatant fluorescence readings were compared to the standard curve to calculate

respective tyrosine production. Tyrosine quantification assays were performed on the yeast strains two more times on subsequent days. Mutant tyrosine production was compared to that of AROp-WT strain.

Intron excision confirmation

The mutant strains producing high tyrosine amounts were selected. DNA was isolated and purified using the Promega miniprep kit. The DNA was eluted in 30 uL of nucleotide-free (NF) water. A confirmation digest with AscI and PvuII restriction enzymes was performed to confirm intron excision. 1.0 uL of each miniprep was mixed with 7 uL of NF water, 1 uL of Cutsmart buffer, and 0.5 uL each of AscI and PvauII. The mixture was incubated in 37°C for over 30 minutes. The samples were loaded onto a 0.5% gel for gel electrophoresis. The desired result was presence of double DNA bands. Samples with two DNA bands indicated presence of original gene and mutated gene. The mutated gene is shorter in length compared to the original gene. This is because the mutated gene lacks an intron as a result of re-insertion after transcription and reverse transcription. Samples with two DNA bands were selected.

Mutant gene isolation

A set of primers surrounding the AROp genes were used to amplify the mutant AROp genes from samples with confirmed intron excision. The PCR reaction mixture consisted of 36.5 uL of NF water, 10 uL of 5X Q5 buffer, 1 uL of dNTP, 0.5 uL each of reverse and forward primers, 1 uL of the sample DNAs, and 0.5 uL of Q5 Hotstart DNA polymerase. After PCR, lower MW bands were extracted, purified, and sequenced. Mutant genes were inserted into the genome of a yeast strain with ARO2, ARO3, and ARO4 triple knockouts. The triple knockout

strain cannot produce high levels of tyrosine. If the strain exhibited increased tyrosine production after mutant gene insertion, the mutant gene was confirmed to be responsible for increased tyrosine production.

Results

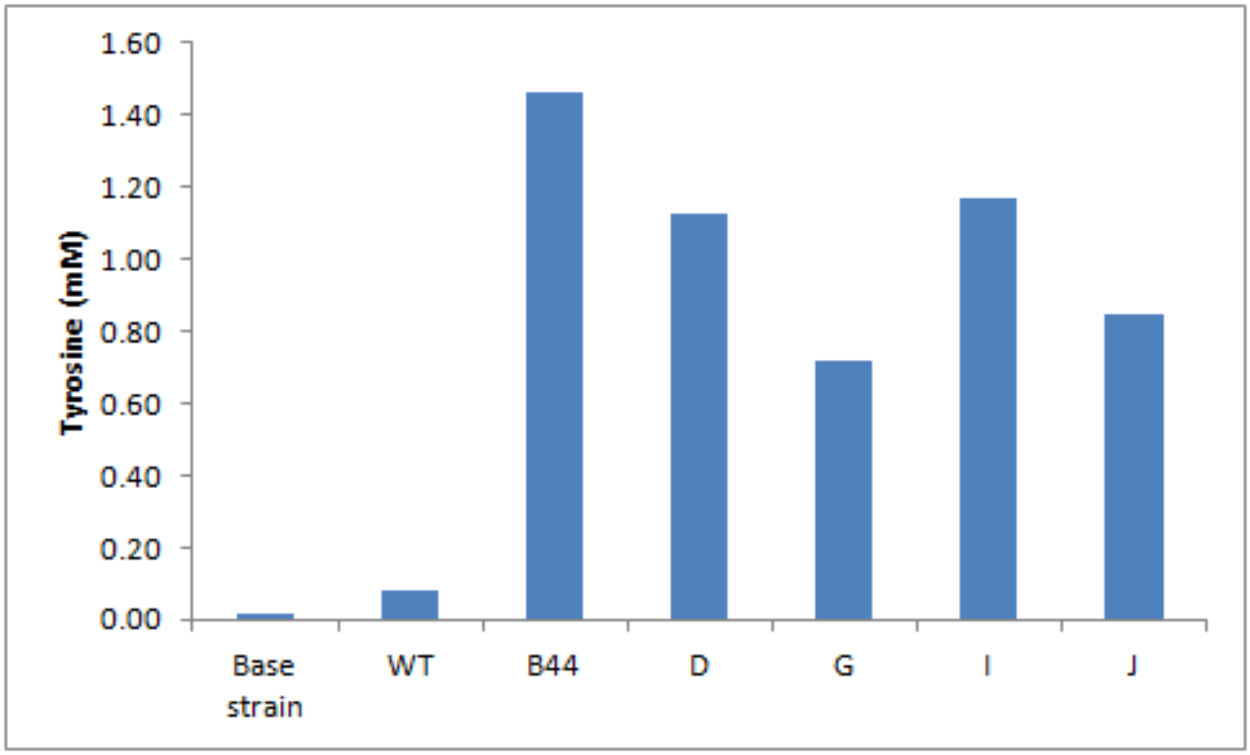

Figure 3. Sort 1 Tyrosine quantification assay results. D, G, I, and J represent the possible mutant strains AROp-D, AROp-G, AROp-I, and AROp-J respectively. Base strain is the yeast strain without the Shikimate genes. WT is the strain with the wild-type Shikimate genes, and B44 is the strain rationally engineered to produce more tyrosine. Successfully isolated mutant genes will eventually be inserted into B44 to further increase tyrosine production.

The tyrosine quantification assay results from the first library sort are shown in figure 3. The B44 strain, which is a yeast strain rationally engineered for high tyrosine production, demonstrated a potential for high tyrosine production. The possible mutant strains AROp-D, AROp-G, AROp-I and AROp-J produced significantly more tyrosine compared to the wild-type (WT) strain. The base strain is a yeast strain without the Shikimate AROp genes and therefore cannot produce any tyrosine.

A PCR reaction was performed for the AROp genes in AROp-D, AROp-G, AROp-I, and AROp-J. PCR results showed double bands for all four possible mutant yeast strains, thus confirming that a successful retrotransposition had taken place in all four strains. The double bands were due to presence of the original gene with the intron and the shorter transposed gene without the intron. However, sequencing results indicated that AROp-D, AROp-I, and AROp-J only had wild-type copies of the AROp genes. On the other hand, sequencing of AROp-G confirmed a mutation in AROp.

The mutant AROp-G gene was purified and then inserted into a plasmid. The plasmid was then transformed into a wild-type yeast strain. The resulting change in tyrosine production, indicated by tyrosine concentration (mM) is shown below in figure 4. AROp-G was once more sequenced to determine what mutation within AROp led to the increased tyrosine production.

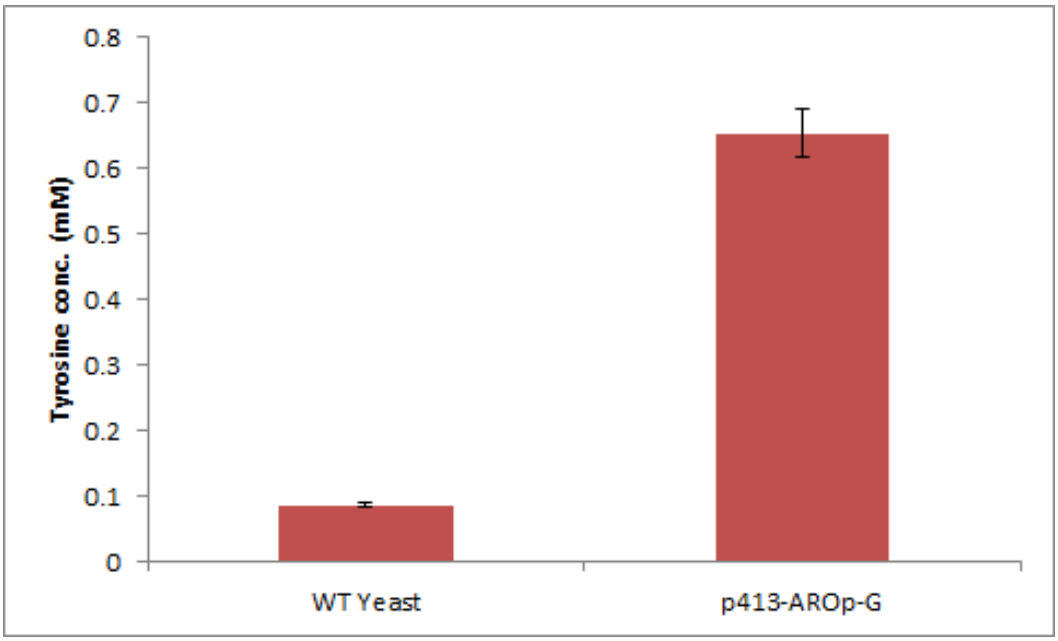

Figure 4. Comparison of tyrosine production between wild-type yeast and yeast strain with mutant gene. The mutant gene was inserted into a plasmid which was then transformed into a wild-type yeast strain. The yeast with the plasmid insertion is labeled as p412-AROp-G on this figure. Wilt-type yeast strain is labeled as WT yeast.

In order to find more mutant sequences, more yeast cells were screened from the mutant library. Eight mutant strains were screened from this Sort 2 library through the use of fluorescent aptamers. Tyrosine production of each of the screened strains were once again determined with tyrosine quantification assay. Tyrosine levels after a 24-hour incubation period is shown below in figure 5.

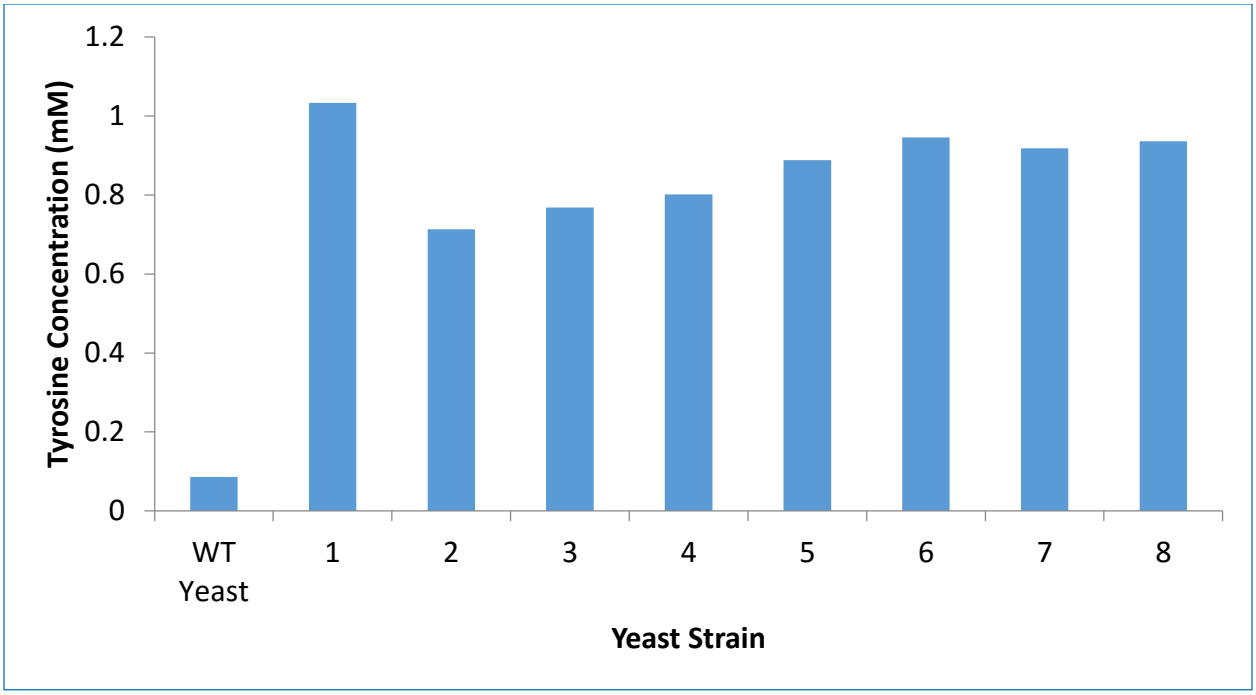

Figure 5. Sort 2 Tyrosine quantification assay results. WT is the wild-type yeast strain. Eight mutant strains were screened from the Sort 2 mutant library.

Discussion

The selected strains from the first sorted library were AROp-D, AROp-G, AROp-I, and AROp-J. All four of these mutants should have had higher tyrosine production than the wild-type based on fluorescent aptamers screening. Tyrosine quantification assay, which is a more reliable means of calculating tyrosine amount, confirmed that the four mutants produced significantly more tyrosine. The base strain served as the negative control. It produced almost no tyrosine, which was expected due to the lack of the Shikimate AROp genes. The wild-type yeast strain, which did have the AROp genes, produced more tyrosine than the base strain while producing less than the screened strains.

However, AROp-G was the only strain confirmed to have a mutation in the AROp genes. PCR and sequencing failed to indicate any mutations in the other three yeast strains. This shows that although retrotransposition did occur for these strains, mutations were perhaps not made during the reverse-transcription step of ICE. Strangely, these strains still produced significantly higher amounts of tyrosine compared to the wild-type yeast strain. Therefore, to conclusively confirm that no mutations had taken place in AROp for these strains, we sought to PCR only the transposed genes and not the original gene. We isolated the mutant gene from the original gene in the AROp-D strain by using the Gibson assembly method. This method used two separate PCR reactions and combined the two PCR products with Gibson assembly to beget the transposed gene without the intron. Gibson assembly is a method that joins multiple DNA strands into a single strand (Gibson et al., 2009). One PCR reaction amplified a part of the gene, while the other PCR reaction amplified the other part of the gene. We used primers that could only bind if introns had been excised. In other words, only the genes that went through transcription, reverse transcription, and reinsertion back into the genome could be amplified through this process. The two PCR products were combined together using Gibson assembly to isolate the transposed gene.

However, sequencing of this isolated gene revealed that it was a wild-type and not a mutant gene, which means that overall, ICE failed to produce a mutant gene in AROp-D, AROp-I, and AROp-J. During reverse transcription of the retroelement, a lack of error by the reverse transcriptase must have led to zero mutations to the target gene. It is very possible that the strains just had several copies of AROp (due to multiple reinsertions of the retroelement during ICE) instead of having one beneficial mutant. Multiple copies of AROp, with introns excised but no mutations present, could still lead to increased tyrosine production. Multiple AROp, and thus

more enzymes in the Shikimate pathway, could have increased general production of aromatic amino acids including tyrosine. Another possibility is that random mutations could have been made outside of our target gene that somehow led to increased tyrosine production for these three strains.

Only the mutant gene from AROp-G was successfully isolated and found to have genomic mutations. The AROp-G mutant gene was inserted into a plasmid which was then transformed into a wild-type yeast strain. In comparison to a wild-type yeast strain, there was almost a seven-fold increase in tyrosine production, as shown in figure 4. The increased tyrosine production confirmed that it was indeed solely due to mutations in AROp. Prior to this step, there were possibilities of increased tyrosine production being influenced by gene copy number. Multiple copies of AROp in AROp-G yeast strain could have been a factor in increased tyrosine production as was the case for the AROp-D yeast strain. However, by inserting a single mutant gene into wild-type yeast and observing for change in tyrosine production, we removed the effects of gene copy number in tyrosine production.

Sequencing revealed that AROp-G strain was a mutant with a truncated C-terminus at the ARO4 gene. The mutation truncated the last 13 amino acids of ARO4. One possible mechanism is that this particular mutation eliminated feedback inhibition within the Shikimate pathway. This has been shown previously in a point mutation at amino acid 226 which is relatively close to the C-terminus (Hartmann et al., 2003).

The mutant strains screened from the Sort 2 library showed promising results. Compared to the wild-type, all mutant yeast strains showed significantly higher tyrosine production. However, just as we did in Sort 1, we will need to find the mutant sequence and insert the mutant

gene into wild-type yeast in order to confirm that a mutation in the Shikimate genes led to the increased tyrosine production.

The next step in this experiment is continuing to find more mutants with elevated tyrosine production. So far, we plan to do so by performing ICE on the AROp-G mutant yeast strain to produce a new second-round library of mutants. The entire process will be repeated again to search for mutants with the highest tyrosine production. After multiple cycles of this process, the mutant AROp genes will eventually be inserted into the B44 yeast strain. As mentioned before, B44 is a yeast strain rationally engineered to produce more tyrosine. This was accomplished by knocking out certain genes while enhancing other genes to divert more resources into tyrosine production. We seek to further increase tyrosine production by inserting a high-performance mutant gene into B44. We will also isolate the mutant enzyme from the AROp-G mutant and measure its activity *in vitro*. We will purify the protein to measure enzyme kinetics.

Future experiments will involve using ICE to produce other aromatic amino acids or compounds of interest. The shikimate pathway also produces other aromatic amino acids in addition to tyrosine such as phenylalanine and tryptophan. Tyrosine production was in essence a test-run of using ICE to produce a desired aromatic amino acid, and the knowledge gained from this experiment will be used to increase production of these other compounds.

Acknowledgements

I like to thank Dr. Alper for allowing me to work in his lab and Joe Abatemarco for teaching and supervising me. I like to thank Dr. Barrick for reviewing this thesis and also advising me throughout my undergraduate years. I also like to thank the Dean's Scholars Honors Program, for challenging me to be involved with research as early as my freshman year and for teaching me to think like a scientist. Finally, I would like to thank my ACTS family for keeping me accountable all these years.

References

Bell, A. A., Wheeler, M. H. (1986). Biosynthesis and functions of fungal melanins. Annu Rev Phytopathol 24:411–451

Cabrera-Valladares, N., Martínez, A., Pinero S., Lagunas-Muñoz, V. H., Tinoco, R., de Anda, R, Vazquez-Duhalt, R, Bolívar, F, Gosset, G. (2006). Expression of the melA gene from Rhizobium etli CFN42 in Escherichia coli and characterization of the encoded tyrosinase. Enzyme Microb Technol 38:772–779

Craig, N. L., Craigie, R., Gellert, M., Lambowitz, A. M. (2002). Mobile DNA II (American Society for Microbiology, Washington)

Curran, K., Leavitt, J., Karim, A., Alper, H. (2013). *Metabolic Engineering of Muconic Acid Production in Saccharomyces cerevisiae*. *Metabolic Engineering* 15(1), 55-66

Curran, K. A., Alper, H.S. (2012). Expanding the chemical palate of cells by combining systems biology and metabolic engineering. Metab. Eng 14:289–297

Deijen, J. B., Orlebeke, J. F. (1994) Effect of tyrosine on cognitive function and blood pressure

under stress. Brain Res Bull 33:319–23

della-Cioppa, G., Garger, S. J., Sverlow, G.G., Turpen, T. H., Grill, L. K. (1990). Melanin production in Escherichia coli from a cloned tyrosinase gene. Biotechnology 8:634–638

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. 3rd, Smith, H. O. (2009). "Enzymatic assembly of DNA molecules up to several hundred kilobases". *Nature Methods* 6 (5): 343–345

Hartmann, M., et al. (2003). Evolution of feedback-inhibited β/α barrel isoenzymes by gene duplication and a single mutation. Proceedings of the National Academy of Sciences 100(3):862-867

Herrmann, K. M. (1995). The shikimate pathway as an entry to aromatic secondary metabolism. Plant Physiol. 107, 7-12

Lütke-Eversloh, T., G. Stephanopoulos. (2007). A semi-quantitative high-throughput screening method for microbial l-tyrosine production in microliter plates. Journal of Industrial Microbiology & Biotechnology 34(12):807-811

Lutke-Eversloh, T., Santos, C. N., Stephanopoulos G. (2007). Perspectives of biotechnological production of L-tyrosine and its applications. Appl microbiol biotechnol. 77(4):751-62

Qi, W. W., Vannelli, T., Breinig, S., Ben-Bassat, A., Gatenby, A. A., Haynie, S. L., Sariaslani, F. S. (2007). Functional expression of prokaryotic and eukaryotic genes in Escherichia coli for conversion of glucose to p-hydroxystyrene. Metab Eng 9:268–76

Rasmussen, D. D., Ishizuka, B., Quigley, M. E., Yen, S. S. (1983). Effects of tyrosine and tryptophan ingestion on plasma catecholamine and 3,4-dihydroxyphenylacetic acid concentrations. J Clin Endocrinal Metab 57(4):760-3

Ricchetti, M., Buc, H. (1990). Reverse transcriptases and genomic variability: the accuracy of

DNA replication is enzyme specific and sequence dependent. The EMBO Journal. 9(5):1583-93

Rohr, F. J., Lobbregt, D., Levy, H. L. (1998). Tyrosine supplementation in the treatment of maternal phenylketonuria. Am J Clin Nutr 67:473–476

Sariaslani, F. S. (2007) Development of a combined biological and chemical process for production of industrial aromatics from renewable resources. Annu Rev Microbiol 61:51–69

Schlessinger, J., Ullrich, A. (1992).Growth factor signaling by receptor tyrosine kinases. Neuron 9:383-91

Vannelli, T., Qi, W. W., Sweigard, J., Gatenby, A. A., Sariaslani, F. S. (2007). Production of p-hydroxycinnamic acid from glucose in Saccharomyces cerevisiae and Escherichia coli by expression of heterologous genes from plants and fungi. Metab Engin 9: 142–151