METABOLIC ENGINEERING OF *Saccharomyces cerevisiae* FOR THE PRODUCTION OF TAXA 4, 5-11, 12 DIENE

Master's Thesis

Farah Dagane GARAAD

Eskişehir, 2019

METABOLIC ENGINEERING OF *Saccharomyces cerevisiae* FOR THE PRODUCTION OF TAXA 4, 5-11, 12 DIENE

Farah Dagane GARAAD

Master of Science Thesis

Department of Advanced Technologies

Program in Biotechnology

Supervisor: Asst. Prof. Dr. Hülya Karaca GENÇER

Eskişehir Anadolu University

Graduate School of Sciences

MAY 2019

FINAL APPROVAL FOR THESIS

This thesis titled "Metabolic Engineering of *Saccharomyces cerevisiae* for The Production of Taxa 4, 5-11, 12 diene" has been prepared and submitted by Farah Dagane GARAAD in partial fulfillment of the requirements in "Anadolu University Directive on Graduate Education and Examination" for the Degree of Master of Science in Department has been examined and approved on 30/05/2019

Committee Members	**Title, Name and Surname**	**Signature**
Member (Supervisor)	: Asst. Prof. Dr. Hülya Karaca GENÇER	.....
Member	: Assoc Prof. Dr. Mehmet Burçin MUTLU	.....
Member	: Prof. Dr. Ahmet ÇABUK	.....

.....
(Title, Name and SURNAME)
Director of Graduate School of Sciences

ÖZET

Yüksek Lisans Tezi

TAXA 4, 5-11, 12 DIENE'NİN *Saccharomyces cerevisiae'da* METABOLİZMA MÜHENDİSLİĞİ İLE ÜRETİLMESİ

Farah Dagane GARAAD

Anadolu Üniversitesi, Fen Bilimleri Enstitüsü

MAYIS 2019

Danışman: Dr. Öğr. Üyesi Hülya KARACA GENÇER

Birçok ilaç ve değerli kimyasal maddesi çoğunlukla bitkilerden elde edilmektedir. Bu değerli kimyasallar arasında taksol; değeri yüksek bir antikanser ilaç olan bir terpenoid olup *Taxus sp.* bitkisinin kabuklarından elde edilmektedir. *Taxus sp.*'nden taxolün ekstraksiyonu oldukça maliyetlidir ve bu durum ilacın maliyetini de arttırmaktadır. Bu problemi çözmek için başvurulan yöntemlerden biri "metabolizma mühendisliği"dir ve bu yöntemle yeni kaynaklardan taksolü üretebilmek için S*accharomyces cerevisiae* ve *Escherichia coli* başta olmak üzere farklı organizmalar kullanılmaktadır. S. *cerevisiae* ve *E. coli* genetik manipülasyona açık olup bu organizmalarla ilgili çok sayıda genetik kaynak bulunması nedeni ile tercih edilmektedir. Ancak taksolün biyosentezi çok basamaklı bir işlem olduğundan ve tüm basamakları tam olarak aydınlatılamadığından taksolün alternatif bir kaynakta üretilebilmesi için öncelikle öncüsü olan taxa 4, 5-11, 12 dienen (taksadien)'in üretimi hedeflenmektedir. Bu çalışmada yüksek miktarda taksadien üretebilen *S. cerevisiae* suşu oluşturmak amacıyla taksadien'nin öncüleri olan Acetyl KoA, Malonyl KoA, Acetocetyl KoA, Mevalonate, ve Geraniol Geraniol Difosfat kodlayan heterolog genler *S. cerevisiae* genomuna entegre edilmiştir. Bu genlerin ifadeleri *TEF1* ve *PGK1* promotorları kullanılarak arttırılmıştır. Bu metabolik değişiklikleri içeren *S. cerevisiae* suşlarının standart şuştan 25 kat daha fazla taksadiene üretebildiği görülmüştür.

Anahtar Sözcükler: Metabolizma mühendisliği, *Saccharomyces cerevisiae*, Taksadien, Taksol, Asetil KoA.

ABSTRACT

Master of Science Thesis

METABOLIC ENGINEERING OF *Saccharomyces cerevisiae* FOR THE PRODUCTION OF TAXA 4, 5-11, 12 DIENE

Farah Dagane GARAAD

Department of Advanced Technologies

Programme in Biotechnology

Graduate School of Sciences, Anadolu University

MAY 2019

Supervisor: Asst. Prof. Dr. Hülya KARACA GENÇER

Production of many drugs and valuable chemicals is mainly obtained from plants. Among these valuable chemicals is taxol; a terpenoid which is an anticancer drug of high value but with low supply due to the low abundance in its natural host; *Taxus sp*. which makes it very difficult and extremely expensive to extract. Metabolic engineering is one of the techniques that is being researched to solve this problem through the production of the taxol in alternative hosts such as *Saccharomyces cerevisiae* and *Escherichia coli* which are easier hosts to modify due to the wealth of knowledge and tools available to genetically modify them. However, before the production of the taxol in the *Saccharomyces cerevisiae* it is important to produce its first known precursor; taxa 4, 5-11, 12 diene (taxadiene), abundant production of which can later be used to produce taxol. In this study the precursors to taxadiene were overproduced in *S. cerevisiae* by boosting the mevalonate precursors; acetyl CoA, malonyl CoA & aceto-acetyl CoA, and mevalonate through the integrated constructs which were expressed using strong constitutive promoters *TEF1* and *PGK1* . The performance of the *S. cerevisiae* strains containing these constructs was tested with plasmids that had been used successfully to produce taxadiene in earlier research and significant increase of taxadiene of almost 25-fold were recorded.

Key words: Metabolic engineering. *Saccharomyces cerevisiae,* Taxadiene, Taxol, Acetyl CoA, Precursor

ACKNOWLEDGEMENTS

I would like to express my unending gratitude to my professor, mentor and supervisor: Asst. Prof. Dr. Hülya KARACA GENÇER who introduced me to the world of metabolic engineering and systems biology, and who has guided me skillfully, and patiently for the past two years through the thesis. She has been inspirational in sharing her experiences and insights always, and her love for her career was the most important support we needed when things got tough in the lab as they most inadvertently do. I take this opportunity to thank you on behalf of all my colleagues in the molecular biology lab for all your support and love. Thank you.

I would also like to thank Mr. Serkan Dereli who made my transition into the molecular biology lab easier and more fun than I could ever imagine. Thank you for the patience you showed in explaining the most complicated protocols with the utmost clarity.

I want to also thank my laboratory colleagues who were very supportive in ensuring our wonderful collaboration to finish our theses successfully. In the same breath, I wish to thank Ömer Abi for his love and care, his tea was a lifesaver, always.

Finally, I would like to thank my family for supporting me all the way to enable me to pursue my academic dreams.

STATEMENT OF COMPLIANCE WITH ETHICAL PRINCIPLES AND RULES

I hereby truthfully declare that this thesis is an original work prepared by me; that I have behaved in accordance with the scientific ethical principles and rules throughout the stages of preparation, data collection, analysis and presentation of my work; that I have cited the sources of all the data and information that could be obtained within the scope of this study, and included these sources in the references section; and that this study has been scanned for plagiarism with "scientific plagiarism detection program" used by Anadolu University, and that "it does not have any plagiarism" whatsoever. I also declare that, if a case contrary to my declaration is detected in my work at any time,
I hereby express my consent to all the ethical and legal consequences that are involved.

.....
(Signature)

.....
(Farah Dagane GARAAD)
10/06/2019

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SYMBOLS AND ABBREVIATIONS;

ALD : Acetaldehyde Dehydrogenase

A-ALD : Acetylating Acetaldehyde Dehydrogenase

ACC : Acetyl CoA Carboxylase

ACL : ATP Citrate Lyase

ACS : Acetyl CoA Synthase

DMAPP : Dimethylallyl Pyrophosphate

FPP : Farnesyl Diphosphate

GGPPS : Geranyl geranyl Diphosphate Synthase

HMG-CoA : 3-hydroxy-3-methylglutaryl-CoA

IPP : Isopentenyl Pyrophosphate

MEP pathway : Methylerythritol 4-phosphate/2-C-methyl-D-erythritol 4-phosphate

PGK : Phosphoglycerate Kinase

TEF : Translation Elongation Factor

tHMG1 : 3-hydroxy-3-methylglutaryl-CoA reductase

TS : Taxadiene Synthase

1. INTRODUCTION

Production of many drugs and valuable chemicals is mainly achieved from plants. Plants secrete these chemicals in small amounts that are only sufficient for their needs. Extraction from these plants is difficult and very expensive thus necessitating the utilization of other methods to produce these valuable chemicals (Atanasov et al., 2015).

The alternative must have the following characteristics: the host has to be a stable, easily modifiable host that multiplies rapidly, is able to be fed on cheap raw materials to give the valuable chemical which should be able to be extracted in an inexpensive process (Y. Chen, Daviet, Schalk, Siewers, & Nielsen, 2013).

The need for the valuable chemicals against the low supply from the plant hosts necessitates the development of different technologies for improving the supply. Chief among these technologies is genetic engineering which involves transferring a gene responsible for the production of the desired product to a new host that has other qualities desirable to the genetic engineer (Woolston, Edgar, & Stephanopoulos, 2013).

Production of the new chemical in a heterologous organism requires the new host's production capacity to compete with existing plant extraction techniques or chemical synthesis of the chemical which might be undesirable in order to be a commercially attractive alternative (Hong & Nielsen, 2012).

This calls for a more holistic view at the production of the valuable chemical from a pathway perspective; that is not the genes that code for the final parts of the chemical but rather at the effect of each gene on the pathway and making modifications based on the more systemic context. Again, isolating product formation to isolated pathways does not provide the full picture and this calls for a more holistic view at the production of the valuable chemical from interconnected pathways or rather a metabolic network perspective (Bailey, 1991; Ostergaard, Olsson, & Nielsen, 2000).

Newer omics studies such as the metabolomics, fluxomics, genomics, proteomics, transcriptomics etc. view the organism in a more holistic manner and demonstrate that the

wider overview is a more instructive way of solving the production issue (Nevoigt, 2008; Roldão, Kim, & Nielsen, 2012).

Metabolic engineering also known as pathway engineering is borne out of this need for the broad overview of the interconnected pathways leading to the desired valuable chemical and the need for improved production levels in the new host based on genetic improvements at different nodes of the pathway(s) simultaneously (Ostergaard et al., 2000).

The earliest definition of metabolic engineering was given in a paper by James E. Bailey (1991), as "*The improvement of cellular activities by manipulation of enzymatic, transport, and regulatory functions of the cell with the use of recombinant DNA technology*". At the time of publishing Bailey's paper, the terms metabolic engineering, synthetic biology, systems biology or pathway engineering were not common. Scientists who were working on research that would today be described as metabolic engineering/pathway engineering simply described their research as 'molecular breeding' (Bailey, 1991).

A similar definition is given by Stephanopoulos et al., (1998) as "*The directed improvement of product formation or cellular properties through the modification of specific biochemical reactions or the introduction of new genes with the use of recombinant DNA technology*"(Stephanopoulos, Aristidou, & Nielsen, 1998).

This gave a more specific focus to the earlier general definition given by Bailey and narrowed it down to specific modifications in known biochemical pathways to achieve a desired specific goal. These objectives could be (Bailey, 1991; Hong & Nielsen, 2012; Ostergaard et al., 2000):

a) Synthesis of product in an organism that does not naturally produce it through expression of heterologous genes/proteins.
b) Improvement of specific product formation (titer, volume) through redirecting of specific metabolites to a desired direction.
c) Creating more robust industrial organisms, tolerant to harsh industrial conditions (high temperature, pH, low oxygen, low nutrient availability etc.)

d) Manipulating the carbon source of organisms to give the organisms ability to utilize cheaper, and more readily available carbon sources (in some cases waste).

1.1. Aim of The Study

The study aimed to boost the precursors to taxadiene by boosting the mevalonate precursors; acetyl CoA through expression of the heterologous *A-ALD* gene, malonyl-CoA through the expression of the heterologous *ACC1* gene and acetoacetyl-CoA through expression of the heterologous *NPHT7* gene, mevalonate itself through the overexpression of HMG-CoA reductase through the integrated constructs. These were aimed to be expressed using strong constitutive promoters *TEF1* and *PGK1*. The performance of these strains containing these constructs was to be tested with plasmids that had been used successfully to produce taxadiene in earlier research so as to measure the increase of taxadiene recorded after the metabolic engineering of the precursor availability.

1.2. The Metabolic Engineering Cycle and Limitations

The metabolic engineering research process generally flows through the steps shown in **Figure 1.1**. The process usually starts with an intense study of the metabolic process and organism to be modified (learn). This is followed by the design of the tools (genetic, enzymatic, etc.) to be used for a specific host (design).

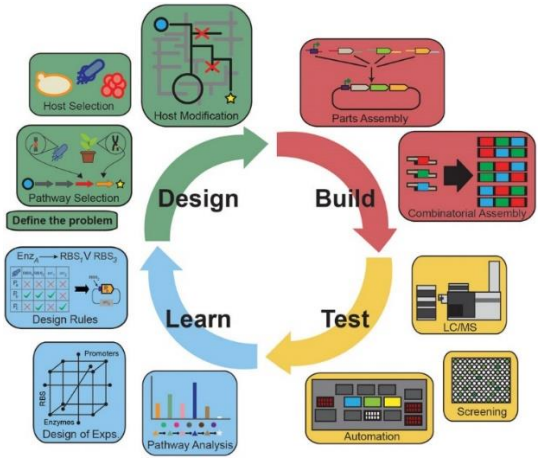

Figure 1.1: *The metabolic engineering cycle*

After the design phase is complete, the next phase involves the assembly of the required tools and modifications (build). The last phase is characterized by execution of the genetic modifications (Petzold, Chan, Nhan, & Adams, 2015).

Following a perturbation in a metabolic pathway after a genetic change, (design, build and test phase) leads to appearance of new phenotypes, new products. This warrants a meticulous recording of the experimental data to plan areas of improvement, and this is the learn phase of the cycle. The significance of this phase is such that it provides the foundation for future studies to base feasible strategies on (Petzold et al., 2015).

The appearance of the desired product also comes with a host of new challenges in; production (titer, yield, and rate), appearance of unwanted (often toxic) products, product toxicity, improper folding of the proteins, proteolysis of the heterologous protein, and most commonly metabolic burden on the host cell (Bailey, 1991; Petzold et al., 2015).

1.3. *Saccharomyces cerevisiae*; Choice as A Microbial Host

1.3.1. Biology/Physiology

Known as the baker's yeast, distiller's yeast, brewer's yeast, wine yeast etc.; *S. cerevisiae* is a yeast of the phylum Ascomycota, and genus Saccharomyces. It has wide applications as an industrial organism in the production of beer, wine, bread, pharmaceuticals and nutraceuticals (Nevoigt, 2008; Ostergaard et al., 2000).

In addition, *S. cerevisiae* has been used extensively for the industrial production of heterologous proteins such as Human Interferon, Hepatitis B Surface Antigens, Human Growth Hormone, Insulin, etc. This is enabled by the ability of the eukaryote to undertake post translational modifications on the proteins – not much unlike mammalian cells. In comparison to mammalian cells, the faster growth rate, abilities and tools to easily modify its genome, and highly developed fermentation processes makes it an even more attractive host in the production of some proteins (Hong & Nielsen, 2012; Paddon et al., 2013).

In industrial applications that require robust organisms, whose biochemistry, physiology, and genetics are well characterized; *S.s cerevisiae* and *E. coli* are unmatched. Furthermore, *S. cerevisiae* being the first eukaryote to be fully sequenced, made the rapid development of molecular tools easier (Goffeau et al., 1996; Woolston et al., 2013).

It is the preferred organism of choice in genetic engineering and more recently – metabolic engineering and synthetic biology, even over *E. coli* due to its; robustness, less risk to contamination through its general tolerance to harsh industrial conditions; low pH, and high ethanol concentrations (Y. Chen et al., 2013; Kocharin, Chen, Siewers, & Nielsen, 2012; Krivoruchko, Serrano-Amatriain, Chen, Siewers, & Nielsen, 2013; Lian, Si, Nair, & Zhao, 2014; Nielsen & Jewett, 2008).

1.3.2. Use as a host in different metabolic engineering applications

Table 1.1: *Production of Different Chemicals in Yeast. Native products are in Blue*

Chemicals Produced in Yeast Cell Factories	
Isoprenoids	Bisabolene, Patchoulol, Cubebol, Farnesol, Sesquiterpenes, Taxadiene, α-santalene, Sterols, Carotenoids, Casbene
Fatty Acids	Linolenic Acid, Stearidonic Acid, Arachionic Acid, Elcosapentaenoic Acid, Hexadecadienoic Acid
Alcohols	Ethanol**,** n-butanol, 1-butanol, Isobutanol, 2-methyl-1-butanol, 1-propanol, 1,2-propanediol, 3-methyl-1-butanol
Organic Acids	Lactic acid, Pyruvic acid, Formic acid, Acetic acid, Succinic acid, Malic acid, 3-hydroxypropionic acid, Fumaric acid**,** muconic acid, Citric acid, mevalonic acid
Others	Non-natural stilbenes, Lycopene, Cinnamoyl anthranilates, **B**-carotene, Polyhydroxyalkanoates, Resveratrol, Se-methyl selenocysteine, Glycerol, Artemisinic acid, non-natural dihydrochalcones, L-ascorbic acid, **B**-amyrin, Ribitol, Non-natural alkaloids

From **Table 1.1**, it is evident that *S. cerevisiae* has seen many applications in metabolic engineering. Excluding ethanol, most of the organic acids and glycerol; the rest of the

chemicals are non-native to the yeasts, with most of them naturally found in plants (M. Li & Borodina, 2015).

Most of these are however still in the optimization stages except for Artemisic acid, Isobutanol, Resveratrol and a few others which have achieved full commercial applications. The application in the industries can be started when the titers, rates and yields are such that they can compete with the preexisting production technologies – whether chemical or biological (Hong & Nielsen, 2012; Paddon et al., 2013).

1.4. Mevalonate Pathway Optimization: Upstream modifications to improve precursor supply

1.4.1. Acetyl CoA

Acetyl CoA is found in four different areas in the yeast cell; the cytosol, the mitochondrion, the nucleus, and the peroxisome (Krivoruchko, Zhang, Siewers, Chen, & Nielsen, 2015) **(Figure 1.2)**.

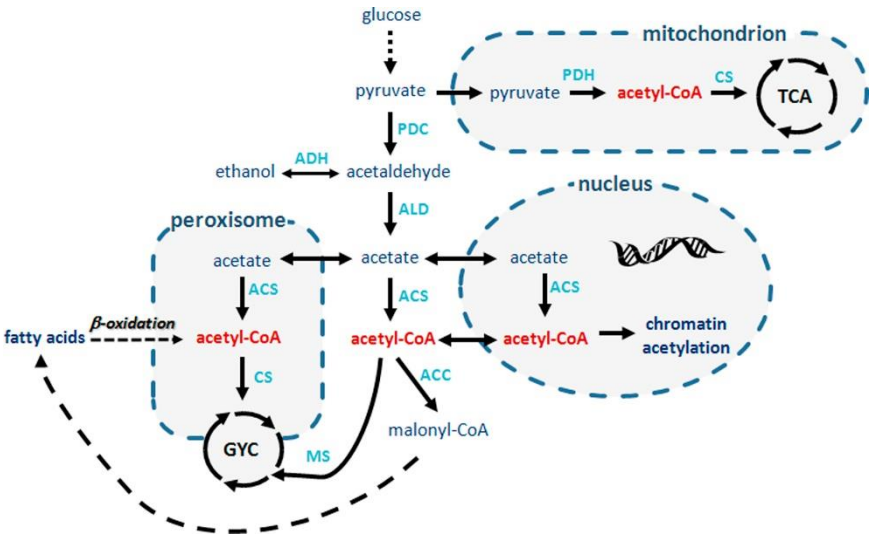

Figure 1.2: S. cerevisiae Acetyl CoA

Acetyl CoA is an important precursor that connects central glucose metabolism, fat metabolism and serves as a precursor to many secondary pathway metabolites. In *S.*

cerevisiae, the acetyl CoA is found in the cytosol, mitochondrion, peroxisome, and the nucleus. The acetyl CoA directs functions ranging from DNA acetylation, and precursor roles in the TCA and glyoxylate cycles (Krivoruchko et al., 2015).

S. cerevisiae being a eukaryotic organism, has a compartmentalized cellular system. This means that acetyl CoA in the different compartments is not available for use in the secondary metabolite pathways which start from the cytosol, with the different membranes impermeable to acetyl CoA (Y. Chen, Siewers, & Nielsen, 2012).

The acetyl CoA formed in the cytosol from acetaldehyde is usually limited and is again consumed in different biochemical reactions; converted to acetate, malonyl CoA, citrate etc. and transported. In pathway modification downstream of Acetyl CoA, a balance always exists for the consumption of Acetyl CoA; for vital primary cellular functions and for secondary metabolism (Krivoruchko et al., 2015).

Therefore, to realize high enough levels to be able to achieve sufficient product formation, it is imperative to improve Acetyl CoA supply, and this is done through implementing different strategies. Strategies entail either; improving production of Acetyl CoA by overexpressing genes involved in its production or; reduction of consumption of Acetyl CoA through deletion of competing pathways. Another strategy applied is a combination of both improving production while reducing consumption (Krivoruchko et al., 2015; Paddon et al., 2013; Rodriguez et al., 2016).

It is easy to note that most methods combine both because; overexpressing the downstream pathway enzymes could mean the disrupting of this balance, i.e. overexpression to a certain optimal level is acceptable, however overexpression beyond this level starts to create a burden for the cell metabolism. On the other hand, complete deletion of Acetyl CoA consuming pathways is also detrimental as the cell might not be able to perform activities vital to its survival and/or growth (Y. Chen, Siewers, et al., 2012).

The different strategies involve deleting routes consuming Acetyl CoA, such as glyoxylate cycle through deletions of genes coding for citrate synthase and malate synthase.

Other deletions might involve alcohol forming pathways e.g. alcohol dehydrogenase enzymes (*ADH*s) which are as a result of the "crab-tree effect" (Kocharin et al., 2012).

On the other hand, acetyl CoA improving is achieved through, as mentioned above, overexpressing the acetyl CoA formation genes in the yeast (*ALD, ACS, PDC* etc.) or heterologously introducing better performing enzymes (in terms of less energy consumption, lesser or dysfunctional feedback regulation, higher substrate to product conversion, etc.) (Y. Chen et al., 2013; Kocharin et al., 2012; Krivoruchko et al., 2013; Song et al., 2016).

Other genes concerned with acetyl CoA in different organisms with better qualities as elucidated before, but which undergo even a lesser number of steps exist, and these are introduced to *S. cerevisiae* through heterologous expression (Krivoruchko et al., 2015).

These genes include acetyl CoA lyase (*ACL*), Pyruvate Formate Lyase (*PFL*), acetylating acetaldehyde dehydrogenase (*A-ALD*), xylulose pathway, and they all proffer some advantages e.g. avoiding the use of ATP, finishing the reaction in a smaller number of steps and avoid producing some unwanted or toxic side products e.g. acetate. All the strategies mentioned above, are either used alone or in different combinations, with varying success (Kozak et al., 2014; Rodriguez et al., 2016; Song et al., 2016).

Table 1.2 summarizes some of the research conducted in the endeavor of improving acetyl CoA, with the improved levels of acetyl CoA tested on a downstream product.

***Table 1.2**: Metabolic engineering strategies to improve Acetyl CoA*

Product of Interest and final product	**Overexpressed Genes/Deleted Genes**	**References**
Polyhydroxy butyrate 16-fold and reduced side products	*ADH2*, *ALD6, ERG10, ACS* (S.E), *phaA*, *phaB*, *phaC/ ΔMLS1* & *ΔCIT2*	(Kocharin et al., 2012)
α-santalene 8.29 mg/l or 4-fold	*ADH2, ALD6, ERG10, ACS* (S.E), *tHMG1, Sansyn, ΔMLS1* & *ΔCIT2*	(Y. Chen et al., 2013)

Butanol, 16.3 mg/l or 6-fold.	*ADH2, ALD6, ACS, ERG10*, Δ*MLS1* & Δ*CIT2*	(Krivoruchko et al., 2013)
>100mg/L n-butanol	Δ*ADH1/4*, *PDC1/5/6*, *ACS, ACL*, Δ*ADH1/4*, Δ*GPD1/2*, Δ*MLS1* & Δ*CIT2*	(Lian et al., 2014)
Lactic acid 142 g/L with yield of 0.89 g/g	Δ*ADH1*, *LDH*, Δ*GPD1*, Δ*CYB2*, Δ*PDC1*+Δ*ALD6*, *A-ALD* genes (*mhpF* and *eutE*)	(Song et al., 2016)

1.4.2. IPP & DMAPP

Production of isoprenoid compounds to a commercial level is contingent upon improvement of the universal isoprenoid precursors Isopentenyl Pyrophosphate (IPP) and DMAPP: Dimethylallyl Pyrophosphate (DMAPP) in both *E. coli* and *S. cerevisiae* (Y. Chen, Zhou, Siewers, & Nielsen, 2015).

Downstream of acetyl CoA, the universal isoprenoid precursors are formed from two pathways; the mevalonate pathway in eukaryotes (discussed above), and the non-mevalonate pathway in prokaryotes (also known as the DXP or MEP pathway). Many strategies have been undertaken to improve the supply of the isoprenoid precursors through overexpression of the local enzymes or heterologous expression of specific enzymes (Y. Chen et al., 2015).

Other strategies attempt to heterologously transfer a whole pathway to an organism that uses a different pathway. This is done to enable the cell to improve the supply of the isoprenoid precursor through utilizing both its normal pathway and the heterologous pathway. This has been done in both *S. cerevisiae*, and *E. coli* with very poor success rates and a host of problems. The problems ranged from toxic pathway intermediates, to difficulty in expressing specific proteins essential for the specific reactions (Carlsen et al., 2013; Kirby et al., 2016; Martin, Piteral, Withers, Newman, & Keasling, 2003; Maury et al., 2008; Partow, Siewers, Daviet, Schalk, & Nielsen, 2012; Pitera, Paddon, Newman, & Keasling, 2007). The summary of the research carried out in this regard is presented in **Table 1.3**:

Table 1.3:Examples of Heterologous transfer of a whole pathway

Organism	Heterologous Pathway Expressed	Remarks/Issues Faced/Strategy
Escherichia coli	Mevalonate	Abundance of isoprenoid precursors caused the cells to cease growth or mutate to overcome this. Accumulation of HMG-CoA could be toxic to cells (Martin et al., 2003).
Escherichia coli	Mevalonate	A follow up to the above research by the same lab. Overexpression of the *tHMG1* to remove accumulation of the HMG CoA, and a 3fold increase in mevalonate production (Pitera et al., 2007).
*Saccharomyces cerevisiae**	MEP/Non-Mevalonate	Heterologous expression of MEP pathway in the yeast while blocking the normal mevalonate pathway using lovastatin (only cells carrying the MEP pathway genes could survive very high concentrations of lovastatin (Maury et al., 2008).
Saccharomyces cerevisiae	MEP/Non-Mevalonate	A repeat of the experiment above by the same team to reconstruct the MEP pathway in the yeast but this time using gene deletions (*ERG13*) to effectively repress the mevalonate pathway instead of lovastatin. No growth occurred in media lacking mevalonate and this was attributed to the lack of Fe-S cluster proteins which are very essential for the last two MEP reactions: lack of activity of *ISPG* and /or *ISPH*. Both enzymes dependent on NADPH and flavodoxin reductase redox system as electron donors to achieve their catalytic activity (Partow et al., 2012).
Saccharomyces cerevisiae	MEP/Non-Mevalonate	Used the same deletions (*ERG13*) to attenuate the mevalonate pathway and thus test the viability of the yeast cell with the MEP pathway alone. Came to the same conclusions as the team above but went ahead to functionally express a bacterial iron-sulfur protein in the cytosol of the yeast. While not entirely a success, the research concluded that the yeast cell has the capability to express at least some bacterial iron-sulfur proteins in the cytosol (Carlsen et al., 2013).

*In this research, the authors assumption that the mevalonate pathway was fully repressed by lovastatin was not accurate. The same lab and other labs devised a better way of fully

repressing the mevalonate pathway through the deletion of *ERG13*, and this gave different results. Other research along similar lines and achieving comparable results include (Kirby et al., 2016).

In summary, while complementing a natural pathway with a heterologous pathway is an attractive option to improve the supply of precursors, it has many challenges that require optimizing before being adopted as a routine laboratory or commercial strategy. Although the last research, (Carlsen et al., 2013) shows that it is indeed possible, it still has a long way before reaching a level where the whole MEP pathway can be successfully integrated into the yeast and used independent of mevalonate. Passing the final hurdle of functionalizing the final two enzymes (encoding the final two steps of the MEP pathway with the Fe-S cluster protein requirement) could be complemented to the natural mevalonate pathway to boost supply of the isoprenoid precursors and their products consequently (Carlsen et al., 2013).

1.4.3. 3-hydroxy-3-methylglutaryl coenzyme A reductase

3-hydroxy-3-methylglutaryl coenzyme A or more commonly known as HMG-CoA reductase (*tHMG1*), is a rate limiting step and one of the most important enzymes of isoprenoid biosynthesis. Due to its impact on the production of isoprenoids, it has been a focus of a lot of metabolic engineering research (Y. Chen et al., 2015; Engels, Dahm, & Jennewein, 2008; Paddon et al., 2013).

In perhaps one of the most striking demonstrations of the significance of HMG-CoA reductase in isoprenoid production; continuous and systemic focus on HMG-CoA reductase led to an increase in amorphadiene titer from 0.5 g/L to 25 g/L. This was done through a series of experiments; use of a truncated HMG-CoA reductase, then changing to a HMG-CoA reductase from *S. aureus* and the integration of a novel fed batch fermentation process led to the highest titers of amorphadiene (Paddon et al., 2013).

Similar expression of a truncated version of yeast HMG-CoA reductase with *Taxus chinensis* Geranyl geranyl Diphosphate Synthase & Taxadiene Synthase resulted in a 50 % increase in Taxadiene production (Engels et al., 2008).

1.5. The Mevalonate Pathway and its Valuable Products

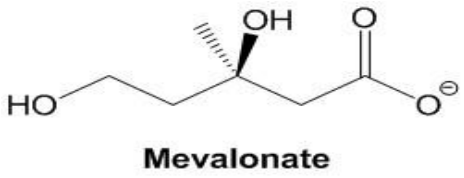

Figure 1.3: Mevalonate Structure (Dewick 2009)

Mevalonate or synonymously known as mevalonic acid is a building block in secondary metabolism to produce a wide array of secondary metabolites, mainly isoprenoids and steroids (Dewick, 2009). Isoprenoids are usually produced to carry out some important functions for the organism such as defense from predators, pollination, communication, light absorption, antioxidant functions, etc. Mevalonate is formed through a series of reactions, from three molecules of acetyl Coenzyme A through two intermediate compounds; Aceto-Acetyl CoA and 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA). The chemical structure of mevalonate is shown in **Figure 1.3** (Dewick, 2009).

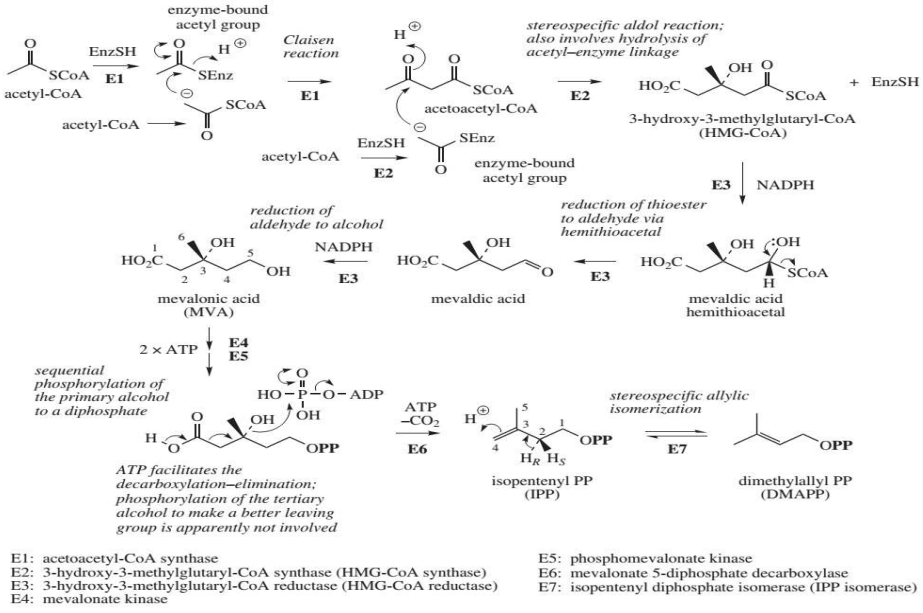

Figure 1.4: The Mevalonate Pathway (Dewick 2009)

The mevalonate pathway is found in eukaryotes, and some bacteria and was actually thought to be the only pathway for isoprenoid production until the recent discovery of the Methylerythritol 4-phosphate/2-C-methyl-D-erythritol 4-phosphate/MEP pathway or simply

known as the non-mevalonate pathway. The mevalonate pathway is the sole isoprenoid producer in animals, and in fungi. Whereas in plants it produces the required IPP in the cytosol while the MEP pathway produces the precursors in the plastids. In prokaryotes the pathway is found in archaea and few eubacteria such as *Staphylococcus aureus*, *Streptococcus pneumoniae* and *Borrelia burgdorferi*, (Boronat & Rodríguez-Concepción, 2015). **Figure 1.4** illustrates the mevalonate pathway adapted from (Dewick, 2009)

Its main role is to produce isoprenoid precursors for the production of other important terpenoids such as cholesterol and is thus the reason why it is targeted using drugs called statins to reduce cholesterol in patients with diseases such as diabetes mellitus through its inhibition. It is an important pathway intermediate (in isoprenoids and generally terpenoids), or a direct precursor to some chemicals (e.g. beta methyl valerolactone which is further transformed into a rubbery polymer in industry (Dewick, 2009).

1.5.1. A Product of the Mevalonate Pathway; Taxol

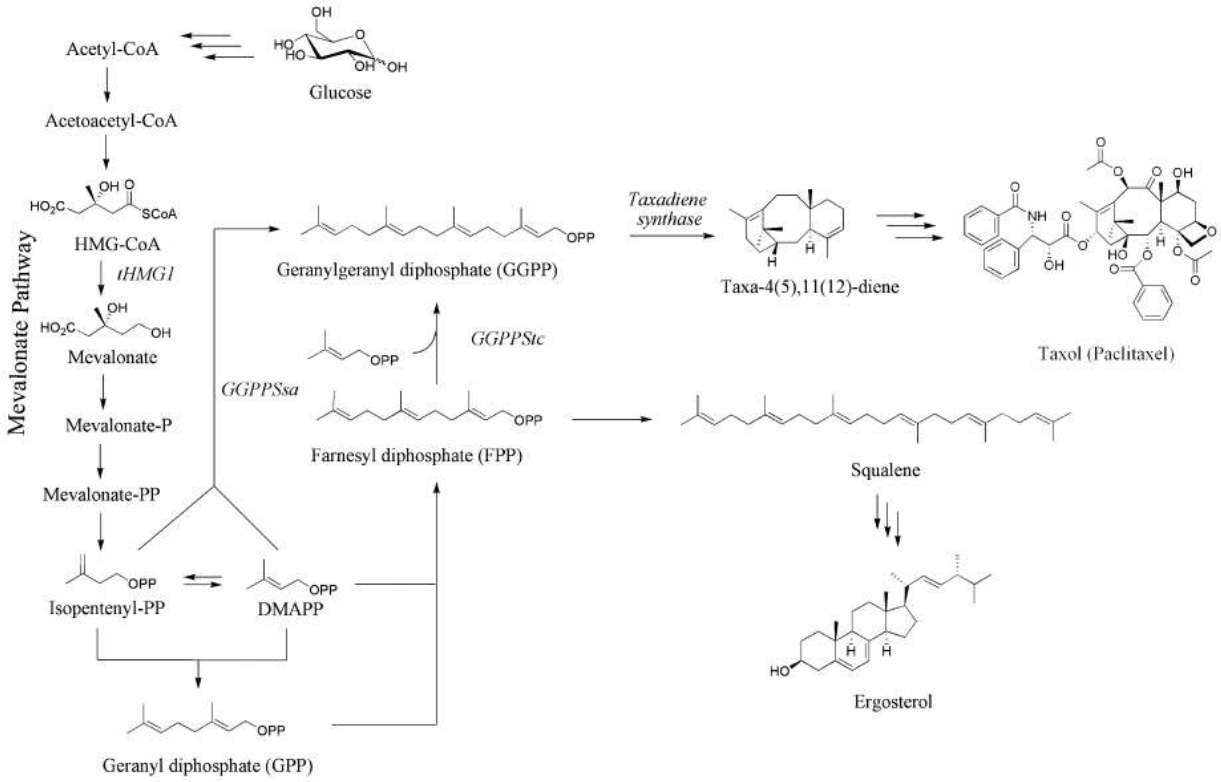

***Figure 1.5**: Taxol Biosynthetic Pathway (Dewick 2009)*

In a letter cited over 4390 times, the isolation and structural characterization of taxol was first reported by Wani and colleagues in 1971. The team extracted the compound from the bark of *Taxus brevifolia*, through a partitioning between water and chloroform, attaining a yield of 0.004 %. Taxol, a diterpene containing 20 carbons, is formed downstream of mevalonate through many intermediates; Isopentenyl Diphosphate and Dimethyl Allyl Diphosphate, Geranyl Diphosphate, Farnesyl Diphosphate, Geranylgeranyl Diphosphate, Taxadiene and finally to Taxol (Mansukhlal C Wani, Taylor, Wall, Coggon, & McPhail, 1971). The simplified diagram and it's pathway from acetyl CoA is shown in **Figure 1.5** (Engels et al., 2008).

The construction of the taxane skeleton is initiated through the cyclization of Geranyl geranyl diphosphate (GGPP) to Taxadiene by taxadiene synthase (*TS*). This is the first committed step towards taxol biosynthesis. The cyclization of Geranyl geranyl diphosphate

yields taxa-4(5),11(12)-diene 94%, taxa-4(20),11(12)-diene 5%, and Verticilene 1% (Croteau, Ketchum, Long, Kaspera, & Wildung, 2006).

Including the *GGPPS* synthase step, it takes around 19 enzymatic steps to complete the biosynthesis of taxol. Majority of the enzymes involved are cytochrome p450 enzymes; cytochrome p450 oxygenases and; cytochrome p450 hydroxylases -catalyzing a series of oxygenations and hydroxylations respectively (Jennewein, Wildung, Chau, Walker, & Croteau, 2004).

It is important to note that even though the Washington State University group led by Rodney Croteau have described most of the enzymatic steps in the biosynthesis of taxol, some enzymatic steps leading to baccatin III remain unknown (Croteau et al., 2006).

1.5.2. Clinical Use

Approval for use of Paclitaxel was initially granted in the treatment of ovarian and breast cancers, later on its activity was tested, and its use started in the treatment of non-small-cell lung cancer, pancreatic cancer, and Kaposi sarcoma associated with AIDS. It functions as an antimitotic by blocking cancer growth through stopping cell division and leading to cell death (Mansukh C Wani & Horwitz, 2014).

1.6. Taxol Production: The Choice of Metabolic Engineering Over Chemical Synthesis; Taxol as an example.

1.6.1. Total synthesis

Total synthesis of taxol, albeit very difficult was achieved in 1992 (Holton et al., 1994; Nicolaou et al., 1994), having remained a challenged for over 20 years since it was first isolated. The total synthesis developed by Holton et al., 1994; Nicolaou et al., 1994) had yields of 91% over 41 steps from (-) patchino, and 85% yield through 51 steps from butene diol respectively.

While this was a breakthrough in organic chemistry, and total synthesis, it was not and is still not a solution to the taxol supply. This is due to the complicated nature, the many steps required to get to taxol, and the insignificant yields from such a process (~0.4% at most), which make it commercially non-viable. This prompts the need to look for other solutions, and a more commercially viable solution when compared to total synthesis; the semi synthesis of taxol (Lee et al., 2014).

1.6.2. The Semi-Synthetic Approach: A Solution to The Difficulty in Supply

After discovering and successfully testing the efficacy of taxol to treat cancer, there resulted a huge ecological problem; The highest amounts of taxol could be harvested from the barks of the yew tree, and this led to the death of the plant. And for a single dose, bark needed to be harvested from between 2-4 fully grown trees. Added to this was the fact that the yew species is a slow growing species, and the extraction process lengthy and very expensive (Mansukh C Wani & Horwitz, 2014).

Therefore, in 1988, a semi-synthetic approach to taxol via 10-Deacetyl baccatin 111 was described by (Denis et al., 1988) the process had a lot of merit in that the 10-Deacetyl baccatin 111 was more abundantly available in the plants, and that it was found in higher amounts in the leaves. This meant that it could be harvested sustainably without leading to the death of the source plants (Denis et al., 1988).

This method though greatly refined is one of the main routes of commercial production of taxol today. What makes it even more attractive is the ability to extract the 10-DAB even from nursery cultivated yew trees. In pursuit of this, large farms were developed in China, Germany etc. where different Taxus species are planted and harvested every two months for their twigs and needles and 10 DAB extracted for a semi-synthetic conversion to taxol (Liu, Gong, & Zhu, 2016).

An alternative route to taxol through 7-β-Xylosyl-10-deacetyltaxol, found in the bark of the yew trees, was described by (Rao, 1997). This compound had been known to be a potential pre-cursor to taxol, but its conversion had not been achieved due to difficulty in

hydrolyzing the xylose residue while maintaining the stability of the compound. Removal of xylosyl moiety from 7- β-xylosyl-10-deacetyltaxol converts it into 10-deacetyltaxol which can be further converted to paclitaxel through acetylation (Rao, 1997).

1.6.3. Production of Taxol in Plants

The commercial production of taxol currently is fully done using plant systems, with slight variations in the techniques. The biotech firm Phyton Biotech produces Taxol commercially using plant cell-suspension cultures through huge bioreactors with a capacity of up to 75000 L (T.-K. Huang & McDonald, 2009).

Callus formation candidates are selected from young needle tissue -young needles that have been collected within two months have higher content of paclitaxel than older ones. Once the callus is established in the solid media, it is suspended in a liquid medium favorable for cell growth (Tabata, 2004).

Production of taxol and other related taxanes is promoted strongly by silver thiosulfate and methyl jasmonates – jasmonates are important in signal transduction processes to regulate genes related to defense. Endogenous levels of jasmonates increase with response to mechanical forces, wounding and pathogen attacks. Maximal induction of methyl jasmonate in baccatin III and taxol accumulation was achieved at the concentration of 100 mmol L and above, growth of the suspension cultures was inhibited by as much as 20-25% (Tabata, 2004).

The plant-based production efforts have been developed through the years, with the improvement in different parts of the operation; callus growth induction, induction of production of secondary metabolites including taxol, different types of media, improvement of bioreactor design etc (Y. Li, Zhang, & Pfeifer, 2015).

Table 1.4 adapted from (Y. Li et al., 2015) has a brief summary of the particular Taxus species used, the different engineering strategies utilized, the type of bioreactor if any and the final titer of Taxol.

***Table 1.4**: Taxol plant cell culture engineering strategies*

Plant cell species	Strategy	Titer
Taxus baccata	Calcium alginate beads immobilization, Methyl Jasmonate, silver thiosulfate as inducers	Maximum of 43.43mg/L on day 16
Taxus chinensis (and endophytes)	Modified separation membrane	25.63mg/L on day 15
T. chinensis	Medium exchange and enhancement agent, silver ion addition, induction with methyl jasmonate and chitosan	43mg/L on day 16
Taxus cuspidata	Modification of the medium composition, and self-immobilization aggregate	431 mg/L on day 55

Production constraints and the unreliability of the plant-based systems have fueled the continued research into other systems for the production of Taxol. These challenges include; each *Taxus* species requiring different conditions, high production costs due to the slow growth of the explants, slow proliferation of suspended cells, low titers leading to unstable yields and sensitivity to shearing (Y. Li et al., 2015).

Other issues reviewed by (Liu et al., 2016) are more or less similar to the challenges mentioned above and are; culture growth inhibition due to secondary metabolite accumulation, aggregation and heterogeneity. These and other constraints lead to costly measures to tackle the challenges such as media optimization, phytohormones, adsorbants, methyl jasmonate, methyl jasmonate & ethylene, precursor and substrate feeding, cellular dispersion.

1.6.4. Production of taxadiene using metabolic engineering in microbial systems

From the literature, the first group to attempt the metabolic engineering of a microorganism for the production of taxadiene overexpressed the native *E. coli* DXP synthase gene (catalyzes the condensation of pyruvate and glyceraldehyde 3-Phosphate to

form 1-deoxy-D-xylulose 5-phosphate i.e. DXP), and IPP isomerase, Geranyl geranyl Diphosphate (*GGPPS*), & Taxadiene synthase (*TS*) (Q. Huang, Roessner, Croteau, & Scott, 2001).

To improve the activity of the enzymes, they linked IPP isomerase and *GGPPS* through a gene fusion and produced a more efficient and soluble truncated form of Taxadiene synthase by deleting a signal peptide that makes up the first 78 amino acids of the protein (Q. Huang et al., 2001).

The group next compared the production of Taxadiene in *in vivo* and *in vitro* (cell-free) systems. The *in vitro* system produced around 10 mg/L of Taxadiene along with its isomer taxa-4(20), 11 (12)-diene. The *in vivo* system produced around 1mg/L of Taxadiene; a considerably lower amount (Q. Huang et al., 2001).

A comparison between the two reveals that even though the in vitro system produces higher amounts of Taxadiene, the process is laborious, time-consuming and expensive. The in vivo system on the other hand is the more attractive option despite its lower yield due to its ability to produce Taxadiene using cheap substrates such as glucose and glycerol (Q. Huang et al., 2001).

In an attempt to engineer baccatin III – a useful precursor in the semi synthesis of Taxol – biosynthesis in *Saccharomyces cerevisiae*; a total of eight taxol biosynthetic genes functionally expressed in *S. cerevisiae* by DeJong et al., 2006). All recombinant proteins were epitope tagged and monitored using monoclonal antibodies with immunoblotting (DeJong et al., 2006).

The aim of the research was to reproduce the sequential steps from primary metabolism to taxadien-5α-acetoxy-10β-ol through overexpressing the respective genes. While yeasts have endogenous *GGPPS*, earlier during the research, it was discovered that the endogenous yeast *GGPPS* was not sufficient to support downstream biosynthesis (DeJong et al., 2006).

Therefore, a total of eight genes were overexpressed; Truncated *GGPPS* and *TS* (from *Taxus* sp.), taxadiene 5α-hydroxylase, Taxoid 10β-hydroxylase, Taxoid 13α-hydroxylase,

Taxadienol 5α-O-acetyl transferase, Taxoid 10β-O-acetyl transferase, and Taxoid 2α-O-benzoyl transferase (DeJong et al., 2006).

Taxadiene production by the engineered *S. cerevisiae* was around 1mg/L (DeJong et al., 2006). This was comparable to earlier titers from *E.* coli by Q. Huang et al., 2001) but lower than titers obtained from *Taxus* cell cultures. Around 25µg taxadien-5α-ol was also produced. Taxadien-5α-yl-acetate and taxadien-5αacetoxy-10β-ol could not be measured due to the elution of endogenous material in the same elution range e.g. geranyl geraniol (DeJong et al., 2006).

Employing a rational metabolic engineering approach, *T. chinensis* Geranylgeranyl Diphosphate synthase and Taxadiene Synthase were heterologously expressed by Engels et al., 2008 resulting in little increase in Taxadiene production. It was therefore hypothesized that this could be due to feedback regulation of HMG CoA (Engels et al., 2008).

To circumvent this, a truncated isoenzyme1 of HMG-CoA which is not inhibited through feedback regulation was expressed to realize 50 % increases in taxadiene production. The *T. chinensis GGPPS* was also replaced with an alternative *GGPPS* from *Sulfolobus acidocaldarius* – with no competition for steroid synthesis, and the *TS* codon optimized to realize a final titer of around 8.7 mg/L (Engels et al., 2008).

The final titer of taxadiene was 8-fold more than the earlier titers from both *S. cerevisiae* and *E. coli*. This notwithstanding, the high amounts of geranylgeraniol –around 33 mg/L produced, suggested that there was further potential for Taxadiene production from its direct precursor through further work on *TS* (Engels et al., 2008).

This research demonstrates the benefits of rational metabolic engineering where iterative work on a single gene, a single metabolic bottleneck and codon optimization are intelligently used to improve production of Taxadiene from microgram (µg) levels to milligram (mg) levels (Engels et al., 2008).

To counteract the non-specific effects of rational metabolic engineering, (Ajikumar et al., 2010) tried a new combinatorial method they termed ‘multivariate-modular pathway

engineering' for taxadiene production in *E. coli*. The objective of this method was to solve issues such as toxicity of the intermediate metabolites to the host cell, hidden pathways and other unknown metabolites that could compete with the synthesis of the target metabolite (Ajikumar et al., 2010).

The approach targeted the native bacterial isoprenoid production pathway known as the MEP or DXP pathway. It comprised the division of the taxadiene pathway into two modules; an upstream module making up the pathway above IPP; and a second downstream module leading to taxadiene. On successful production of taxadiene, they expressed a cytochrome p450 dependent enzyme; taxadiene 5α-hydroxylase which catalyzes the oxygenation of taxadiene to taxadien-5α-ol (Ajikumar et al., 2010).

Modifications were done to target the distinct modules as a whole rather than the conventional rational metabolic engineering approach of targeting single genes iteratively. In the upper module, 4 genes that were found to be rate limiting were targeted while on the lower module – changing the order of the fusion genes between *GGPPS* and *TS* was found to be beneficial (Ajikumar et al., 2010).

Since the research also took into consideration metabolomic studies of the 16 strains under study; they discovered an inverse correlation between taxadiene production and indole concentration. On further study, it became clear that taxadiene synthesis was inhibited by indole at levels above 100 mg/L – even though this was dependent on the strain (Ajikumar et al., 2010).

The best taxadiene producers in the reactors yielded a final titer of around 1020 mg/L taxadiene and around 24 mg/L taxadien-5α-ol, which is to date the highest yield of taxadiene obtained in a microbial host (Ajikumar et al., 2010). The taxadien-5α-ol yield is also the highest produced so far and was at the time estimated by the researchers to be more than 2400-fold of the metabolite produced in *S. cerevisiae* by earlier research (DeJong et al., 2006).

Based on the hypothesis that improved terpene production is achieved through increased precursor supply and that endogenous *S. cerevisiae GGPPS* was insufficient to

support efficient terpene biosynthesis; (Ding et al., 2014) used a protein modelling and docking strategy to select the best *GGPPS* based on fit with its substrate Farnesyl Diphosphate (FPP) (Ding et al., 2014).

The study compared *GGPPS* enzymes from up to 7 different organisms using a docking strategy for catalytic efficiency and for the best fit with Farnesyl Diphosphate (FPP). The best producing strain was developed using *GGPPS* from *Taxus baccata x Taxus cuspidata*, with taxadiene production of up to 72.8 mg/L. This study demonstrates the great importance of *GGPPS* enzyme on taxadiene and consequently other terpenoids (Ding et al., 2014).

2. MATERIALS AND METHODS

2.1. Strain

The background strain (SCIG22a) that was used to test the metabolic changes carried by the construct was a *S. cerevisiae* yeast previously engineered in another study: *MATa MAL2-8^{c} SUC2 URA3-52 lpp1Δ:: loxP dpp1Δ :: loxP* P_{ERG9} *Δ :: loxP-* P_{HXT1} *gdh1Δ :: loxP* P_{TEF1} *-ERG20*P_{PGK1} *-GDH2* + P_{TEF1} *-tHMG1* (López et al., 2015; Scalcinati et al., 2012). **Table 5** lists all the different strains used in the study and their sources.

***Table 2.1**: List of Strains used in the study*

Strain Name	**Construct added**	**Origin**
SCIG22a	None	(López et al., 2015; Scalcinati et al., 2012)
SCIG22a B2	Construct 2	This study
SCIG22a B2-B1	Construct 2 + Construct 1	This study
SCIG22a B2-B1-E	Construct 2 + Construct 1 + Construct E	This study
SCIG22a B2-B1-SWF	Construct 2 + Construct 1 + Construct SWF	This study
SCIG22a	Construct SWF	This study
SCIG22a B2-B1-G	Construct 2 + Construct 1 + Construct G	This study

2.1.1. Growth Conditions and Media

All yeast cultures were conducted at a rotation per minute (RPM) of 200 (liquid cultures) and at a standard temperature of 30 ºC. *E. coli* DH5α was cultured in Luria-Bertani (LB broth/agar) at 37 ºC for 12-16 hours and at an RPM of 150 with the addition of ampicillin for strains carrying a plasmid and no ampicillin for negative strains.

Prior to the transformation, the background strains were grown on Yeast-Peptone - Dextrose Agar (YPD agar) made up of; Yeast extract 10 g/l; peptone from animal tissue 20 g/l; glucose 20 g/l and 20 g/l agar. A 5 mL overnight culture of YPD broth was used to inoculate 50 mL of fresh YPD broth from a starting OD600 of between 0.05 to 0.1 prior to

the transformation and the transformation conducted at an OD600 of between 0.5 to 1.2 (Gietz, 2014).

After transformation, the strains with the URA selectable marker (construct 2 & 3 were selected using a synthetic media minus uracil made of; CSM-URA DOB 0.77 g/L, YNB without amino acids 6.9 g/L with glucose and agar added at 20 g/L. After verification through colony PCR and sequencing, the marker was looped out by selecting on a media containing 5-FOA; YNB without amino acids 6.9 g/L, CSM-URA DOB 0.77 g/L, uracil 50 mg/L, 5-FOA 750 mg/L, glucose and agar added at 20 g/L (Tang, Feng, & Chen, 2013).

The strains with the acetamide selectable marker (construct 1) were selected on a synthetic medium with acetamide as the sole Nitrogen source as follows; KH_2P0_4 3g/L, $MgSO_4.7H_2O$, 0.5 g/L, acetamide 0.6g/L, K_2SO_4 6.6g/L, 1ml/l trace element solution, 1mL/L vitamin solution, 20g/L glucose and 20g/L agar. On selection and confirmation through colony PCR and sequencing, the marker was looped out through the use of a medium containing fluoroacetamide; KH_2P0_4 3g/L, $MgSO_4.7H_2O$, 0.5 g/L, fluoroacetamide 2.3g/L, $(NH_4)_2SO_4$ 6.6g/L, 1mL/L trace element solution, 1mL/L vitamin solution, 20g/L glucose and 20g/L agar (Solis-Escalante et al., 2013).

The strains containing all the constructs were cultured in delft media made up of KH_2PO_4 14.4 g/L, $(NH4)_2SO_4$ 7.5 g/L, $MgSO_4 \cdot 7H_2O$ 0.5 g/L, vitamin 1 mL/L, metal solution 1mL/L, 20 g/L sugar which was divided into 20 mL shake flasks and overlaid with dodecane to collect taxadiene for 72 hours (Y. Chen et al., 2013).

2.1.2. GC Analysis

GC-MS (Shimadzu GCMS-QP2010 ultra) was used to analyze 0.2 µL of the sample solution by equipping with the capillary column DB-5MS (30 m x 0.25 mm x 0.25 µm) under the following conditions: Injection port temperature, 250 °C; Carrier gas, Helium (99.999%); Injection mode, Splitless; Pressure, 49 kPa; Total flow, 13.4 mL/min; Oven Temperature, 50 °C to 300 °C at a speed of 10 °C/min. MS conditions; Ion Source Temperature, 200 °C; Interface Temp., 320 °C; Acquisition Mode, SIM; Selected m/z, 272, 122, 107 (Tippmann, Scalcinati, Siewers, & Nielsen, 2016)

2.2. Plasmid vectors preparation

2.3. Making *E. coli* competent; and the transformation and purification of the plasmid vectors

Commercial *E. coli* DH5α cells were made chemically competent through the $CaCl_2$ protocol, transformed with the plasmids using heat shock at 42 ºC for two minutes and then plated on antibiotic containing media. The positive colonies were selected from the plates and grown again in LB broth containing antibiotic for 12-16 hours after which the plasmids were extracted and purified using the GeneJet Plasmid purification kit (#K0503).

Extraction and the purification of the plasmids were carried out according to the manufacturer's direction except for the elution step where MQ water was used instead of the supplied elution buffer. **Table 2.2** lists all the vectors from which the genes were amplified.

***Table 2.2**: Plasmids used in the study*

Plasmid name	**Genes contained**
PIYC08	tADH, Tcyc1
EutE	*A-ALD*
GM1	p*TEF1*
B1	*tHMG1*
X2	corresponding integration sites
XI-1	corresponding integration sites and selectable marker for *URA3*
NPH	*NPHT7*
pAD	*ACC*

2.4. Construction of DNA Cassettes

The PCR's were performed using Prime Star HS DNA polymerase (Takara), a high-fidelity polymerase with robust performance. The reactions were carried out according to the manufacturer's protocol with some few but minimal modifications for the sizes, template concentrations, primer concentrations, and DNTP concentrations. The reactions were carried out under general conditions used for most of the genes but were modified too for some cases) 98ºC: 30 seconds, then 35 cycles of 98ºC: 10 seconds 56ºC: 8 seconds 72ºC: 1000 bp/min and a final extension of 72ºC: 10 min and 4C: ∞

On amplification and purification of the gene fragments, overlap extension PCR (Chaaim, Fukunaga, Hoshida, & Akada, 2009; Urban, Neukirchen, & Jaeger, 1997) was performed in two steps; a first step consisting of 15 cycles without addition of primers to facilitate the joining of the overlapping ends, and a second amplification step of 35 cycles of the full construct with the addition of the forward primer of the first gene fragment and the reverse primer of the last gene fragment in the construct. All fusions were carried out using Prime Star HS DNA Polymerase (Takara) and Prime Star GXL (Takara) Polymerase, both of which are high fidelity polymerases.

2.4.1. Cassette 1

Table 2.3 lists all the different genes that make up construct 1 and their respective sizes.

Table 2.3: Cassette 1 genes

Name of gene/fragment	Vector from which it's amplified	Primers	Size (BP)
X-2	X-2	T1-7 B1-15	531
TADH1	piYC08	K B1-12	195
A-ALD	EutE	B1-2 B1-1	1404
pTEF-pPGK	GM-1	B1-8 B1-10	1412
THMG1 -tHMG1	B1	B1-14 B1-13	1775
X2 Down DR	X-2	P R	204
AMD	Synthesized Gene	B1-3 B1-4	2196
DR2 X2	X-2	B Z	490

2.4.2. Cassette 2

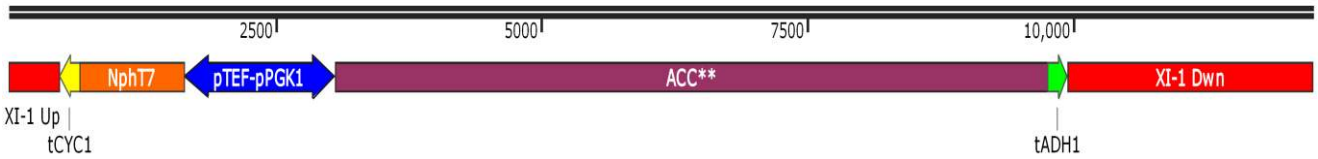

Table 2.4 lists all the different genes that make up construct 2 and their respective size

Table 2.4: Cassette 2 genes

Name of gene/fragment	Vector from which it's amplified	Primers	Size (BP)
XI-1	XI-1	B2-11, B2-13	531
Tcyc	piYC08	L, M	190
NPHT7	NPH	B2-3, B2-4	990
pTEF-p*PGK1*	GM-1	B2-6, B2-5	1412
*ACC1***	PAD	B2-2, B2-1	1775
TADH1	piYC08	B2-7, B2-8	204
URA XI-1 Down	XI-1	B2-14, B2-12	2196

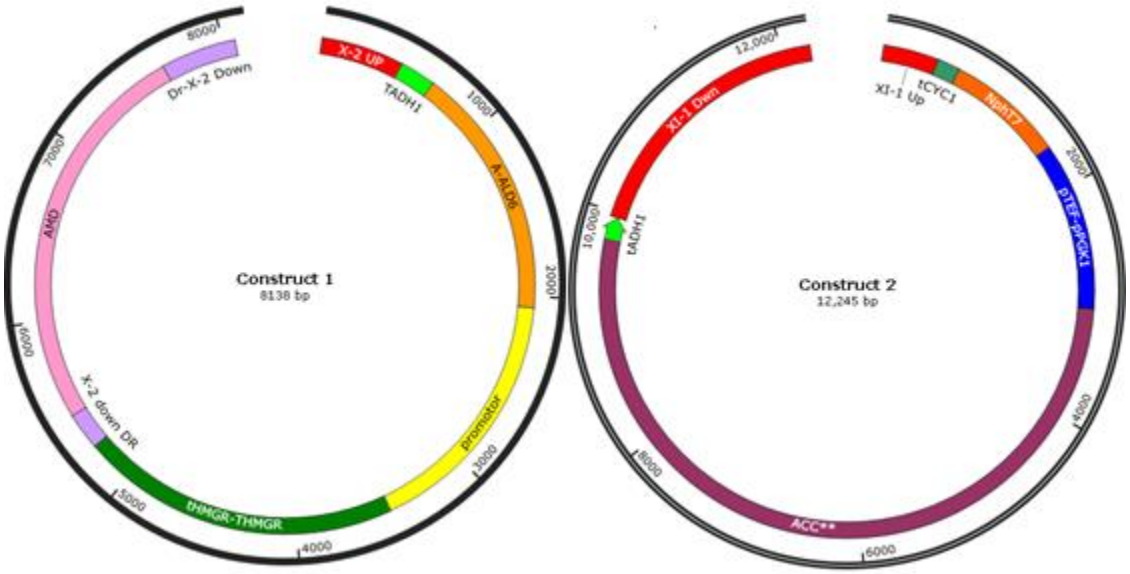

Figure 2.1: Cassettes 1 & 2 for genomic integration

Figure 2.1: The full constructs' map showing all the different genes; the constructs are linear and have only been illustrated in a circular map. These constructs make up the upper part of the mevalonate pathway.

2.4.3. Plasmids E, F(SW) and G

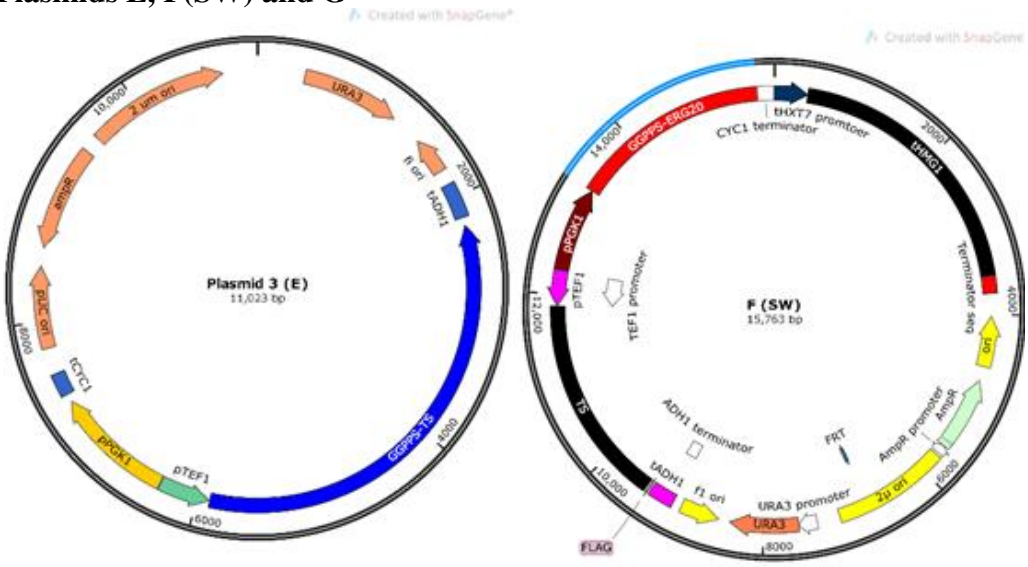

Figure 2.2: *Episomal Plasmids E and F (SW)*

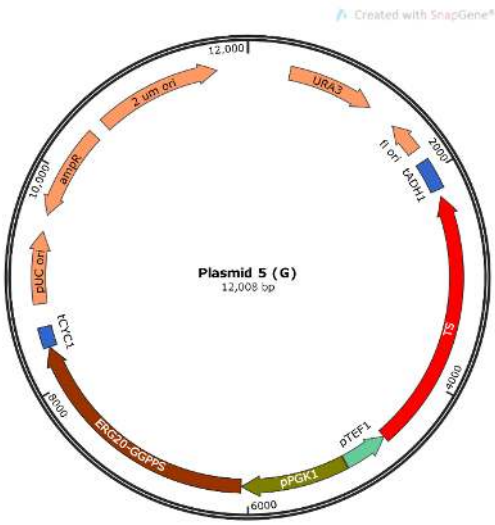

Figure 2.3: *Episomal Plasmid G*

Figures 2.2 & 2.3; The episomal plasmids that complete the remaining part of the pathway to taxadiene and therefore all contain *TS* which encodes the gene that catalyzes the first committed step to taxadiene.

2.5. Gel Electrophoresis and Preparation of DNA

50X TAE Electrophoresis (#B49 thermo) was diluted to a working concentration of 1X which contains 40 mM Tris, 20 mM Acetate and 1mM EDTA at a pH of around 8 and run at

80V/85V for 30 minutes to resolve the bands, with time increased if more resolution was needed.

Constructs or gene fragments were run alongside a GeneRuler 1 kb DNA Ladder marker (#SM0311 thermo) and fragments that did not contain other extra bands (primers only) were purified using a PCR purification kit (#K0701) while those that had extra bands were purified by a GeneJet gel extraction kit (#K0692).

Finally, the quantity of DNA in the constructs were measured using a nanodrop 1000c and a minimum total DNA quantity of 2000 ng in ≤ 5µL was used for each transformation experiment. DNA was concentrated to the minimal required volume using a CentriVap DNA Vacuum Concentrator.

2.6. Transformation

The electroporation protocol we employed was one first described by (Becker & Guarente, 1991) with the modifications suggested by (Thompson, Register, Curotto, Kurtz, & Kelly, 1998) which substantially improve transformation efficiencies. The electroporation method was used together with the lithium acetate/PEG method (Gietz, 2014). The three constructs were sequentially transformed through electroporation and the lithium acetate/PEG method into the background yeast strain; SCIG22a (López et al., 2015).

Prior to the transformation cells were plated on YPD agar and single colonies selected for an overnight culture in 5mL YPD. The overnight culture was then used to inoculate a 50 mL YPD which was then left to divide until an optical density (OD600) of between 0.8 and 1 which signified the exponential phase of the yeast cells.

Construct 2 was transformed first, and presumptive positive colonies were subjected to colony PCR. After confirming that the construct was indeed integrated in the correct site and contained no disruptive mutations, the *URA3* marker was looped out using 5-FOA media containing uracil and both colony PCR & growth on the SDA minus uracil were carried out to confirm the loop out.

The first construct was then transformed into the new strains carrying construct 2, and colony PCR subjected too to confirm the positive colonies carried the constructs. A second sequencing was done for the part containing the construct and confirmed to be correct. Since

this construct carried an AMD marker, the marker was looped out using Fluoro acetamide (FAC) media. The loop out was similarly confirmed through colony PCR and through growth on AMD agar.

Once these were confirmed, the strain containing both constructs (1 and 2) of SCIG22a were transformed with the third constructs (episomal plasmids) to complete the pathway to taxadiene. The constructs (episomal plasmids) had an *URA3* selectable marker and no loop out was done.

2.7. Colony PCR and sequencing

Primers for colony PCR were separately designed using the SnapGene software & NCBI BLAST and ordered from Oligomer Biyoteknoloji. Primers were designed to begin from the area upstream of the integration site in the yeast genome, inside the constructs, and the area outside the construct on the other side of the integration sites so as to verify both the location and the correctness of the construct.

The presumptive positive colonies were picked from the selective plates and dissolved in 200 mM LiAc 1% SDS and incubated at 70 ºC for 5 minutes. Two successive washes with 96 % alcohol, and then 70 % alcohol respectively were then conducted. This was followed by a final resuspension with 50-80 µL of MQ water, a short centrifugation of 1 minute at 15000 g and collection of the supernatant. 1µL of this supernatant was used for each colony PCR experiment (Lõoke, Kristjuhan, & Kristjuhan, 2017).

Colony PCR was conducted using DreamTaq Green DNA Polymerase (thermo #EP0714) according the manufacturer recommended protocol with the only modification of using a touchdown setting from 60-50 ºC with a decrease of 0.3º at each cycle as follows: 95 ºC: 3 minutes, then 35 cycles of 95 ºC: 30 seconds 60-50 ºC: 30 seconds 72 ºC: 1000 bp/min and a final extension of 72 ºC: 10 min and 4C: ∞

Sequencing primers were designed on the SnapGene software in 800 bp overlapping fragments for the three constructs. Colony PCR products of a minimum total DNA quantity of 50 ng in 100 µL were purified and sent for Sanger sequencing through Oligomer Biyoteknoloji. on completion the sequence results were sent as *.ab1* file reads with raw sequence data.

Analysis of sequencing results was carried out using; CLC Main workbench 8 to align and assemble the sequencing reads into contigs, and MEGA 7 to manually do the alignments using the ClustalW default settings of; gap opening penalty of 15 and a gap extension penalty of 6.66 for both pairwise alignment and multiple alignment with gaps.

3. RESULTS

3.1. Amplification of the gene fragments

3.1.1. Construct 1

Gene Fragments 1 – 8 include: X-2 (531bp), TADH1 (195bp), A-ALD (1404bp), pTEF-pPGK (1412bp), THMG1 -tHMG1 (1775bp), X2 Down DR (20bp), AMD (2196bp), DR2 X2 (250bp) shown in **Figure 3.1**, which were purified before being fused into one construct.

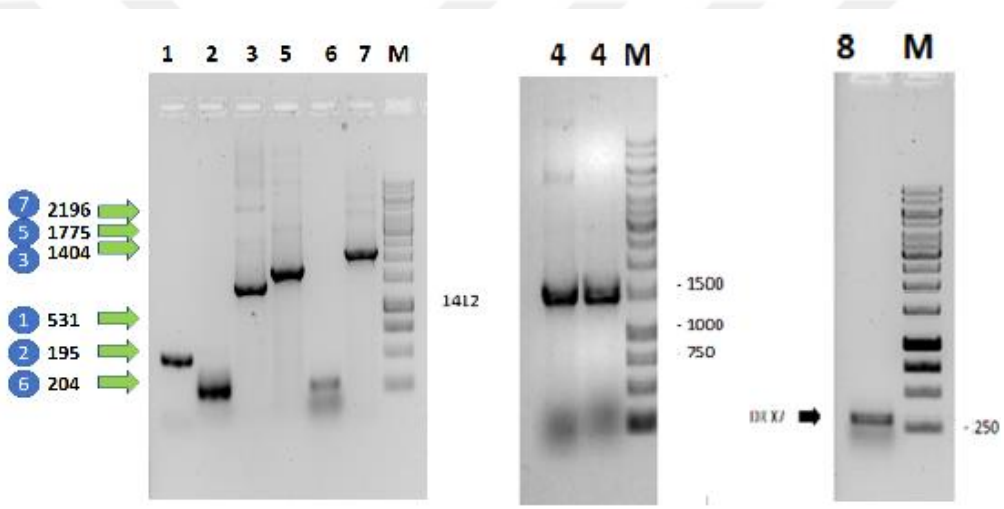

Figure 3.1: Cassette 1 gene fragments PCR product

Full construct; Genes 1 – 8 fused into a 8.2KB construct

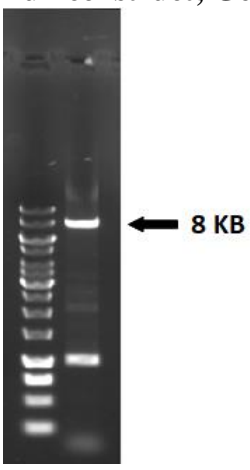

Figure 3.2: Cassette 1 fusion PCR product

Gene fragments shown in **Figure 3.2**, were fused into an 8 KB construct which was further purified from the gel before transforming into the background yeast strain.

3.1.2. Construct 2:

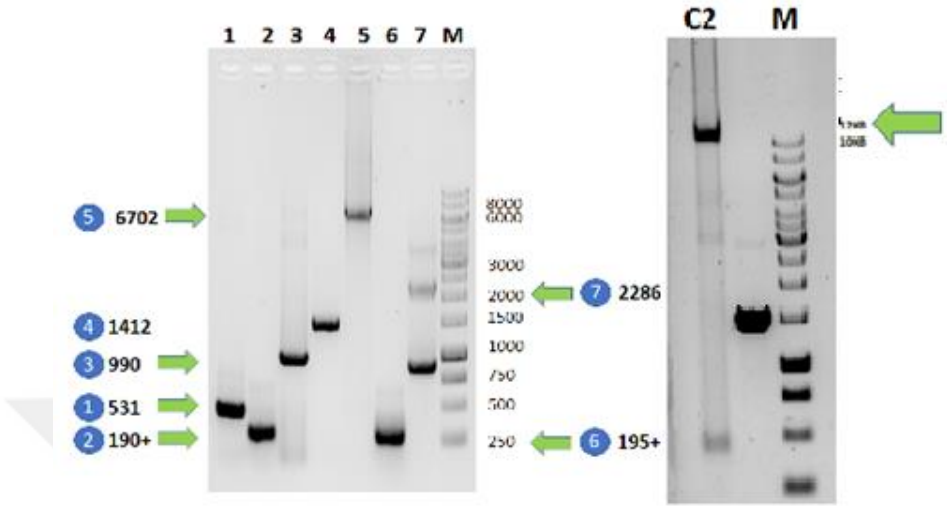

Figure 3.3: Cassette 2 gene fragments PCR product & Cassette 2 fusion PCR product

Figure 3.3 shows Gene Fragments 1 – 7 which included XI-1 up (531bp), TCYC1 (190bp), NPHT7 (990bp), pTEF-pPGK1 (1412bp), ACC1**(6702bp), TADH1 (195bp), URA XI-1 Down (2286). After purification, these were then fused into the full 12 KB construct shown as C2.

3.2. Colony PCR images for SCIG22a Construct 2

5 Fragments in total: fragment 1, 2, 3, 4 & 5

Fragment 1 Fragment 2:

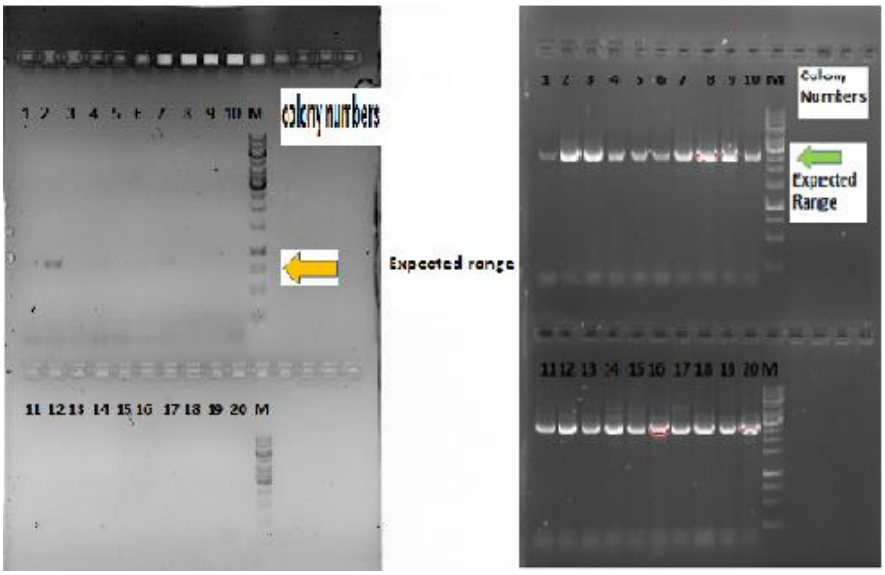

Figure 3.4: Colony PCR Products

Fragment one started upstream of the XI-1 integration site until an area inside the second gene (tCYC1) in the construct so as to confirm the correct integration of the cassette in the targeted region. The second fragment started from the first gene to the end of the promoter. Fragment one and two are shown in **Figure 3.4**.

For Fragments 3, 4 & 5: only the colonies that were both positive and gave strong bands in fragment 1 and 2> 1, 2, 4, 6, &7

For Fragment 4: The two colonies that worked in fragment 3: 2 & 4 , for Fragment 5: both colonies:2&4

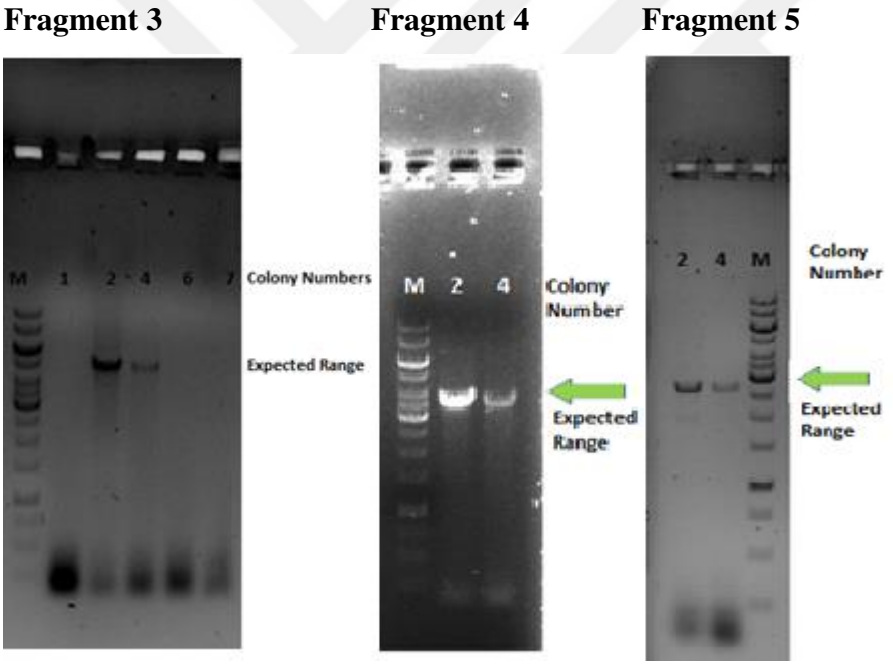

Figure 3.5: *Colony PCR images for SCIG22a Construct 2 (Fragment 3, 4, & 5)*

In **Figure 3.5**, fragment 3 was made of colonies 1, 2, 4, 6, &7, fragment 4 from colonies 2 & 4 and fragment 5 made of colonies 2 & 4. Fragment 3, 4, & 5 continued all the way to the end of the construct to confirm the accuracy of the construct's recombination into the yeast genome and the correct integration into the targeted region.

3.3. Colony PCR images for SCIG22a Construct 1

3 Fragments in total, fragment 1, 2 & 3

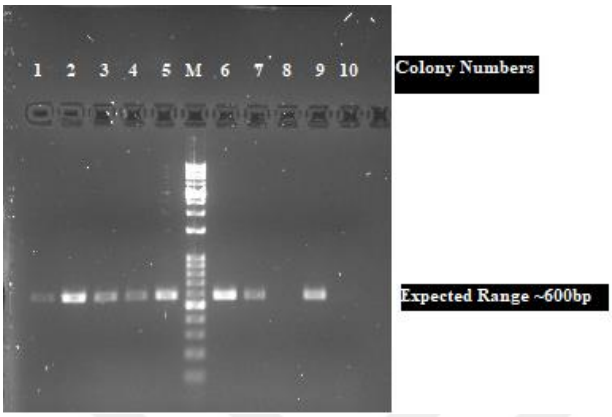

Figure 3.6: *Fragment 1 (Around 600bp)*

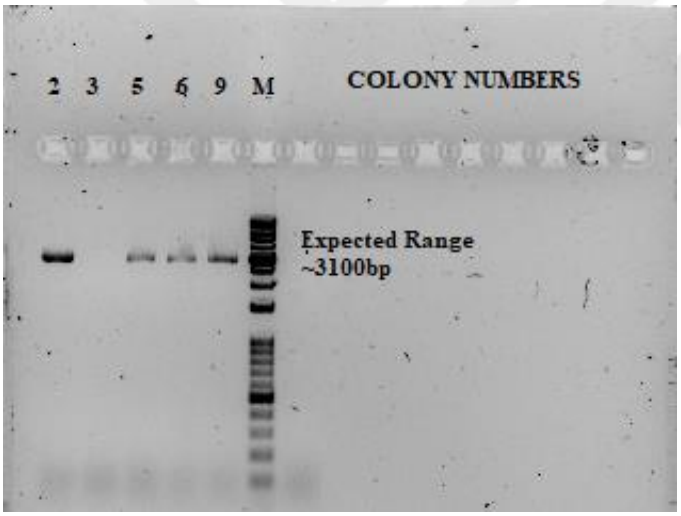

Figure 3.7*: Fragment 2 around 3100 bp*

Colonies 1,3,4,5&7: 1, 4 & 7 are positive at expected range of 5100bp

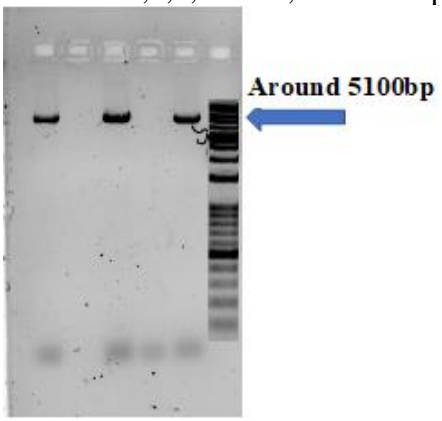

Figure 3.8*: Fragment 3 Around 5100bp*

Similar to the way construct one was confirmed through colony PCR; Fragment one started upstream of the respective integration site i.e. X-2 site until an area inside the second gene in the construct so as to confirm the correct integration of the cassette in the targeted region. The second fragment confirmed the rest of the construct while the third fragment went outside the construct into the expected area beyond the integration site i.e. X-2 down site (an area of around 300 bp beyond the integration site of the construct). These are shown in **Figures 3.6, 3.7 & 3.8**.

3.4. Colony Images

After transformation with Construct 1 on SDA-URA and After Loop out with 5-FOA

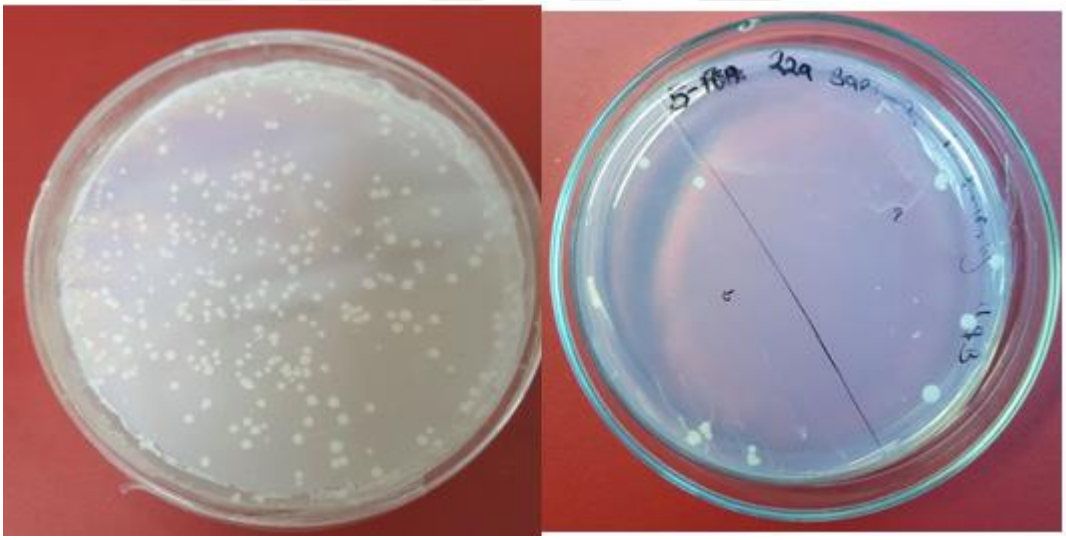

***Figure 3.9**: SDA-URA transformation plate and 5-FOA plate*

Clones were selected successfully on synthetic dextrose agar or SDA without uracil and counterselection also known as loop out was performed to remove the marker gene so as to reduce the metabolic burden on the cells. The counterselection is done through (5-FOA) or 5-fluoroorotic acid which selects against the *URA3* gene and through homologous recombination aided by flanking of repetitive sequences before and after the *URA3* gene (Boeke, La Croute, & Fink, 1984).

The growth on the 5-FOA for clones carrying the *URA3* gene leads to the formation of the toxic 5-flurouracil which forces the cell through a homologous recombination to remove

the *URA3* gene and thus selects against *URA3* which can then be reused in another transformation.

After transformation with Construct 2 on AMD and After Loop out with FAC

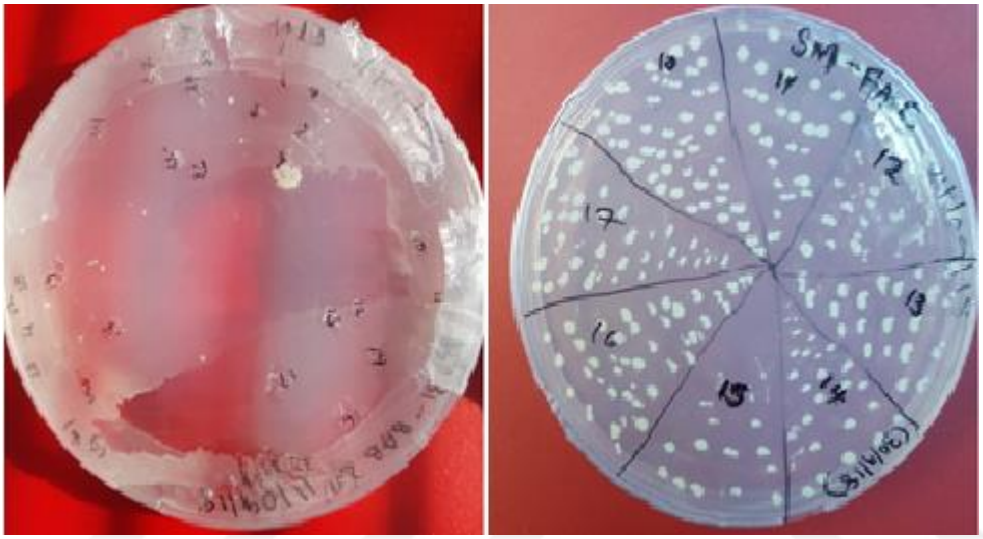

***Figure 3.10**: AMD transformation plate and FAC plate*

Construct 1 was selected on a growth media with acetamide as the sole Nitrogen source since it had the *amdS* gene from which codes for acetamidase enzyme; that catalyzes the breakdown of acetamide to ammonia and acetate and thus enables the cells carrying the gene to grow on acetamide as the only Nitrogen source (Solis-Escalante et al., 2012).

Like the earlier construct, the cells carrying this construct were counter selected on a medium containing fluoroacetamide which effectively removed the marker gene through homologous recombination with the marker flanked on both sides by direct repeats. Counterselection with fluoroacetamide leads to the formation of the toxic fluoroacetate which ensured the cells that grew on the media were negative for the *amdS* gene and thus the gene could be recycled (Solis-Escalante et al., 2013).

3.5. Episomal Plasmids

The strains were further transformed with three episomal plasmids (E, F (SW), and G) after which 5 colonies were selected from the transformants and subjected to colony PCR to check for the presence of specific genes in the plasmids; *GGPPS* and/or *TS*.

The LiAc protocol was used to transform around 500 ng of the respective episomal plasmid DNA containing the genes coding for the remainder of the pathway to Taxadiene; the first intermediate to Taxol.

The genes included *GGPPS*; an extra heterologous copy from *T. canadensis* since the yeast already has another copy, Taxadiene synthase; a novel gene heterologously expressed, and *ERG20*; overexpressed too.

After transformation with Plasmid F(SW)

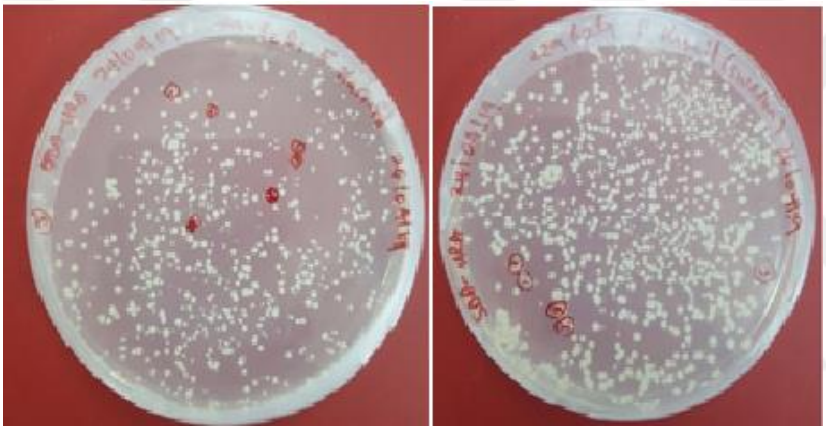

Figure 3.11*:From left to right After transformation with Plasmid E and F(SW)*

After transformation with Plasmid G

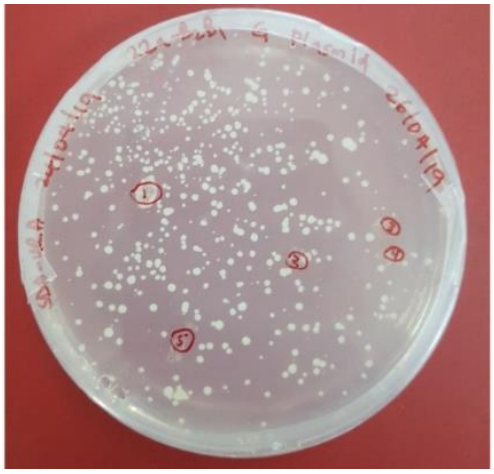

Figure 3.12*: After transformation with Plasmid G*

Clones were selected successfully on synthetic dextrose agar or SDA without uracil and colony PCR performed to confirm correct plasmid vector transformation.

Plasmid E

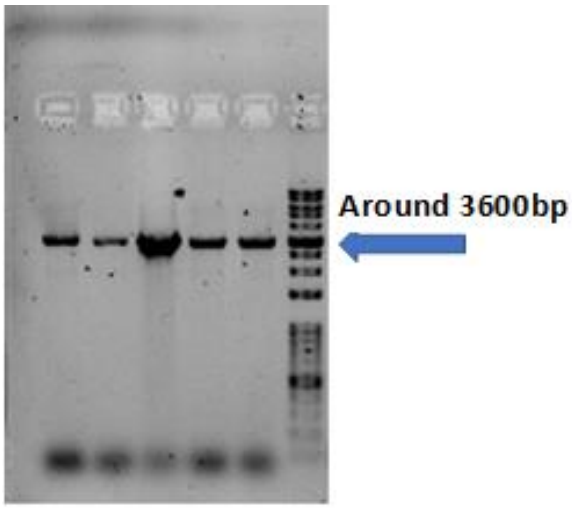

Figure 3.13*: Plasmid E transformants: Colony PCR for all 5 selected plasmids positive*

Five colonies were selected, and colony PCR performed to confirm the expression of the fusion of *GGPPS* and *TS* (GPPS-*TS*) which was 3656 bp and amplified from the *GGPPS* forward primer and the *TS* reverse primers.

Plasmid F (SW)

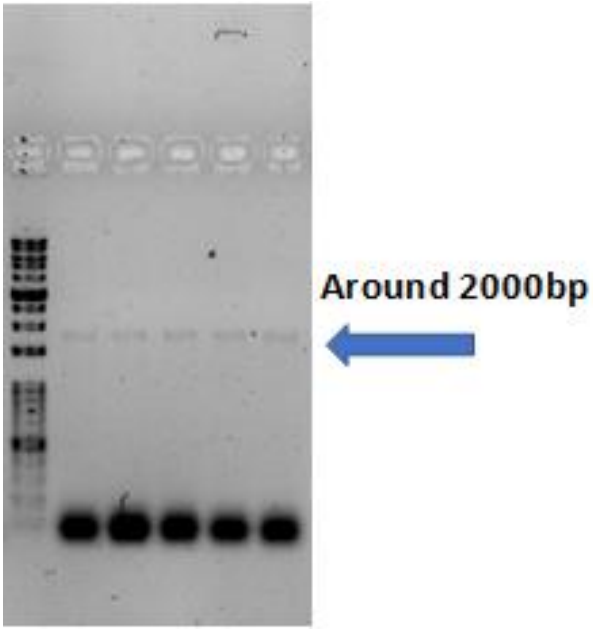

Figure 3.14*: Plasmid F(SW): Colony PCR for all 5 selected plasmids positive*

Five colonies were selected, and colony PCR performed to confirm the expression of *TS* which was around 2000bp and amplified from the *TS* forward primer and the *TS* reverse primers.

Plasmid G

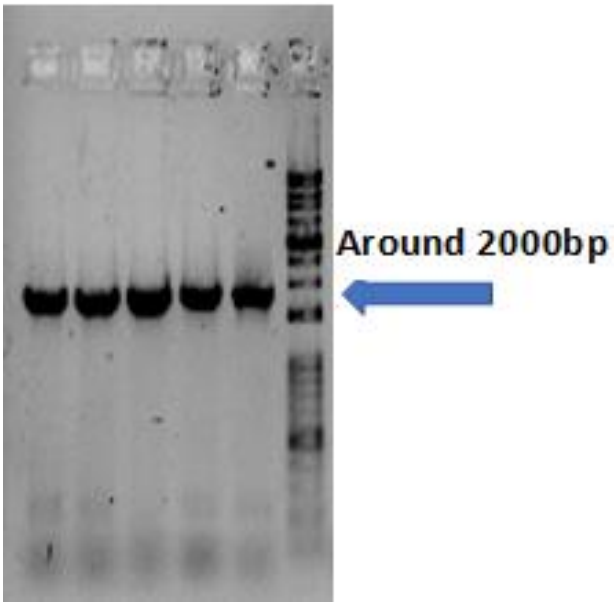

Figure 3.15*: Plasmid G transformants: Colony PCR for all 5 selected plasmids positive*

Five colonies were selected, and colony PCR performed to confirm the expression of *TS* which was around 2000bp and amplified from the *TS* forward primer and the *TS* reverse primers.

3.6. Gas Chromatography Results

Table 3.1: GC results of the three strains analyzed, and how they relate to the standard

Sample Colony Number	**Area one**	**Area two**	**Average**	**Compared to the standard average**
B2B1-E1	22220855	23108050	22664453	26.085
B2B1-E2	9416877	8309258	8863068	10.20
B2B1-E3	17835763	21838563	19837163	22.83
B2B1-E4	28361567	30987427	29674497	34.15
B2B1-E5	23846508	10810833	17328671	19.94
		Average of 5	**19673570**	**22.64**
B2B1-F1 (SW)	20554866	19839536	20197201	23.245
B2B1-F2 (SW)	22435895	16794671	19615283	22.575
B2B1-F3 (SW)	24852299	13110018	18981159	21.845
B2B1-F4 (SW)	20781617	12711071	16746344	19.27
B2B1-F5 (SW)	26444398	18015788	22230093	25.585
		Average of 5	**19554016**	**22.505**
B2B1-G1	12345700	13077976	12711838	14.63
B2B1-G2	7007298	7403214	7205256	8.29
B2B1-G3	38978	247025	143001.5	0.16
B2B1-G4	22100588	16148485	19124537	22.01
B2B1-G5	1668020	1416437	1542229	1.77
		Average of 5	**8145372**	**9.37**
Standard (SCIG22a-F(SW)				
	Area-1	**Area-2**	**Average**	
F1	169817	82579	126198/**868867.8**	
F2	1125918	1329152	1227535	
F3	919529	100872	510200.5	
		Average of 2	**868867.8**	

Strains for F(SW) had the highest taxadiene compared to the standard. In general, all the strains performed better than the standard strain. The standard strain was the SCIG22a strain transformed with only the F plasmid and didn't contain any other constructs from this study. The other strains B2B1; E, F(SW) and G have the two other integrated constructs and in addition have either plasmid E, G or F(SW). Five colonies were selected from each and thus the five different results for each of them. The column on the far right shows how many more times they performed better than the standard strain.

4. DISCUSSION

Construct two which is the lower pathway was the first one to be transformed using the *URA3* selection marker gene; a prototrophic gene complementing the uracil biosynthesis pathway. The auxotrophic background strain that the study employed had a deletion in the URA biosynthetic pathway and was therefore unable to grow without the *URA3* gene from *Kluyveromyces lactis* that codes for Orotidine-5′-phosphate decarboxylase which was included in the DNA construct to be transformed.

The transformation method chosen for construct 1 and 2 was integration which targets homologous ends in both the DNA construct to be transformed and in the yeast genome that are homologously recombined by the yeast DNA repair system (Orr-Weaver, Szostak, & Rothstein, 2006).

Integration offers very stable strains that don't require addition of antibiotic or other methods of maintaining the construct as compared to episomal plasmids. Another advantage of integration is that once the linear DNA constructs are stably transformed, the selectable marker gene can be removed without any disruption to the other genes in the construct. This achieves three things; metabolic burden is reduced as the extra prototrophic gene reduced, the strain is reversed to its earlier auxotrophic state again which means the marker gene can be reused, and thirdly, the strain does not have to be kept in the chemically defined media to ensure selection pressure is maintained or even its survival (Siewers, Mortensen, & Nielsen, 2010).

With the complete sequencing of the yeast genome (Goffeau et al., 1996), it became increasingly possible to consider integrating DNA constructs into the yeast genome and (Flagfeldt, Siewers, Huang, & Nielsen, 2009) characterized 20 integration sites in different parts of the yeast genome for any differences in expression levels. A few sites were reported to be very unstable and others with low expression levels. The sites that were reported to have good stability and the best expression levels were selected for this research.

In this research, DNA homologous to the integration sites in the yeast was selected and the genes designed in between the integration site i.e. between the up and down region ensuring the recombination of the construct inside the yeast genome through the homologous recombination and DNA repair mechanism unique to the yeast (Orr-Weaver et al., 2006).

The third constructs were transformed through episomal plasmids that completed the pathway to taxadiene through the expression of *TS*, *GGPPS*, and (in some) an extra copy of *tHMG1*. Since the construct was contained in plasmids that provided an autotrophy for uracil biosynthesis, constant selection pressure was maintained through continuous culture in a synthetic medium deficient of uracil.

Since the taxadiene production results are results of the final strain with all the constructs; this means that the final strain is a result of all the metabolic engineering techniques that this research employed in achieving those results. Therefore, the different strategies that were employed in the research to lead to the thirty-fold results will include the integration of DNA cassettes into the yeast genome rather than expression of the cassettes as episomal plasmids, choice of the promoters, the choice of terminators, the different genes that made up the engineering strategies, their choice and how they have impacted similar research, and finally the influence of the episomal plasmids and the genes they contained.

4.1. Choice of Promoters

One of the methods to control transcriptional activity is through promoter modulation and especially using constitutive promoters. Promoters that are constitutively expressed are very useful in metabolic engineering to ensure constant production of the desired product as compared to inducible promoters that require addition of specific compounds or depletion of a particular feed to start transcription. Constitutive promoters are usually found in genes whose expression is independent of the phase of growth or specific environmental conditions (Hubmann, Thevelein, & Nevoigt, 2014).

The *TEF1* encoding translation elongation factor 1 (Gatignol, Dassain, & Tiraby, 1990) and *PGK1* encoding phosphoglycerate kinase (Ogden et al., 2015), have been identified as very strong constitutive promoters that can be used to promote superior gene expression and consequently product formation. In a study aimed at characterizing various endogenous promoters to test their strengths while using a reporter gene; *Lac Z* to measure the activity of ß-galactosidase. *TEF1* and *PGK1* were the strongest and most consistent even at different glucose concentrations (Partow, Siewers, Bjørn, Nielsen, & Maury, 2010).

Another important element in this research was the bidirectional orientation of the two constitutive promoters; since promoter activity is not affected by a change in orientation (Ogden et al., 2015). (Y. Chen, Partow, Scalcinati, Siewers, & Nielsen, 2012) also applied the same bidirectional orientation of two constitutive primers; *TEF* and *PGK1* to improve patchoulol production – which is a very important compound in the perfume industry, by three-fold.

The current research therefore employed the bidirectional approach in construct one to overexpress; *A-ALD* which encodes acetylating acetaldehyde dehydrogenase using *TEF1* in one direction and *PGK1* to overexpress the truncated version of *tHMG1* which encodes 3-hydroxy-3-methylglutaryl-CoA reductase in the other direction. Construct two similarly used *TEF1* to overexpress *NPHT7* which encodes acetoacetyl-CoA synthase while *PGK1* was used to overexpress *ACC* which encodes acetyl-coA carboxylase in the other direction.

4.2. Choice of Terminators

One of the limitations to transcriptional control that even strong promoters can't solve include mRNA stability, processivity of RNA polymerase, and protein stability (Hubmann et al., 2014).

mRNA stability and efficiency of translation which might directly impact protein production are correlated with the terminator sequence used (Curran, Karim, Gupta, & Alper, 2013). (YAMANISHI, KATAHIRA, & MATSUYAMA, 2011) sought to demonstrate this by coupling the green fluorescent protein encoding gene with several terminators to investigate changes in protein yield depending on the terminator sequences. Some terminators greatly enhanced mRNA stability and this resulted in improved yields of the desired GFP (YAMANISHI et al., 2011).

In a subsequent research, relative to a powerful promoter *PGK1*; different terminator regions were measured, and a correlation was found between the length of the 3' Untranslated Region. The shorter the 3'UTR, the more stable the mRNA and the more abundant the mRNA. They hypothesized that the longer 3' UTR could encode active degradation mechanisms which shorten the mRNA half-life and thus led to reduced mRNA (Yamanishi et al., 2013).

These studies highlight the importance of terminator choice in metabolic engineering studies that target overexpression of genes and consequent overproduction of proteins which increase production of desired products.

(Curran et al., 2013) investigated on how to prolong mRNA half-life and have better gene expression for applications in metabolic engineering through modulations on various terminator sequences against different strength promoters. The study discovered that coupling of high capacity terminator sequences with a low strength promoter achieved the same expression achieved when using a high strength promoter only.

They discovered too that as the strength of the promoter increased the level of gene expression decreased. Therefore, low capacity terminator sequences coupled with a strong constitutive promoter would yield the highest gene expression and highest protein yield and vice versa (Curran et al., 2013).

This was employed in this research which coupled *A-ALD* to a strong constitutive promoter and a low capacity terminator whereas *NPHT7* was coupled to a strong constitutive promoter and a low capacity terminator and *ACC* coupled to a low capacity to achieve highest gene expressions possible.

4.3. Precursor Boosting Genetic Modifications

The two constructs for integration were made up of the upper part of the mevalonate pathway before and after acetyl CoA. The genes for the acetyl CoA improvement on construct 1 and 2 constitute genes that are vital in boosting acetyl CoA levels in the cytoplasm and include acetylating acetaldehyde dehydrogenase (*A-ALD*), and Acetyl CoA Carboxylase (*ACC1*).

4.3.1. Acetyl CoA

The strategy of this research in using *A-ALD* was based on the ability of the enzyme to convert acetaldehyde directly to acetyl CoA while skipping the ATP dependent conversion of acetaldehyde to acetate by acetaldehyde dehydrogenase (*ALD*). The non-dependency on ATP serves to reduce the metabolic burden on the cell while increasing the abundance of

acetyl CoA in the cytosol, as the endogenous *ALD* is still active. The simplified reaction is illustrated below:

Acetaldehyde + CoA + NAD^+ = Acetyl CoA + NADH + H^+

This strategy was investigated earlier by a research that sought to fully replace the yeast endogenous acetyl CoA synthases and acetaldehyde dehydrogenases with *A-ALD* through complete deletion of the former enzymes. The research tested several genes coding *A-ALD* from bacterial sources and codon optimized them for expression in *Saccharomyces cerevisiae.* The best performing strains were the *A-ALD* from *E. coli* and *Listeria innocua* which exhibited growth rates of over 70 % above the negative control strain (Kozak et al., 2014).

4.3.2. Malonyl CoA

Immediately after acetyl CoA production in the cytosol, to pull down its flux down to the desired product; Acetyl CoA carboxylase (*ACC1*) was employed in this research. *ACC* catalyzes the ATP dependent carboxylation of Acetyl CoA to malonyl CoA as shown in the reaction below.

Acetyl-CoA + ATP + Hydrogen Carbonate = ADP + H^+ + Malonyl-CoA + Phosphate

The carboxylation of *ACC* of acetyl CoA forms malonyl CoA which is a precursor involved in fatty acid production and in the production of other valuable malonyl CoA derived chemicals such as 3-hydroxypropionic acid. However, the activity of *ACC* is tightly regulated through transcriptional regulation and through snf1 protein kinase repression. In order to free the enzyme's potential for application in metabolic engineering, (Shi, Chen, Siewers, & Nielsen, 2014) through a site directed mutagenesis removed post translational modification and thus greatly the production of 3-hydroxypropionic acid and fatty acid ethyl esters (Shi et al., 2014). Similar research has been carried out to improve malonyl CoA availability as a precursor to other valuable chemicals such as the fatty acid ethyl esters which

are essential in biodiesel production, 3-hydroxypropionic acid and resveratrol (X. Chen, Yang, Shen, Hou, & Bao, 2018; S. Li, Si, Wang, & Zhao, 2015).

4.3.3. Acetoacetyl CoA

In addition to the endogenous acetoacetyl-CoA thiolase or Acetyl-CoA C-acetyltransferase which is encoded by *ERG10*, this study also heterologously expressed another acetoacetyl-CoA thiolase encoded by the gene *NPHT7* in a soil strain of *Streptomyces* that catalyzes the condensation of malonyl CoA and acetyl CoA to give acetoacetyl CoA; the first compound of the mevalonate pathway (Okamura, Tomita, Sawa, Nishiyama, & Kuzuyama, 2010). The *NPHT7* gene used in this research was chosen as a strategy to redirect all the acetyl CoA overproduced from the effect of the *A-ALD* and the malonyl CoA generated by the *ACC1* into an irreversible step; essentially committing the chemicals to the mevalonate pathway where they are desired (Okamura et al., 2010).

Acetyl-CoA + H^+ + Malonyl-CoA = Acetoacetyl-CoA + CO_2 + CoA

In this strategy therefore, the *nphhT7* gene was expected to work in tandem with the endogenous *ERG10* which uses two molecules of acetyl CoA through a Claisen reaction while the *NPHT7* was expected to condense any excess acetyl CoA to the malonyl CoA generated by *ACC1* and thereby increase the flux down to mevalonate. This is in line with other research such as (Okamura et al., 2010) where the enzyme was used to increase the production of mevalonate by 3 fold. While (Tippmann, Ferreira, Siewers, Nielsen, & Chen, 2017) couldn't get any improvement of the product titers through the expression of the enzyme, this study employed a different approach; whereby instead of expressing through an episomal plasmid, the *NPHT7* was integrated into the genome and the endogenous *ERG10* was not deleted as was done in (Tippmann et al., 2017). This was due to their findings that the sole reliance of the organism on the malonyl CoA bypass became detrimental to the growth of the organism and for the production of farnesene.

4.3.4. *tHMG1*

Further down the pathway *tHMG1* which codes for HMG-CoA reductase has been known as a bottleneck whereby buildup of other intermediates especially HMG-CoA occurs due to the low activity or feedback regulation of the enzyme (Paddon et al., 2013). This and other research points to the integration of an extra copy of the enzyme in the genome to be effective in raising product titers. In addition, the use of a truncated *HMG1*, denoted as *tHMG1* has been demonstrated to be very effective in improving the production of different isoprenoid compounds such as artemisinin, α-santalene, β-ionone and farnesene (Y. Chen et al., 2013; López et al., 2015; Paddon et al., 2013; Tippmann et al., 2017).

2 NADPH + 3-hydroxy-3-methyl-glutaryl-CoA <=> 2 NADP + mevalonate + Coenzyme A

The background strain employed in this research already had an extra copy of the truncated *tHMG1* gene (López et al., 2015) and in addition, an extra copy was integrated together with construct 2 into the yeast genome since it was expected that there would be a heavy flux of HMG-CoA for the enzyme to reduce to mevalonate and since it was known from transcriptomics and metabolomics studies that HMG-CoA accumulation led to growth inhibition and thereby less product (Paddon et al., 2013).

In this research addition to the integrated copy of the *tHMG1* , one of the episomal plasmids contained an extra copy of the *tHMG1* to investigate the change this would have on the overall production in comparison to the control strain which had one copy, no construct, and one copy in its genome and the other strains which had one copy in its genome, another with the construct and one with the plasmid, and a final strain(s) with one copy in its genome, another with the construct and no *tHMG1* in the plasmid.

The performance of the strains with the *tHMG1* integrations plus an extra *tHMG1* in the episomal plasmid (FSW) in comparison to the strains without an extra *tHMG1* were not that different except for the G plasmid. E plasmid and F(SW) were almost the same in their production of Taxadiene while G plasmid had less production even though it didn't have an extra *tHMG1* like F(SW). This might point to another reason other than the obvious reason which could be that the *tHMG1* exerts metabolic stress on the cell and thus reduces its growth rate. Plasmid F(SW) strain had the highest production of Taxadiene even though it had 2

copies of integrated *tHMG1* and an extra copy of *tHMG1* in the episomal plasmid. This shows that the *tHMG1*(s) were capable of dealing with the increased flux down the mevalonate pathway and little or no accumulation was encountered as the growth rates were similar to the other strains.

4.4. Episomal Plasmids

The final strategy of this research was the genes that were contained in the episomal plasmids that were included to complete the final genes that encode the enzymes to convert geranyl geranyl diphosphate to taxadiene and an overexpression of the genes *ERG20* and *GGPPS*. *GGPPS* codon optimized from *Taxus canadensis* was fused with the endogenous *ERG20* and the *TS* codon optimized from *Taxus brevifolia* in one plasmid. The plasmids used were designed by other researchers in the lab to test the optimized acetyl CoA construct and had the said genes in different orders. Plasmid E had *GGPPS-TS* fused with no *ERG20*, Plasmid F(SW) had *GGPPS-ERG20* fused, plus *TS* and an extra copy of *tHMG1* separate while Plasmid G had an *ERG20-GGPPS* fusion (different order from FSW), and a separate *TS* on the same plasmid.

From the similar performance of the strains containing the F(SW) and E plasmids, this goes to show how different engineering techniques can be used to promote higher flux to the desired product. Plasmid E employed a *GGPPS-TS* fusion which has been shown earlier to improve the product yield due to the proximity of the enzymes to the substrate (Ajikumar et al., 2010). On the other hand Plasmid F(SW) employed a different approach of not only fusing *GGPPS-ERG20* but also overexpressing *GGPPS* gene that is known to be a bottleneck in the pathway flux; *GGPPS* from *T. canadensis* which is superior in its activity and doesn't face the regulation that the endogenous *GGPPS* undergoes, and *TS* from *T. brevifolia*.

The G plasmid had the lowest performance even though the only difference from the F(SW) plasmid was the change in order of the fusion of the genes from *GGPPS-ERG20* to *ERG20-GGPPS* and the extra *tHMG1*. The hypothesis for these results could be that the *GGPPS-ERG20* was more superior in handling the metabolic flux compared to the *GGPPS-ERG20* fusion which resulted in the disparities. This is in line with the findings by (Ajikumar et al., 2010) where changing the order from *GGPPS-TS* to *TS-GGPPS* increased product by

3 fold. Another reason for the this could be that *tHMG1* was a bottleneck and therefore the extra *tHMG1* in Plasmid F(SW) was key in its superior performance.

Similarly, the high performance of the E plasmid which only had a *GGPPS-TS* with no other overexpression from other endogenous genes other than the heterologous *GGPPS* and *TS* proves that the extra *tHMG1* and *ERG20* could be detrimental and might only serve to balance out their performance. This is an important result as it shows that how much the order of genes might affect the product formation regardless of the pathway optimization from the top. It also shows how more product yield can be obtained with boosting of the precursors as was done in this research for Acetyl CoA, Malonyl CoA, Mevalonate and finally GGPP and FPP which resulted in more product compared to the standard strain with no engineering for precursor overproduction. These results also call for increased pathway optimization as any deleterious effects of metabolic stress could be balanced out by the inclusion of other genes to ease the metabolic flux.

Finally, these results demonstrate the importance of gene regulation and how it can be circumvented through the use of heterologous genes that are not regulated by the host organism. This is exemplified by the use of heterologous genes such as *NPHT7* from *Streptomyces* and the *GGPPS* from *T. canadensis*. Gene regulation was also circumvented or reduced through the use of truncated gene such as *ACC1* and *tHMG1* which greatly improved their activities. This research shows the use of many elements in combination to increase the product more than 20-fold compared to the standard strain. In addition to the gene modifications and overexpression, the research utilized a balanced approach of powerful promoters with low capacity terminators for best results. Genetic integration was also employed to ensure no loss of plasmid while it was complemented with the counterselection method so as to enable the recycling of the same selection markers. All these strategies show the many parts of the metabolic engineering research which touches on a lot of aspects of the metabolic pathway; balancing of the pathway products, regulation and so on.

5. CONCLUSION

Using the same strain optimized for overproducing the precursors of the mevalonate pathway; other valuable terpenoids such as, fatty acid ethyl esters, flavors, etc. could be produced. This would be done by simply expressing the genes that should complete the desired pathway from GGPP in an episomal plasmid or a repeat integration in one of the many other sites not targeted by this research. The ease of reuse of an optimized strain to produce a totally different product relying on the same precursors is a desirable phenomenon in metabolic engineering as it ensures novelty in the production of the new compound while only modulating specific genes downstream. It also increases the speed at which valuable compounds can be produced in new hosts through fewer changes. While some products could be toxic and limit growth in the host organism, the change to another product might work a lot better than the original product that it was optimized to produce. Therefore, it would be valuable to express other pathways to compare performance with the native hosts as this might open other opportunities.

The results obtained in the shake flask results need further culture in bioreactors to find the highest potential yield, titer and volume as well as to test the performance in an industrial setting. This would increase the product titers much further as the bioreactors are optimized for more production. It would also show the tolerance of the recombinant yeast strains to the industrial temperatures, pH and other stress conditions. The bioreactor yields on glucose and titers would give a more accurate result on the commercial viability of the taxadiene production.

The research had various limitations among which was the lack of a taxadiene standard to use for the GC analysis. This was solved through the extraction of taxadiene from known taxadiene producing strains (standard strain in the GC results) and then using this as the standard against which all other strains were compared. Another limitation was the inability to transform 3 other plasmids which would have shed more light on the performance of the optimized precursors. The inability could be due to an unknown metabolic burden that did not allow the uptake of the plasmids and thus the transformation. A final limitation to the research was the limited time which could have been used to study further the full metabolic profile or metabolomics, which would give more understanding to the effects of the gene

overexpression(s) on the cellular metabolic machinery and guide future engineering approaches on possible nodes for modulation.

Another possible gene that could be targeted in future includes the ATP mediated acetyl CoA lyase (*ACL*) from oleaginous yeasts which is an important gene in the conversion of citrate to acetyl CoA. *ACL* has been used in the increase the production of mevalonate in *S. cerevisiae* (Rodriguez et al., 2016). Other genes that could be heterologously expressed in yeast to optimize the mevalonate pathway include the *mvaE* that encodes the enzyme that catalyzes both the conversion of acetyl CoA to acetoacetyl CoA and the conversion of HMG-CoA to mevalonate, and *mvaS* that converts acetoacetyl CoA to mevalonate (Paddon et al., 2013; Rodriguez et al., 2016).

In the industrial setting cheaper feedstocks are more attractive and therefore attention should be given to further research on lignocellulosic biomass to produce taxadiene and other valuable chemicals to solve sustainability and waste issues.

REFERENCES

Ajikumar, P. K., Xiao, W.-H., Tyo, K. E. J., Wang, Y., Simeon, F., Leonard, E., … Stephanopoulos, G. (2010). Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science (New York, N.Y.)*, *330*(6000), 70–74. https://doi.org/10.1126/science.1191652

Atanasov, A. G., Waltenberger, B., Pferschy-Wenzig, E. M., Linder, T., Wawrosch, C., Uhrin, P., … Stuppner, H. (2015). Discovery and resupply of pharmacologically active plant-derived natural products: A review. *Biotechnology Advances*. https://doi.org/10.1016/j.biotechadv.2015.08.001

Bailey, J. E. (1991). Toward a science of metabolic engineering. *Science*, *252*(5013), 1668 LP – 1675. https://doi.org/10.1126/science.2047876

Becker, D. M., & Guarente, L. (1991). High-Efficiency Transformation of Yeast by Electroporation. *Methods in Enzymology*. https://doi.org/10.1016/0076-6879(91)94015-5

Boeke, J. D., La Croute, F., & Fink, G. R. (1984). A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *MGG Molecular & General Genetics*. https://doi.org/10.1007/BF00330984

Boronat, A., & Rodríguez-Concepción, M. (2015). *Terpenoid Biosynthesis in Prokaryotes BT - Biotechnology of Isoprenoids* (J. Schrader & J. Bohlmann, Eds.). https://doi.org/10.1007/10_2014_285

Carlsen, S., Ajikumar, P. K., Formenti, L. R., Zhou, K., Phon, T. H., Nielsen, M. L., … Stephanopoulos, G. (2013). Heterologous expression and characterization of bacterial 2-C-methyl-d-erythritol-4-phosphate pathway in Saccharomyces cerevisiae. *Applied Microbiology and Biotechnology*. https://doi.org/10.1007/s00253-013-4877-y

Cha-aim, K., Fukunaga, T., Hoshida, H., & Akada, R. (2009). Reliable fusion PCR mediated by GC-rich overlap sequences. *Gene*.

https://doi.org/10.1016/j.gene.2008.12.014

Chen, X., Yang, X., Shen, Y., Hou, J., & Bao, X. (2018). Screening phosphorylation site mutations in yeast acetyl-CoA carboxylase using malonyl-CoA sensor to improve malonyl-CoA-derived product. *Frontiers in Microbiology*. https://doi.org/10.3389/fmicb.2018.00047

Chen, Y., Daviet, L., Schalk, M., Siewers, V., & Nielsen, J. (2013). Establishing a platform cell factory through engineering of yeast acetyl-CoA metabolism. *Metabolic Engineering*, *15*(1), 48–54. https://doi.org/10.1016/j.ymben.2012.11.002

Chen, Y., Partow, S., Scalcinati, G., Siewers, V., & Nielsen, J. (2012). Enhancing the copy number of episomal plasmids in Saccharomyces cerevisiae for improved protein production. *FEMS Yeast Research*. https://doi.org/10.1111/j.1567-1364.2012.00809.x

Chen, Y., Siewers, V., & Nielsen, J. (2012). Profiling of cytosolic and peroxisomal acetyl-CoA metabolism in Saccharomyces cerevisiae. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0042475

Chen, Y., Zhou, Y. J., Siewers, V., & Nielsen, J. (2015). Enabling technologies to advance microbial isoprenoid production. *Advances in Biochemical Engineering/Biotechnology*. https://doi.org/10.1007/10_2014_284

Croteau, R., Ketchum, R. E. B., Long, R. M., Kaspera, R., & Wildung, M. R. (2006). Taxol Biosynthesis and Molecular Genetics. *Phytochemistry Reviews*, *5*(1), 75–97. https://doi.org/10.1007/s11101-005-3748-2

Curran, K. A., Karim, A. S., Gupta, A., & Alper, H. S. (2013). Use of expression-enhancing terminators in Saccharomyces cerevisiae to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metabolic Engineering*. https://doi.org/10.1016/j.ymben.2013.07.001

DeJong, J. H. M., Liu, Y., Bollon, A. P., Long, R. M., Jennewein, S., Williams, D., & Croteau, R. B. (2006). Genetic engineering of taxol biosynthetic genes in

Saccharomyces cerevisiae. *Biotechnology and Bioengineering*. https://doi.org/10.1002/bit.20694

Denis, J. N., Greene, A. E., Guenard, D., Gueritte-Voegelein, F., Mangatal, L., & Potier, P. (1988). Highly efficient, practical approach to natural taxol. *Journal of the American Chemical Society*, *110*(17), 5917–5919. https://doi.org/10.1021/ja00225a063

Dewick, P. M. (2009). The Mevalonate and Methylerythritol Phosphate Pathways: Terpenoids and Steroids. In *Wiley Online Books*. *Medicinal Natural Products*. https://doi.org/doi:10.1002/9780470742761.ch5

Ding, M. Z., Yan, H. F., Li, L. F., Zhai, F., Shang, L. Q., Yin, Z., & Yuan, Y. J. (2014). Biosynthesis of taxadiene in saccharomyces cerevisiae: Selection of geranylgeranyl diphosphate synthase directed by a computer-aided docking strategy. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0109348

Engels, B., Dahm, P., & Jennewein, S. (2008). Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. *Metabolic Engineering*, *10*(3), 201–206. https://doi.org/https://doi.org/10.1016/j.ymben.2008.03.001

Flagfeldt, D. B., Siewers, V., Huang, L., & Nielsen, J. (2009). Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. *Yeast*. https://doi.org/10.1002/yea.1705

Gatignol, A., Dassain, M., & Tiraby, G. (1990). Cloning of Saccharomyces cerevisiae promoters using a probe vector based on phleomycin resistance. *Gene*. https://doi.org/10.1016/0378-1119(90)90159-O

Gietz, R. D. (2014). Yeast transformation by the LiAc/SS carrier DNA/PEG method. *Methods in Molecular Biology*. https://doi.org/10.1007/978-1-4939-0799-1_4

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., … Oliver, S. G. (1996). Life with 6000 Genes. *Science*, *274*(5287), 546 LP – 567.

https://doi.org/10.1126/science.274.5287.546

Holton, R. A., Somoza, C., Kim, H. B., Liang, F., Biediger, R. J., Boatman, P. D., … Kim, S. (1994). First total synthesis of taxol. 1. Functionalization of the B ring. *Journal of the American Chemical Society*, *116*(4), 1597–1598. https://doi.org/10.1021/ja00083a066

Hong, K.-K., & Nielsen, J. (2012). Metabolic engineering of Saccharomyces cerevisiae: a key cell factory platform for future biorefineries. *Cellular and Molecular Life Sciences*, *69*(16), 2671–2690. https://doi.org/10.1007/s00018-012-0945-1

Huang, Q., Roessner, C. A., Croteau, R., & Scott, A. I. (2001). Engineering Escherichia coli for the synthesis of taxadiene, a key intermediate in the biosynthesis of taxol. *Bioorganic and Medicinal Chemistry*. https://doi.org/10.1016/S0968-0896(01)00072-4

Huang, T.-K., & McDonald, K. A. (2009). Bioreactor engineering for recombinant protein production in plant cell suspension cultures. *Biochemical Engineering Journal*, *45*(3), 168–184. https://doi.org/https://doi.org/10.1016/j.bej.2009.02.008

Hubmann, G., Thevelein, J. M., & Nevoigt, E. (2014). Natural and modifi ed promoters for tailored metabolic engineering of the yeast saccharomyces cerevisiae. *Methods in Molecular Biology*. https://doi.org/10.1007/978-1-4939-0563-8_2

Jennewein, S., Wildung, M. R., Chau, M., Walker, K., & Croteau, R. (2004). Random sequencing of an induced Taxus cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(24), 9149–9154. https://doi.org/10.1073/pnas.0403009101

Kirby, J., Dietzel, K. L., Wichmann, G., Chan, R., Antipov, E., Moss, N., … Zhao, L. (2016). Engineering a functional 1-deoxy-D-xylulose 5-phosphate (DXP) pathway in Saccharomyces cerevisiae. *Metabolic Engineering*. https://doi.org/10.1016/j.ymben.2016.10.017

Kocharin, K., Chen, Y., Siewers, V., & Nielsen, J. (2012). Engineering of acetyl-CoA metabolism for the improved production of polyhydroxybutyrate in Saccharomyces cerevisiae. *AMB Express*, *2*(1), 52. https://doi.org/10.1186/2191-0855-2-52

Kozak, B. U., van Rossum, H. M., Benjamin, K. R., Wu, L., Daran, J. M. G., Pronk, J. T., & Van Maris, A. J. A. (2014). Replacement of the Saccharomyces cerevisiae acetyl-CoA synthetases by alternative pathways for cytosolic acetyl-CoA synthesis. *Metabolic Engineering*. https://doi.org/10.1016/j.ymben.2013.11.005

Krivoruchko, A., Serrano-Amatriain, C., Chen, Y., Siewers, V., & Nielsen, J. (2013). Improving biobutanol production in engineered Saccharomyces cerevisiae by manipulation of acetyl-CoA metabolism. *Journal of Industrial Microbiology & Biotechnology*, *40*(9), 1051–1056. https://doi.org/10.1007/s10295-013-1296-0

Krivoruchko, A., Zhang, Y., Siewers, V., Chen, Y., & Nielsen, J. (2015). Microbial acetyl-CoA metabolism and metabolic engineering. *Metabolic Engineering*, *28*, 28–42. https://doi.org/10.1016/J.YMBEN.2014.11.009

Lee, Y.-H., Lee, Y.-R., Park, C.-S., Im, S.-A., Song, S., Hong, J. T., … Lee, C.-K. (2014). Baccatin III, a precursor for the semisynthesis of paclitaxel, inhibits the accumulation and suppressive activity of myeloid-derived suppressor cells in tumor-bearing mice. *International Immunopharmacology*, *21*(2), 487–493. https://doi.org/https://doi.org/10.1016/j.intimp.2014.06.012

Li, M., & Borodina, I. (2015). Application of synthetic biology for production of chemicals in yeast Saccharomyces cerevisiae. *FEMS Yeast Research*. https://doi.org/10.1111/1567-1364.12213

Li, S., Si, T., Wang, M., & Zhao, H. (2015). Development of a Synthetic Malonyl-CoA Sensor in Saccharomyces cerevisiae for Intracellular Metabolite Monitoring and Genetic Screening. *ACS Synthetic Biology*. https://doi.org/10.1021/acssynbio.5b00069

Li, Y., Zhang, G., & Pfeifer, B. A. (2015). *Current and Emerging Options for Taxol Production BT - Biotechnology of Isoprenoids* (J. Schrader & J. Bohlmann, Eds.).

https://doi.org/10.1007/10_2014_292

Lian, J., Si, T., Nair, N. U., & Zhao, H. (2014). Design and construction of acetyl-CoA overproducing Saccharomyces cerevisiae strains. *Food, Pharmaceutical and Bioengineering Division 2014 - Core Programming Area at the 2014 AIChE Annual Meeting*. https://doi.org/10.1016/j.ymben.2014.05.010

Liu, W. C., Gong, T., & Zhu, P. (2016). Advances in exploring alternative Taxol sources. *RSC Advances*, *6*(54), 48800–48809. https://doi.org/10.1039/C6RA06640B

Lõoke, M., Kristjuhan, K., & Kristjuhan, A. (2017). Extraction of genomic DNA from yeasts for PCR-based applications. *BioTechniques*. https://doi.org/10.2144/000114497

López, J., Essus, K., Kim, I., Pereira, R., Herzog, J., Siewers, V., … Agosin, E. (2015). Production of $β$-ionone by combined expression of carotenogenic and plant CCD1 genes in Saccharomyces cerevisiae. *Microbial Cell Factories*, *14*(1), 84. https://doi.org/10.1186/s12934-015-0273-x

Martin, V. J. J., Piteral, D. J., Withers, S. T., Newman, J. D., & Keasling, J. D. (2003). Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nature Biotechnology*. https://doi.org/10.1038/nbt833

Maury, J., Asadollahi, M. A., Møller, K., Schalk, M., Clark, A., Formenti, L. R., & Nielsen, J. (2008). Reconstruction of a bacterial isoprenoid biosynthetic pathway in Saccharomyces cerevisiae. *FEBS Letters*. https://doi.org/10.1016/j.febslet.2008.10.045

Nevoigt, E. (2008). Progress in metabolic engineering of Saccharomyces cerevisiae. *Microbiology and Molecular Biology Reviews : MMBR*, *72*(3), 379–412. https://doi.org/10.1128/MMBR.00025-07

Nicolaou, K. C., Yang, Z., Liu, J. J., Ueno, H., Nantermet, P. G., Guy, R. K., … Sorensen, E. J. (1994). Total synthesis of taxol. *Nature*, *367*(6464), 630–634. https://doi.org/10.1038/367630a0

Nielsen, J., & Jewett, M. C. (2008). Impact of systems biology on metabolic engineering of Saccharomyces cerevisiae. *FEMS Yeast Research*, *8*(1), 122–131. https://doi.org/10.1111/j.1567-1364.2007.00302.x

Ogden, J. E., Stanway, C., Kim, S., Mellor, J., Kingsman, A. J., & Kingsman, S. M. (2015). Efficient expression of the Saccharomyces cerevisiae PGK gene depends on an upstream activation sequence but does not require TATA sequences. *Molecular and Cellular Biology*. https://doi.org/10.1128/mcb.6.12.4335

Okamura, E., Tomita, T., Sawa, R., Nishiyama, M., & Kuzuyama, T. (2010). Unprecedented acetoacetyl-coenzyme A synthesizing enzyme of the thiolase superfamily involved in the mevalonate pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(25), 11265–11270. https://doi.org/10.1073/pnas.1000532107

Orr-Weaver, T. L., Szostak, J. W., & Rothstein, R. J. (2006). Yeast transformation: a model system for the study of recombination. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.78.10.6354

Ostergaard, S., Olsson, L., & Nielsen, J. (2000). Metabolic engineering of Saccharomyces cerevisiae. *Microbiology and Molecular Biology Reviews : MMBR*, *64*(1), 34–50. https://doi.org/10.1128/mmbr.64.1.34-50.2000

Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., … Newman, J. D. (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature*, *496*, 528. Retrieved from https://doi.org/10.1038/nature12051

Partow, S., Siewers, V., Bjørn, S., Nielsen, J., & Maury, J. (2010). Characterization of different promoters for designing a new expression vector in Saccharomyces cerevisiae. *Yeast*. https://doi.org/10.1002/yea.1806

Partow, S., Siewers, V., Daviet, L., Schalk, M., & Nielsen, J. (2012). Reconstruction and Evaluation of the Synthetic Bacterial MEP Pathway in Saccharomyces cerevisiae. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0052498

Petzold, C. J., Chan, L. J. G., Nhan, M., & Adams, P. D. (2015). Analytics for Metabolic Engineering. *Frontiers in Bioengineering and Biotechnology*, *3*, 135. https://doi.org/10.3389/fbioe.2015.00135

Pitera, D. J., Paddon, C. J., Newman, J. D., & Keasling, J. D. (2007). Balancing a heterologous mevalonate pathway for improved isoprenoid production in Escherichia coli. *Metabolic Engineering*. https://doi.org/10.1016/j.ymben.2006.11.002

Rao, K. V. (1997). Semi-synthesis of paclitaxel from naturally occurring glycosidic precursors. *Journal of Heterocyclic Chemistry*, *34*(2), 675–680. https://doi.org/10.1002/jhet.5570340255

Rodriguez, S., Denby, C. M., Van Vu, T., Baidoo, E. E. K., Wang, G., & Keasling, J. D. (2016). ATP citrate lyase mediated cytosolic acetyl-CoA biosynthesis increases mevalonate production in Saccharomyces cerevisiae. *Microbial Cell Factories*, *15*, 48. https://doi.org/10.1186/s12934-016-0447-1

Roldão, A., Kim, I.-K., & Nielsen, J. (2012). *Bridging Omics Technologies with Synthetic Biology in Yeast Industrial Biotechnology BT - Systems Metabolic Engineering* (C. Wittmann & S. Y. Lee, Eds.). https://doi.org/10.1007/978-94-007-4534-6_9

Scalcinati, G., Partow, S., Siewers, V., Schalk, M., Daviet, L., & Nielsen, J. (2012). Combined metabolic engineering of precursor and co-factor supply to increase α-santalene production by Saccharomyces cerevisiae. *Microbial Cell Factories*. https://doi.org/10.1186/1475-2859-11-117

Shi, S., Chen, Y., Siewers, V., & Nielsen, J. (2014). Improving Production of Malonyl Coenzyme A-Derived Metabolites by Abolishing Snf1-Dependent Regulation of Acc1. *MBio*. https://doi.org/10.1128/mbio.01130-14

Siewers, V., Mortensen, U. H., & Nielsen, J. (2010). Genetic engineering tools for Saccharomyces cerevisiae. *Man. Ind. Microbiol. Biotechnol. (3rd Ed.)*.

Solis-Escalante, D., Kuijpers, N. G. A., Bongaerts, N., Bolat, I., Bosman, L., Pronk, J. T.,

… Daran-Lapujade, P. (2013). amdSYM, A new dominant recyclable marker cassette for Saccharomyces cerevisiae. *FEMS Yeast Research*. https://doi.org/10.1111/1567-1364.12024

Song, J. Y., Park, J. S., Kang, C. D., Cho, H. Y., Yang, D., Lee, S., & Cho, K. M. (2016). Introduction of a bacterial acetyl-CoA synthesis pathway improves lactic acid production in Saccharomyces cerevisiae. *Metabolic Engineering*. https://doi.org/10.1016/j.ymben.2015.09.006

Stephanopoulos, G. N., Aristidou, A. A., & Nielsen, J. (1998). *CHAPTER 1 - The Essence of Metabolic Engineering* (G. N. Stephanopoulos, A. A. Aristidou, & J. B. T.-M. E. Nielsen, Eds.). https://doi.org/https://doi.org/10.1016/B978-012666260-3/50002-9

Tabata, H. (2004). *Paclitaxel Production by Plant-Cell-Culture Technology BT - Biomanufacturing* (J.-J. Zhong, Ed.). https://doi.org/10.1007/b13538

Tang, X., Feng, H., & Chen, W. N. (2013). Metabolic engineering for enhanced fatty acids synthesis in Saccharomyces cerevisiae. *Metabolic Engineering*. https://doi.org/10.1016/j.ymben.2013.01.003

Thompson, J. R., Register, E., Curotto, J., Kurtz, M., & Kelly, R. (1998). An improved protocol for the preparation of yeast cells for transformation by electroporation. *Yeast*. https://doi.org/10.1002/(SICI)1097-0061(19980430)14:6<565::AID-YEA251>3.0.CO;2-B

Tippmann, S., Ferreira, R., Siewers, V., Nielsen, J., & Chen, Y. (2017). Effects of acetoacetyl-CoA synthase expression on production of farnesene in Saccharomyces cerevisiae. *Journal of Industrial Microbiology & Biotechnology*, *44*(6), 911–922. https://doi.org/10.1007/s10295-017-1911-6

Tippmann, S., Scalcinati, G., Siewers, V., & Nielsen, J. (2016). Production of farnesene and santalene by Saccharomyces cerevisiae using fed-batch cultivations with RQ-controlled feed. *Biotechnology and Bioengineering*, *113*(1), 72–81. https://doi.org/10.1002/bit.25683

Urban, A., Neukirchen, S., & Jaeger, K. E. (1997). A rapid and efficient method for site-directed mutagenesis using one-step overlap extension PCR. *Nucleic Acids Research*. https://doi.org/10.1093/nar/25.11.2227

Wani, Mansukh C, & Horwitz, S. B. (2014). Nature as a remarkable chemist: a personal story of the discovery and development of Taxol. *Anti-Cancer Drugs*, *25*(5), 482–487. https://doi.org/10.1097/CAD.0000000000000063

Wani, Mansukhlal C, Taylor, H. L., Wall, M. E., Coggon, P., & McPhail, A. T. (1971). Plant antitumor agents. VI. Isolation and structure of taxol, a novel antileukemic and antitumor agent from Taxus brevifolia. *Journal of the American Chemical Society*, *93*(9), 2325–2327. https://doi.org/10.1021/ja00738a045

Woolston, B. M., Edgar, S., & Stephanopoulos, G. (2013). Metabolic Engineering: Past and Future. *Annual Review of Chemical and Biomolecular Engineering*, *4*(1), 259–288. https://doi.org/10.1146/annurev-chembioeng-061312-103312

Yamanishi, M., Ito, Y., Kintaka, R., Imamura, C., Katahira, S., Ikeuchi, A., … Matsuyama, T. (2013). A genome-wide activity assessment of terminator regions in saccharomyces cerevisiae provides a “terminatome” toolbox. *ACS Synthetic Biology*. https://doi.org/10.1021/sb300116y

YAMANISHI, M., KATAHIRA, S., & MATSUYAMA, T. (2011). TPS1 Terminator Increases mRNA and Protein Yield in a Saccharomyces cerevisiae Expression System . *Bioscience, Biotechnology, and Biochemistry*. https://doi.org/10.1271/bbb.110246

Curriculum Vitae for Farah Dagane GARAAD

PERSONAL DETAILS

Names : Farah Dagane

Surname : GARAAD

Date of Birth : 25/03/1992

Gender : Male

Language(S) : Fluent in English, Swahili, Somali and Turkish

Nationality : Kenyan

Email contacts :

daganefarah02@gmail.com/daganefarah@yahoo.com/farah.dg@anadolu.edu.tr

EDUCATIONAL BACKGROUND

September 2016 – June 2019 : Master of Science in Biotechnology at Eskişehir Technical University

September 2015- June 2016 : Advanced Turkish Language Proficiency Course at Osmangazi University

October 2010-April 2014 : Bachelors of Science in Microbiology and Biotechnology

WORKING EXPERIENCE: Internships

September 2014 - December 2014: National Biosafety Authority, Kenya's GMO regulatory Authority

As an intern, duties included : screening Applications to conduct GMO research for consistency and coherence, assisting in compilation of reports for review by the Board of management, assist in review of expert risk assessment reports on current applications etc.

July - September 2012 : Kenya Bureau of Standards (Industrial attachment)

Worked as a student Intern at Kenya Bureau of Standards Headquarters, skills gained included; general microbiology laboratory operations; maintenance of sterility and aseptic methods; techniques valuable to detection of pathogens and spoilage microorganisms in food; handling and processing of samples for analysis; preparation and maintenance of cultures;

reading and interpretation of results; building and maintaining good working relationships with colleagues to ensure assigned tasks are completed.

June - September 2011: Islamic Medical Relief Agency (Industrial attachment)

Worked as a medical lab intern during this period in this hospital where work involved handling and analysis of samples for local infections and diseases including but not limited to, Malaria, typhoid fever, Hepatitis, Syphilis, Urinalysis etc.

Link to LinkedIn Profile:

https://www.linkedin.com/in/farah-dagane-74886845/