Gene source selection as a metabolic engineering tool for naringenin production in *Saccharomyces cerevisiae*

Mark, Rita

2019

Mark, R. (2019). Gene source selection as a metabolic engineering tool for naringenin production in *Saccharomyces cerevisiae*. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/137008>

<https://doi.org/10.32657/10356/137008>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 21:26:22 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**GENE SOURCE SELECTION AS A METABOLIC
ENGINEERING TOOL FOR NARINGENIN PRODUCTION IN
*SACCHAROMYCES CEREVISIAE***

**MARK RITA
SCHOOL OF CHEMICAL AND BIOMEDICAL
ENGINEERING
2019**

**GENE SOURCE SELECTION AS A METABOLIC
ENGINEERING TOOL FOR NARINGENIN PRODUCTION IN
*SACCHAROMYCES CEREVISIAE***

MARK RITA

School of Chemical and Biomedical Engineering

A thesis submitted to Nanyang Technological University in partial
fulfilment of the requirement for the degree of
Doctor of Philosophy

2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

30 July 2019.

Date

Rita Mark

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

30 July 2019

Date

Professor William Chen
Director
Food Science and Technology Programme
Nanyang Technological University

Prof. William, Chen Wei Ning

Authorship Attribution Statement

This thesis contains material from 2 papers published in the following peer-reviewed journals in which I am listed as an author.

Chapter 2 is published as Mark, R.; Lyu, X.; Lee, J. J. L.; Parra-Saldívar, R.; Chen, W. N., Sustainable production of natural phenolics for functional food applications. *Journal of Functional Foods* **2019**, 57, 233-254.

The contributions of the authors are as follows:

- I and Lyu, X. had researched and compiled prior data.
- Lee, J. J. L. and Parra-Saldívar, R. had interpreted the results.
- Prof Chen, W. N. designed the work and revised the manuscript.

Chapter 4 and 5 are published as Mark, R.; Lyu, X.; Ng, K. R.; Chen, W. N., Gene Source Screening as a Tool for Naringenin Production in Engineered *Saccharomyces cerevisiae*. *ACS Omega* **2019**, 4 (7), 12872-12879.

The contributions of the co-authors are as follows:

- I carried out the experiments and wrote the manuscript.
- Prof Chen, W. N. designed the research and edited the manuscript.
- Lyu, X. and Ng, K. R. commented on the experiments and manuscript

30 July 2019.

Date

.....

Rita Mark

ACKNOWLEDGMENTS

This work would not have been possible without the unwavering support and patience of my supervisor, Prof William, Chen Wei Ning. His help and encouragement were profound when I faced obstacles during my research. I feel privileged to have had this opportunity to work under his prudent supervision.

I would like to convey my appreciation towards Asst Prof Miao Yansong and Assoc Prof Thirumaran s/o Thanabalu from the School of Biological Sciences. Their guidance as part of the TAC meeting was insightful and thoughtful. Furthermore, I am grateful to Dr. Xiaomei Lyu and Dr. Jaslyn Lee Jie Lin, for their continuous advice and support. They provided many constructive suggestions, which were essential for my project. Additionally, I would like to extend my gratitude towards Kuan Rei Ng, who taught me the necessary experiment protocols and encouraged me during my research.

A very special thanks to Sulagna, Felix, Sanda and Nimesh for their friendship and support during the dark days. To Sulagna, you are the most amazing best friend I could ever imagine. To Felix, you had supported me mentally, motivated me not to give up and without you, I would still struggle with my scientific writing. To Sandra, thank you for always listening to my complaints and help to solve my experimental riddles. To Nimesh, you would be always by my side and help me to stand up after falling.

I would like to thank the rest of my lab mates, whose help and support were a vital part of my research and my life in Singapore. Additionally, I appreciate NTUGSA for

making my student life in NTU memorable. Most importantly, I am grateful for all my friends who have helped me to enjoy my life outside of school.

Lastly, heartfelt gratitude goes to my family for their love, care and continuous support. Köszönöm az összes segítséget, támogatást és szeretetet!

TABLE OF CONTENTS

ACKNOWLEDGMENTS	I
TABLE OF CONTENTS.....	III
LIST OF FIGURES	VI
LIST OF TABLES.....	IX
LIST OF ABBREVIATIONS	XI
THESIS SUMMARY	XIV
1. INTRODUCTION	1
1.1 Background.....	1
1.2 Aims and objectives.....	5
1.3 Outline of the thesis	8
2. LITERATURE REVIEW	10
2.1 Introduction.....	10
2.2 Phenolic acids	14
2.2.1 <i>p-Coumaric acid</i>	16
2.2.2 <i>Caffeic and ferulic acids</i>	18
2.3 Flavonoids.....	20
2.3.1 <i>Flavanones</i>	21
2.3.2 <i>Flavones</i>	28
2.3.3 <i>Flavanols</i>	29
2.3.4 <i>Anthocyanins</i>	31
2.4 Conclusion	33
3. GENE SOURCE SELECTION FOR 4-COUMARATE: COA LIGASE AND CHALCONE SYNTHASE IN BY4741 <i>S. CEREVISIAE</i> STRAIN.....	35
3.1 Introduction.....	36
3.2 Experimental Details.....	39
3.2.1 <i>Gene source comparison</i>	39
3.2.2 <i>Microbial strains, media, plasmid and genes</i>	39
3.2.3 <i>Construction of plasmids, yeast transformation and cultivation</i>	40

3.2.4 Protein extraction and Western blot	42
3.2.5 Naringenin extraction and HPLC analysis.....	43
3.2.6 Growth curve	44
3.3 Results and Discussion	44
3.3.1 Gene source selection	44
3.3.2 Expression of new combinations in <i>S. cerevisiae</i>	49
3.4 Conclusion	56
4. ENHANCEMENT OF NARINGENIN PRODUCTION IN <i>S. CEREVISIAE</i> BY DIVERSE METABOLIC APPROACHES	57
4.1 Introduction.....	58
4.2 Experimental Details.....	60
4.2.1 Gene source comparison	60
4.2.2 Microbial strains, media, plasmid and genes	61
4.2.3 Construction of plasmids	62
4.2.4 Naringenin extraction and HPLC analysis.....	63
4.2.5 Growth curve	64
4.2.6 Intracellular metabolite extraction and GC-MS analysis.....	65
4.2.7 Site-directed mutagenesis	66
4.3 Results and Discussion	67
4.3.1 Metabolite profiling	67
4.3.2 Expression of <i>CHI</i> in <i>S. cerevisiae</i>	70
4.3.3 Y-28 <i>S. cerevisiae</i> strain	77
4.3.4 Site-directed mutagenesis for enzyme <i>CHS</i>	80
4.4 Conclusion	86
5. POTENTIAL CARBON AND NITROGEN SOURCES SCREENED FOR OPTIMIZED MEDIA AND GROWTH CONDITIONS.....	88
5.1 Introduction.....	89
5.2 Experimental details	90
5.2.1 Microbial strains, plasmid and genes.....	90
5.2.2. Preparation and usage of alternative carbon and nitrogen sources	91
5.2.3 Naringenin extraction and HPLC analysis.....	91

5.2.4 <i>Growth curve</i>	92
5.3 Results and Discussion	93
5.3.1 <i>Potential carbon and nitrogen source screening</i>	93
5.3.2 <i>Galactose as a carbon source</i>	95
5.3.3 <i>Sucrose as a carbon source</i>	97
5.3.4 <i>Sucrose and glycerol mixture as a carbon source</i>	100
5.3.5 <i>Okara as a nitrogen source</i>	105
5.4 Conclusion	108
6. CONCLUSION AND FUTURE DIRECTIONS.....	109
6.1 Conclusion	109
6.2 Future directions	113
6.2.1 <i>Genome integration</i>	113
6.2.2 <i>Protein fusion</i>	114
6.2.3 <i>Food waste as carbon source substitute</i>	115
REFERENCES	117

LIST OF FIGURES

Figure 1 Main death causes in the world from 2017.	2
Figure 2 Main risk factors for death causes from 2017.	2
Figure 3 Main dietary risk factors for death causes from 2017	3
Figure 4 Graphical presentation of this thesis`s main aim.....	7
Figure 5 Biosynthesis paths and classification of phenolic compounds.....	11
Figure 6 Biosynthetic pathways of diverse phenolic acids and the most effective strategies for their production improvement.....	15
Figure 7 Biosynthetic pathways of flavonoids and the most effective strategies for their production improvement.....	25
Figure 8 Biosynthetic pathway of naringenin	38
Figure 9 Plasmid pCEV-G1-Km structure with inserted 4CL and CHS	41
Figure 10 Gene screening by tertiary structures and phylogenetic tree.....	48
Figure 11 Confirmation of the new combinations by gel electrophoresis	50
Figure 12 Naringenin production in BY4741 <i>S. cerevisiae</i>	51
Figure 13 Extracellular <i>p</i> -coumaric acid, naringenin chalcone and naringenin in BY4741 <i>S.</i> <i>cerevisiae</i>	53
Figure 14 Western blot results of the sixteen strains expressed in <i>S. cerevisiae</i>	54
Figure 15 Growth curve of the naringenin producing <i>S. cerevisiae</i> BY4741 strains.	55
Figure 16 Catalysis of naringenin chalcone to naringenin by CHI.....	59
Figure 17 Plasmid pCEV-G1-Ph structure with inserted CHI.....	62

Figure 18 GC-MS chromatography overlay example for two naringenin producing strains and a control.....	68
Figure 19 Graphical interpretation of the three key metabolites' abundance comparison in the naringenin producing and non-producing strains.....	69
Figure 20 Metabolic changes induced by naringenin production in <i>S. cerevisiae</i>	70
Figure 21 Gene screening by tertiary structures and phylogenetic tree.	73
Figure 22 Confirmation of pCEV-G1-Ph and CHI enzymes ligation by gel electrophoresis	75
Figure 23 Naringenin production in BY4741 <i>S. cerevisiae</i> , expressing enzymes 4CL, CHS and CHI.....	76
Figure 24 Comparison of naringenin production in BY4741 and Y-28 <i>S. cerevisiae</i> strains	78
Figure 25 Growth curve of the naringenin producing <i>S. cerevisiae</i> Y-28 strains.	79
Figure 26 Multiple amino acid sequence alignment of CHS genes.....	82
Figure 27 Naringenin production in BY4741 <i>S. cerevisiae</i> , expressing enzymes <i>M. truncatula</i> 4CL and mutated <i>V. vinifera</i> CHS.	84
Figure 28 Growth curve of the naringenin producing <i>S. cerevisiae</i> BY4741 strains, using mutated CHS.....	85
Figure 29 Comparison of the Y-28 <i>S. cerevisiae</i> strain`s naringenin production.....	94
Figure 30 Naringenin and <i>p</i> -coumaric acid production of Y-28 <i>S. cerevisiae</i> strains utilizing galactose as a carbon source	96
Figure 31 Growth curve of the naringenin producing Y-28 <i>S. cerevisiae</i> strains, utilizing galactose as carbon source.....	97

Figure 32 Naringenin and <i>p</i> -coumaric acid production of Y-28 <i>S. cerevisiae</i> strains utilizing sucrose as a carbon source	98
Figure 33 Growth curve of the naringenin producing Y-28 <i>S. cerevisiae</i> strains, utilizing sucrose as a carbon source	100
Figure 34 Comparison of naringenin production in strains BY4741, Y-28 utilizing glucose, Y-28 utilizing sucrose and Y-28 utilizing sucrose and glycerol as carbon source	103
Figure 35 Growth curve of the naringenin producing <i>S. cerevisiae</i> Y-28 strains utilizing 1% sucrose and 1% glycerol as the carbon source.....	104
Figure 36 Naringenin production using okara as nitrogen source.	106
Figure 37 Growth curve of naringenin production using okara as nitrogen source.....	107
Figure 38 Overall workflow of this thesis	110
Figure 39 Key achievements of this thesis	112
Figure 40 Tertiary structure of potential 4CL-CHS bifunctional protein created by in-frame gene fusion	115

LIST OF TABLES

Table 1 Metabolic engineering of <i>E. coli</i> and <i>S. cerevisiae</i> for production of <i>p</i> -coumaric acid.....	17
Table 2 Production of flavanones in <i>E. coli</i> and <i>S. cerevisiae</i>	22
Table 3 Sequence blast results for the 4CL gene from different sources	45
Table 4 Sequence blast results for the CHS gene from different sources.....	45
Table 5 Amino acid sequence comparison results for the 4CL gene from different sources	46
Table 6 Amino acid sequence comparison results for the CHS gene from different sources.	46
Table 7 Percentage identity for 4CL protein tertiary structures.	47
Table 8 Percentage identity for CHS protein tertiary structures.....	47
Table 9 The effect of expressing six different gene combinations on naringenin production in the <i>S. cerevisiae</i> BY4741 strain.....	52
Table 10 Site-directed mutagenesis primers..	67
Table 11 Comparison of the three key metabolites in the naringenin producing and non-producing strains.....	69
Table 12 Sequence blast results for the CHI gene from different sources.....	71
Table 13 Amino acid sequence comparison results for the CHI gene from different source..	72
Table 14 Percentage identity for CHI protein tertiary structures.....	72
Table 15 Six gene combinations expressed in the <i>S. cerevisiae</i> Y-28 strain.....	77

Table 16 Obtained naringenin yield produced in the <i>S. cerevisiae</i> Y-28 strains grown on sucrose and glycerol solution as a carbon source.	101
--	-----

LIST OF ABBREVIATIONS

3GT	Anthocyanidin 3-O-glucosyltransferase
4CL	4-Coumarate: coenzyme ligase
4HPA3H	4-Hydroxyphenylacetate 3-hydroxylase
AA	Amino Acid
ACC	Acetyl-CoA carboxylase
AMP	Adenosine monophosphate
ANS	Anthocyanin synthase
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
C3H	p-Coumarate 3-hydroxylase
C4H	Cinnamate 4-hydroxylase
CHI	Chalcone isomerase
CHR	Chalcone reductase
CHS	Chalcone synthase
COMT	Catechol-O-methyltransferase
CPR	Cytochrome P450 reductase
DAHP	3-Deoxy-D-arabinoheptulosonate 7-phosphate
DFR	Dihydroflavonol 4-reductase
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EFSA	European Food Safety Authority

F3'5'H	Flavonoid 3',5'-hydroxylase
F3'H	Flavonoid 3'-hydroxylase
F3H	Flavanone-3-hydroxylase
FCC	Frozen competent cell
FDA	U.S. Food and Drug Administration
FLS	Flavonol synthase
FMO	Flavin-containing monooxygenase
FNS	Flavone synthase
FSI	Flavone synthase I
G418	Geneticin
GAPD	Glyceraldehyde 3-phosphate dehydrogenase
GC-MS	Gas chromatography–mass spectrometry
GDH1	NADPH-dependent glutamate dehydrogenase
GDH2	NAD ⁺ -linked glutamate dehydrogenase
GLN1	Glutamine synthetase
GLT1	NADH-dependent glutamate synthase
GRAS	Generally regarded as safe
HpaBC	<i>E. coli</i> hydroxylase complex
HPLC	High performance liquid chromatography
HRP	Horseradish peroxidase
ICS	Isochorismate synthase
IFS	Isoflavone synthase
IPL	Isochorismate lyase
KS	β -keto synthase

LAR	Leucoanthocyanidin reductase
LB	Luria-Bertani broth
LDL	Low-density lipoprotein
MSTFA	N-methyl-N-(trimethylsilyl)-trifluoroacetamide
NCBI	National Center for Biotechnology Information
PAL	Phenylalanine ammonia lyase
PCR	Polymerase chain reaction
PDC	Pyruvate decarboxylase
PDH	Pyruvate dehydrogenase
PGK	Phosphoglycerate kinase
PKS	Polyketide synthase enzymes
PUFA	Polyunsaturated fatty acid
PVDF	Polyvinylidene difluoride
RNA	Ribonucleic acid
SDS-PAGE	Sodium dodecyl sulphate - Polyacrylamide gel electrophoresis
ssDNA	Single-stranded DNA
STS	Stilbene synthase
TAL	Tyrosine ammonia lyase
TCA cycle	Citric acid cycle
TMB	3,3',5,5'-Tetramethylbenzidine
TMCS	Trimethylchlorosilane
UDP	Uridine diphosphate

THESIS SUMMARY

Mankind has evolved from hunting and gathering lifestyle to be a good and service provider society. This cultural and social evolution has demanded fast developments in food science and technology. However dietary risk was still a major death cause in 2017. Functional foods were designed to tackle these dietary risks. Natural plant secondary metabolites, such as flavonoids have great potential in the food industry as functional food components.

Flavonoids, a sub-group of the phenolic compounds have high antimicrobial and antioxidant activities, which presents the opportunity to apply them as food additives and preservatives. Metabolic engineering of *Saccharomyces cerevisiae* is a sustainable flavonoid production technique. However, current naringenin production yield is low due to inefficient enzymatic activity.

This thesis has constructed the naringenin pathway in *S. cerevisiae* BY4741 strain and used gene source screening as a tool to identify the best gene source for enzymes 4-coumarate: coenzyme ligase (4CL) and chalcone synthase (CHS). For the first time, the 4CL gene from *Medicago truncatula* and CHS gene from *Vitis vinifera* were expressed in *S. cerevisiae* and this combination provided the highest yield of naringenin. It was 28-fold higher as compared to the reference strain. The results have demonstrated, that gene source selection has a high impact on naringenin production in *S. cerevisiae*.

Metabolite profiling has studied the cellular and enzymatic behavior during naringenin production. Three metabolites, L-aspartate, glycine and serine were identified

as important cellular compounds. Furthermore, site-directed mutagenesis of CHS demonstrated the significance of amino acids Glu61, Lys151, Ser184 and the Ser388-Ala389-Pro390 triplet. Mutation of Glu61/Asp61 and Ser184/Ala184 have significantly decreased naringenin production, due to their location at the CoA-binding tunnel and at the coumaroyl-binding pocket.

A genetically modified Y-28 *S. cerevisiae* strain was also utilized for the expression of the naringenin pathway. This resulted in an increased yield with the highest naringenin production reaching 28.68 mg/L. In addition, the similarity pattern of the combinations' production yields was compared between the BY4741 and Y-28 strains and were found to be matching. Alternative carbon and nitrogen sources were compared and analyzed in order to decrease the cost of culturing media. The 1% sucrose and 1% glycerol mixture was found to be a potential carbon source. Furthermore, okara was identified as an ideal nitrogen source.

The results demonstrated that the selection and combination of enzymes from the correct gene source could greatly improve naringenin production. For the future, this could help commercialize flavonoid production, which would result in natural food preservatives and additives.

1. INTRODUCTION

1.1 Background

Food has been an important aspect of human civilization since the existence of mankind. From a hunting and gathering lifestyle, we have evolved to be a good and service provider society. This cultural and social evolution have demanded fast developments in food science and technology. Food preservation, longer shelf-life, proper nutrition have topped the list of research areas in the food industry ¹. These evolutions have been on-going for the past hundreds of years. However, during the past few years, the development has significantly leaped forward due to the advancements in other science and engineering fields such as microbiology, biotechnology, chemistry, biology and physics. The combination of new knowledge has enabled the recent fast growth in the food industry ².

During the first civilizations food availability was a big concern. In contrast today's society incorporated food into the social, emotional and mental aspects of our life. However, food is still an important death cause risk factor. Looking at the main death causes in the world from 2017 (Figure 1) cardiovascular diseases, cancer, digestive diseases, diarrheal diseases and diabetes are in the top 10 positions ³. The common factor in these diseases and sicknesses is the main risk factor, which is dietary risk. Dietary risk was found to be the biggest risk factor for death causes in 2017 (Figure 2). More than 19% of all death is due to dietary related problems. The list also includes other food-related risk factors, such as high fasting plasma glucose, high LDL, high body-mass index and malnutrition ³. This information shows that even though society perceives food as a comfort and enjoyment source, in reality, it is still highly influential on our health and survival.

Figure 1 Main death causes in the world from 2017. Pie chart representation of the major death causes based on data collected from IHME³.

Figure 2 Main risk factors for death causes from 2017. Bar chart representation of main risk factors based on data collected from IHME³.

The main dietary risks are originating from high amount of sodium consumption and low level of whole grain, fruit, nuts, vegetables and omega 3 intakes (Figure 3)³. Salt plays an important flavoring role in processed food, which is responsible for 75% of the global sodium consumption. Additionally our sodium intake is also the consequence of the high amount of salt-based food preservatives⁴. Functional foods were designed to tackle these problems. Functional foods emerged in the 1980s and they are highly beneficial against the dietary risk factors due to their special functions. Functional foods are food products incorporating a special constituent which has beneficial health properties. These properties can improve general health, decrease the risks to develop certain diseases and help to cure specific illnesses. However, as functional food has become available and widespread, new harmful products, competing with the same title had appeared.

Figure 3 Main dietary risk factors for death causes from 2017. Column chart representation of main dietary risk factors based on data collected from IHME³.

The new trend for health-conscious mindset and health-promoting diets from the society has not only intrigued scientists but businesses as well. In order to gain higher sells, products promoted as highly health beneficial often contained only minimal amount of beneficial compounds. Lack of labeling rules often allows companies to exaggerate their potential advantages. This led to a confusing market, where consumers required to do their own investigation to find the ideal food products.

Food agencies have developed approval policies in order to scientifically prove the beneficial constituents and eliminate the usage of harmful compounds. The U.S. Food and Drug Administration (FDA) and the European Food Safety Authority (EFSA) are the main food agencies with the aim of protecting the public health by ensuring the security of drugs, and the safety of the food supply ⁵. They successfully identified several harmful food additives and banned their usage. However, in several cases, the approval decisions differ between the agencies as their approving criteria are different. For example, there are ongoing discussions between the scientist community and EFSA to progress with antioxidant related claims ⁶⁻⁷. While FDA had already passed an antioxidant claim for dietary supplements. Even though the amount of compounds, which can use the term “antioxidant” is very limited ⁸.

Both agencies have raised concerns about artificial food additives, which had been proven to be the underlying reason for allergies, asthma and anaphylaxis ⁹⁻¹⁰. Addressing this issue, natural food additives were found to be favored over the synthetic ones and researches on natural food additives have increased. Phenolics were an ideal target group, as they are a diverse group of plant secondary metabolites with high antimicrobial and antioxidant activities. They are mainly found in vegetables, fruits, and cereals. Furthermore,

phenolics-enriched diets were discovered to be highly preventative towards diseases such as cancers, diabetes and heart disease ¹¹⁻¹².

Flavonoids, a sub-group of the phenolic compounds, are an especially interesting and widely available plant secondary metabolites. The most common phenolics in food are flavonoids along with phenolic acids ¹³. In plants, flavonoids provide protection against UV radiation, physical damage, and infections ¹⁴. While in humans, their anti-oxidant, anti-allergenic, anti-inflammatory and anti-cancer properties have been showed to help prevent the development of neurodegenerative and cardiovascular diseases, cancer, Parkinson's disease, obesity, diabetes and aging¹⁵⁻¹⁷. Flavonoids have great potential as food additives, and they can be used in food or beverages to increase their nutritional value and create a value-added product¹². Furthermore, flavonoids possess highly active anti-microbial properties, which presents the opportunity to apply flavonoids as food preservatives.

1.2 Aims and objectives

Even though there is a significant potential for the use of flavonoids in the food industry, the current production, which is mainly based on plant extraction, has many limitations. The amount of available plants is restricted by small plantation sizes, unfavorable climate, long growing time and cycles. At the same time, the concentration of flavonoids is relatively low in these plants¹⁸. Chemical synthesis is an alternative flavonoid production method, but due to the fact that certain flavonoids' structures are complex, the synthesis is expensive, time and energy demanding¹⁹.

A potential alternative for sustainable production of flavonoids is the application of metabolic engineering on food-grade microorganisms, such as *Saccharomyces cerevisiae*. In this production method, a newly established enzymatic system in *S. cerevisiae* mimics the plants' flavonoid producing pathway, and reproduces it, utilizing the already existing precursors in the host microorganism²⁰. Metabolic engineering food-grade microorganisms provides a faster, energy and labor efficient, environmental-friendly production process as compared to the available alternative production techniques.

There had been several studies conducted on flavonoid pathway construction in *E. coli* and the pathway had also been successfully reconstructed in *S. cerevisiae*²¹. Several strategies had been applied to create a commercially applicable production yield. These strategies include malonyl-CoA and acetyl-CoA supply optimization, increased precursor, phenylalanine and tyrosine production and reduction of by-product formation²². Overall, the improvement in the obtained flavonoid yields was not satisfying. However, one key aspect of efficient metabolic engineering production is the high enzyme activity. This factor had not been studied in regards of the flavonoid pathway in *S. cerevisiae*. As non-native genes are expressed in the host microorganisms, compatibility problems and expression difficulties can arise. In order to avoid these issues, non-compatible genes need to be screened out and the ideal genes need to be identified.

This thesis aims to establish gene source selection as a metabolic engineering tool and demonstrates its effect on the flavonoid production in *S. cerevisiae*. With the above aims in mind, a systematically and comprehensive gene source screening was performed for enzymes 4-coumarate: coenzyme A ligase (4CL) and chalcone synthase (CHS), which are responsible for naringenin production. Four 4CL and four CHS gene sources have been

chosen as the main component for this research and have been expressed in two *S. cerevisiae* strains (Figure 4). In order to enhance the naringenin production, chalcone isomerase (CHI) gene source screening and culture condition optimization based on carbon and nitrogen source were performed. Furthermore, metabolite profiling and site-directed mutagenesis had provided novel information about the effect of gene source selection on the cell function.

Figure 4 Graphical presentation of this thesis's main aim ²³. Investigating the effect of gene source screening on naringenin production in *S. cerevisiae*.

1.3 Outline of the thesis

The thesis has been subdivided into the following sections. Following introduction, a literature review is presented in Chapter 2. The review focuses natural phenolics produced by metabolic engineering for functional food applications, emphasizing flavonoids and their precursors, phenolic acids. Additionally, the review incorporated the recent progress for microbial production of phenolic compounds in *S. cerevisiae* and *Escherichia coli*. Chapter 3 introduces the molecular cloning and expression methods used for pathway development. Moreover, analyzation techniques such as Western Blot, HPLC and GC-MS settings and procedure details are described along with the required extraction methods.

Chapter 4 demonstrates the construction of the naringenin pathway in *S. cerevisiae* BY4741 strain by the expression of enzymes 4CL and CHS. Gene source selection was performed targeting enzymes with higher enzymatic activities. The chosen 4CL gene sources were *Arabidopsis thaliana*, *Medicago truncatula*, *Petroselinum crispum* and *Streptomyces coelicolor*. The CHS gene sources were *A. thaliana*, *Glycyrrhiza uralensis*, *Hypericum androsaemum* and *Vitis vinifera*. HPLC analysis detected the naringenin production and identified a unique combination as the highest producer. The results had demonstrated, that gene source selection has a high impact on naringenin production in *S. cerevisiae*.

Chapter 5 investigates the cellular and enzymatic behavior during naringenin production in *S. cerevisiae* by metabolite profiling. Three metabolites were identified as important cellular compounds during naringenin production. Additionally, this chapter focuses on the optimization of the naringenin production by performing gene source selection for enzyme CHI and expressing it along with enzymes 4CL and CHS.

Furthermore, CHS was mutated by site-directed mutagenesis that identified four key amino acids in the *V. vinifera* CHS sequence that has important structural and functional properties.

Chapter 6 focuses on the enhancement of naringenin production. A new Y-28 *S. cerevisiae* strain was utilized for the expression of the naringenin pathway as an alternative host microorganism. The effect of gene source selection in a different *S. cerevisiae* strain was also explored. Furthermore, culturing conditions were optimized by substituting the carbon and nitrogen sources in the culturing media. The findings had led to higher naringenin production rate and lower production cost.

Lastly, Chapter 7 summarizes the key results of this study, followed by the identification of potential future works.

2. LITERATURE REVIEW

2.1 Introduction

Recently, the consumption behavior of general population has shifted towards a more health-conscious behavior. Healthier lifestyle, value-added products, and health-boosting foods and drinks, with potential to lower the risk of specific illness, are becoming increasingly popular²⁴⁻²⁵. On the other hand, food quality suffers easily from oxidation and microbiological spoilage, which would lead to changes in food color, odor, texture, and flavor, and give rise to food poisoning. Food preservatives provide an effective solution to address this food-safety issue. However, current chemical preservatives are associated with potential health risks, such as allergy and cancer²⁶⁻²⁷. Therefore, the search for functional food additives as well as natural, effective, and non-toxic food preservatives have aroused great interest¹².

Phenolic compounds are a diverse family of plant secondary metabolites found in fruits, vegetables, and oats. In plants, phenolics are synthesized by the phenylpropanoid, pentose phosphate, and shikimate pathways. Based on their carbon skeleton, different phenolics can be categorized into the following groups: flavonoids, tannins, stilbenes, lignans, phenolic acids, curcuminoids, coumarins, and quinones (Figure 5)²⁸. Phenolics play important roles in the growth and reproduction system, and can also protect the plants against ultraviolet radiation and aggression from pathogens or predators²⁹. In recent years, it was discovered that phenolics-enriched diets can help to prevent a broad array of diseases, like cancers, cardiovascular diseases, diabetes, neurodegenerative diseases, heart sickness and aging¹¹⁻¹². Especially the flavonoid compounds attracted attention, as isoflavonoids could be used

for cancer chemotherapy, proanthocyanidins could tackle urinary tract infections and some flavonoids possess anti-diabetic properties ³⁰.

Figure 5 Biosynthesis paths and classification of phenolic compounds. ³¹ Metabolic pathways connecting the shikimate pathways to different phenolic compounds, classification shown by different colored boxes.

Anthocyanins could relieve the oxidative stress related to chronic diseases and decrease cardiovascular related death ³², and ferulic acid was officially classified as an antioxidant in the list of food additives by FDA ³³. Furthermore, phenolic compounds are reported to possess hepatoprotective activities to improve liver damage and treatment of digestive problems ³⁴, anti-proliferative, DNA scission inhibitory and anti-obesity activities ³⁵.

More interestingly, phenolic compounds have a great potential to be natural food preservatives as the result of their high antioxidant and antimicrobial activities. They can donate hydrogen to terminate the free radical production, to stop the chain reaction which would lead to cell destruction during oxidation. This process not only prevents the spoilage of vegetables and fruits to extend the food products' shelf-life but can also safeguard the human physique against the harm of the responsive oxygen species ¹². In addition, numerous *in vitro* assays and food tests have revealed their high antimicrobial properties, especially towards pathogenic strains ³⁶⁻³⁸.

Food consumption is the main way for humans to intake phenolic compounds. Fruits, vegetables and beverages are all rich in phenolics, especially coffee, green tea, and wine. The average phenolics intake from food is highly dependent on geological locations. Statistically, each Japanese consume an average of 1492 mg of phenolic compounds from food per day, while Brazilians consume 460 mg of phenolic compounds per day ³⁹⁻⁴¹. The most commonly available phenolics in food are phenolic acids and flavonoids, especially flavanols and anthocyanins ¹³. Apart from natural food sources, phenolic compounds are available from plant extraction and chemical synthesis for usage as supplements and food preservatives. In supplements, the total phenolic content can range between 60 to 200 mg/g and fruit extracts top the list ⁴². However, plants are subjected to long growth cycles and

are affected by the season and weather. Thus the low concentration of phenolic compounds obtained from plants, are not able to meet the demands of the food industry¹⁸. For chemical synthesis, many complex and tedious steps are required. The large number of chemicals used also causes environmental stress¹⁹. Metabolic engineering of microorganisms provides an alternative approach to sustainably produce phenolics from renewable feedstocks in larger amounts. The expression of heterologous enzymes participating in the plant phenolics biosynthetic pathway in microorganisms endows them the capacity to produce phenolic compounds. In addition, the faster growth rate of microorganisms, as compared to plants, makes them more appealing for commercial use⁴³. Furthermore, the use of microorganisms potentially results in less environmental damage and toxicity to humans compared to chemical biosynthesis methods⁴⁴.

Bacteria *E. coli* and yeast *S. cerevisiae* are the most used model organisms for metabolic engineering, due to their genetic tractability and mature high-density fermentation technologies. *S. cerevisiae* is especially suitable for food preservative production, as it is a generally regarded as safe (GRAS) microorganism and has been utilised in breweries and other food industries for thousands of years (bread and wine).

As the main aim of this thesis was to give a systematic and comprehensive gene source screening for naringenin production in *S. cerevisiae*, this review incorporated the recent progress for microbial production of flavonoids and their precursor, phenolic acids. As during the research, the pathway construction was performed in *E. coli*, the literature review also included the key compounds' metabolic engineering development in *E. coli*.

2.2 Phenolic acids

Phenolic acids are a class of substances which consist of a phenolic ring and at least one organic carboxylic acid function and are found in plants. They are the simplest class of phenolics and are the precursors for the synthesis of many others. As reported in many studies⁴⁵⁻⁴⁷, phenolic acids have antioxidant, antimicrobial, anti-inflammatory, antimutagenic, anticarcinogenic and other biological properties. For example, phenolic acids are able to inhibit many pathogenic bacteria and fungi, such as *Staphylococcus aureus*, *E. coli*, *Aspergillus parasiticus*, *Klebsiella pneumonia*, *Aspergillus flavus* and *Bacillus cereus*, making them attractive food preservatives^{12,48-49}. Advanced progress on their health beneficial properties and more details about their antimicrobial properties are comprehensively reviewed by Guzman et al.⁵⁰, Park et al.⁵¹, and Sova et al.⁵².

In the phenolic acid biosynthetic pathway, chorismate (the ending product of the shikimate pathway) or cinnamic acid (an intermediate of the phenylpropanoid pathway) are the main substrates (Figure 6). As per the quantity of carboxylic acids, natural phenolic acids are classified into two main groups: hydroxycinnamic acids (C6-C3) derived from cinnamic acid and hydroxybenzoic acids (C6-C1) derived from benzoic acid (Figure 6). The main hydroxybenzoic acids are: benzoic acids, salicylic acid, gallic acid, quinic acid, and shikimic acid; and the main hydroxycinnamic acids are: ferulic acid, coumaric acids and caffeic acid. Hydroxycinnamic acids are the precursors for flavonoid production. The status of the microbial synthesis using *E. coli* and *S. cerevisiae* for the hydroxycinnamic acids is presented below.

Figure 6 Biosynthetic pathways of diverse phenolic acids and the most effective strategies for their production improvement ³¹ Hydroxycinnamic acids are labeled in blue. Hydroxybenzoic acids are labeled in yellow. For improvement of phenolic acids production, up red arrows indicate the key precursors for enhancement; crosses indicate genes for knock-out; black boxes indicate key enzymes for overexpression.

2.2.1 *p*-Coumaric acid

p-Coumaric acid is the precursor for the biosynthesis of the downstream flavonoids, lignans, stilbenes, condensed tannins, flavonoids, and curcuminoids (Figure 5). It is generated from the compound phenylalanine via phenylalanine ammonia lyase (PAL) and cinnamate 4-hydroxylase (C4H, a membrane-bound P450 monooxygenase), or produced from tyrosine via tyrosine ammonia lyase (TAL) ⁵³⁻⁵⁵ (Figure 6). There have been many efforts in its synthesis in *E. coli* and *S. cerevisiae*, as summarized in Table 1. To date, the highest production of *p*-coumaric acid in *E. coli* was 974 mg/L obtained by Kang et al. ⁵⁶, and 1.93 g/L in *S. cerevisiae* achieved by Rodriguez et al ⁵⁷⁻⁵⁸.

In summary, the main effective strategies (Figure 6) are: 1) optimizing the conversion rate from phenylalanine or tyrosine to *p*-coumaric acid, by gene screening of PAL-4CL and TAL, or gene codon optimization. For this, PAL from *Rhodotorula rubra*, *A. thaliana*, *Populus kitakamiensis*, *Rhodotorula glutinis* were the most popular sources studied, while TAL from *Flavobacterium johnsoniae* and *Herpetosiphon aurantiacus* were found to have better performance in converting tyrosine ^{59 56, 60-63}; 2) Improving the supply of precursor phenylalanine or tyrosine, via overexpression of key enzymes engaged in the shikimate pathway, elimination of feedback inhibition, and knock-out of competing branches. For example, in *E. coli*, enzymatic catalysis by *tyrA*, *ppsA*, *tktA*, and *Aarⁿ* were found to be the main barrier in the shikimate pathway and aromatic acid biosynthetic pathway, in which *tyrA*, *aroF*, and *aroG* were feedback inhibited by tyrosine.

Table 1 Metabolic engineering of *E. coli* and *S. cerevisiae* for production of *p*-coumaric acid ³¹ Main strategies and obtained yields are collected based on target compounds and host microorganisms.

Target compound	Strategies	Host microorganism	Yield	Reference
<i>p</i> -Coumaric acid	Overexpression of <i>Rhodotorula rubra</i> PAL	<i>E. coli</i>		64
<i>p</i> -Coumaric acid	Substitution of PAL and C4H for TAL	<i>E. coli</i>	2.3 mg/L	59
<i>p</i> -Coumaric acid	Overexpression of codon-optimized TAL in tyrosine overproducing strain	<i>E. coli</i>	974 mg/L	56
<i>p</i> -Coumaric acid	Determination of TAL and PAL enzyme kinetics and expression in <i>E. coli</i>	<i>E. coli</i>	630 μ M	62
<i>p</i> -Coumaric acid	Overexpression of TAL, <i>aroG</i> ^{fbr} and <i>tktA</i> , in a <i>pheA</i> inactivated <i>E. coli</i> strain, which secretes p-CoA to the media	<i>E. coli</i>	22 mg/L	65
<i>p</i> -Coumaric acid	Optimization of yeast extract concentration in a TAL overexpressed <i>E. coli</i> strain	<i>E. coli</i>	475 mg/L	66
<i>p</i> -Coumaric acid	Comparison of TAL enzyme sources	<i>E. coli</i>	213 mg/L	60
<i>p</i> -Coumaric acid	Comparison of TAL, PAL and SeSam8 enzyme sources	<i>E. coli</i>	0.54 mM	63
<i>p</i> -Coumaric acid	Mechanism study of carbon allocation into phenylpropanoid metabolism	<i>S. cerevisiae</i>		61
<i>p</i> -Coumaric acid	Overexpression of PAL and optimization of precursor feed	<i>S. cerevisiae</i>		67
<i>p</i> -Coumaric acid	Determination of TAL and PAL enzyme kinetics and expression in <i>E. coli</i>	<i>S. cerevisiae</i>	200 μ M	62
<i>p</i> -Coumaric acid	Optimization of phenylalanine feed concentration	<i>S. cerevisiae</i>	650 μ M	68
<i>p</i> -Coumaric acid	Reduction of by-product formation with ARO10 and PDC5 knock-out, overexpression of feedback-resistant DAHP synthase and chorismate mutase, identification of the shikimate kinases importance in flux-controlling steps Aro2p, optimization of transporters	<i>S. cerevisiae</i>	1.93 g/L	58 57
<i>p</i> -Coumaric acid	Targeted metabolomics study and implementation via global engineering of central metabolism	<i>S. cerevisiae</i>		69
<i>p</i> -Coumaric acid	Identification of new enzymatic bottlenecks by combinatorial metabolic engineering	<i>S. cerevisiae</i>	21 mg/L	70

As reported in several studies, over-expression of *ppsA* and *tktA*, and introduction of feedback-inhibition-resistant *aroG^{fbr}* and *tktA^{fbr}* led to significant improvement of *p*-coumaric acid^{56, 62, 65-66}. Elimination of phenylalanine consumption, via knock-out of gene *pheA*, provided an effective complementary approach to further improve its production⁶⁵. Similarly, in *S. cerevisiae*, over-expression of its feedback-resistant 3-deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) synthases *ARO4^{K229L}* and chorismate mutase *ARO7^{G141S}* were performed to eliminate their feedback inhibition from tyrosine. The knock-out of phenylpyruvate decarboxylase *ARO10* and pyruvate decarboxylase 5 (*PDC5*) eliminated the by-product phenylethanol. These efforts strengthened the supply of tyrosine and obtained more *p*-coumaric acid^{57-58, 69-70}; 3) Discovery and elimination of potential limiting steps by omics analysis. In the study by Rodriguez, A, seven different key genes (*tat1*, *tpo1*, *alp1*, *agp3*, *ady2*, *gal2*, and *bap2*) were identified when comparing metabolic profiling with different *S. cerevisiae* strains. The various deletions improved *p*-coumaric acid production⁵⁷. For more detailed information, please refer to Table 1.

2.2.2 Caffeic and ferulic acids

In plants, *p*-coumaric acid is transformed to caffeic acid and ferulic acid by subsequent methylation reactions and hydroxylation, via a plant-specific cytochrome P450 dependent monooxygenase-*p*-coumarate 3-hydroxylase (*C3H*) and *O*-methyltransferase (*COMT*), respectively (Figure 6).

As the first enzyme, *C3Hs* belong to P450 enzymes in plants, it is difficult to be functionally expressed since they are unstable in bacteria. To address this issue, recent efforts focus on the expression of microbial hydroxylases in *E. coli* to produce caffeic and ferulic acids. The earliest work was reported by Choi et al.⁷¹, in which 4-coumarate 3-

hydroxylase originating from *Saccharothrix espanaensis* was introduced in *E. coli*. As a result, caffeic acid was detected (the titer was not reported) and ferulic acid was obtained (7.1 mg/L) when COMT from *A. thaliana* was expressed ⁷¹. As reported in another study, expressing C3H from *S. espanaensis* and COMT from *Medicago sativa* resulted in 28.8 mg/L of ferulic acid and 70 mg/L of caffeic acid ⁷². Additionally, P450 CYP199A2 from *Rhodopseudomonas palustris* was also found to possess hydroxylation activity of *p*-coumaric acid. Through the analysis of its crystal structure and use of site-directed mutagenesis, a mutant (F185L) with 5.5-fold higher activity was developed. A total of 2.8 g/L of caffeic acid was obtained with whole-cell catalysis process ⁷³. Besides, recently discovered hydroxylases are the native *E. coli* hydroxylase complexes, 4-hydroxyphenylacetate 3-hydroxylase (4HPA3H), and 4HPA3H from other bacteria (*Pseudomonas aeruginosa* and *Thermus thermophilus*) ⁷⁴.

The production of caffeic and ferulic acids were further increased by improving precursor supply like tyrosine, phenylalanine or *p*-coumaric acid. The most common strategies are over-expression of gene *tyrA^{fbr}*, *ppsA*, *tktA* and *aroG^{fbr}* for the synthesis of aromatic acids, knock-out of metabolic branches consuming tyrosine or phenylalanine, over-expression of optimized TAL (Figure 6) ^{56, 75-77}. For example, over-expression of TAL from *Rhodobacter capsulatus* along with *tyrA^{fbr}*, *ppsA*, *tktA*, and *aroG^{fbr}* resulted in the creation of 50.2 mg/L of caffeic acid, presented by Yan et al ⁷⁸. In another study, inhibiting *pheA* and *tyrA* and expressing *tyrA^{fbr}*, *ppsA*, *tktA*, and *aroG^{fbr}*, led to a total of 766 mg/L of caffeic acid produced from the precursors tyrosine and phenylalanine ⁷⁹.

2.3 Flavonoids

Flavonoids are a large unit of the phenolic compound group, made up of more than 8000 compounds. In plants, flavonoids provide protection against UV radiation, physical damage, and infections¹⁴. In recent years, it was shown they have antioxidant properties and can lower the danger of cardiovascular, Parkinson's and Alzheimer's diseases, cancer and diabetes¹⁶. Their antimicrobial activity was demonstrated using assays with flavonoid-rich plant extracts and pure flavonoids, which inhibited *E. coli*, *Salmonella enteritidis*, *S. aureus* and *Micrococcus luteus*⁸⁰. Additionally, quercetin and apigenin were identified as potential DNA gyrase inhibitors⁸¹. Some flavonoids have great antifungal activity. For example, galangin can inhibit *Cladosporium sphaerospermum*, *Aspergillus tamaritii*, *Penicillium digitatum*, *A. flavus* and *Penicillium italicum* and coxsackie B virus⁸². Additionally, quercetin, dihydroquercetin, leucocyanidin, dihydrofisetin, catechin and pelargonidin chloride can be helpful against herpes simplex virus, respiratory syncytial virus, Sindbis virus and poliovirus³⁶.

Flavonoids are synthesized by the phenylpropanoid pathway, using starting precursor phenylalanine or tyrosine. Through the conversion of *p*-coumaric acid, naringenin is produced, which is at the central branch point of the pathway. The main branches are flavones, flavanones, flavones, flavanols, anthocyanins, and isoflavonoids (Figure 5)¹⁸. The research status regarding the biosynthesis of these above compounds in microbes is discussed. Their glycosylated, methylated, hydroxylated and prenylated forms are in the reviews by Pandey et al.⁴³ and Cao et al.⁸³.

2.3.1 Flavanones

Flavanones is the most important flavonoid group, as they are the precursors for many other flavonoids. They are generally responsible for preventing cardiovascular, neurodegenerative and cancer diseases ⁸⁴. Additionally, eriodictyol could reduce the risk of diabetes ⁸⁵. The main flavanones which are well studied are pinocembrin, naringenin, eriodictyol, and liquiritigenin ⁴³. The endeavors in the generation of these compounds are condensed in Table 2 and are discussed below.

2.3.1.1 Pathway construction

Pathway construction is commonly the first step for heterologous biosynthesis of biochemicals. To date, the pathway for biosynthesis of pinocembrin, naringenin, eriodictyol, and liquiritigenin are well studied ⁸⁵⁻⁸⁷. Naringenin is synthesized through the phenylpropanoid pathway, under the catalysis of PAL, C4H, 4CL, CHS and CHI using phenylalanine as the substrate (Figure 7). Naringenin can also be synthesized by TAL ⁵³, 4CL, CHS and CHI from tyrosine ⁵⁹. For the production of pinocembrin, 4CL from *S. coelicolor* was found to convert cinnamic acid into cinnamoyl-CoA, which sidesteps the C4H enzyme for the generation of pinocembrin chalcone ⁶⁴.

Table 2 Production of flavanones in *E. coli* and *S. cerevisiae*³¹ Main strategies, substrates and obtained yields are collected based on target compounds and host microorganisms.

Target compound	Substrate	Strategies	Host microorganism	Yield	Reference
naringenin	tyrosine and phenylalanine	Overexpression of three gene clusters with different promoter location	<i>E. coli</i>	0.27 µg/L	64
pinocembrin				0.17 µg/L	
naringenin	tyrosine and phenylalanine	Substitution of PAL with TAL	<i>E. coli</i>	20.8 mg/L	59
naringenin	Tyrosine	Fermentative production of flavanones	<i>E. coli</i>	57 mg/L	88
pinocembrin	phenylalanine			58 mg/L	
pinocembrin	<i>p</i> -coumaric acid	Metabolic engineering of the central metabolism by the enhancement of the malonyl-CoA availability, analysis of macromolecule complexes, amplification of acetate assimilation pathways, engineering alternative carbon assimilation pathway and inhibition of competitive metabolic pathways	<i>E. coli</i>	710 mg/L	89 90
naringenin				186 mg/L	
eriodictyol				54 mg/L	
naringenin	<i>p</i> -coumaric acid			33.7 mg/L	
liquiritigenin				16.9 mg/L	
pinocembrin	cinnamic acid	Analysis of four-step metabolic circuit recombinant strains in shake flasks	<i>E. coli</i>	28.5 mg/L	91
7-hydroxyflavanone				1.9 mg/L	
eriodictyol	caffeic acid			5.2 mg/L	
butin				4.2 mg/L	
naringenin	<i>p</i> -coumaric acid			118.9 mg/L	
liquiritigenin		Enhanced intracellular malonyl-CoA and higher cytochrome P450 activity	<i>S. cerevisiae</i>	13.5 mg/L	92
pinocembrin	cinnamic acid			57.2 mg/L	
eriodictyol	caffeic acid			7.9 mg/L	
eriodictyol	tyrosine		<i>E. coli</i>	107 mg/L	
eriodictyol	naringenin	Optimization of culturing media, induction point and temperature with substrate delay time	<i>E. coli</i>	62.7 mg/L	93
pinocembrin	phenylalanine	Overexpression of fatty acid synthases to enhance malonyl-CoA availability	<i>E. coli</i>	29.9 mg/L	94

pinocembrin	Glucose	Optimization of metabolic balance by modular and stepwise modular metabolic strategy. Identification of ideal gene copy number, promoter strength and pH. Construction of clustered regularly interspaced short palindromic repeats interference (CRISPR) for malonyl-CoA enhancement.	<i>E. coli</i>	525 mg/L	95-97
naringenin	Glucose	Optimization of enzyme gene sources and expression levels in a tyrosine overproducing strain	<i>E. coli</i>	84 mg/L	60
naringenin	<i>p</i> -coumaric acid	Increase in intracellular malonyl-CoA availability by an integrated computational and experimental technique	<i>E. coli</i>	474 mg/L	98
naringenin	<i>p</i> -coumaric acid	Enhancement of intracellular malonyl-CoA other cofactor supply by a cipher of evolutionary design	<i>E. coli</i>	215 mg/L	99
eriodictyol				114 mg/L	
naringenin	<i>p</i> -coumaric acid	Development and application of ePathBrick vectors	<i>E. coli</i>	68 mg/L	100
naringenin	<i>p</i> -coumaric acid	Development and application of a CRISPathBrick vector and pCRISPReporter	<i>E. coli</i>	18.9 mg/L	101
naringenin	Glucose	Mutation of targeted locations across the genome and application of sensory proteins for chemical concentration optimization	<i>E. coli</i>	61 mg/L	102
naringenin	<i>p</i> -coumaric acid	Enhancement of malonyl-CoA by synthetic antisense RNAs	<i>E. coli</i>	91.3 mg/L	103
naringenin	Glucose	Enhancement of malonyl-CoA by CRISPR interference system and antisense RNA	<i>E. coli</i>	421 mg/L	104-106
naringenin	Glucose	Optimization of metabolic balance by modifying plasmid gene copy numbers and promoter strengths			
naringenin	Glucose	Optimization of co-culture system by IPTG induction timing, inoculation ratio, carbon source and strain selection	<i>E. coli</i> – <i>E. coli</i>	41.5 mg/L	107
naringenin	Methanol	Utilization of methanol as a substrate	<i>E. coli</i>		108
naringenin	Tyrosine	Overexpression of PAL, 4CL and CHS in an AH22	<i>S. cerevisiae</i>	7 mg/L	67
Pinocembrin		and pad1 knock out strain		0.8 mg/L	

naringenin	<i>p</i> -coumaric acid	Overexpression of C4H, 4CL, CHS and CHI	<i>S. cerevisiae</i>	28.3 mg/L	109
pinocembrin	cinnamic acid			16.3 mg/L	
eriodictyol	caffeic acid	Optimization of phenylalanine feed concentration	<i>S. cerevisiae</i>	6.5 mg/L	68
naringenin	<i>p</i> -coumaric acid phenylalanine			15.6 mg/L 8.9 mg/L	
naringenin	glucose	Development of yeast chassis strain with alleviating feedback inhibition and reduced byproduct formation	<i>S. cerevisiae</i>	108 mg/L	21 110
naringenin	glucose	Reduction of by-product formation, overexpression of feedback-resistant genes,	<i>S. cerevisiae</i>	1.55 mg/L	111
liquiritigenin		identification of important flux-controlling genes		5.31 mg/L	
naringenin	glycerol and sucrose	Application of modified GAL system, enhancement of malonyl-CoA supply, elimination of feedback inhibition and down- regulation of competing pathways	<i>S. cerevisiae</i>	90 mg/L	18

Leguminous plants were found to have a chalcone reductase (CHR), which can catalyze the intermediate of the CHS reaction to yield 6'-deoxychalcone. Over-expression of CHI from *M. sativa*, CHS from *Petunia hybrida* or *M. sativa*, 4CL from *P. crispum* and CHR from *M. sativa* reportedly obtained 16.9 mg/L of liquiritigenin from *p*-coumaric acid and 5.2 mg/L eriodictyol from caffeic acid in *E. coli*, as well as 13.5 mg/L of liquiritigenin and 7.9 mg/L of eriodictyol in *S. cerevisiae* ⁹¹. For biosynthesis of eriodictyol, flavonoid 3'-hydroxylase (F3'H) - cytochrome P450 reductase (CPR) complex (from *P. hybrida* and *Catharanthus roseus*), as well as *E. coli* hydroxylase complex (HpaBC), provide another feasible approach, which uses naringenin as the substrate (Figure 7) ⁹³.

Figure 7 Biosynthetic pathways of flavonoids and the most effective strategies for their production improvement. ³¹ Detailed biosynthesis branches for malonyl-CoA and relative regulation are present in the red box in the bottom right corner. Up red arrows and black boxes indicate the key precursors and key enzymes for production of flavonoids respectively, and red crosses indicate the main enzymes, in competing branches, for the knockout to increase flavonoid accumulation.

2.3.1.2 Improving malonyl-CoA supply

Malonyl-CoA is the precursor of flavanones (Figure 7). Its low concentration is a bottleneck for flavonoid production. Hwang et al. reported that over-expression of acetyl-CoA carboxylase (ACC), which catalyzes the synthesis of malonyl-CoA from acetyl-CoA, showed significant improvements in the pool of malonyl-CoA (Figure 7)⁸⁸⁻⁸⁹. Another effective approach was the establishment of a recombinant malonate assimilation pathway from *Rhizobium trifolii*, which enabled the conversion of supplemented malonate to malonyl-CoA under the catalysis of *matB* and *matC* (Figure 7)⁹⁷. However, these methods gave rise to an extra metabolic load to the cells and required supplementation of malonate. Therefore, other strategies focused on genomic regulation. In *E. coli*, fatty acid biosynthesis is the only metabolic pathway consuming malonyl-CoA. This is a bottleneck which prevents efficient flavonoids accumulation. Increasing the malonyl-CoA supply have been reported. This was achieved by the inhibition of the fatty acid biosynthesis with the supplementation of the fatty acid synthesis inhibitor cerulenin (Figure 7)^{60, 94}, or repressing the genes *FabD* and *FabB/FabF* (encoding for β -ketoacyl-ACP synthase II) responsible for fatty acid synthesis via the synthetic antisense RNAs^{90, 103, 106}. In addition, acetyl-CoA serves as the precursor for malonyl-CoA, which is consumed by the TCA cycle and glycolysis. Simultaneous deletion or silencing of the related genes *eno*, *mdh*, *adhE*, *fumC* and *sucC* which are involved in the TCA cycle and glycolysis efficiently channeled carbon flux towards malonyl-CoA and generated higher amount of naringenin or pinocembrin^{96, 105}.

In addition to macro-control, Xu et al.⁹⁸ displayed a combined computational and experimental approach to boost the malonyl-CoA supply in *E. coli*. It was shown that the

OptForce methodology¹¹² was used to forecast the lowest number of genetic changes for improved yield of malonyl-CoA. Thus, the knock-out of *fumC* (encoding fumarate hydratase) and *sucC* (encoding succinyl-CoA synthetase), and overexpression of target ACC, pyruvate dehydrogenase (PDH), glyceraldehyde 3-phosphate dehydrogenase (GAPD), and phosphoglycerate kinase (PGK) synergistically channeled carbon flux towards malonyl-CoA and obtained a sum of 474 mg/L of naringenin⁹⁸. Similarly to the OptForce methodology, a CiED program⁹⁹ was developed by the same group to identify key genetic targets involved in the malonyl-CoA synthesis. The engineered *E. coli* strain, which overexpressed ACC and *BPL* and had deletions of TCA cycle and glycolysis genes, *sdhA*, *adhE*, *brnQ* and *cite*, produced 114 mg/L eriodictyol and 215 mg/L naringenin⁹⁹.

2.3.1.3 Comprehensive regulations

In addition to regulating the precursor supply, studies were performed to further optimize the metabolism of target flavanones. For example, a novel sensor-directed evolution of the whole endogenous pathway of naringenin was performed in *E. coli* using MAGE for mutagenesis¹¹³, the *ttgR* gene (encoding HTH-type transcriptional regulator) was used as the sensor for naringenin, and *TolC*, outer membrane channel was used as the selector (sodium dodecyl-sulfate was used for positive selection, and colicin E1 was used for negative selection). After 4 rounds of evolution, 36 times more naringenin was produced as compared to the parental strain¹⁰². In another study, modular engineering was employed to balance the different modules, by arranging copy numbers and promoters so as to maximize the metabolism of naringenin¹⁰⁴. More recently, in the study by Ganesan et al.¹⁰⁷, an *E. coli* - *E. coli* co-culture system was introduced for naringenin synthesis. This included a tyrosine over-producing strain¹¹⁴ which expressed TAL from *R. glutinis*, and a

second strain containing the TAL and 4CL genes from *P. crispum*, CHS from *P. hybrida* and CHI from *M. sativa*. The co-culture system improved the production of naringenin compared to the monoculture, due to the reduction of metabolic stress and flexible adjustment between the two engineered strains ¹⁰⁷.

The engineering works in *S. cerevisiae* are fewer as compared to *E. coli*. In addition to the over-expression of key enzymes, as well as the elimination of feedback inhibition from tyrosine, the most effective strategies for pathway optimization were to inhibit the competing branches, such as phenyl ethanol and phloretic acid. The knock-out of active PDC5 and ARO10 resulted in a 3 times higher naringenin production by repressing the phenylethanol branch ²¹. Further, the replacement of TSC13 (responsible for fatty acid and phloretic acid production) with ECR from *Musca domestica* was a useful approach to decrease the production of the side product phloretic acid ¹¹⁰.

2.3.2 Flavones

Flavones is another group of flavonoids with beneficial health properties for humans. The main compounds well studied are apigenin, luteolin, and chrysin. Apigenin has high anticancer and anti-viral activity, luteolin possesses anti-inflammatory properties, while chrysin has anti-aging activity ⁴³.

As reported in several studies, an introduction of flavone synthase (FNS) could enable microbes to produce apigenin ¹¹⁵⁻¹¹⁶. For example, expression of FNS from *P. crispum* along with 4CL from *P. crispum*, CHS and CHI from *P. hybrida*, resulted in a production of 415 µg/L of apigenin and 10µg/L of luteolin, with *p*-coumaric and caffeic acid as precursors ¹¹⁷. This strain was further optimized by inhibiting fatty acid synthesis. This was carried out by repressing the genes FabB and FabF, which resulted in 5 mg/L of

chrysin, 110 mg/L of apigenin and 4 mg/L of luteolin⁹⁰. Similarly, FNS from *P. crispum* was over-expressed in a naringenin and pinocembrin producing strain (Figure 7)⁸⁸, resulted in a production of 13 mg/L of apigenin and 9.4 mg/L of chrysin¹¹⁸⁻¹¹⁹.

2.3.3 Flavanols

Flavanols are mainly known for their anticancer properties¹²⁰. Additionally, they can also minimize the possibility of coronary heart disease and prevent age-induced vascular injuries¹⁶. The main flavanol compounds are kaempferol, galangin, quercetin, myricetin, catechin, afzelechin, and fisetin (Figure 7). Amongst these, kaempferol and quercetin are able to limit the risk of several chronic diseases, especially cancer. In addition fisetin has anti-inflammatory activity and can diminish the likelihood of Alzheimer's disease¹²¹⁻¹²².

The first flavanol synthesized in *E. coli* was kaempferol. This was obtained by expression of the flavonol synthase (FLS) from *Citrus unshiu* grown on naringenin as the substrate (Figure 7)¹²³. Flavanone-3-hydroxylase (F3H) is necessary for higher kaempferol production, which works together with FLS to convert more naringenin into kaempferol. In the study by Miyahisa et al., over-expression of F3H from *Citrus sinensis* and FLS in a naringenin producing strain⁸⁸ resulted in 15 mg/L of kaempferol from tyrosine and 1.1 mg/L of galangin from phenylalanine¹¹⁸. Production of kaempferol in a recombinant *S. cerevisiae* strain was demonstrated by over-expressing PAL, CPR, C4H, 4CL, CHS and CHI, and F3H & FLS^{68, 124}. In combination with gene screening, the supplement of *p*-coumaric acid and optimization of fermentation conditions, 66.3 mg/L of kaempferol was obtained in the engineered *S. cerevisiae* strain¹²⁵.

For biosynthesis of quercetin and myricetin, another two enzymes, flavonoid 3',5'-hydroxylase (F3'5'H) and CPR are required in addition to the above enzymes (Figure 7). The fusion of F3'5'H with CPR, from *C. roseus*, along with the expression of 4CL from *P. crispum*, CHS and CHI from *P. hybrida*, F3H from *M. domestica* and FLS from *A. thaliana* resulted in 300 µg/L of kaempferol and 50 µg/L of quercetin when grown on *p*-coumaric acid. With the supplementation of eriodictyol, 1100 µg/L of quercetin and 10 µg/L of myricetin were produced ¹²⁶.

Catechin and afzelechin synthesis have also been accomplished in an engineered *E. coli* strain. This was carried out with the expression of 4CL from *P. crispum*, CHS from *P. hybrida*, CHI from *M. sativa*, F3H from *M. domestica*, dihydroflavonol 4-reductase (DFR) from *Anthurium andraeanum* and leucoanthocyanidin reductase (LAR) from *Desmodium uncinatum* (Figure 7). With culture condition optimization, 8.8 mg/L of catechin was obtained from eriodictyol while 0.74 mg/L of afzelechin was produced from naringenin ¹¹⁶. Furthermore, catechin production was further improved by deletion of genes *pgi*, *ppc* and *pldA*, based on a CiED model. By using dihydroquercetin as a substrate, 36 mg/L of catechin was produced, with DFR from *A. andraeanum* and LAR from *D. uncinatum* ¹²⁷. In order to enhance catechin production, an ePathBrick system ¹⁰⁰ was proposed and applied to screen for the best multi-enzymatic pathway construction. The combination of the following enzymes was created: F3H from *C. sinensis*, *M. domestica* and *P. crispum*, DFR from *A. andraeanum*, *C. sinensis* and *Fragaria ananassa*, LAR from *C. sinensis* and *D. uncinatum*. The combination with the highest catechin (129-130 mg/L) contained F3H from *C. sinensis*, LAR from *D. uncinatum* and DFR from *C. sinensis*/*M. domestica*, by utilizing eriodictyol as the substrate. The strain containing DFR from *M. domestica* produced the

highest afzelechin (39.5 mg/L) from naringenin. With further optimization using gene copy-number and an introduction of scaffold strategies, 910 mg/L of catechin and 68 mg/L afzelechin were obtained in the *pgi* and *ppc* knock-out strain ¹²⁸. In addition, the *E. coli* hydroxylase complex, HpaBC, could likewise empower the transformation of afzelechin into catechin. With this approach, 34.7 mg/L catechin was obtained ⁹³.

2.3.4 Anthocyanins

Anthocyanins are known to be mainly responsible for the color of flowers. Additionally, they have been studied in connection with cardiovascular and neurodegenerative diseases, anti-cancer or anti-inflammatory activities ¹⁶. The main representatives are pelargonidin, cyanidin, and delphinidin ¹²⁹.

Anthocyanins are generated from naringenin under the catalysis of flavanone 3-hydroxylase (F3H), dihydroflavonol 4-reductase (DFR), anthocyanin synthase (ANS), and anthocyanidin 3-O-glucosyltransferase (3GT) (Figure 7). According to Yan et al. (2005), 5.6 µg/L of pelargonidin 3-O-glucoside and 6 µg/L of cyanidin 3-O-glucoside were obtained from naringenin and eriodictyol, by introducing F3H from *M. domestica*, DFR from *A. andraeanum*, ANS from *M. domestica* and 3GT from *P. hybrida* ¹³⁰. In the study, ANS was identified as the key enzyme for biosynthesis of anthocyanins. To eliminate this bottleneck, ANS enzymes from *Gerbera hybrida*, *Antirrhinum majus*, *P. hybrida*, and *M. domestica* were screened, amongst which *P. hybrida* was found to have the highest activity. The resultant pelargonidin 3-O-glucoside yield was improved to 0.98 mg/L, while the cyanidin 3-O-glucoside amount reached 2.07 mg/L, in combination with cofactor optimization. Furthermore, the addition of 3MaT from *Dahlia variabilis* was found to enable the synthesis of pelargonidin 3-O-6"-O-malonylglucoside and cyanidin 3-O-6"-O-

malonylglucoside. By overexpressing of *pgm*, *galU* (encoding phosphoglucomutase and glucose-1-phosphate uridylyltransferase), additionally fusing of 3GT and ANS, 78.9 mg/L of pelargonidin 3-*O*-glucoside and 70.7 mg/L of cyanidin 3-*O*-glucoside were obtained from afzelechin and catechin, respectively ¹³¹, in which pH was found playing important roles in the production of anthocyanins. More modifications have been performed, including the overexpression of transporter protein yadH, knockout of *tolC* ⁹⁰, 3GT and ANS fusion, and overexpression of phosphoglucomutase, CMP kinase and nucleoside-diphosphate kinase. The last manipulation resulted in 350 mg/L of cyanidin 3-*O*- glucoside ¹³². Even though the production of the anthocyanin has been achieved in *E. coli* as mentioned, they mainly rely on supplementing intermediates as the substrates. Recently, Jones et al reported the *de novo* production of anthocyanin in *E. coli* through the development of a polyculture system, which consisted of 4-strains of *E. coli* expressing 15 exogenous enzymes from different plants and microorganisms. This enabled the construction of the whole pathway of anthocyanin while efficiently controlling the resulting metabolic burden. Overall, 9.5 mg/L of pelargonidin 3-*O*-glucoside was produced from glucose ¹³³.

The synthesis of anthocyanins in *S. cerevisiae* was shown recently. A total of 61 mg/L of naringenin was formed via expression of PAL from *A. thaliana*, C4H from *A. majus*, native CPR1, 4CL from *A. thaliana*, CHS from *H. androsaemum*, and CHI from *M. sativa*. Next, a series of genes, including F3'H, F3'5'H, F3H, DFR, 3GT, and ANS, were screened with the purpose of obtaining the enzyme with the best activity. F3'H from *P. hybrida* and F3'5'H from *Solanum lycopersicum* were found to be the best combination amongst the various F3'5'Hs and F3Hs, while all other F3Hs performed efficiently.

Amongst the five DFRs, DFR from *A. andraeanum* produced the highest afzelechin, which was then used for pelargonidin 3-*O*-glucoside production. The DFR from *Populus trichocarpa* and *Iris hollandica* were chosen for cyanidin 3-*O*-glucoside and delphinidin 3-*O*-glucoside synthesis, respectively. Amongst the nine 3GTs, 3GT from *Dianthus caryophyllus* was the best one for production of pelargonidin 3-*O*-glucoside while 3AGT from *F. ananassa* was the best one for production of cyanidin 3-*O*-glucoside and delphinidin 3-*O*-glucoside. Moreover, 14 ANSs were screened although no large differences were displayed. Based on the above efforts, ~1.25 mg/L of pelargonidin 3-*O*-glucoside, ~1.9 mg/L of cyanidin 3-*O*-glucoside, and ~2 mg/L of delphinidin 3-*O*-glucoside were obtained in the recombinant *S. cerevisiae* strain ¹³⁴.

2.4 Conclusion

The phenolic compounds' natural health-beneficial properties, as well as antioxidant/antimicrobial activities, make them perfect candidates for functional food additives and food preservatives. In order to produce them in the most sustainable way, metabolic engineering has emerged as an alternative biosynthesis technique. Genetic engineering can be used to increase commercial attractiveness, thereby reducing the cost of production. In the past decade, big improvements have been achieved for production of phenolic compounds in both *S. cerevisiae* and *E. coli*. Several new techniques were successfully created, demonstrated and established. These techniques have helped to create more efficient production processes and made it possible to utilize more-cost-effective substrates and approaches. Even though the reviewed studies have shown that *S. cerevisiae*

and *E. coli* are able to efficiently synthesize different phenolic compounds, none of them had been commercially applied.

The main limiting factor for commercial use of phenolics is the low productivity of microbial cell factories, which is currently far from industrial applications. Several strains did not utilize the substrates well or generated unwanted side products. In addition, genes were poorly expressed, or the culture conditions were not favorable for the host microorganisms. Addressing these problems, my research screened potential enzyme gene sources in order to increase enzymatic expression, which leads to a higher naringenin production. An ideal gene source combination of 4CL and CHS was found, which provides higher naringenin production as compared to the current mainstream *A. thaliana* ones.

3. GENE SOURCE SELECTION FOR 4-COUMARATE: COA LIGASE AND CHALCONE SYNTHASE IN BY4741 *S.* *CEREVISIAE* STRAIN

Abstract

Flavonoids are plant secondary metabolites with great potential in the food industry. Metabolic engineering of *Saccharomyces cerevisiae* is a sustainable production technique. However, the current naringenin production yield is low due to inefficient enzymatic activity. Hence this study uses gene source screening as a tool to identify the best gene source for enzymes 4-coumarate: coenzyme ligase (4CL) and chalcone synthase (CHS). For the first time, the 4CL gene from *Medicago truncatula* and CHS gene from *Vitis vinifera* were expressed in *Saccharomyces cerevisiae* BY4741 strain and this combination provided the highest yield of naringenin. It was 28-fold higher as compared to the reference strain. Our results demonstrated that the selection and combination of enzymes from the correct gene source could greatly improve naringenin production. For the future, this could provide a better and more cost-effective alternative sustainable method to produce flavonoids.

3.1 Introduction

Flavonoids are natural compounds with highly advantageous health benefits. Their anti-cancer, anti-inflammatory, anti-allergenic and anti-oxidant properties have been showed to help prevent the development of neurodegenerative and cardiovascular diseases, Parkinson's disease, obesity, diabetes and aging¹⁵⁻¹⁷. Flavonoids can be found in plants, especially in berries. Consumption of flavonoid-rich food has been related to improved health. Due to this, flavonoids have the potential as food additives, and they can be used in food or beverages to increase their nutritional value and create a value-added product¹². Furthermore, flavonoids possess highly active anti-microbial properties, which presents the opportunity to apply flavonoids as food preservatives. As increasing disadvantages of using synthetic compounds in the food industry are emerging, the use of natural food preservatives and additives is a better solution⁴⁴.

Even though there is a significant potential for the use of flavonoids in the food industry, the current production, which is mainly based on plant extraction, has many limitations. Firstly, many of these plants are also food sources. Hence their usage for flavonoid extraction means a decrease in the overall food availability. Secondly, the amount of available plants is restricted by small plantation sizes, unfavorable climate, long growing time and cycles. At the same time, the concentration of flavonoids is relatively low in these plants. All these limitations add up to an expensive, high energy demanding preparation process, which leaves place for further improvements¹⁸.

An alternative solution is chemical synthesis, but due to the fact that certain flavonoids' structures are complex, synthesis is expensive, time and energy demanding¹⁹. Recently more sustainable options have been investigated, such as the potential use of food

waste. The possibility to extract antioxidants from food wastes, fruit and vegetable manufacturing residues had been investigated¹³⁵. However, the extraction process is not optimal, which is very laborious and the obtained amount of flavonoids is relatively low.

A potential alternative for sustainable production of flavonoids is the application of metabolic engineering on food-grade microorganisms, such as *S. cerevisiae*. Metabolic engineering alters the function of the microorganism by the insertion of non-native DNA sequences, which produces enzymes that are capable of producing the target compound. This newly established enzymatic system mimics the plants' flavonoid producing pathway, and reproduces it, utilizing the already existing precursors in the host microorganism²⁰. This method saves time, labor and energy, additionally does not compete with food source. Thus, it could provide a better and more cost-effective alternative sustainable method to produce flavonoids.

Naringenin is the central precursor of the flavonoid pathway. It is produced through the phenylpropanoid metabolic pathway. As it is a central compound in this pathway, it is vital for the production of other flavonoids such as isoflavonoids, flavonols, dihydroflavonols and anthocyanidins¹⁴. The pathway for naringenin production (Figure 8) consists of five enzymes, which are phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate: coenzyme (CoA) ligase (4CL), chalcone synthase (CHS) and chalcone isomerase (CHI). Alternatively, tyrosine ammonia lyase (TAL) is able to convert tyrosine to *p*-coumaric acid (Figure 8)⁵³. In this way, the pathway can bypass C4H. Additionally, naringenin chalcone is able to convert to naringenin without CHI under certain conditions. This opens up the possibility to bypass CHI and shorten the necessary changes in the host microorganism⁵⁹.

Figure 8 Biosynthetic pathway of naringenin ²³. Upstream pathway compounds are shown in red boxes, downstreams ones in purple boxes. Two key enzymes are highlighted in red.

Naringenin was first produced with the aid of metabolic engineering tools in *E. coli* by Hwang et al⁶⁴. Then the pathway was introduced into *S. cerevisiae* by Jiang et al⁶⁷. For the sake of boosting the naringenin production, *p*-coumaric acid overexpressing strains were applied and gene overexpression, deletion and replacement were performed^{18, 68, 109, 111}. Further engineering steps made it possible to use glucose as the feedstock, which resulted in the yield of 108 mg/L naringenin^{21, 110}. However, the current enzymatic activity was generally low, which leaves room for further improvements. While the genetic changes

were well-explored, the enzyme sources were unexplored, even though it is known that the origin of the enzyme greatly influences its enzymatic activities and could affect the production rate. Hence, gene screening should be carried out. This research focused on identifying the source of the 4CL and CHS enzymes, which lead to the highest naringenin production.

3.2 Experimental Details

3.2.1 Gene source comparison

The gene sequences used for the sequence comparison, phylogenetic tree and protein tertiary structure were obtained from the NCBI Nucleotide database, based on the accession code taken from published literature. The DNA and amino acid sequences were compared using NCBI's Basic Local Alignment Search Tool (BLAST), where multiple sequences were compared at the same time. The tree was prepared in MEGA7¹³⁶ using Maximum Likelihood statistical method and Tamura-Nei model on MUSCLE aligned sequences¹³⁷. The tertiary structures were designed with Phyre2¹³⁸, visualized, aligned and analyzed in Chimera¹³⁹.

3.2.2 Microbial strains, media, plasmid and genes

E. coli Top 10 (Novagen, USA) was chosen as the strain for gene manipulation, while *S. cerevisiae* BY4741⁵³ was utilised for pathway engineering. *E. coli* cell cultures were grown in Luria-Bertani broth (LB) at 37°C and supplemented with 100 µg/ml of ampicillin for recombinants harboring pUC57 and 50 mg/ml kanamycin for pCEV-G1-Km derived plasmids. Yeast was cultured in YPD media (2% glucose, 1% yeast extract, 2%

peptone) supplemented with either 200 µg/ml of geneticin (G418) at 30°C. The chemicals and antibiotics were purchased from Sigma (Sigma Aldrich, USA).

The plasmids used for the experiments are pCEV-G1-Km¹⁴⁰ and it expressed genes 4CL and CHS. All genes used in this study were synthesized by GenScript, with codon-optimization according to *S. cerevisiae* and cloned in pUC57. The DNA sequences were obtained from the NCBI GenBank with the following reference numbers for 4CL: *A. thaliana* (NM_113019.3), *M. truncatula* (XM_003610795.2), *P. crispum* (KF765780.1), *S. coelicolor* (NC_003888.3), for CHS: *A. thaliana* (NM_121396.3), *G. uralensis* (EF026979.1), *H. androsaemum* (AF315345.1), *V. vinifera* (NM_001280950.1).

3.2.3 Construction of plasmids, yeast transformation and cultivation

During the construction of pCEV-G1-Km-4CL-CHS recombinant plasmids, each 4CL was digested from the pUC57 plasmid using NotI and SpeI. The pCEV-G1-Km plasmid was opened up with the same restriction enzymes. 4CL genes were ligated to the open section of the plasmid, which is using TEF1 promoter followed with the NotI restriction site. The section also involves an ADH1 terminator, which is after the SpeI restriction site. The ligation results were checked by gel electrophoresis, which confirmed that all the 4CL genes were successfully placed into the plasmid. All four pCEV-G1-Km-4CL recombinant plasmids were further digested with BamHI and SacII. The CHS genes were digested from pUC57 plasmids using the same restriction enzymes. CHS genes were ligated to the open section of the pCEV-G1-Km-4CL recombinant plasmids. This section is using PGK1 promoter followed with the BamHI restriction site. The section also involves a CYC1 terminator, which is after the SacII restriction site. The ligation results were

checked by gel electrophoresis, which confirmed all 16 new recombinant plasmids containing 4CL and CHS genes were successfully generated (Figure 9) .

Figure 9 Plasmid pCEV-G1-Km structure with inserted 4CL and CHS. Promoters, terminators are indicated with grey and dark blue arrows. Restriction sites locations are shown in orange boxes.

The constructed plasmids were transformed into BY4741 *S. cerevisiae* strain using the lithium acetate/single-stranded carrier DNA/polyethylene glycol method¹⁴¹, in order to generate the recombinant strains, for naringenin production. The positive strains for BY4741 were screened on YPD plates supplemented with 200 µg/ml geneticin. Single colony, which contained the pCEV-G1-Km-4CL-CHS, was inoculated in 5ml YPD with

geneticin and shaken at 200 rpm for 20 hours at 30°C. Then 2% of the overnight culture was added into 50 ml fresh YPD media supplemented with an antibiotic and incubated for 72 hours. For the BY4741 strains, at 12 hours of fermentation, 10 mg *p*-coumaric acid was added to the culture as a precursor.

3.2.4 Protein extraction and Western blot

In order to extract the proteins, the yeast cultures were centrifuged at 5000 rpm for 5 minutes. Then accordingly to the cell weight Y-PER™ Yeast Protein Extraction Reagent from ThermoFisher and EDTA-free Protease Inhibitor from Roche were added to the pellets. The samples were agitated on a shaker for 20 minutes, then it was centrifuged on 14 000 rpm for 10 minutes. The supernatant accomodating the protein extract was carefully collected. 20 µl of protein extract were resolved using a 15% SDS-PAGE gel. The resolved proteins were electro-transferred to PVDF membranes obtained from Bio-Rad at 25V for 50 minutes. The membrane was blocked with 5% milk in TBS-T solution (10% TBS and 1% Tween-20) for 1 hour at room temperature. TBS-T solution was used for washing the membrane three times for 15 minutes. Then the membrane was probed with HRP conjugated anti-6X His tag antibody (Rabbit polyclonal) obtained from abcam in 1:4000 dilution. The antibody was added to the 3% milk TBS-T solution and the membrane was rolled in it for 1 hour at room temperature. Again, TBS-T solution was used for washing the membrane three times for 15 minutes. The results were visualized using Pierce™ 1-step ultra TMB blotting solution obtained from ThermoFisher. 2 ml TMB solution was spread on the membrane and was incubated for 10 minutes on room temperature. The signal bands then appeared on the membrane.

3.2.5 Naringenin extraction and HPLC analysis

10 ml BY4741 fermented culture was centrifuged at 10 000 rpm for 1 minute, then the supernatant was taken for extracellular naringenin and naringenin chalcone analysis. To the supernatant, ethyl acetate was added in a 1:1 ratio, which was mixed by 30 seconds vortexing and 2 hours rotating at room temperature. The mixture was centrifuged at 10000 rpm for 10 minutes, after which the upper layer was separated and gathered. After all liquid has evaporated on 50 °C, 500 µl ethyl acetate was added, and the sample was filtered before HPLC analysis.

Naringenin, naringenin chalcone and *p*-coumaric acid were analyzed by HPLC system (Agilent 1100) equipped with a variable wavelength detector and C18 column (4.6 mm x 150 mm, RESTEK). The process was performed with a gradient method involving 2 solvents: methanol and water. The program included several steps, the first one 25% of methanol, in the second one the concentration increased to 75% in 10 minutes, then in the following 20 minutes it reached 100%, and it was held there for 10 minutes. Finally, methanol content returned to 25% in 2 minutes and it was held there for 13 minutes. Triplicates were run for each sample. The analysis was performed with the flow rate of 0.5 ml/min and the signal was detected at 280 nm.

The obtained HPLC peaks were converted into production rates based on the standards` concentration calibration curve. The mean and standard deviation of the triplet results were calculated by Excel formulas. Statistical significance was established within groups and between two groups using homoscedastic t-tests. Between more than two groups one-way ANOVA test was applied. Alpha was set for 0.05 for both calculations. The obtained p-values are featured on the graphs in star format.

3.2.6 Growth curve

The sampling for the growth curve preparation started, when from the 5 ml overnight culture, 1 ml was taken and inoculated into 50 ml fresh YPD media supplemented with corresponding antibiotics. This inoculation step had defined the 0 hours. The cell cultures' OD₆₀₀ values were collected in every 3 or 6 hours in the initial growth phase, then after the first day, the sampling interval was changed to 12 hours. The OD₆₀₀ measurements were performed by NanoDrop™ 2000c Spectrophotometer.

3.3 Results and Discussion

3.3.1 Gene source selection

4CL and CHS gene sequences were screened in the NCBI database based on their origin, which was compared by their evolutionary background, gene sequences and protein tertiary structures. Based on the results four 4CL and four CHS gene sources were selected and expressed in *S. cerevisiae*. The 4CL gene sources were *A. thaliana*, *M. truncatula*, *P. crispum* and *S. coelicolor*. The CHS gene sources were *A. thaliana*, *G. uralensis*, *H. androsaemum* and *V. vinifera*. *A. thaliana* is a suitable reference as it is the most frequently used gene source for the expression of 4CL and CHS genes by research groups. Most of the other enzymes have been expressed before, but this is the first time 4CL from *M. truncatula* and CHS from *V. vinifera* are used for naringenin production. Additionally, many unique combinations have been expressed for the first time.

3.3.1.1 DNA sequence comparison

Based on the gene sequence blast results (Table 3 and 4), 4CL sources were very diverse. It was found that *S. coelicolor* showed the least similarities, as compared with the other gene sequences. This was reasonable as *S. coelicolor* is a bacterium, while the others are plants. For CHS there was less diversity. The gene sequence from *G. uralensis* was more distinguished, with a slightly less similarity percentage. The relation between *A. thaliana*, *H. androsaemum* and *V. vinifera* was very close as close to 75% of their sequence is identical.

Table 3 Sequence blast results for the 4CL gene from different sources²³. Similarities are shown in percentage.

4CL	Amount of nucleotides	Similarity		
		<i>A. thaliana</i>	<i>M. truncatula</i>	<i>P. crispum</i>
<i>A. thaliana</i>	1671			
<i>M. truncatula</i>	1764	62.5%		
<i>P. crispum</i>	1635	63%	57.1%	
<i>S. coelicolor</i>	1569	51.4%	46.8%	54.7%

Table 4 Sequence blast results for the CHS gene from different sources²³. Similarities are shown in percentage.

CHS	Amount of nucleotides	Similarity		
		<i>A. thaliana</i>	<i>G. uralensis</i>	<i>H. androsaemum</i>
<i>A. thaliana</i>	1188			
<i>G. uralensis</i>	1170	68%		
<i>H. androsaemum</i>	1173	73.9%	72.8%	
<i>V. vinifera</i>	1182	74.7%	69.7%	74.7%

3.3.1.2 Amino acid sequence comparison

Amino acid sequences provide a better insight into gene source relations, as they are more conserved as compared to DNA sequences. The shorter sequence increases the sensitivity of the comparison. Less noise is obtained due to the different DNA triplets that provide the same amino acid, which makes individual, third-position mutations irrelevant¹⁴². The amino acid blast results (Table 5 and 6) were very similar to the DNA sequence blast results. For 4CL gene sources, the difference widened between *S. coelicolor* and the plant-based sequences. At the same time *A. thaliana*, *M. truncatula* and *P. crispum* were found to have closer relation. As for the CHS gene sources, the sequences were even more similar with similarity values above 84%.

Table 5 Amino acid sequence comparison results for the 4CL gene from different sources. Similarities are shown in percentage.

4CL	Mass	Similarity		
		<i>A. thaliana</i>	<i>M. truncatula</i>	<i>P. crispum</i>
<i>A. thaliana</i>	61.6 kDa			
<i>M. truncatula</i>	64.5 kDa	64.1%		
<i>P. crispum</i>	60.6 kDa	69.2%	64.6%	
<i>S. coelicolor</i>	56 kDa	44%	41%	43.2%

Table 6 Amino acid sequence comparison results for the CHS gene from different sources. Similarities are shown in percentage.

CHS	Mass	Similarity		
		<i>A. thaliana</i>	<i>G. uralensis</i>	<i>H. androsaemum</i>
<i>A. thaliana</i>	43.9 kDa			
<i>G. uralensis</i>	43.4 kDa	83.2%		
<i>H. androsaemum</i>	43.5 kDa	86%	87.6%	
<i>V. vinifera</i>	43.5 kDa	86.1%	84.4%	87.5%

3.3.1.3 Tertiary structure comparison

Protein tertiary structure comparison is a more realistic, 3-D analogy as compared to sequence blasts. The percentage identity of the proteins' tertiary structures was shown to be similarly aligned with the gene DNA and amino acid sequence blast results (Table 7 and 8). The 4CL structures were more diverse as compared to the CHS structures. In particular, the results from *S. coelicolor* were shown to be the most noticeably different. However, the tertiary structure of *M. truncatula* 4CL was unlike the others, as the percentage identity was only around 50%, which was lower than the DNA sequence similarity value. An aligned tertiary structure comparison is shown in Figure 10 (a).

Table 7 Percentage identity for 4CL protein tertiary structures ²³. Similarities are based on structure alignment and comparison.

4CL	Percentage identity		
	<i>A. thaliana</i>	<i>M. truncatula</i>	<i>P. crispum</i>
<i>A. thaliana</i>			
<i>M. truncatula</i>	51.4%		
<i>P. crispum</i>	68.75%	48.8%	
<i>S. coelicolor</i>	37.92%	42.16%	37.92%

Table 8 Percentage identity for CHS protein tertiary structures ²³. Similarities are based on structure alignment and comparison.

CHS	Percentage identity		
	<i>A. thaliana</i>	<i>G. uralensis</i>	<i>H. androsaemum</i>
<i>A. thaliana</i>			
<i>G. uralensis</i>	83.12%		
<i>H. androsaemum</i>	86.49%	87.53%	
<i>V. vinifera</i>	86.75%	84.94%	87.92%

Figure 10 Gene screening by tertiary structures and phylogenetic tree ²³. (a) Alignment of *A. thaliana* and *M. truncatula* 4CL tertiary structure. (b) Phylogenetic tree of known 4CL and CHS gene sources.

3.3.1.4 Evolutionary relationship

A phylogenetic tree, which displayed the evolutionary relationship between the enzymes was created for 4CL and CHS. As shown from Figure 10 (b), the 4CL gene from *A. thaliana* and *P. crispum* were more similar as they were closer, whereas *M. truncatula* and *S. coelicolor* were at two branches on different ends of the tree. Similarly, the CHS gene from *A. thaliana* and *H. androseamum* were shown to sit in the middle of the branch, while towards two different ends were *G. uralensis* and *V. vinifera*. These findings were found to be in alignment with the DNA, amino acid sequence and protein structure similarities.

3.3.2 Expression of new combinations in *S. cerevisiae*

The pCEV-G1-Km plasmids containing the different 4CL-CHS combinations were constructed in *E. coli* and then expressed in *S. cerevisiae*. The correct plasmid ligation was confirmed by gel electrophoresis, the naringenin production was quantified by HPLC and the protein expression was verified by western blot.

3.3.2.1 Gel electrophoresis for confirmation of new combinations

First, the pCEV-G1-Km plasmid and 4CL ligation was confirmed by gel electrophoresis. The verification digestion has shown bands between 1500 and 2000 base pairs, that demonstrated the correct ligation. As a second step, the pCEV-G1-Km-4CL plasmid was ligated with CHS genes. This created sixteen combinations. The ligation was confirmed by digestion and the gene fragment was run on gel electrophoresis. The obtained band between 1000 and 1200 base pairs had verified the correct insertion (Figure 11). All sixteen combinations were correctly ligated, expressed in *E. coli* and confirmed by gel electrophoresis.

Figure 11 Confirmation of the new combinations by gel electrophoresis. Verification of the CHS gene ligation to pCEV-G1-Km-4CL by obtaining the band around 1100 base pairs. Example samples on the left side are pCEV-G1-Km-at4CL-atCHS, pCEV-G1-Km-at4CL-guCHS, pCEV-G1-Km-at4CL-haCHS, pCEV-G1-Km-at4CL-vvCHS, pCEV-G1-Km-pc4CL-atCHS, pCEV-G1-Km-pc4CL-guCHS, and on the right side pCEV-G1-Km-mt4CL-vvCHS.

3.3.2.2 HPLC results

After the confirmed combinations were expressed in *S. cerevisiae*, the naringenin production was quantified by HPLC (Figure 12). The highest yield was reached by expressing 4CL from *M. truncatula* and CHS from *V. vinifera*. This is a novel combination as neither of these enzymes have been expressed in *S. cerevisiae*. Their naringenin

production was 28-fold higher, as compared to the reference strain which expressed the 4CL gene and CHS gene from *A. thaliana* (Table 9). Interestingly, even when paired up with different enzyme sources, the 4CL gene from *M. truncatula* provided the highest naringenin production rate. In comparison, the CHS gene from *V. vinifera* was inferior as compared to the other CHS sources. However, combinations with 4CL from *P. crispum* and *S. coelicolor* did not result in naringenin production. Similarly, pCEV-G1-Km-mt4CL-atCHS and pCEV-G1-Km-mt4CL-haCHS were unable to synthesize naringenin.

Figure 12 Naringenin production in BY4741 *S. cerevisiae*²³. Comparison of naringenin production utilizing 4CL gene sources from *A. thaliana* and *M. truncatula* with CHS gene sources from *A. thaliana*, *G. uralensis*, *H. androseamum* and *V. vinifera*. The obtained *p* value is 0.007.

The HPLC analysis had also confirmed the produced naringenin chalcone content and the remaining *p*-coumaric acid amount in the cell cultures (Table 9, Figure 13). The detected naringenin chalcone amounts were neglectable in most of the samples, additionally the statistical significance was also found not significant. The highest naringenin chalcone content was 0.078 mg/L, which was produced by pCEV-G1-Km-at4CL-haCHS. The low amount of naringenin chalcone is due to its efficient conversion to naringenin. While studies had shown that CHI catalyzes the conversion of naringenin chalcone to naringenin, it was also proven that under certain conditions this reaction can happen spontaneously⁵⁹. In our case, we did not require CHI, as the detected amount of naringenin chalcone was neglectable. The remaining *p*-coumaric acid content in the cell cultures demonstrated the efficiency of the enzyme 4CL. The results can be seen in Table 9 and Figure 13. In average, 4CL from *M. truncatula* consumed more *p*-coumaric acid than 4CL from *A. thaliana*. However, all the samples contained a significant amount of *p*-coumaric acid at the end of the fermentation process, which leaves place for further improvements.

Table 9 The effect of expressing six different gene combinations on naringenin production in the *S. cerevisiae* BY4741 strain ²³

BY4741 strain	Naringenin amount (mg/L)	<i>p</i> -Coumaric acid amount (mg/L)
pCEV-G1-Km-mt4CL-vvCHS	7.3	3.32
pCEV-G1-Km-mt4CL-guCHS	6.6	3.16
pCEV-G1-Km-at4CL-vvCHS	3.95	4.67
pCEV-G1-Km-at4CL-haCHS	2.95	5.09
pCEV-G1-Km-at4CL-guCHS	2.7	3.15
pCEV-G1-Km-at4CL-atCHS	0.26	7.65

Figure 13 Extracellular *p*-coumaric acid, naringenin chalcone and naringenin in BY4741 *S. cerevisiae*. Comparison of the six naringenin producing strain containing 4CL from *A. thaliana* and *M. truncatula* with CHS from *A. thaliana*, *G. uralensis*, *H. androseamum* and *V. vinifera*. The obtained *p* value for *p*-coumaric acid is 0.0007, for naringenin chalcone is 0.15 and for naringenin is 0.007. The *p* value in between groups is 0.0012.

3.3.2.3 Western blot

A western blot analysis was performed to confirm the efficiency of 4CL genes from *P. crispum* and *S. coelicolor* and detect why their combination did not produce naringenin in the *S. cerevisiae* BY4741 strain. As the results shown in Figure 14, the 4CL proteins from *P. crispum* and *S. coelicolor* were not expressed in yeast. As *S. coelicolor* is a bacterial gene source, it was possible that it would not be compatible with the *S. cerevisiae*. However, the reason for the expression problem with 4CL from *P. crispum* requires further studies.

It was also discovered that the 4CL gene from *M. truncatula* was not compatible with the CHS gene from *A. thaliana* and *H. androsaemum*.

Figure 14 Western blot results of the sixteen strains expressed in *S. cerevisiae*. Verification of 4CL and CHS genes expression. 4CL expression represented by bands at 60 kDa, while the bands at 40kDa demonstrated the expression of CHS.

3.3.2.4 Growth curve

A growth curve had been constructed to analyze the growth phases for the naringenin producing BY4741 yeast strains (Figure 15). The obtained results were compared to the control, wild type strain (wt). The cell cultures' OD₆₀₀ values were collected in every 3 hours in the initial growth phase, then after the first day, the sampling interval was changed to 12 hours. *p*-Coumaric acid was added to the cultures after 12 hours as a precursor. During the exponential phase (approx. 0 – 15 hours), the control strain along with pCEV-G1-Km-mt4CL-guCHS had reached the highest OD₆₀₀ of 16. The growth of

pCEV-G1-Km-at4CL-vvCHS was slower than the other strains. After reaching the stationary phase (approx. 18-60 hours), the cell cultures' OD₆₀₀ values were above 13 with the highest average of 21.7 at 48 hours. Strains containing pCEV-G1-Km-at4CL-atCHS and pCEV-G1-Km-mt4CL-vvCHS were the first one entering to death phase after 48 hours. The obtained growth curves were very alike to the control, wild type strain. This indicated that the cells were not excessively burdened by the naringenin pathway.

Figure 15 Growth curve of the naringenin producing *S. cerevisiae* BY4741 strains. Wild type BY4741 *S. cerevisiae* strain (wt) used as control for the optical density measurements.

3.4 Conclusion

In summary, several 4CL and CHS gene source combinations were analyzed and expressed in *S. cerevisiae* BY4741 strain. For the first time, the 4CL gene from *M. truncatula* and the CHS gene from *V. vinifera* were expressed in *S. cerevisiae* and their combination resulted in the highest naringenin yield. The production was 28-fold higher as compared to the reference BY4741 strain expressing the 4CL gene and CHS gene from *A. thaliana*. The results had clearly demonstrated, that gene source selection has a high impact on naringenin production in *S. cerevisiae*.

Overall, gene screening was shown to be a necessary tool for optimal flavonoid production. The correct selection and combination of gene sources could improve enzyme efficiency, leading to more efficient and sustainable production platform for flavonoid production. This could help increase the possibility to commercialize flavonoid production for future works. The flavonoids produced could also be used as a natural, healthier food preservative and could be incorporated into food and beverages products to add value.

4. ENHANCEMENT OF NARINGENIN PRODUCTION IN *S. CEREVISIAE* BY DIVERSE METABOLIC APPROACHES

Abstract

Flavonoids have great health benefits and they are highly advantageous. However, currently their commercialization is limited due to unclear official health statements and by non-optimized and non-profitable production processes. In order to address these problems, the cellular and enzymatic behavior were analyzed, and the flavonoid pathway production was improved. Gene source selection for enzyme CHI was performed to identify possible optimization variables. A genetically modified Y-28 *S. cerevisiae* strain was also utilized for the expression of the naringenin pathway. This resulted in an increased yield with the highest naringenin production reaching 28.68 mg/L. Metabolite profiling has studied the cellular and enzymatic behavior during naringenin production. Three metabolites, L-aspartate, glycine and serine were identified as important cellular compounds. Furthermore, site-directed mutagenesis of CHS demonstrated the significance of amino acids Glu61, Lys151, Ser184 and the Ser388-Ala389-Pro390 triplet. Mutation of Glu61/Asp61 and Ser184/Ala184 have significantly decreased naringenin production, due to their location at the CoA-binding tunnel and at the coumaroyl-binding pocket. These findings could help commercialize flavonoid production, which would result in natural food preservatives and additives.

4.1 Introduction

During the last decades, flavonoids had emerged as a hot topic for chemical, genetic and metabolic researchers. Even though flavonoids are highly advantageous, currently their usage is very limited^{15, 17, 21, 143}. The biggest concern with the commercial application of phenolic extracts, including flavonoids is not having a clear health statement. For example, phenolic extracts in olive oil have been certified as safe (GRAS No. 726), but phenolics' health benefits have not yet been approved (EU No. 1924/2006)¹⁴⁴⁻¹⁴⁵. In addition, it needs to be noted that excessive amount of certain natural substrates can be fatal to humans as well. However, to date, neither FDA nor the EFSA have identified the ideal dietary intake of phenolic diets¹⁴⁶. Phenolic compounds, including flavonoids, should be thoroughly tested to determine their safety in food and to assess regulatory limits for their daily intake prior to food use.

Another issue for flavonoid commercialization is not optimized and profitable production processes. Flavonoid extraction from plants is limited by seasonal growth cycles, plantation sizes and low yield¹⁸. Additionally, most of the plants are food sources, which makes the extraction process less favorable. Chemical synthesis is also inadequate as the result of the flavonoids' complex chemical structure, which makes the synthesis expensive¹⁹. Currently, there is no profitable production process available. Metabolic engineering food-grade microorganisms, such as *S. cerevisiae* had emerged as a potential alternative, that saves time, labor and energy, additionally does not compete with food source²⁰. This makes the method ideal for flavonoid production and enables the potential for commercialization. However current production yields still require improvement.

In order to address these problems, the cellular and enzymatic behavior must be analyzed, and the flavonoid pathway production needs to be optimized. Enzyme CHI is responsible for the isomerization of naringenin chalcone into naringenin via intramolecular and stereospecific cyclization (Figure 16). Even though this step can happen spontaneously, the turnover rate can increase 10^7 -fold using CHI^{88, 147}. As the speed of conversion is faster, more naringenin is produced. This demonstrates the need to express CHI along with 4CL and CHS. Furthermore, a new *S. cerevisiae* strain, named Y-28, was utilized to enhance naringenin production¹⁸. The Y-28 strain overexpressed TAL from *F. johnsoniae*, 4CL from *A. thaliana* and CHS from *H. androsaemum*, as they were integrated to the *S. cerevisiae* genome. The strain was engineered to enable higher naringenin yields.

Figure 16 Catalysis of naringenin chalcone to naringenin by CHI. Enzymatic cyclization by CHI converts naringenin chalcone into naringenin.

Metabolite profiling is a highly accurate tool for the analysis of highly complex mixtures, that can identify and quantify more than 300 metabolites in a single extract. This enables the exploration of the key metabolic pathways in the cells and provides an insight

into the complex regulatory processes ¹⁴⁸⁻¹⁴⁹. On the other hand, site-directed mutagenesis is a microbiological tool to modify genes. This tool helps to identify and study the structure-function relationship and biological activity of the protein. The method includes an intentional single or combinational mutation in the protein's DNA sequence, in the form of change, insertion or deletion ¹⁵⁰. As a result, the effects of the specific amino acid change can be quantified and further analyzed, and the obtained structural change can be linked to functional differences ¹⁵¹. Site-directed mutagenesis is an important protein engineering technique that allows the optimization of natural enzymes' properties.

Overall, this study focused on the optimization of naringenin production by utilizing genetically modified Y-28 *S. cerevisiae* strain, performing gene source selection for enzyme CHI and mutating enzyme CHS by site-directed mutagenesis. Additionally, metabolite profiling was conducted to identify the key cellular changes during naringenin production.

4.2 Experimental Details

4.2.1 Gene source comparison

The gene sequences used for the sequence comparison, phylogenetic tree and protein tertiary structure were obtained from the NCBI Nucleotide database, based on the accession code taken from published literature. The DNA and amino acid sequences were compared using NCBI's Basic Local Alignment Search Tool (BLAST), where multiple sequences were compared at the same time. The tree was prepared in MEGA7¹³⁶ using Maximum Likelihood statistical method and Tamura-Nei model on MUSCLE aligned

sequences¹³⁷. The tertiary structures were designed with Phyre2¹³⁸, visualized, aligned and analyzed in Chimera¹³⁹.

4.2.2 Microbial strains, media, plasmid and genes

E. coli Top 10 (Novagen, USA) was chosen as the strain for gene manipulation, while *S. cerevisiae* BY4741⁵³ and Y-28¹⁸ were used for pathway engineering. The plasmids used for the experiments are pCEV-G1-Km and pCEV-G1-Ph¹⁴⁰. Genes 4CL and CHS were cloned and expressed in pCEV-G1-Km, while the pCEV-G1-Ph plasmid expressed TAL and CHI. *E. coli* cell cultures were grown in Luria-Bertani broth (LB) at 37°C and supplemented with 100 µg/ml of ampicillin for recombinants harboring pUC57, 50 mg/ml kanamycin for pCEV-G1-Km derived plasmids and 40 µg/mL of phleomycin for pCEV-G1-Ph plasmids. Yeast was cultured in YPD media (1% yeast extract, 2% peptone, 2% glucose) supplemented with either 200 µg/ml of geneticin (G418) or with additional 40 µg/mL of phleomycin antibiotics at 30°C. The chemicals and antibiotics were provided by Sigma (Sigma Aldrich, USA).

All genes utilized in this project were synthesized by GenScript, with codon-optimization according to *S. cerevisiae* and cloned in pUC57. The DNA sequences were obtained from the NCBI GenBank with the following reference numbers for 4CL: *A. thaliana* (NM_113019.3), *M. truncatula* (XM_003610795.2), *P. crispum* (KF765780.1), *S. coelicolor* (NC_003888.3), for CHS: *A. thaliana* (NM_121396.3), *G. uralensis* (EF026979.1), *H. androsaemum* (AF315345.1), *V. vinifera* (NM_001280950.1) and for CHI: *A. thaliana* (NP_191072), *Lotus japonicas* (AJ548840), *M. sativa* (KF765782), *Pueraria lobata* (D63577.1). TAL from *F. johnsoniae* was gifted by Lyu et al¹⁸.

4.2.3 Construction of plasmids

During the preparation of the pCEV-G1-Ph-CHI recombinant plasmids, each CHI was digested from the pUC57 plasmid using BamHI and SacII. The pCEV-G1-Ph plasmid was opened with the same restriction enzymes. CHI genes were ligated to the open section of the plasmid, which is using PGK promoter followed with the BamHI restriction site. The section also involves a CYC terminator, which is after the SacII restriction site. The ligation results were checked by gel electrophoresis, which confirmed that all the CHI genes were successfully placed into the plasmid (Figure 17).

Figure 17 Plasmid pCEV-G1-Ph structure with inserted CHI. Promoters, terminators are indicated with blue arrows. Restriction sites locations are shown in orange boxes.

The constructed plasmids were transformed into BY4741 *S. cerevisiae* strain using the lithium acetate/single-stranded carrier DNA/polyethylene glycol method¹⁴¹, in order to generate the recombinant strains, for naringenin production. The positive strains for BY4741 were screened on YPD plates supplemented with 200 µg/ml geneticin and 40 µg/mL of phleomycin. Single colony, which contained the confirmed pCEV-G1-Ph-CHI, was inoculated in 5ml YPD with geneticin and phleomycin, and it was shaken at 200 rpm for 20 hours at 30°C. Then 2% of the overnight culture was inoculated into 50 ml fresh YPD media supplemented with antibiotics and incubated for 72 hours. For the BY4741 strains, at 12 hours of fermentation, 10 mg *p*-coumaric acid was added to the culture as a precursor.

4.2.4 Naringenin extraction and HPLC analysis

10 ml BY4741 or 1 ml Y-28 culture at the end of the fermentation was centrifuged at 10 000 rpm for 1 minute, then the supernatant was taken for extracellular naringenin and naringenin chalcone analysis. To the supernatant, ethyl acetate was added in a 1:1 ratio, which was mixed by 30 seconds vortexing and 2 hours rotating at room temperature. The sample was centrifuged at 10000 rpm for 10 minutes, after which the upper layer was separated and collected. After all liquid has evaporated on 50 °C, 500 µl ethyl acetate was added, and the sample was filtered before HPLC analysis.

Naringenin was analyzed by HPLC system (Agilent 1100) equipped with a variable wavelength detector and C18 column (4.6 mm x 150 mm, RESTEK). The process was performed with a gradient method involving 2 solvents: methanol and water. The program included several steps, the first one 25% of methanol, in the second one the concentration increased to 75% in 10 minutes, then in the following 20 minutes it reached 100%, and it

was held there for 10 minutes. Finally, methanol content returned to 25% in 2 minutes and it was held there for 13 minutes. Each sample was run in triplicates. The analysis was performed with the flow rate of 0.5 ml/min and the signal was detected at 280 nm.

The obtained HPLC peaks were converted into production rates based on the standards' concentration calibration curve. The mean and standard deviation of the triplet results were calculated by Excel formulas. Statistical significance was established within groups and between two groups using homoscedastic t-tests. Between more than two groups one-way ANOVA test was applied. Alpha was set for 0.05 for both calculations. The obtained p-values are featured on the graphs in star format.

4.2.5 Growth curve

The sampling for the growth curve preparation started, when from the 5 ml overnight culture, 1 ml was taken and inoculated into 50 ml fresh YPD media supplemented with corresponding antibiotics. This inoculation step had defined the 0 hours. The cell cultures' OD₆₀₀ values were collected in every 3 or 6 hours in the initial growth phase, then after the first day, the sampling interval was changed to 12 hours. The OD₆₀₀ measurements were performed by NanoDrop™ 2000c Spectrophotometer. The analysis was performed under the cell culture function using cuvettes. For the analysis up to 1 ml cultures were pipetted to a cuvette at the specified time intervals. As the machine's OD₆₀₀ measurement range is between 0 and 2, the cultures with higher than 2 OD₆₀₀ value were diluted. For each measurement, a reference wild type BY4741 or Y-28 was also measured.

4.2.6 Intracellular metabolite extraction and GC-MS analysis

From the cultivated samples, pellets equalling 10 of OD₆₀₀ were collected as the cultures were centrifuged on 12 000 rpm for 2 minutes. The pellets were washed with 700 µl water three times and were centrifuged on 12 000 rpm for 2 minutes. 0.5 ml water, 0.5 ml methanol and 10 µl 2mg/ml ribitol and water solution were added to the pellets along with glass beads. The pellet cells were broken in a bead mill by using 20 seconds intervals for six times. The broken cells were centrifuged on 12000 rpm for 2 minutes. The supernatant was collected and the methanol was evaporated at 37°C overnight. The samples were completely dried in freeze drier after 12 hours. 50 µl 20 mg/ml methoxyamine hydrochloride dissolved in pyridine was added to the dried samples and mixed by the vortex. They were incubated at 37°C for 1 hour. 100 µl of N-methyl-N-(trimethylsilyl)-trifluoroacetamide (MSTFA) with 1% trimethylchlorosilane (TMCS) were added for sialylation and the samples were gently turned up and down a few times for adequate mixing. The mixture was centrifuged on 5000 rpm for 20 seconds and was placed on 70°C heater for 30 minutes. Before GC-MS measurements the samples were centrifuged on 14000 rpm for 20 minutes, and the upper layer was collected for analysis¹⁵².

Metabolite profiling was performed by using GC-MS system (Agilent Technologies 7890A-5975C) equipped with an HP-5MS, 5% Phenyl-Methyl-Silox capillary column (30 m × 250 µm × 0.25 µm Agilent J&W Scientific, Folsom, CA, USA), using Triple-Axis HED-EM detector. 1 µl of the sample was injected by splitless mode. Each sample was run in triplicates. The carrier gas used was helium with the constant flow rate of 1.1 ml/min. The inlet temperature was set to 250°C, while the ion source temperature was 230°C. The oven temperature was set to start on 75°C for 4 minutes, and with 4°C/min increase it

reached 280°C at the end of the run, where it was kept for 2 minutes. The obtained compounds were analyzed with Agilent MSD Chemstation Data Analysis and identified with NIST08 mass spectral library. The selected compounds' auto integrated peaks had been aligned and compared based on their abundance.

4.2.7 Site-directed mutagenesis

Site-directed mutagenesis was executed by PCR-based approach, where two overlapping oligonucleotides were directly used as primers in a PCR reaction. These primers contained a part of the perfectly complementary section to the target region except for a specific mismatch at the target DNA, where the mutation was introduced. The PCR was executed based on the Takara PrimeSTAR GXL DNA polymerase protocol ¹⁵³. 5 µl 5X PrimeSTAR GXL Buffer and 2 µl dNTP mixture were mixed with 1 µl 0.2 µM solution of each primers. The primers introducing the mutations are shown in Table 10. 1 µl pCEV-G1-Km-mt4cl-vvCHS recombinant plasmid was added as a template to the mixture along with 0.5 µl PrimeSTAR GXL DNA polymerase. Distilled water was used to fill the mixture up to 25 µl.

The reaction mixture was run on a 3-step PCR method with 12 cycles. The initial stage of a cycle was 98°C for 10 seconds, that was followed by 60°C for 15 seconds and ended on 68°C for 10 minutes. After the reaction finished the wild type methylated template was degraded by 1 µl DpnI digestion enzyme on 37 °C for 3 hours. The obtained enzyme sequence was confirmed by gel electrophoresis and the mutation by DNA sequencing. The DNA sequencing was executed by 1st BASE Asia. The confirmed sequence was transformed into *S. cerevisiae* with electroporation ¹⁵⁴.

Table 10 Site-directed mutagenesis primers. Mutated sections are in bold.

Mutation		Primers
AA61	Forward	GTTTAAAAGAATGTGTGATAAATCAATGATCAATAAG
	Reverse	CTTATTGATCATTGATTTATCACACATTCTTTTAAAC
AA151	Forward	CAAAATTGTTGGGTTTGAG ACC ATCAGTTAAGAGATTG
	Reverse	CAATCTCTTAAC TGAGGTCT CAAACCCAACAATTTG
AA184	Forward	CAGAAAACAACGCTGGT GCT AGAGTTTTGGTTGTTTG
	Reverse	CAAACAACCAAACTCTAG CACC AGCGTTGTTTTCTG
AA388	Forward	GTTGTTTTGCATTCTGTTGCTGCACATCACCACCATC
	Reverse	GATGGTGGTGATGTGCAGCAACAGAATGCAAAACAAC

4.3 Results and Discussion

4.3.1 Metabolite profiling

In order to understand the key mechanisms for naringenin production and how this mechanism influenced cell behavior, metabolite profiling was performed using GC-MS. The obtained peaks were aligned and compared based on their abundance (Figure 18). Three key metabolites were found L-aspartate, glycine and serine. These three metabolites were shown to be present in very different amounts, in the naringenin producing and non-producing strains (Table 11). In the naringenin producing strain, the amount of L-aspartate and serine was found to be only half, while the glycine amount was almost doubled (Figure 19).

The decrease in the L-aspartate amount indicated that more oxaloacetic acid was produced, which then enhanced gluconeogenesis.¹⁵⁵⁻¹⁵⁶ Through the fatty acid synthesis pathway, it could lead to higher malonyl-CoA amount¹⁵⁷, which could then have been consumed by the CHS enzyme during naringenin production. Similarly, the decrease in the serine amount might be due to higher conversion towards malonyl-CoA, which has been consumed by pyruvate, acetyl-CoA and fatty acid synthesis¹⁵⁸. On the other hand, the

increased glycine amount showed that it was not converted into serine, even though the amount of serine was depleted¹⁵⁹. This indicated that the amino acid catabolic pathway was inhibited (Figure 20). However, more research is required to further understand the connection between these pathways to the naringenin producing pathway.

Figure 18 GC-MS chromatography overlay example for two naringenin producing strains and a control ²³ A: Serine (retention time: 14.35 mins) B: L-Aspartate (retention time: 19.7 mins) C: Glycine (retention time: 19.95 mins) Black chromatograph: pCEV-G1-Km-at4CL-haCHS. Blue chromatograph: pCEV-G1-Km-mt4CL-vvCHS. Red chromatograph: Control.

Table 11 Comparison of the three key metabolites in the naringenin producing and non-producing strains ²³. L-aspartate, glycine and serine abundance production shown according to naringenin production.

Abundance	L-aspartate	Glycine	Serine
pCEV-G1-Km-at4CL-atCHS	1 515 237	9 851 703	0
pCEV-G1-Km-at4CL-guCHS	1 071 890	9 833 160	1 682 530
pCEV-G1-Km-at4CL-haCHS	1 475 277	9 840 346	1 278 530
pCEV-G1-Km-at4CL-vvCHS	1 178 761	4 187 608	1 453 023
pCEV-G1-Km-mt4CL-guCHS	1 214 060	10 096 217	1 223 493
pCEV-G1-Km-mt4CL-vvCHS	1 698 566	11 824 813	1 629 505
Average abundance:			
Naringenin producing strains	1 358 465	9 272 308	1 211 172
Naringenin non-producing strains	2 757 179	5 007 451	2 417 988

Figure 19 Graphical interpretation of the three key metabolites' abundance comparison in the naringenin producing and non-producing strains. L-aspartate, glycine and serine production shown according to naringenin production.

Figure 20 Metabolic changes induced by naringenin production in *S. cerevisiae*²³. The speculative enhanced pathway by naringenin producing strains (shown by orange arrows) mapped onto metabolic pathways (shown by gray arrows). Blue and red boxes are demonstrating the decrease and the increase of the targeted amino acids' amount.

4.3.2 Expression of *CHI* in *S. cerevisiae*

After identifying the ongoing metabolic processes in the cell, the first strategy to enhance the naringenin production was to express *CHI* along with *4CL* and *CHS*. *CHI* had been shown to increase the speed of the conversion between naringenin chalcone to naringenin⁸⁸, which can result in higher naringenin production. First *CHI* gene sources were screened in the NCBI database based on their origin, which was compared by their evolutionary background, gene sequences and protein tertiary structures. Based on the results four *CHI* gene sources were selected and expressed in *S. cerevisiae* BY4741 strain. The *CHI* gene sources were *A. thaliana*, *L. japonicus*, *M. sativa* and *P. lobata*. *A. thaliana*

was chosen as a reference as it was chosen also for 4CL and CHS. Along with *A. thaliana*, CHI from *M. sativa* and *P. lobata* had been expressed in *S. cerevisiae* before. However, this is the first time CHI from *L. japonicus* was used for naringenin production.

4.3.2.1 DNA sequence and amino acid sequence comparison

Based on the gene sequence blast results (Table 12), CHI sources were quite diverse. It was found that *A. thaliana* showed the least similarities, as compared with the other gene sequences. The relation between *L. japonicus*, *M. sativa* and *P. lobata* was very close, as 76% - 77% of their sequence was identical. The amino acid blast result (Table 13) was very similar to the DNA sequence blast results. The difference widened between *A. thaliana* and the other sequences. At the same time, *M. sativa* and *P. lobata* were found to have a closer relationship with the similarity percentage of 80%.

Table 12 Sequence blast results for the CHI gene from different sources. Similarities are shown in percentage.

CHI	Amount of nucleotides	Similarity		
		<i>A. thaliana</i>	<i>L. japonicus</i>	<i>M. sativa</i>
<i>A. thaliana</i>	741			
<i>L. japonicus</i>	678	54.9%		
<i>M. sativa</i>	669	56%	77.9%	
<i>P. lobata</i>	675	54.5%	77.5%	76.4%

Table 13 Amino acid sequence comparison results for the CHI gene from different source. Similarities are shown in percentage.

CHI	Mass	Similarity		
		<i>A. thaliana</i>	<i>L. japonicus</i>	<i>M. sativa</i>
<i>A. thaliana</i>	27.1 kDa			
<i>L. japonicus</i>	24.8 kDa	47.3%		
<i>M. sativa</i>	24.3 kDa	49%	78.1%	
<i>P. lobata</i>	24.3 kDa	49.3%	76.3%	80%

4.3.2.2 Tertiary structure comparison

The percentage identity of the proteins' tertiary structures was shown to be similarly aligned with the gene DNA and amino acid sequence blast results (Table 14) In particular, the tertiary structure of *A. thaliana* CHI was unlike the others, as the percentage identity was only around 46%, which was lower than the DNA sequence similarity value. The similarity value for *M. sativa* and *P. lobata* (81.13%) was even higher than during the amino acid sequence comparison An aligned tertiary structure comparison is shown in Figure 21 (a).

Table 14 Percentage identity for CHI protein tertiary structures. Similarities are based on structure alignment and comparison.

4CL	Percentage identity		
	<i>A. thaliana</i>	<i>L. japonicus</i>	<i>M. sativa</i>
<i>A. thaliana</i>			
<i>L. japonicus</i>	46.23%		
<i>M. sativa</i>	46.70%	77.83%	
<i>P. lobata</i>	48.11%	76.89%	81.13%

(a)

(b)

Figure 21 Gene screening by tertiary structures and phylogenetic tree (a) Alignment of *A. thaliana* and *L. japonicus* CHI tertiary structures. (b) Phylogenetic tree of known CHI gene sources.

4.3.2.3 Evolutionary relationship

A phylogenetic tree, which displayed the evolutionary relationship between the enzymes was created for CHI. The NCBI database for CHI has significantly increased in the last few years, due to the isolation of the genes from new sources. As shown from Figure 21 (b), all the chosen CHI sources are located towards the mid-section. However, when the sources were selected the CHI gene from *A. thaliana* was more significantly towards the upper end. The distance between *A. thaliana* and the other three gene sources is still notable. On the other hand, the other gene sources are closely related, especially *P. lobata* and *M. sativa*. These findings were found to be in alignment with the DNA, amino acid sequence and protein structure similarities.

4.3.2.4 Gel electrophoresis for confirmation of the CHI expression in *S. cerevisiae*

The pCEV-G1-Ph plasmids containing the different CHI enzymes were constructed in *E. coli* and then expressed in *S. cerevisiae* along with the pCEV-G1-Km combinations from the earlier study. The pCEV-G1-Ph plasmid and CHI ligation was confirmed by gel electrophoresis. The verification digestion has shown bands close to 750 base pairs, that demonstrated the correct ligation (Figure 22). All the four enzymes were correctly ligated, expressed in *E. coli* and confirmed by gel electrophoresis.

Figure 22 Confirmation of pCEV-G1-Ph and CHI enzymes ligation by gel electrophoresis. Verification of the CHI gene ligation to pCEV-G1-Ph by obtaining the band around 700 base pairs. Example samples on the upper part are pCEV-G1-Ph-atCHI, pCEV-G1-Ph-ljCHI, and on the below part: pCEV-G1-Ph-plCHI and pCEV-G1-Ph-msCHI.

4.3.2.5 HPLC results

After the confirmed combinations were expressed in *S. cerevisiae*, the naringenin production was quantified by HPLC (Figure 23). Contradicting our expectation, the naringenin production did not increase with the expression of CHI. The highest production was 1.66 mg/L by pCEV-G1-Km-mt4CL-guCHS and pCEV-G1-Ph-plCHI, which is a four-fold decrease as compare when the pCEV-G1-Km-mt4CL-guCHS was expressed alone in *S. cerevisiae* BY4741 strain. There was no defined pattern obtained based on gene source selection, additionally, there were no big differences between the obtained yields. The only consistency found was the lower naringenin production as compared when only 4CL and

CHS were expressed in *S. cerevisiae* BY4741 strain. This seems to be due to the expression of the two plasmids had overloaded the *S. cerevisiae* BY4741 cells and limited its growth and function. After looking into the utilized plasmids, it was clear that pCEV-G1-Km and pCEV-G1-Ph were a less satisfying choice for naringenin production. The main reason for this might be that the antibiotic used had a negative effect on the growth rate of *S. cerevisiae*.¹⁴⁰ As the growth rate decreased, less naringenin was produced, that could explain the obtained lower yields.

Figure 23 Naringenin production in BY4741 *S. cerevisiae*, expressing enzymes 4CL, CHS and CHI. Comparison of the six naringenin producing strain based on CHI gene sources from *A. thaliana*, *L japonicus*, *M. sativa* and *P. lobata*. The obtained *p* value is 7.6E-10.

4.3.3 Y-28 *S. cerevisiae* strain

4.3.3.1 Combinations expression in Y-28 *S. cerevisiae* strain

In order to increase the naringenin production and study how a different type of *S. cerevisiae* strain affects gene source selection, the six pCEV-G1-Km plasmids were expressed in the genetically engineered *S. cerevisiae* Y-28 strain. Along with the recombinant pCEV-G1-Km plasmids, a pCEV-G1-Ph-fjTAL-msCHI plasmid was also expressed in the Y-28 strain. The pCEV-G1-Ph-fjTAL-msCHI recombinant plasmid was obtained from Lyu et al and it contained overexpressed TAL from *F. johnsoniae* and CHI from *M. sativa*¹⁸. The six combinations expressed in the Y-28 strain used glucose as the carbon source and the naringenin production was then analyzed by HPLC (Table 15).

Table 15 Six gene combinations expressed in the *S. cerevisiae* Y-28 strain ²³. Comparison of the obtained naringenin production using different 4CL and CHS gene sources.

Y-28 strain	Naringenin amount (mg/L)
pCEV-G1-Km-mt4CL-vvCHS	24.41
pCEV-G1-Km-mt4CL-guCHS	24.51
pCEV-G1-Km-at4CL-vvCHS	22.31
pCEV-G1-Km-at4CL-haCHS	20.46
pCEV-G1-Km-at4CL-guCHS	20.72
pCEV-G1-Km-at4CL-atCHS	19.15

The results showed a similar pattern for gene source selection as compared to the wildtype BY4741 strain, and the expected increased yield was achieved (Figure 27). The

lowest naringenin production was obtained by the 4CL and CHS genes from *A. thaliana*. However, the production value significantly increased from 0.26 mg/L in BY4741 to 19.15 mg/L in Y-28. The highest production was 24 mg/L naringenin, which was three times higher than the highest production in the BY4741 strain. This was reached by two strains expressing the 4CL gene from *M. truncatula* and the CHS genes from *V. vinifera* and *G. uralensis*.

Figure 24 Comparison of naringenin production in BY4741 and Y-28 *S. cerevisiae* strains. Column chart representation of the obtained naringenin yields by the six efficient genen combinations of 4CL from *A. thaliana* and *M. truncatula* with CHS from *A. thaliana*, *G. uralensis*, *H. androseamum* and *V. vinifera*. The obtained *p* value for BY4741 strains is 0.007 and for the Y-28 strains is 2.05E-6. The *p* value in between groups is 1.3E-7.

4.3.3.2 Growth curve

A growth curve had been constructed to analyze the growth phases for the naringenin producing yeast Y-28 strains (Figure 28). The obtained results were compared to the control, wild type Y-28 strain. The cell cultures' OD₆₀₀ values were collected in every 6 hours in the initial growth phase, then after the first day, the sampling interval was changed to 12 hours.

Figure 25 Growth curve of the naringenin producing *S. cerevisiae* Y-28 strains. Wild type Y-28 *S. cerevisiae* strain used as control for the optical density measurements.

As compared to the growth curve obtained in the BY4741 strains, the phases were delayed and not as well defined for the Y-28 strain. The control, Y-28 strain not containing

any plasmid, had achieved significantly higher OD₆₀₀ values and reached stationary phase faster than the other samples. While the exponential phase for the control lasted 36 hours, the other strains were there for around 48 hours. After reaching the stationary phase, the cell cultures' OD₆₀₀ values were above 20, with the highest control value approaching 60. Several strains had started entering to death phase after 60 hours. There was a large gap between the obtained growth curves for the wild-type and plasmid expressing strains. This indicated that the cells were burdened by the expression of the plasmid and the naringenin pathway has affected the cell growth and function.

4.3.4 Site-directed mutagenesis for enzyme CHS

4.3.4.1 CHS structure and mechanism

CHS is part of the polyketide synthase enzymes (PKS) and it is responsible for the alteration of 4-coumaroyl-CoA with malonyl-CoA to naringenin chalcone and CO₂. It is the simplest PKS enzyme and it is part of the type III PKS group ¹⁶⁰. CHS is a homodimer protein with each monomer approximately 40-45 kDa in size. It has two separate active sites where the decarboxylation, condensation, and cyclization reactions take place ¹⁶¹. Type III PKSs are unique as they can use free CoA thioesters as substrates without acyl carrier proteins during the catalysis reactions. CHS's typical CoA thioester is 4-coumaroyl-CoA, which undergoes three condensations with malonyl-CoA to produce naringenin chalcone ¹⁶². Each monomer's structure has a five-layer $\alpha\beta\alpha\beta\alpha$ core, that surrounds a β -keto synthase (KS) active site and a dimerization interface. The dimerization interface is a six-residue loop, which separates the two monomers' active sites. The active site is buried within an interior cavity except for a 16 Å CoA-binding tunnel. The tunnel links the catalytic site to the external environment ¹⁶¹.

There are four conserved amino acid residues, which define the catalytic machinery of CHS. These are Cys164, Phe215, His303, and Asn336¹⁶¹. Cys164 is the loading site for the starter molecule, *p*-coumaroyl moiety and it initiates the reaction series. Asn336 orients the moiety near His303, where the malonyl-CoA is decarboxylated and the polyketide chain is elongated. This is followed by the Claisen-type condensation of three acetate units derived from malonyl -CoA with the enzyme-bound *p*-coumaroyl moiety. The two units go through cyclization and aromatization to form naringenin chalcone¹⁶³. Phe215 is located between the active site and CoA-binding tunnel, and it is an important amino acid. Phe215's main task is to orient the substrates during the reaction, additionally, it blocks the unneeded proteins from entering the active site¹⁶⁰.

The data on the structure and components of CHS has been widely researched and the available information is well-established.. The structure is simpler, and the function is more straightforward as compare to 4CL. For these reasons, this study focused on the mutation of CHS. *V. vinifera* was selected as the target gene source as it produced the highest amount of naringenin in Chapter 4.

4.3.4.2 Mutation sites

Amino acid sequent alignment of the four CHS gene sources from Chapter 4 was performed in order to analyze the four conserved amino acids and their surroundings (Figure 24). Cys164, Phe215, His303, and Asn336 had been found in all the four CHS at the exact same location. When the four genes were expressed in *S. cerevisiae*, they had produced a different amount of naringenin. We assume that there are specific amino acids in the *V. vinifera* CHS sequence, which provides an higher yield. *V. vinifera* CHS sequence was compared to the other three sequences, and four amino acid residues were identified,

where the sequence of *A. thaliana*, *G. uralensis*, and *H. androsaemum* aligned and only *V. vinifera* differed.

Figure 26 Multiple amino acid sequence alignment of CHS genes. Conserved sites are highlighted with red. Mutated amino acid residues are shown in blue.

These four amino acid residues were chosen as mutation site for the site-directed mutagenesis. Glu61 is part of the CoA-binding tunnel, and it is unique amino acid as the other three CHSs have aspartic acid at this location. Lys151 is situated near C164, which area is highly conserved. The other genes contain Arg151 instead of Lys151. Ser184 is located very close to the center of the coumaroyl-binding pocket, where the rest of the genes contain alanine instead of serine. As for the final site, it was observed that CHS from *V. vinifera* contains an additional three amino acids (Ser, Ala, Pro) at the end of the amino acid sequence (388-390). Overall, Glu61 was mutated into Asp61, Lys151 into Arg151, Ser184 into Ala184 and the Ser388-Ala389-Pro390 triplet was removed.

4.3.4.3 Naringenin production

After the confirmed mutated enzymes were expressed in *S. cerevisiae*, the naringenin production was quantified by HPLC (Figure 25). All the four mutations resulted in decreased naringenin production, which demonstrated the significance of Glu61, Lys151, Ser184 and the Ser388-Ala389-Pro390 triplet.

The lowest naringenin productions were 3.17 mg/L and 3.2 mg/L by the mutations of Glu61/Asp61 and Ser184/Ala184 respectively. This yield was less than half of the non-mutated enzyme's and it was more similar to the production of CHS from *A. thaliana*. Lys151/Arg151 and Ser388-Ala389-Pro390 triplet had produced 4 mg/L naringenin. This demonstrated that Lys151 and the Ser388-Ala389-Pro390 triplet were also responsible for the higher efficiency of CHS from *V. vinifera*. However, their impact is less as compare to Glu61 and Ser184. This can be due to the location of Glu61 and Ser184 at the CoA-binding

tunnel and at the coumaroyl-binding pocket, that play important roles in the CHS mechanism.

Figure 27 Naringenin production in BY4741 *S. cerevisiae*, expressing enzymes *M. truncatula* 4CL and mutated *V. vinifera* CHS. Comparison of the mutated enzymes efficiency to the non-mutated (wt) CHS. The obtained *p* value is 0.002.

4.3.4.4 Growth curve

A growth curve had been constructed to analyze the growth phases for the naringenin producing BY4741 yeast strains using mutated CHS (Figure 26). The obtained results were compared to the control strain containing the non-mutated CHS (WT). The cell cultures' OD₆₀₀ values were collected in every 6 hours in the initial growth phase, then after the first day, the sampling interval was changed to 12 hours. *p*-Coumaric acid was added to the cultures after 12 hours as a precursor.

Figure 28 Growth curve of the naringenin producing *S. cerevisiae* BY4741 strains, using mutated CHS. BY4741 *S. cerevisiae* strain containing the non-mutated CHS used as control for the optical density measurements.

The five growth curves were almost identical with very similar OD₆₀₀ values. During the exponential phase (approx. 0 – 24 hours), the control non-mutated strain had a faster growth rate. However, after reaching the stationary phase (approx. 24-60 hours), the cell cultures' OD₆₀₀ values were stagnating between 20-25. This indicated that the recombinant plasmids containing the mutated CHS had the same cellular burden as the non-mutated pCEV-G1-Km-mt4CL-vvCHS.

4.4 Conclusion

In conclusion, metabolite profiling explored the cellular mechanism behind naringenin production and concluded that L-aspartate, glycine and serine amounts were significantly altered. These three metabolites are related to the consumption of malonyl-CoA. However, further investigations are expected to unwind the definite mechanism and relation of these metabolites to naringenin.

Contradicting our expectation, the naringenin production did not increase with the expression of CHI along with 4CL and CHS. The highest production was 1.66 mg/L by pCEV-G1-Km-mt4CL-guCHS and pCEV-G1-Ph-plCHI, which is a four-fold decrease as compare when the pCEV-G1-Km-mt4CL-guCHS was expressed alone in *S. cerevisiae* BY4741 strain. There was no defined pattern obtained based on gene source selection, additionally, there were no big differences between the obtained yields.

The 4CL and CHS gene source combinations were also expressed in *S. cerevisiae* Y-28 strains. This resulted in an increased naringenin production as compared to the BY4741 strain. In addition, the similarity pattern of the combinations' production yields was compared between the BY4741 and Y-28 strains and were found to be matching. These results had shown that gene source selection had a defined pattern for naringenin production in *S. cerevisiae*.

Site-targeted mutagenesis had demonstrated the significance of amino acids Glu61, Lys151, Ser184 and the Ser388-Ala389-Pro390 triplet in *V. vinifera* CHS. Mutation of Glu61/Asp61, Lys151/Arg151, Ser184/Ala184 and the removal of Ser388-Ala389-Pro390 triplet had resulted in decreased naringenin production. Mutation of Glu61 and Ser184 had

halved the naringenin production. This can be due to their location at the CoA-binding tunnel and at the coumaroyl-binding pocket.

These findings addressed the cellular and enzymatic behavior of *S. cerevisiae* during naringenin production and the optimization of the flavonoid pathway expression. These are important steps towards the safe, optimized and profitable commercial application of flavonoids by metabolic engineering. The commercial production would enable the usage of flavonoids as natural food additives and food preservatives, which could bring extra health-benefits to the society.

5. POTENTIAL CARBON AND NITROGEN SOURCES SCREENED FOR OPTIMIZED MEDIA AND GROWTH CONDITIONS

Abstract

The industrial production and commercial usage of yeast had been available since the 19th century. The advantageous properties and acceptance in the food industry made *S. cerevisiae* a model microorganism for metabolic engineering food additives, which also enabled flavonoid production in yeast. *S. cerevisiae* can metabolize various carbon and nitrogen sources, even though the preferences for carbon sources are sugars and for nitrogen sources are ammonium ions and urea. Based on growth conditions and requirements, unconventional carbon and nitrogen substrates can also be utilized. In this study, yeast culturing conditions were optimized by testing galactose, sucrose, fructose, mannitol, potato starch, wheat flour and potato peel waste as potential carbon sources. Furthermore, peptone, yeast nitrogen without amino acid supplement and okara were screened as alternative nitrogen sources. The 1% sucrose and 1% glycerol mixture was found to be a reasonable option to replace glucose as carbon source. Additionally, okara was identified as an ideal nitrogen source. The usage of alternative carbon and nitrogen sources can decrease the cost of culturing media and provide a more sustainable production platform.

5.1 Introduction

Yeasts have been known to mankind for hundreds of years. They have been part of traditional food processes, such as wine, beer and bread making. Yeasts are eukaryotic unicellular microorganisms belonging to the group of ascomycetous or basidiomycetous fungi. They are capable of reproducing by budding or fission with spores non-enclosed in a fruiting body ¹⁶⁴. Since the end of the 19th century, the industrial production and commercial use of yeast had been available. Today, yeast is also used as high nutritional value proteins, enzymes and vitamins source, additionally genetic and metabolic engineering enabled further food industrial application such as food additives, preservatives and flavoring components ¹⁶⁵.

Fast growth, high tolerance for environmental stresses, well-established metabolic tools, simple nutritional requirements and well-known genomic data are several key properties making *S. cerevisiae* an outstanding yeast for commercial usage. *S. cerevisiae* is widely accepted for food production due to its GRAS status and long history in the food industry ¹⁶⁶⁻¹⁶⁷. For example, yeast had been part of cheese and bread production for hundreds of years. The advantageous properties and acceptance in the food industry made *S. cerevisiae* a model microorganism for metabolic engineering food additives, such as flavonoids.

Yeast growth requires carbon and nitrogen sources, which are included in the culturing media. The main carbon source is glucose, which is converted via the glycolytic pathway to pyruvate. The pyruvate enters the Krebs cycle and exits as anabolites and energy in the shape of ATP. The conversion happens in the presence of oxygen. In anaerobic conditions, yeast undergoes fermentation, which is more commonly utilized in alcohol

making. In this study, yeast growth was supplied with air, which enhanced the decarboxylation of pyruvate to acetyl-CoA. Then, acetyl-CoA was oxidized to CO₂, energy and intermediate compounds that promoted the yeast growth ¹⁶⁵. Nitrogen source and metabolism is critical for the metabolic reactions of proteins, amino acids, and other nitrogenous compounds. In *S. cerevisiae*, nitrogenous parts for amino acids are commonly synthesized from glutamate and glutamine. These two compounds are connected to the TCA cycle and nitrogen metabolism through α -ketoglutarate. This conversion is made possible by NADPH-dependent glutamate dehydrogenase (GDH1), glutamine synthetase (GLN1), NADH-dependent glutamate synthase (GLT1), and NAD⁺-linked glutamate dehydrogenase (GDH2) ¹⁶⁸.

Yeast can metabolize various carbon and nitrogen sources, even though the preferences for carbon sources are sugars, such as glucose, sucrose and maltose and for nitrogen sources are ammonium ions and urea. Based on growth conditions and requirements, unconventional carbon and nitrogen substrates can be utilized as well. In this study, yeast culturing conditions were optimized by testing galactose, sucrose, fructose, mannitol, potato starch, wheat flour and potato peel waste as potential carbon sources and utilizing peptone, yeast nitrogen without amino acid supplement and okara as a nitrogen source.

5.2 Experimental details

5.2.1 Microbial strains, plasmid and genes

S. cerevisiae Y-28¹⁸ strain was selected for carbon and nitrogen sources screening. The plasmids used for the experiments are pCEV-G1-Km and pCEV-G1-Ph¹⁴⁰. Genes 4CL and CHS were expressed in pCEV-G1-Km, while the pCEV-G1-Ph plasmid expressed TAL

and CHI. The reference yeast was cultured in YPD media (1% yeast extract, 2% peptone, 2% glucose) supplemented with 200 µg/ml of G418 and 40 µg/mL of phleomycin antibiotics at 30°C. The chemicals and antibiotics were purchased from Sigma (Sigma Aldrich, USA).

5.2.2. Preparation and usage of alternative carbon and nitrogen sources

During the potential carbon and nitrogen sources screening the same ratios were held. From the commercial carbon sources, 2% galactose, 2% sucrose, the mixture of 1% sucrose and 1% glycerol, 2% fructose, 2% glycerol and 2% mannitol media were prepared. Similarly, wheat flour was diluted into 2% media. Potato starch and potato peel waste were first dried and hydrolyzed with 0.7 M HCl for 3 hours on 100 °C. After hydrolyzation the pH was adjusted to 6 and was diluted with distilled water in order to make the final potato starch and potato peel waste content 2%. All the media were autoclaved before usage, additionally, the wheat flour, potato starch and potato peel waste media were filtered.

From the commercial nitrogen sources 3% yeast nitrogen with amino acids and 3% peptone media were prepared. Okara was valorized by solid-state fermentation using *Rhizopus oligosporus*¹⁶⁹. The nutrients are extracted and diluted with distilled water in order to make the final okara content 3%. All the media were autoclaved before usage. During culture preparation, the media was complemented with 200 µg/ml of G418 and 40 µg/mL of phleomycin antibiotics.

5.2.3 Naringenin extraction and HPLC analysis

1 ml Y-28 culture at the end of the fermentation was centrifuged at 10 000 rpm for 1 minute, then the supernatant was taken for extracellular naringenin and naringenin

chalcone analysis. To the supernatant, ethyl acetate was added in a 1:1 ratio, which was mixed by 30 seconds vortexing and 2 hours rotating at room temperature. The sample was centrifuged at 10000 rpm for 10 minutes, after which the upper layer was separated and collected. After all liquid has evaporated on 50 °C, 500 µl ethyl acetate was added, and the sample was filtered before HPLC analysis.

Naringenin and *p*-coumaric acid were analyzed by HPLC system (Agilent 1100) equipped with a variable wavelength detector and C18 column (4.6 mm x 150 mm, RESTEK). The process was performed with a gradient method involving 2 solvents: methanol and water. The program included several steps, the first one 25% of methanol, in the second one the concentration increased to 75% in 10 minutes, then in the following 20 minutes it reached 100%, and it was held there for 10 minutes. Finally, methanol content returned to 25% in 2 minutes and it was held there for 13 minutes. Each sample was run in triplicates. The analysis was performed with the flow rate of 0.5 ml/min and the signal was detected at 280 nm.

The obtained HPLC peaks were converted into production rates based on the standards' concentration calibration curve. The mean and standard deviation of the triplet results were calculated by Excel formulas. Statistical significance was established within groups and between two groups using homoscedastic t-tests. Between more than two groups one-way ANOVA test was applied. Alpha was set for 0.05 for both calculations. The obtained p-values are featured on the graphs in star format.

5.2.4 Growth curve

The sampling for the growth curve preparation started, when from the 5 ml overnight culture, 1 ml was taken and inoculated into 50 ml fresh YPD media supplemented

with corresponding antibiotics. This inoculation step had defined the 0 hours. The cell cultures' OD₆₀₀ values were collected in every 3 or 6 hours in the initial growth phase, then after the first day, the sampling interval was changed to 12 hours. The OD₆₀₀ measurements were performed by NanoDrop™ 2000c Spectrophotometer. The analysis was performed under the cell culture function using cuvettes. For the analysis up to 1 ml cultures were pipetted to a cuvette at the specified time intervals. As the machine's OD₆₀₀ measurement range is between 0 and 2, the cultures with higher than 2 OD₆₀₀ value were diluted. For each measurement, a reference Y-28 was also measured.

5.3 Results and Discussion

5.3.1 Potential carbon and nitrogen source screening

Nine alternative carbon sources and three potential nitrogen sources were compared based on Y-28 naringenin production in order to replace glucose and YP media. The chosen carbon sources were galactose, sucrose, glycerol, fructose, mannitol, sucrose and glycerol mixture, wheat flour, potato starch and potato peel waste. Commercial alternatives were chosen in order to enhance the naringenin production, while food leftover and waste products were selected to lower production cost. As Figure 29 demonstrated galactose, sucrose and the sucrose-glycerol mixture had competitive production yields to glucose, while fructose had produced significantly less naringenin. Mannitol, glycerol and the three food-related carbon sources had very limited naringenin production.

Comparison of the Y-28 *S. cerevisiae* strain's naringenin production

Figure 29 Comparison of the Y-28 *S. cerevisiae* strain's naringenin production. Analysis of the obtained *p*-coumaric acid and naringenin yields utilizing different carbon and nitrogen and sources. The obtained *p* value for naringenin is 0.007 and for *p*-coumaric acid is 0.006. The *p* value in between groups is 0.1.

The selected potential nitrogen sources were peptone, yeast nitrogen without amino acids and okara. Only okara had the capacity to generate more than 10 mg/L naringenin. Utilizing okara as a nitrogen source is highly advantageous as it decreases the accumulated waste, reduces the culturing media cost and promotes the upcycling of food waste. In order

to further analyze the production efficiency of the alternative carbon and nitrogen sources, galactose, sucrose and the sucrose-glycerol mixture were chosen as potential carbon sources and okara as a nitrogen source.

5.3.2 Galactose as a carbon source

Galactose can break down into glucose 1-phosphate through the Leloir pathway. The first step is the conversion of β -D-galactose into α -D-galactose by galactose mutarotase, followed by galactokinase to produce galactose 1-phosphate. The third reaction is the generation of glucose 1-phosphate and UDP-galactose using galactose-1-phosphate uridylyl transferase. As the final step, the leftover UDP-galactose is converted to UDP-glucose by UDP-galactose 4-epimerase.¹⁷⁰ In order to investigate the effect of utilizing galactose as carbon sources on naringenin production, the six plasmids, which were expressed in the Y-28 strains before, were also grown on galactose (Figure 30).

As Figure 30 demonstrated, galactose was not an ideal carbon source for the pCEV-G1-Km and pCEV-G1-Ph plasmids expression. All the six samples containing plasmids obtained lower naringenin yield as compare to the Y-28 control strain. However, pCEV-G1-Km-mt4CL-guCHS and pCEV-G1-Km-mt4CL-vvCHS had still higher production than the strains containing 4CL from *A. thaliana*. These results had aligned with the earlier findings on gene source selection.

Figure 30 Naringenin and *p*-coumaric acid production of Y-28 *S. cerevisiae* strains utilizing galactose as a carbon source. Comparison of the 4CL genes from *A. thaliana* and *M. truncatula* with CHS gene sources from *A. thaliana*, *G. uralensis*, *H. androseamum* and *V. vinifera* grown on galactose. The obtained *p* value for naringenin is 1.5E-5 and for *p*-coumaric acid is 5.9E-5. The *p* value in between groups is 0.0011.

The main reason that the strains containing plasmids had performed poorly can be due to the lower growth rate (Figure 31). The obtained growth rates were compared to the control, wild type Y-28 strain, which also used galactose as carbon source. The cell cultures' OD₆₀₀ values were collected in every 6 hours in the initial growth phase, then after the first day, the sampling interval was changed to 12 hours. The control Y-28 strain had achieved significantly higher OD₆₀₀ values and reached the stationary phase faster than the other samples. While the exponential phase for the control lasted 36 hours, the other strains

were there at around 60 hours. After reaching the stationary phase, the cell cultures' OD₆₀₀ values were above 30, with the highest control value approaching 60. The significant difference between the obtained growth curves for the control and plasmid expressing strains indicated that the cells were burdened by the expression of the plasmid and the naringenin pathway has affected the cell growth and function.

Figure 31 Growth curve of the naringenin producing Y-28 *S. cerevisiae* strains, utilizing galactose as carbon source. Y-28 *S. cerevisiae* strain utilizing galactose as carbon source used as control for the optical density measurements.

5.3.3 Sucrose as a carbon source

Sucrose can breakdown into glucose through the conversion to sucrose-6-phosphate, and then to glucose-6-phosphate and fructose by sucrose hydrolyzing enzyme. Furthermore, fructose can be converted to glucose-6-phosphate by glucose-6-phosphate

isomerase¹⁷¹. This is an entirely different pathway as galactose takes to glucose conversion. The six recombinant plasmids containing Y-28 strains were also grown on sucrose and the naringenin production along with the growth curve were analyzed.

Figure 32 Naringenin and *p*-coumaric acid production of Y-28 *S. cerevisiae* strains utilizing sucrose as a carbon source. Comparison of the 4CL genes from *A.thaliana* and *M. truncatula* with CHS gene sources from *A. thaliana*, *G. uralensis*, *H. androseamum* and *V. vinifera* grown on sucrose. The obtained *p* value for naringenin is 1.8E-5 and for *p*-coumaric acid is 2.5E-5. The *p* value in between groups is 0.005.

Figure 32 contains the obtained results from HPLC, that demonstrated the efficient expression of pCEV-G1-Km and pCEV-G1-Ph recombinant plasmids in the Y-28 strain grown on sucrose. The obtained results displayed a similar pattern as the Y-28 strains grown on glucose. The 4CL gene from *M. truncatula* with CHS from *V. vinifera* achieved the most elevated naringenin amount with 22.26 mg/L. It was followed by the combinations of 4CL from *M. truncatula* with CHS from *G. uralensis* and 4CL from *A. thaliana* with CHS from *V. vinifera*, that produced 20.29 mg/L and 20.09 mg/L naringenin respectively.

The obtained growth rates also aligned with the naringenin production rates. The cell cultures' OD₆₀₀ values were collected in every 6 hours in the initial growth phase, then after the first day, the sampling interval was changed to 12 hours. As compared to the control, wild type Y-28 strain also grown on sucrose, the samples' growth rates were slower (Figure 33). The control Y-28 strain had achieved significantly higher OD₆₀₀ values and reached the stationary phase faster than the other samples. The strain containing pCEV-G1-Km-mt4CL-vvCHS had followed it with about 12 hours delay. After reaching the stationary phase, the cell cultures' OD₆₀₀ values were above 20, with the highest control value approaching 60. Overall, the control Y-28 strain obtained higher OD₆₀₀ values as compared to the other strains, especially pCEV-G1-Km-at4CL-atCHS and pCEV-G1-Km-at4CL-guCHS had a lower growth rate. The difference between the obtained growth curves for the control and plasmid expressing strains indicated that the cells were burdened by the plasmids expression.

Growth curve for Y-28 strains utilizing sucrose as carbon source

Figure 33 Growth curve of the naringenin producing Y-28 *S. cerevisiae* strains, utilizing sucrose as a carbon source. Y-28 *S. cerevisiae* strain utilizing sucrose as carbon source used as control for the optical density measurements.

5.3.4 Sucrose and glycerol mixture as a carbon source

Growing the Y-28 recombinant plasmids strains on sucrose had very similar production rate as compared when glucose served as a carbon source. However, it was previously shown that the use of 1% sucrose and 1% glycerol as the carbon source, instead of 2% glucose or sucrose could increase the naringenin yield in the Y-28 strains¹⁸. For this reason, the mixture of 1% sucrose and 1% glycerol was tested as a carbon source.

The different production speeds and yields are due to the alternative breakdown processes of glucose, sucrose and glycerol. Glycerol converts to glycerol 3-phosphate and then to dihydroxyacetone phosphate by glycerol kinase and glycerol 3-phosphate

dehydrogenase, which allows easy access to carbon source ¹⁷². On the other hand, sucrose takes a longer time to breakdown into glucose through the conversion to sucrose-6-phosphate, and then to glucose-6-phosphate and fructose by sucrose hydrolyzing enzyme. Furthermore, fructose can be converted to glucose-6-phosphate by glucose-6-phosphate isomerase ¹⁷¹. When the cultures were grown on glucose, the carbon source would be immediately consumed at the initial growth stage. However, the use of glycerol and sucrose would provide carbon source both at the initial stage by glycerol breakdown and in later stages by sucrose breakdown.

5.3.4.1 Naringenin production

The usage and effectiveness of 1% sucrose and 1% glycerol as carbon source were investigated. The six recombinant plasmids, which were expressed in the earlier Y-28 strains, were also grown on this sucrose and glycerol solution. Overall, a general increase had been detected in the naringenin yields (Table 16).

Table 16 Obtained naringenin yield produced in the *S. cerevisiae* Y-28 strains grown on sucrose and glycerol solution as a carbon source ²³. Efficiency analysis of the six strains expressing six different gene source combinations on sucrose and glycerol solution as a carbon source.

Y-28 strain	Naringenin amount (mg/L)
pCEV-G1-Km-mt4CL-vvCHS	28.68
pCEV-G1-Km-mt4CL-guCHS	28.10
pCEV-G1-Km-at4CL-vvCHS	23.29
pCEV-G1-Km-at4CL-haCHS	28.66
pCEV-G1-Km-at4CL-guCHS	19.42
pCEV-G1-Km-at4CL-atCHS	26.17

The obtained results displayed a similar pattern as before, where the use of the 4CL gene from *M. truncatula* resulted in the highest naringenin production (Table 16). However, it was surprising to note that the strain expressing the pCEV-G1-Km-at4CL-haCHS recombinant plasmid, had the same naringenin production as the strain expressing the 4CL gene from *M. truncatula*.

The naringenin production of the BY4741 strain, Y-28 utilizing glucose, sucrose and the mixture of sucrose and glycerol was compared in Figure 34. The best yield was accomplished in the Y-28 strain by the 1% sucrose and 1% glycerol carbon source, followed by glucose. The performance pattern based on gene origin was similar in all carbon sources. The best combination was 4CL from *M. truncatula* and CHS from *V. vinifera*, that confirmed the importance of gene source selection.

In this study, the highest obtained naringenin yield from a Y-28 strain was 28.68 mg/L, which was lower as compared to the 90 mg/L naringenin production obtained by Lyu et al.¹⁸ After looking into the utilized plasmids, it was clear that pCEV-G1-Km was a less satisfying choice for naringenin production, as compared to pks plasmid. The main reason for this might be that the antibiotic used for pCEV-G1-Km, was G418. G418 had a negative effect on the growth rate of *S. cerevisiae*.¹⁴⁰ As the growth rate decreased, less naringenin was produced, which might explain the lower yields obtained. The results had clearly demonstrated, that gene source selection has a high impact on naringenin production. However, as the naringenin production increased, the gene source selection was shown to have less impact on the final amount of naringenin.

Figure 34 Comparison of naringenin production in strains BY4741, Y-28 utilizing glucose, Y-28 utilizing sucrose and Y-28 utilizing sucrose and glycerol as carbon source. Analysis of the different 4CL and CHS gene sources` efficiency grown on three different carbon sources. The obtained p value for BY4741 strains is 0.007, for Y-28 grown on glucose is $2\text{E-}6$, for Y-28 grown on sucrose is $1.8\text{E-}5$ and for Y-28 grown on sucrose and glycerol is $7.4\text{E-}7$. The p value in between groups is $1.8\text{E-}10$.

5.3.4.2 Growth curve

A growth curve had been constructed to analyze the growth phases for the naringenin producing yeast Y-28 strains utilizing 1% sucrose and 1% glycerol as the carbon

source (Figure 35). The obtained results were compared to the control, wild type Y-28 strain, which also used the mixture as a carbon source. The cell cultures' OD₆₀₀ values were collected in every 6 hours in the initial growth phase, then after the first day, the sampling interval was changed to 12 hours.

Figure 35 Growth curve of the naringenin producing *S. cerevisiae* Y-28 strains utilizing 1% sucrose and 1% glycerol as the carbon source. Y-28 *S. cerevisiae* strain utilizing sucrose and glycerol as carbon source used as control for the optical density measurements.

The obtained growth curve was similar to the earlier Y-28 growth curve utilizing glucose. The control Y-28 strain had achieved significantly higher OD₆₀₀ values and reached the stationary phase faster than the other samples. The strain containing pCEV-G1-Km-mt4CL-guCHS had followed it with about 6 hours delay. While the exponential phase

for the control lasted 24 hours, the other strains were there for around 48 hours. After reaching the stationary phase, the cell cultures' OD₆₀₀ values were above 20, with the highest control value approaching 60. The strains had started entering to death phase after 60 hours. There was a large gap between the obtained growth curves for the control and plasmid expressing strains. This indicated that the cells were burdened by the expression of the plasmid and the naringenin pathway has affected the cell growth and function. The similarity between the two growth curves indicates, the different carbon source had burdened the cells alike.

5.3.5 Okara as a nitrogen source

Okara is a soybean residue created during soymilk and tofu production. It is a major waste product that is mainly disposed into landfills. Worldwide 14 million tons of okara are generated annually and it is one of the main agro-wastes in Southeast Asia ¹⁶⁹. However, okara has highly nutritious components that can be hydrolyzed into simple sugars, amino acids and minerals ¹⁷³. Using solid-state fermentation, the precious nutritious components were obtained from okara and was utilized as a nitrogen source in the Y-28 culturing media.

In order to explore the potential of utilizing okara as nitrogen sources on naringenin production, the control Y-28 strain and the highest naringenin producing recombinant plasmid pCEV-G1-Km-mt4CL-vvCHS were grown on a media consisting okara and the sucrose-glycerol mixture. The obtained results had shown that the control Y-28 had produced slightly more naringenin utilizing okara (22.14 mg/L) as compared to YP media (20.25 mg/L) (Figure 36). On the other hand, the Y-28 strain containing pCEV-G1-Km-mt4CL-vvCHS had better production rate grown on YP (28.68 mg/L) as compared to the okara nitrogen source (22.98 mg/L).

Figure 36 Naringenin production using okara as nitrogen source. The control Y-28 strain and the Y-28 strain containing pCEV-G1-Km-mt4CL-vvCHS recombinant plasmid utilizing the sucrose and glycerol mixture as carbon source, and YP or okara as a nitrogen source are compared. The obtained p value is 0.0005.

The growth curves of the control Y-28 strain and the Y-28 strain containing pCEV-G1-Km-mt4CL-vvCHS recombinant plasmid utilizing the sucrose and glycerol mixture as carbon source, and YP or okara as nitrogen source had been compared (Figure 37). The cell cultures' OD₆₀₀ values were collected in every 6 hours in the initial growth phase, then after the first day, the sampling interval was changed to 12 hours. The control Y-28 strains had reach stationary phase faster as compared to the samples containing pCEV-G1-Km-mt4CL-vvCHS. Y-28 grown on YP has the steepest exponential growth, followed by Y-28 grown

on okara with about 6 hours delay. The growth of the pCEV-G1-Km-mt4CL-vvCHS strains utilizing YP and okara were similar, but the okara one reached the stationary phase faster as compared to the YP one. After reaching the stationary phase, the cell cultures' OD₆₀₀ values were above 30, with the highest control value approaching 60.

Figure 37 Growth curve of naringenin production using okara as nitrogen source. The control Y-28 strain and the Y-28 strain containing pCEV-G1-Km-mt4CL-vvCHS recombinant plasmid utilizing the sucrose and glycerol mixture as carbon source, and YP or okara as a nitrogen source are compared.

Overall, okara was able to produce a similar amount of naringenin as YP media. The growth curves had demonstrated that the cell is not burdened more by utilizing okara as utilizing YP media. Taking into account that YP media is a commercial media costing 38 SGD/L, okara would be a free replacement that helps to reduce the generated agro-waste. This demonstrated the potential to use food-waste product for naringenin production.

5.4 Conclusion

Alternative carbon and nitrogen sources were compared and analyzed in order to decrease the cost of culturing media. Galactose and sucrose were identified as a potential substitutes for glucose. However the media growth rate significantly decreased by plasmid expression when galactose was used as carbon source. Utilizing sucrose as a media component achieved similar results as glucose. Furthermore, the 1% sucrose and 1% glycerol mixture was found to further enhance the naringenin production.

As for the nitrogen source, okara, a soybean waste residue was identified as an ideal replacement for YP media. Okara is an ideal culturing media ingredient, as it is a major waste source, that mainly ends up on landfills. This decreases the production cost, and helps to reduce the waste production around the world. The combination of the 1% sucrose and 1% glycerol mixture as carbon source and okara as nitrogen source is a very promising alternative for YPD media. The obtained naringenin yield grown on the media consisting of the 1% sucrose and 1% glycerol mixture as carbon source and okara as nitrogen source were very similar to the production on YPD media.

Optimization of the *S. cerevisiae* strain`s utilized carbon and nitrogen sources had led to higher naringenin production rate and lower production cost. This improvement addressed the problem of low productivity of microbial cell factories and high production cost, which are posing as a limiting factor for commercial use of flavonoids. The flavonoids` natural health-beneficial properties, as well as antioxidant/antimicrobial activities, make them perfect candidates for functional food additives and food preservative.

6. CONCLUSION AND FUTURE DIRECTIONS

6.1 Conclusion

This thesis set out to establish gene source selection as a metabolic engineering tool and demonstrate its effect on the flavonoid production in *S. cerevisiae*. With the above aims in mind, a systematically and comprehensive gene source screening was performed for enzymes 4-coumarate: coenzyme A ligase (4CL) and chalcone synthase (CHS), which are responsible for naringenin production. In order to enhance the naringenin production, chalcone isomerase (CHI) gene source screening and culture condition optimization based on carbon and nitrogen source were performed. Furthermore, metabolite profiling and site-directed mutagenesis had provided novel information about the effect of gene source selection on the cell function. The overall workflow is summarized in Figure 38.

Based on the data presented in this thesis it was concluded that gene source selection has a high impact on naringenin production in *S. cerevisiae*. Gene screening was shown to be a necessary tool for optimal flavonoid production. Several 4CL and CHS gene source combinations were analyzed and expressed in *S. cerevisiae* BY4741 strain. For the first time, the 4CL gene from *M. truncatula* and the CHS gene from *V. vinifera* were expressed in *S. cerevisiae* and their combination resulted in the highest naringenin yield. The production was 28-fold higher as compared to the reference BY4741 strain expressing the 4CL gene and CHS gene from *A. thaliana*.

Figure 38 Overall workflow of this thesis to establish gene source selection as a metabolic engineering tool and demonstrate its effect on the flavonoid production in *S. cerevisiae*.

Metabolite profiling explored the cellular mechanism behind naringenin production and concluded that L-aspartate, glycine and serine amounts were significantly altered. These three metabolites are related to the consumption of malonyl-CoA. However further studies are needed to unravel the exact mechanism of these metabolites on naringenin production. Site-targeted mutagenesis had demonstrated the significance of amino acids Glu61, Lys151, Ser184 and the Ser388-Ala389-Pro390 triplet in *V. vinifera* CHS. Mutation

of Glu61/Asp61, Lys151/Arg151, Ser184/Ala184 and the removal of Ser388-Ala389-Pro390 triplet had resulted in decreased naringenin production. Mutation of Glu61 and Ser184 had halved the naringenin production. This was due to their location at the CoA-binding tunnel and at the coumaroyl-binding pocket.

Contradicting our expectation, the naringenin production did not increase with the expression of CHI along with 4CL and CHS. The highest production was 1.66 mg/L by pCEV-G1-Km-mt4CL-guCHS and pCEV-G1-Ph-plCHI, which is a four-fold decrease as compare when the pCEV-G1-Km-mt4CL-guCHS was expressed alone in *S. cerevisiae* BY4741 strain. There was no defined pattern obtained based on gene source selection, additionally, there were no big differences between the obtained yields.

The 4CL and CHS gene source combinations were also expressed in *S. cerevisiae* Y-28 strains. This resulted in an increased naringenin production as compared to the BY4741 strain. In addition, the similarity pattern of the combinations' production yields was compared between the BY4741 and Y-28 strains and were found to be matching. These results had shown that gene source selection had a defined pattern for naringenin production in *S. cerevisiae*. Alternative carbon and nitrogen sources were compared and analyzed in order to decrease the cost of culturing media. The 1% sucrose and 1% glycerol mixture was found to be a potential carbon source to replace glucose. Additionally, okara, a soybean waste residue was identified as an ideal replacement for YP media as a nitrogen source.

Overall, these results had achieved the aim of this thesis. The main achievements were collected in Figure 39. Key gene sources were identified, that enabled the enhancement of naringenin production. Additionally, the effect of the naringenin production on the yeast's cellular mechanism was discovered along with two important amino acids in

the CHS sequence, that takes part in the enzymatic catalytic processes. Furthermore the production cost was reduced by the replacement of YPD media with the combination of the 1% sucrose and 1% glycerol mixture as carbon source and okara as nitrogen source.

Figure 39 Key achievements of this thesis. The key points achieved in this thesis are collected and shown on basic radial graphic.

The enhanced naringenin production and the reduced costs are important steps towards an improved and optimized production platform. While the cellular and enzymatic

findings are mainly contribution to the scientific community that requires further studies before they are implemented for production optimization. This study enables the safe and profitable commercialization of flavonoids, and encourages its wider-ranged usage as natural food preservatives and as components of value-added products.

6.2 Future directions

6.2.1 Genome integration

The complete genome sequence of *S. cerevisiae* is well-known and all the genes are identified. This knowledge makes the yeast genome an easy target for manipulation, insertion and deletion¹⁷⁴. Insertion of TAL, 4CL, CHS and CHI genes into the yeast genome enables naringenin production without vectors and markers. Genome integrated metabolic pathways are more stable and less influential on cell growth as compared to when they are expressed in vectors.

Integration of the novel combination including 4CL from *M. truncatula* and CHS from *V. vinifera* could further enhance naringenin production as the current Y-28 genome engineered *S. cerevisiae* strain contains 4CL from *A. thaliana* and CHS from *H. androsaemum*. It has been proven that the combination of 4CL from *M. truncatula* and CHS from *V. vinifera* is able to produce higher amount of naringenin as compared to the combination of 4CL from *A. thaliana* and CHS from *H. androsaemum*. This demonstrates the potential of genome integration for the novel gene combination to further increase the naringenin yield.

6.2.2 Protein fusion

Enzyme expression and efficiency depend on protein expression. In order to increase the expression, metabolic and biosynthetic efficiency can be increased by the creation of macromolecular complexes¹⁷⁵. Creation of macromolecular complexes originates from plants and metabolic engineers copy this phenomenon to increase production yield. This is done by linking genes together, which creates a fused working complex. Bifunctional proteins are a type of these complexes, which includes only two fused proteins.¹⁷⁶⁻¹⁷⁷ Metabolic engineers discovered several advantages of this technique, as the amount of needed vectors are reduced and the distance between the proteins are shorter, which increases the speed of the metabolite processes. This advantage is especially useful in sequential reactions, where the substrates more specifically and rapidly reach the target protein, as the distance between them are way shorter. By this, the activity of the competing pathways are reduced and the cell feedback reaction of the enzymatic complex is increased.¹⁷⁸

Earlier studies have already demonstrated that protein fusion can enhance flavonoid production in *S. cerevisiae*. Tian and Dixon had created an artificial bifunctional enzyme containing CHI and IFS by in-frame gene fusion. The expression had successfully enhanced the genistein production¹⁷⁹. These findings demonstrate the potential of protein fusion for naringenin production. As the two most important enzymes for the flavonoid pathway are 4CL and CHS, they should be chosen as target proteins. A bifunctional protein consisting of 4CL from *M. truncatula* and CHS from *V. vinifera* could be created by in-frame fusion, which connects the two enzymes by a Gly–Ser–Gly linker (Figure 40). This study could aim to investigate the effect of gene fusion on flavonoids production in *S. cerevisiae*.

Figure 40 Tertiary structure of potential 4CL-CHS bifunctional protein created by in-frame gene fusion. Green protein tertiary structure: 4CL from *M. truncatula*, Blue protein tertiary structure: CHS from *V. vinifera*, Red linker structure: Gly–Ser–Gly linker.

6.2.3 Food waste as carbon source substitute

Carbon source is a significant constituent of the flavonoid producing yeast cultures. At the same time, it is one of the main production cost sources¹⁸⁰. Commercial carbon sources, such as glucose, galactose and sucrose can cost up to 97.9 - 454 SGD/kg. An alternative solution is to substitute commercial carbon sources with food waste. More than 1.3 billion tons of edible food was discarded in the world in 2018. General food waste contains 46.8% of carbon content, additionally, the amount of sugar in the waste is 53% of the food waste dry weight¹⁸¹. Utilizing food waste as a carbon source would substantially decrease the production cost and it would help to upcycle waste products, which lessen pollution¹⁸². Food waste contains a high amount of sugar and nutrients.

Therefore, they can be inexpensive energy-rich resources, which makes them an ideal carbon source for commercialization ¹⁸³.

In this study, potato peel waste was tested as a carbon source, but the naringenin production rate was not satisfactory. Earlier studies had tested several other food waste products as carbon sources for *S. cerevisiae*. Some examples are sauerkraut waste ¹⁸⁴, whey lactose ¹⁸³, cassava waste ¹⁸⁵, pineapple waste ¹⁸⁶, mango waste ¹⁸⁷, orange peel ¹⁸⁸ and sugarcane molasses ¹⁸⁹. While potato, corn, wheat, pineapple waste ¹⁹⁰, and even cafeteria food waste ¹⁸¹ had been successfully utilized as a carbon source for a different type of yeast strains. As sugarcane and sugarcane waste is widely available in Singapore, they could be tried as an alternative food waste carbon source. Sugarcane was successfully used as a media component for biodiesel production with *Rhodospiridium toruloides* ¹⁹¹, additionally, sugarcane molasses and wild sugarcane were actively used as a carbon source for ethanol production using *S. cerevisiae* ¹⁹²⁻¹⁹³. These demonstrated the potential of sugarcane as a carbon source and provide a great opportunity to explore.

REFERENCES

1. Lund, D., Food processing: from art to engineering. *Food technology* **1989**, 43 (9), 242-247.
2. Titus, D.; James Jebaseelan Samuel, E.; Mohana Roopan, S., Importance of Food Science and Technology- Way to Future. *Bioorganic Phase in Natural Food: An Overview* **2018**, 11-23.
3. Collaborators, G. R. F., Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **2018**, 392 (10159), 1923-1994.
4. Dötsch, M.; Busch, J.; Batenburg, M.; Liem, G.; Tareilus, E.; Mueller, R.; Meijer, G., Strategies to Reduce Sodium Consumption: A Food Industry Perspective. *Critical Reviews in Food Science and Nutrition* **2009**, 49 (10), 841-851.
5. Hamburg, M. A.; Sharfstein, J. M., The FDA as a Public Health Agency. *New England Journal of Medicine* **2009**, 360 (24), 2493-2495.
6. EFSA Panel on Dietetic Products, N. A., Guidance for the scientific requirements for health claims related to antioxidants, oxidative damage and cardiovascular health. *EFSA Journal* **2018**, 16 (1), 5136.
7. de Boer, A.; Vos, E.; Bast, A., Implementation of the nutrition and health claim regulation – The case of antioxidants. *Regulatory Toxicology and Pharmacology* **2014**, 68 (3), 475-487.
8. Ross, S., Functional foods: the Food and Drug Administration perspective. *The American Journal of Clinical Nutrition* **2000**, 71 (6), 1735S-1738S.
9. Carocho, M.; Barreiro, M. F.; Morales, P.; Ferreira, I. C. F. R., Adding Molecules to Food, Pros and Cons: A Review on Synthetic and Natural Food Additives. *Comprehensive Reviews in Food Science and Food Safety* **2014**, 13 (4), 377-399.
10. Amchova, P.; Kotolova, H.; Ruda-Kucerova, J., Health safety issues of synthetic food colorants. *Regulatory Toxicology and Pharmacology* **2015**, 73 (3), 914-922.
11. Balasundram, N.; Sundram, K.; Samman, S., Phenolic compounds in plants and agri-industrial by-products: Antioxidant activity, occurrence, and potential uses. *Food Chemistry* **2006**, 99 (1), 191-203.
12. Shahidi, F.; Ambigaipalan, P., Phenolics and polyphenolics in foods, beverages and spices: Antioxidant activity and health effects – A review. *Journal of Functional Foods* **2015**, 18, 820-897.
13. Scalbert, A.; Williamson, G., Dietary Intake and Bioavailability of Polyphenols. *The Journal of Nutrition* **2000**, 130 (8), 2073S-2085S.
14. Forkmann, G.; Martens, S., Metabolic engineering and applications of flavonoids. *Current Opinion in Biotechnology* **2001**, 12 (2), 155-160.
15. Harborne, J. B.; Williams, C. A., Advances in flavonoid research since 1992. *Phytochemistry* **2000**, 55 (6), 481-504.
16. Vauzour, D.; Rodriguez-Mateos, A.; Corona, G.; Oruna-Concha, M. J.; Spencer, J. P. E., Polyphenols and Human Health: Prevention of Disease and Mechanisms of Action. *Nutrients* **2010**, 2 (11), 1106-1131.

17. Dixon, R. A., Phytoestrogens. *Annual Review of Plant Biology* **2004**, 55 (1), 225-261.
18. Lyu, X.; Ng, K. R.; Lee, J. L.; Mark, R.; Chen, W. N., Enhancement of Naringenin Biosynthesis from Tyrosine by Metabolic Engineering of *Saccharomyces cerevisiae*. *Journal of Agricultural and Food Chemistry* **2017**, 65 (31), 6638-6646.
19. Quideau, S.; Deffieux, D.; Douat-Casassus, C.; Pouységu, L., Plant Polyphenols: Chemical Properties, Biological Activities, and Synthesis. *Angewandte Chemie International Edition* **2011**, 50 (3), 586-621.
20. Cress, B. F.; Linhardt, R. J.; Koffas, M., Isoflavonoid Production by Genetically Engineered Microorganisms. *Natural Products: Phytochemistry, Botany and Metabolism of Alkaloids, Phenolics and Terpenes* **2013**, 1647-1681.
21. Koopman, F.; Beekwilder, J.; Crimi, B.; van Houwelingen, A.; Hall, R. D.; Bosch, D.; van Maris, A. J.; Pronk, J. T.; Daran, J.-M., De novo production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae*. *Microbial Cell Factories* **2012**, 11 (1), 155-170.
22. Krivoruchko, A.; Nielsen, J., Production of natural products through metabolic engineering of *Saccharomyces cerevisiae*. *Current Opinion in Biotechnology* **2015**, 35, 7-15.
23. Mark, R.; Lyu, X.; Ng, K. R.; Chen, W. N., Gene Source Screening as a Tool for Naringenin Production in Engineered *Saccharomyces cerevisiae*. *American Chemical Society Omega* **2019**, 4 (7), 12872-12879.
24. Siró, I.; Kápolna, E.; Kápolna, B.; Lugasi, A., Functional food. Product development, marketing and consumer acceptance—A review. *Appetite* **2008**, 51 (3), 456-467.
25. Bigliardi, B.; Galati, F., Innovation trends in the food industry: The case of functional foods. *Trends in Food Science & Technology* **2013**, 31 (2), 118-129.
26. Etemadi, A.; Sinha, R.; Ward, M. H.; Graubard, B. I.; Inoue-Choi, M.; Dawsey, S. M.; Abnet, C. C., Mortality from different causes associated with meat, heme iron, nitrates, and nitrites in the NIH-AARP Diet and Health Study: population based cohort study. *British Medical Journal* **2017**, 357, 1957-1967.
27. Varraso, R.; Camargo, C. A., Processed meat consumption and lung health: more evidence for harm. *European Respiratory Journal* **2014**, 43 (4), 943-946.
28. Basli, A.; Belkacem, N.; Amrani, I., *Health Benefits of Phenolic Compounds Against Cancers*. InTech: 2017; Vol. 10.
29. Pandey, K. B.; Rizvi, S. I., Plant polyphenols as dietary antioxidants in human health and disease. *Oxidative Medicine and Cellular Longevity* **2009**, 2 (5), 270-278.
30. Shetty, K., Role of proline-linked pentose phosphate pathway in biosynthesis of plant phenolics for functional food and environmental applications: a review. *Process Biochemistry* **2004**, 39 (7), 789-804.
31. Mark, R.; Lyu, X.; Lee, J. J. L.; Parra-Saldívar, R.; Chen, W. N., Sustainable production of natural phenolics for functional food applications. *Journal of Functional Foods* **2019**, 57, 233-254.
32. Tsuda, T., Dietary anthocyanin-rich plants: Biochemical basis and recent progress in health benefits studies. *Molecular Nutrition & Food Research* **2012**, 56 (1), 159-170.
33. Fazary, A. E.; Ju, Y.-H., Feruloyl Esterases as Biotechnological Tools: Current and Future Perspectives. *Acta Biochimica et Biophysica Sinica* **2007**, 39 (11), 811-828.

34. Kumar, S.; Pandey, A. K., Chemistry and Biological Activities of Flavonoids: An Overview. *The Scientific World Journal* **2013**, *2013*, 16.
35. Baboota, R. K.; Bishnoi, M.; Ambalam, P.; Kondepudi, K. K.; Sarma, S. M.; Boparai, R. K.; Podili, K., Functional food ingredients for the management of obesity and associated co-morbidities – A review. *Journal of Functional Foods* **2013**, *5* (3), 997-1012.
36. Cushnie, T. P. T.; Lamb, A. J., Antimicrobial activity of flavonoids. *International Journal of Antimicrobial Agents* **2005**, *26* (5), 343-356.
37. Cowan, M. M., Plant Products as Antimicrobial Agents. *Clinical Microbiology Reviews* **1999**, *12* (4), 564-582.
38. Puupponen-Pimiä, R.; Nohynek, L.; Meier, C.; Kähkönen, M.; Heinonen, M.; Hopia, A.; Oksman-Caldentey, K.-M., Antimicrobial properties of phenolic compounds from berries. *Journal of Applied Microbiology* **2001**, *90* (4), 494-507.
39. Corrêa, V. G.; Tureck, C.; Locateli, G.; Peralta, R. M.; Koehnlein, E. A., Estimate of consumption of phenolic compounds by Brazilian population. *Revista de Nutrição* **2015**, *28*, 185-196.
40. Karam, J.; Bibiloni, M. D. M.; Tur, J. A., Polyphenol estimated intake and dietary sources among older adults from Mallorca Island. *PLOS ONE* **2018**, *13* (1).
41. Taguchi, C.; Fukushima, Y.; Kishimoto, Y.; Suzuki-Sugihara, N.; Saita, E.; Takahashi, Y.; Kondo, K., Estimated Dietary Polyphenol Intake and Major Food and Beverage Sources among Elderly Japanese. *Nutrients* **2015**, *7* (12), 10269-10281.
42. Bresciani, L.; Calani, L.; Cossu, M.; Mena, P.; Sayegh, M.; Ray, S.; Del Rio, D., (Poly)phenolic characterization of three food supplements containing 36 different fruits, vegetables and berries. *PharmaNutrition* **2015**, *3* (2), 11-19.
43. Pandey, R. P.; Parajuli, P.; Koffas, M. A. G.; Sohng, J. K., Microbial production of natural and non-natural flavonoids: Pathway engineering, directed evolution and systems/synthetic biology. *Biotechnology Advances* **2016**, *34* (5), 634-662.
44. Ng, K. R.; Lyu, X.; Mark, R.; Chen, W. N., Antimicrobial and antioxidant activities of phenolic metabolites from flavonoid-producing yeast: Potential as natural food preservatives. *Food Chemistry* **2019**, *270*, 123-129.
45. Dai, J.; Mumper, R. J., Plant Phenolics: Extraction, Analysis and Their Antioxidant and Anticancer Properties. *Molecules* **2010**, *15* (10), 7313-7352.
46. Birosova, L.; Mikulasova, M.; Vaverkova, S., Antimutagenic effect of phenolic acids. *Biomedical Papers of Medical Faculty, University Palacky* **2005**, *149* (2), 489-91.
47. Stich, H. F.; Rosin, M. P., Naturally Occurring Phenolics as Antimutagenic and Anticarcinogenic Agents. In *Nutritional and Toxicological Aspects of Food Safety*, Friedman, M., Ed. Springer US: Boston, MA, 1984; pp 1-29.
48. Stojković, D.; Petrović, J.; Soković, M.; Glamočlija, J.; Kukić-Marković, J.; Petrović, S., In situ antioxidant and antimicrobial activities of naturally occurring caffeic acid, p-coumaric acid and rutin, using food systems. *Journal of the Science of Food and Agriculture* **2013**, *93* (13), 3205-3208.
49. Aziz, N. H.; Farag, S. E.; Mousa, L. A.; Abo-Zaid, M. A., Comparative antibacterial and antifungal effects of some phenolic compounds. *Microbios* **1998**, *93* (374), 43-54.
50. Guzman, J. D., Natural Cinnamic Acids, Synthetic Derivatives and Hybrids with Antimicrobial Activity. *Molecules* **2014**, *19* (12), 19292-19349.

51. Park, E.-S.; Moon, W.-S.; Song, M.-J.; Kim, M.-N.; Chung, K.-H.; Yoon, J.-S., Antimicrobial activity of phenol and benzoic acid derivatives. *International Biodeterioration & Biodegradation* **2001**, 47 (4), 209-214.
52. Sova, M., Antioxidant and Antimicrobial Activities of Cinnamic Acid Derivatives. *Mini Reviews in Medicinal Chemistry* **2012**, 12 (8), 749-767.
53. Winzeler, E. A., et al., Functional Characterization of the *S. cerevisiae* Genome by Gene Deletion and Parallel Analysis. *Science* **1999**, 285 (5429), 901-906.
54. Hotze, M.; Schröder, G.; Schröder, J., Cinnamate 4-hydroxylase from *Catharanthus roseus* and a strategy for the functional expression of plant cytochrome P450 proteins as translational fusions with P450 reductase in *Escherichia coli*. *Federation of European Biochemical Societies Letters* **1995**, 374 (3), 345-350.
55. Mathias, B.; Georg E, S., Overexpression of a designed 2.2 kb gene of eukaryotic phenylalanine ammonia-lyase in *Escherichia coli*. *Federation of European Biochemical Societies Letters* **1999**, 457 (1), 57-60.
56. Kang, S.-Y.; Choi, O.; Lee, J. K.; Hwang, B. Y.; Uhm, T.-B.; Hong, Y.-S., Artificial biosynthesis of phenylpropanoic acids in a tyrosine overproducing *Escherichia coli* strain. *Microbial Cell Factories* **2012**, 11 (1), 153.
57. Rodriguez, A.; Chen, Y.; Khoomrung, S.; Özdemir, E.; Borodina, I.; Nielsen, J., Comparison of the metabolic response to over-production of p-coumaric acid in two yeast strains. *Metabolic Engineering* **2017**, 44, 265-272.
58. Rodriguez, A.; Kildegaard, K. R.; Li, M.; Borodina, I.; Nielsen, J., Establishment of a yeast platform strain for production of p-coumaric acid through metabolic engineering of aromatic amino acid biosynthesis. *Metabolic Engineering* **2015**, 31, 181-188.
59. Watts, K. T.; Lee, P. C.; Schmidt-Dannert, C., Exploring Recombinant Flavonoid Biosynthesis in Metabolically Engineered *Escherichia coli*. *ChemBioChem* **2004**, 5 (4), 500-507.
60. Santos, C. N. S.; Koffas, M.; Stephanopoulos, G., Optimization of a heterologous pathway for the production of flavonoids from glucose. *Metabolic Engineering* **2011**, 13 (4), 392-400.
61. Ro, D.-K.; Douglas, C. J., Reconstitution of the Entry Point of Plant Phenylpropanoid Metabolism in Yeast (*Saccharomyces cerevisiae*). *Journal of Biological Chemistry* **2004**, 279 (4), 2600-2607.
62. Vannelli, T.; Wei Qi, W.; Sweigard, J.; Gatenby, A. A.; Sariaslani, F. S., Production of p-hydroxycinnamic acid from glucose in *Saccharomyces cerevisiae* and *Escherichia coli* by expression of heterologous genes from plants and fungi. *Metabolic Engineering* **2007**, 9 (2), 142-151.
63. Jendresen, C. B.; Stahlhut, S. G.; Li, M.; Gaspar, P.; Siedler, S.; Förster, J.; Maury, J.; Borodina, I.; Nielsen, A. T., Novel highly active and specific tyrosine ammonia-lyases from diverse origins enable enhanced production of aromatic compounds in bacteria and yeast. *Applied and Environmental Microbiology* **2015**, 81 (13), 4458-4476.
64. Hwang, E. I.; Kaneko, M.; Ohnishi, Y.; Horinouchi, S., Production of Plant-Specific Flavanones by *Escherichia coli* Containing an Artificial Gene Cluster. *Applied and Environmental Microbiology* **2003**, 69 (5), 2699-2706.
65. Camacho-Zaragoza, J. M.; Hernández-Chávez, G.; Moreno-Avitia, F.; Ramírez-Iñiguez, R.; Martínez, A.; Bolívar, F.; Gosset, G., Engineering of a microbial coculture of

Escherichia coli strains for the biosynthesis of resveratrol. *Microbial Cell Factories* **2016**, *15* (1), 163.

66. Chen, Z.; Sun, X.; Li, Y.; Yan, Y.; Yuan, Q., Metabolic engineering of *Escherichia coli* for microbial synthesis of monolignols. *Metabolic Engineering* **2017**, *39*, 102-109.

67. Jiang, H.; Wood, K. V.; Morgan, J. A., Metabolic Engineering of the Phenylpropanoid Pathway in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* **2005**, *71* (6), 2962-2969.

68. Trantas, E.; Panopoulos, N.; Ververidis, F., Metabolic engineering of the complete pathway leading to heterologous biosynthesis of various flavonoids and stilbenoids in *Saccharomyces cerevisiae*. *Metabolic Engineering* **2009**, *11* (6), 355-366.

69. Gold, N. D.; Gowen, C. M.; Lussier, F.-X.; Cautha, S. C.; Mahadevan, R.; Martin, V. J. J., Metabolic engineering of a tyrosine-overproducing yeast platform using targeted metabolomics. *Microbial Cell Factories* **2015**, *14* (1), 73.

70. Mao, J.; Liu, Q.; Song, X.; Wang, H.; Feng, H.; Xu, H.; Qiao, M., Combinatorial analysis of enzymatic bottlenecks of l-tyrosine pathway by p-coumaric acid production in *Saccharomyces cerevisiae*. *Biotechnology Letters* **2017**, *39* (7), 977-982.

71. Choi, O.; Wu, C.-Z.; Kang, S. Y.; Ahn, J. S.; Uhm, T.-B.; Hong, Y.-S., Biosynthesis of plant-specific phenylpropanoids by construction of an artificial biosynthetic pathway in *Escherichia coli*. *Journal of Industrial Microbiology & Biotechnology* **2011**, *38* (10), 1657-1665.

72. Wang, S.; Zhang, S.; Xiao, A.; Rasmussen, M.; Skidmore, C.; Zhan, J., Metabolic engineering of *Escherichia coli* for the biosynthesis of various phenylpropanoid derivatives. *Metabolic Engineering* **2015**, *29*, 153-159.

73. Furuya, T.; Arai, Y.; Kino, K., Biotechnological Production of Caffeic Acid by Bacterial Cytochrome P450 CYP199A2. *Applied and Environmental Microbiology* **2012**, *78* (17), 6087-6094.

74. Takeo, M.; Murakami, M.; Niihara, S.; Yamamoto, K.; Nishimura, M.; Kato, D.-i.; Negoro, S., Mechanism of 4-Nitrophenol Oxidation in *Rhodococcus* sp. Strain PN1: Characterization of the Two-Component 4-Nitrophenol Hydroxylase and Regulation of Its Expression. *Journal of Bacteriology* **2008**, *190* (22), 7367-7374.

75. Santos, C. N. S.; Xiao, W.; Stephanopoulos, G., Rational, combinatorial, and genomic approaches for engineering L-tyrosine production in *Escherichia coli*. *Proceedings of the National Academy of Sciences* **2012**, *109* (34), 13538-13543.

76. Zhang, H.; Stephanopoulos, G., Engineering *E. coli* for caffeic acid biosynthesis from renewable sugars. *Applied Microbiology and Biotechnology* **2013**, *97* (8), 3333-3341.

77. Rodrigues, J. L.; Araújo, R. G.; Prather, K. L. J.; Kluskens, L. D.; Rodrigues, L. R., Heterologous production of caffeic acid from tyrosine in *Escherichia coli*. *Enzyme and Microbial Technology* **2015**, *71*, 36-44.

78. Lin, Y.; Yan, Y., Biosynthesis of caffeic acid in *Escherichia coli* using its endogenous hydroxylase complex. *Microbial Cell Factories* **2012**, *11* (1), 42.

79. Huang, Q.; Lin, Y.; Yan, Y., Caffeic acid production enhancement by engineering a phenylalanine over-producing *Escherichia coli* strain. *Biotechnology and Bioengineering* **2013**, *110* (12), 3188-3196.

80. Borrás-Linares, I.; Fernández-Arroyo, S.; Arráez-Roman, D.; Palmeros-Suárez, P. A.; Del Val-Díaz, R.; Andrade-González, I.; Fernández-Gutiérrez, A.; Gómez-Leyva, J. F.; Segura-Carretero, A., Characterization of phenolic compounds, anthocyanidin, antioxidant

and antimicrobial activity of 25 varieties of Mexican Roselle (*Hibiscus sabdariffa*). *Industrial Crops and Products* **2015**, 69, 385-394.

81. Ohemeng, K. A.; Schwender, C. F.; Fu, K. P.; Barrett, J. F., DNA gyrase inhibitory and antibacterial activity of some flavones(1). *Bioorganic & Medicinal Chemistry Letters* **1993**, 3 (2), 225-230.

82. Meyer, J. J. M.; Afolayan, A. J.; Taylor, M. B.; Erasmus, D., Antiviral activity of galangin isolated from the aerial parts of *Helichrysum aureonitens*. *Journal of Ethnopharmacology* **1997**, 56 (2), 165-169.

83. Cao, H.; Chen, X.; Jassbi, A. R.; Xiao, J., Microbial biotransformation of bioactive flavonoids. *Biotechnology Advances* **2015**, 33 (1), 214-223.

84. Chanet, A.; Milenkovic, D.; Manach, C.; Mazur, A.; Morand, C., Citrus Flavanones: What Is Their Role in Cardiovascular Protection? *Journal of Agricultural and Food Chemistry* **2012**, 60 (36), 8809-8822.

85. Marín, L.; Gutiérrez-del-Río, I.; Yagüe, P.; Manteca, Á.; Villar, C. J.; Lombó, F., De Novo Biosynthesis of Apigenin, Luteolin, and Eriodictyol in the Actinomycete *Streptomyces albus* and Production Improvement by Feeding and Spore Conditioning. *Frontiers in Microbiology* **2017**, 8 (921).

86. Song, M. C.; Kim, E. J.; Kim, E.; Rathwell, K.; Nam, S.-J.; Yoon, Y. J., Microbial biosynthesis of medicinally important plant secondary metabolites. *Natural Product Reports* **2014**, 31 (11), 1497-1509.

87. Kaneko, M.; Hwang, E. I.; Ohnishi, Y.; Horinouchi, S., Heterologous production of flavanones in *Escherichia coli*: potential for combinatorial biosynthesis of flavonoids in bacteria. *Journal of Industrial Microbiology and Biotechnology* **2003**, 30 (8), 456-461.

88. Miyahisa, I.; Kaneko, M.; Funa, N.; Kawasaki, H.; Kojima, H.; Ohnishi, Y.; Horinouchi, S., Efficient production of (2S)-flavanones by *Escherichia coli* containing an artificial biosynthetic gene cluster. *Applied Microbiology and Biotechnology* **2005**, 68 (4), 498-504.

89. Leonard, E.; Lim, K.-H.; Saw, P.-N.; Koffas, M. A. G., Engineering Central Metabolic Pathways for High-Level Flavonoid Production in *Escherichia coli*. *Applied and Environmental Microbiology* **2007**, 73 (12), 3877-3886.

90. Leonard, E.; Yan, Y.; Fowler, Z. L.; Li, Z.; Lim, C. G.; Lim, K. H.; Koffas, M. A., Strain improvement of recombinant *Escherichia coli* for efficient production of plant flavonoids. *Molecular Pharmaceutics* **2008**, 5 (2), 257-265.

91. Yan, Y.; Huang, L.; Koffas, M. A. G., Biosynthesis of 5-deoxyflavanones in microorganisms. *Biotechnology Journal* **2007**, 2 (10), 1250-1262.

92. Zhu, S.; Wu, J.; Du, G.; Zhou, J.; Chen, J., Efficient Synthesis of Eriodictyol from l-Tyrosine in *Escherichia coli*. *Applied and Environmental Microbiology* **2014**, 80 (10), 3072-3080.

93. Jones, J. A.; Collins, S. M.; Vernacchio, V. R.; Lachance, D. M.; Koffas, M. A. G., Optimization of naringenin and p-coumaric acid hydroxylation using the native *E. coli* hydroxylase complex, HpaBC. *Biotechnology Progress* **2015**, 32 (1), 21-25.

94. Cao, W.; Ma, W.; Zhang, B.; Wang, X.; Chen, K.; Li, Y.; Ouyang, P., Improved pinocembrin production in *Escherichia coli* by engineering fatty acid synthesis. *Journal of Industrial Microbiology & Biotechnology* **2016**, 43 (4), 557-566.

95. Wu, J.; Zhang, X.; Dong, M.; Zhou, J., Stepwise modular pathway engineering of *Escherichia coli* for efficient one-step production of (2S)-pinocembrin. *Journal of Biotechnology* **2016**, *231*, 183-192.
96. Wu, J.; Zhang, X.; Zhou, J.; Dong, M., Efficient biosynthesis of (2S)-pinocembrin from d-glucose by integrating engineering central metabolic pathways with a pH-shift control strategy. *Bioresource Technology* **2016**, *218*, 999-1007.
97. Wu, J.; Du, G.; Zhou, J.; Chen, J., Metabolic engineering of *Escherichia coli* for (2S)-pinocembrin production from glucose by a modular metabolic strategy. *Metabolic Engineering* **2013**, *16*, 48-55.
98. Xu, P.; Ranganathan, S.; Fowler, Z. L.; Maranas, C. D.; Koffas, M. A. G., Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA. *Metabolic Engineering* **2011**, *13* (5), 578-587.
99. Fowler, Z. L.; Gikandi, W. W.; Koffas, M. A. G., Increased Malonyl Coenzyme A Biosynthesis by Tuning the *Escherichia coli* Metabolic Network and Its Application to Flavanone Production. *Applied and Environmental Microbiology* **2009**, *75* (18), 5831-5839.
100. Xu, P.; Vansiri, A.; Bhan, N.; Koffas, M. A. G., ePathBrick: A Synthetic Biology Platform for Engineering Metabolic Pathways in *E. coli*. *American Chemical Society Synthetic Biology* **2012**, *1* (7), 256-266.
101. Cress, B. F.; Toparlak, Ö. D.; Guleria, S.; Lebovich, M.; Stieglitz, J. T.; Englaender, J. A.; Jones, J. A.; Linhardt, R. J.; Koffas, M. A. G., CRISPathBrick: Modular Combinatorial Assembly of Type II-A CRISPR Arrays for dCas9-Mediated Multiplex Transcriptional Repression in *E. coli*. *American Chemical Society Synthetic Biology* **2015**, *4* (9), 987-1000.
102. Raman, S.; Rogers, J. K.; Taylor, N. D.; Church, G. M., Evolution-guided optimization of biosynthetic pathways. *Proceedings of the National Academy of Sciences* **2014**, *111* (50), 17803-17808.
103. Yang, Y.; Lin, Y.; Li, L.; Linhardt, R. J.; Yan, Y., Regulating malonyl-CoA metabolism via synthetic antisense RNAs for enhanced biosynthesis of natural products. *Metabolic Engineering* **2015**, *29*, 217-226.
104. Wu, J.; Zhou, T.; Du, G.; Zhou, J.; Chen, J., Modular Optimization of Heterologous Pathways for De Novo Synthesis of (2S)-Naringenin in *Escherichia coli*. *PLOS ONE* **2014**, *9* (7).
105. Wu, J.; Du, G.; Chen, J.; Zhou, J., Enhancing flavonoid production by systematically tuning the central metabolic pathways based on a CRISPR interference system in *Escherichia coli*. *Scientific Reports* **2015**, *5*.
106. Wu, J.; Yu, O.; Du, G.; Zhou, J.; Chen, J., Fine-Tuning of the Fatty Acid Pathway by Synthetic Antisense RNA for Enhanced (2S)-Naringenin Production from l-Tyrosine in *Escherichia coli*. *Applied and Environmental Microbiology* **2014**, *80* (23), 7283-7292.
107. Ganesan, V.; Li, Z.; Wang, X.; Zhang, H., Heterologous biosynthesis of natural product naringenin by co-culture engineering. *Synthetic and Systems Biotechnology* **2017**, *2* (3), 236-242.
108. Whitaker, W. B., et al., Engineering the biological conversion of methanol to specialty chemicals in *Escherichia coli*. *Metabolic Engineering* **2017**, *39*, 49-59.
109. Yan, Y.; Kohli, A.; Koffas, M. A. G., Biosynthesis of Natural Flavanones in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* **2005**, *71* (9), 5610-5613.

110. Lehka, B. J., et al., Improving heterologous production of phenylpropanoids in *Saccharomyces cerevisiae* by tackling an unwanted side reaction of Tsc13, an endogenous double-bond reductase. *Federation of European Microbiological Societies Yeast Research* **2017**, *17* (1).
111. Rodriguez, A.; Strucko, T.; Stahlhut, S. G.; Kristensen, M.; Svenssen, D. K.; Forster, J.; Nielsen, J.; Borodina, I., Metabolic engineering of yeast for fermentative production of flavonoids. *Bioresource Technology* **2017**, *245*, 1645-1654.
112. Ranganathan, S.; Suthers, P. F.; Maranas, C. D., OptForce: An Optimization Procedure for Identifying All Genetic Manipulations Leading to Targeted Overproductions. *PLOS Computational Biology* **2010**, *6* (4), 1000744-1000755.
113. Wang, H. H.; Isaacs, F. J.; Carr, P. A.; Sun, Z. Z.; Xu, G.; Forest, C. R.; Church, G. M., Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **2009**, *460*, 894-898.
114. Santos, C. N. S., *Combinatorial search strategies for the metabolic engineering of microorganisms*. Massachusetts Institute of Technology, PhD thesis: Cambridge, MA, 2010.
115. Thuan, N. H.; Chaudhary, A. K.; Van Cuong, D.; Cuong, N. X., Engineering co-culture system for production of apigenin in *Escherichia coli*. *Journal of Industrial Microbiology & Biotechnology* **2018**, *45* (3), 175-185.
116. Leonard, E.; Yan, Y.; Lim, K. H.; Koffas, M. A. G., Investigation of Two Distinct Flavone Synthases for Plant-Specific Flavone Biosynthesis in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* **2005**, *71* (12), 8241-8248.
117. Leonard, E.; Chemler, J.; Lim, K. H.; Koffas, M. A. G., Expression of a soluble flavone synthase allows the biosynthesis of phytoestrogen derivatives in *Escherichia coli*. *Applied Microbiology and Biotechnology* **2006**, *70* (1), 85-91.
118. Miyahisa, I.; Funa, N.; Ohnishi, Y.; Martens, S.; Moriguchi, T.; Horinouchi, S., Combinatorial biosynthesis of flavones and flavonols in *Escherichia coli*. *Applied Microbiology and Biotechnology* **2006**, *71* (1), 53-58.
119. Lee, H.; Kim, B. G.; Kim, M.; Ahn, J.-H., Biosynthesis of Two Flavones, Apigenin and Genkwanin, in *Escherichia coli*. *Journal of Microbiology and Biotechnology* **2015**, *25* (9), 1442-1448.
120. Liu, H. L.; Jiang, W. B.; Xie, M. X., Flavonoids: Recent Advances as Anticancer Drugs. *Recent Patents on Anti-Cancer Drug Discovery* **2010**, *5* (2), 152-164.
121. Chen, A. Y.; Chen, Y. C., A review of the dietary flavonoid, kaempferol on human health and cancer chemoprevention. *Food Chemistry* **2013**, *138* (4), 2099-2107.
122. Zheng, L. T.; Ock, J.; Kwon, B.-M.; Suk, K., Suppressive effects of flavonoid fisetin on lipopolysaccharide-induced microglial activation and neurotoxicity. *International Immunopharmacology* **2008**, *8* (3), 484-494.
123. Lukačín, R.; Wellmann, F.; Britsch, L.; Martens, S.; Matern, U., Flavonol synthase from *Citrus unshiu* is a bifunctional dioxygenase. *Phytochemistry* **2003**, *62* (3), 287-292.
124. Naesby, M., et al., Yeast artificial chromosomes employed for random assembly of biosynthetic pathways and production of diverse compounds in *Saccharomyces cerevisiae*. *Microbial Cell Factories* **2009**, *8* (45).
125. Duan, L.; Ding, W.; Liu, X.; Cheng, X.; Cai, J.; Hua, E.; Jiang, H., Biosynthesis and engineering of kaempferol in *Saccharomyces cerevisiae*. *Microbial Cell Factories* **2017**, *16* (1).

126. Leonard, E.; Yan, Y.; Koffas, M. A. G., Functional expression of a P450 flavonoid hydroxylase for the biosynthesis of plant-specific hydroxylated flavonols in *Escherichia coli*. *Metabolic Engineering* **2006**, *8* (2), 172-181.
127. Chemler, J. A.; Fowler, Z. L.; McHugh, K. P.; Koffas, M. A. G., Improving NADPH availability for natural product biosynthesis in *Escherichia coli* by metabolic engineering. *Metabolic Engineering* **2010**, *12* (2), 96-104.
128. Zhao, S.; Jones, J. A.; Lachance, D. M.; Bhan, N.; Khalidi, O.; Venkataraman, S.; Wang, Z.; Koffas, M. A. G., Improvement of catechin production in *Escherichia coli* through combinatorial metabolic engineering. *Metabolic Engineering* **2015**, *28*, 43-53.
129. Li, D.; Wang, P.; Luo, Y.; Zhao, M.; Chen, F., Health benefits of anthocyanins and molecular mechanisms: Update from recent decade. *Critical Reviews in Food Science and Nutrition* **2017**, *57* (8), 1729-1741.
130. Yan, Y.; Chemler, J.; Huang, L.; Martens, S.; Koffas, M. A., Metabolic engineering of anthocyanin biosynthesis in *Escherichia coli*. *Applied and Environmental Microbiology* **2005**, *71* (7), 3617-3623.
131. Yan, Y.; Li, Z.; Koffas, M. A. G., High-yield anthocyanin biosynthesis in engineered *Escherichia coli*. *Biotechnology and Bioengineering* **2008**, *100* (1), 126-140.
132. Lim, C. G.; Wong, L.; Bhan, N.; Dvora, H.; Xu, P.; Venkiteswaran, S.; Koffas, M. A. G., Development of a Recombinant *Escherichia coli* Strain for Overproduction of the Plant Pigment Anthocyanin. *Applied and Environmental Microbiology* **2015**, *81* (18), 6276-6284.
133. Jones, J. A., et al., Complete Biosynthesis of Anthocyanins Using *E. coli* Polycultures. *mBio* **2017**, *8* (3), 00621-17.
134. Eichenberger, M.; Hansson, A.; Fischer, D.; Dürr, L.; Naesby, M., De novo biosynthesis of anthocyanins in *Saccharomyces cerevisiae*. *Federation of European Microbiological Societies Yeast Research* **2018**, *18* (4), 1-13.
135. Pérez-Bonilla, M.; Salido, S.; van Beek, T. A.; Waard, P. d.; Linares-Palomino, P. J.; Sánchez, A.; Altarejos, J., Isolation of antioxidative secoiridoids from olive wood (*Olea europaea* L.) guided by on-line HPLC–DAD–radical scavenging detection. *Food Chemistry* **2011**, *124* (1), 36-41.
136. Kumar, S.; Stecher, G.; Tamura, K., MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* **2016**, *33* (7), 1870-1874.
137. Edgar, R. C., MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **2004**, *32* (5), 1792-1797.
138. Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E., The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **2015**, *10*, 845-858.
139. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera-a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004**, *25* (13), 1605-12.
140. Vickers, C. E.; Bydder, S. F.; Zhou, Y.; Nielsen, L. K., Dual gene expression cassette vectors with antibiotic selection markers for engineering in *Saccharomyces cerevisiae*. *Microbial Cell Factories* **2013**, *12* (1).

141. Daniel Gietz, R.; Woods, R. A., Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods in Enzymology* **2002**, *350*, 87-96.
142. Feng, D. F.; Johnson, M. S.; Doolittle, R. F., Aligning amino acid sequences: Comparison of commonly used methods. *Journal of Molecular Evolution* **1985**, *21* (2), 112-125.
143. Heim, K. E.; Tagliaferro, A. R.; Bobilya, D. J., Flavonoid antioxidants: chemistry, metabolism and structure-activity relationships. *The Journal of Nutritional Biochemistry* **2002**, *13* (10), 572-584.
144. EFSA Panel on Dietetic Products, N.; Allergies, General scientific guidance for stakeholders on health claim applications. *European Food Safety Authority Journal* **2016**, *14* (1), 4367-38.
145. Reboredo-Rodríguez, P.; Valli, E.; Bendini, A.; Di Lecce, G.; Simal-Gándara, J.; Gallina Toschi, T., A widely used spectrophotometric assay to quantify olive oil biophenols according to the health claim 432/2012. *European Journal of Lipid Science and Technology* **2016**, *118* (10), 1593-1599.
146. Harnly, J. M.; Bhagwat, S.; Lin, L.-Z., Profiling methods for the determination of phenolic compounds in foods and dietary supplements. *Analytical and Bioanalytical Chemistry* **2007**, *389* (1), 47-61.
147. Park, S.-H.; Lee, C. W.; Cho, S. M.; Lee, H.; Park, H.; Lee, J.; Lee, J. H., Crystal structure and enzymatic properties of chalcone isomerase from the Antarctic vascular plant *Deschampsia antarctica* Desv. *PLOS ONE* **2018**, *13* (2), 0192415-0192432.
148. Lisec, J.; Schauer, N.; Kopka, J.; Willmitzer, L.; Fernie, A. R., Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nature Protocols* **2006**, *1* (1), 387-396.
149. Fiehn, O.; Kopka, J.; Dörmann, P.; Altmann, T.; Trethewey, R. N.; Willmitzer, L., Metabolite profiling for plant functional genomics. *Nature Biotechnology* **2000**, *18* (11), 1157-1161.
150. Yang, H.; Li, J.; Du, G.; Liu, L., Microbial Production and Molecular Engineering of Industrial Enzymes: Challenges and Strategies. In *Biotechnology of Microbial Enzymes*, Brahmachari, G., Ed. Academic Press: 2017.
151. Cavicchioli, R.; Curmi, P. M. G.; Siddiqui, K. S.; Thomas, T., Proteins from Psychrophiles. *Methods in Microbiology* **2006**, *35*, 395-436.
152. Wang, M.; Bai, J.; Chen, W. N.; Ching, C. B., Metabolomic Profiling of Cellular Responses to Carvedilol Enantiomers in Vascular Smooth Muscle Cells. *PLOS ONE* **2010**, *5* (11), 15441-15449.
153. Jia, H.; Guo, Y.; Zhao, W.; Wang, K., Long-range PCR in next-generation sequencing: comparison of six enzymes and evaluation on the MiSeq sequencer. *Scientific Reports* **2014**, *4*, 5737-5745.
154. Becker, D. M.; Guarente, L., High-efficiency transformation of yeast by electroporation. In *Methods in Enzymology*, Academic Press: 1991; Vol. 194.
155. Forest, J. C.; Wightman, F., Amino Acid Metabolism in Plants. III. Purification and Some Properties of a Multispecific Aminotransferase Isolated from Bushbean Seedlings (*Phaseolus vulgaris* L.). *Canadian Journal of Biochemistry* **1972**, *50* (7), 813-829.
156. Izui, K.; Matsumura, H.; Furumoto, T.; Kai, Y., Phosphoenolpyruvate carboxylase: A New Era of Structural Biology. *Annual Review of Plant Biology* **2004**, *55* (1), 69-84.

157. Zu, X., et al., Chemical Genetics of Acetyl-CoA Carboxylases. *Molecules* **2013**, *18* (2), 1704-1719.
158. Bornaes, C.; Petersen, J. G.; Holmberg, S., Serine and threonine catabolism in *Saccharomyces cerevisiae*: the CHA1 polypeptide is homologous with other serine and threonine dehydratases. *Genetics* **1992**, *131* (3), 531-539.
159. Akhtar, M.; El-Obeid, H. A., Inactivation of serine transhydroxymethylase and threonine aldolase activities. *Biochimica et Biophysica Acta (BBA) - Enzymology* **1972**, *258* (3), 791-799.
160. Abe, I.; Sano, Y.; Takahashi, Y.; Noguchi, H., Site-directed Mutagenesis of Benzalacetone Synthase: The Role of Phe215 in Plant Type III Polyketide Synthases. *Journal of Biological Chemistry* **2003**, *278* (27), 25218-25226.
161. Jez, J. M.; Ferrer, J.-L.; Bowman, M. E.; Austin, M. B.; Schröder, J.; Dixon, R. A.; Noel, J. P., Structure and mechanism of chalcone synthase-like polyketide synthases. *Journal of Industrial Microbiology and Biotechnology* **2001**, *27* (6), 393-398.
162. Abe, I.; Utsumi, Y.; Oguro, S.; Morita, H.; Sano, Y.; Noguchi, H., A Plant Type III Polyketide Synthase that Produces Pentaketide Chromone. *Journal of the American Chemical Society* **2005**, *127* (5), 1362-1363.
163. Jiang, C.; Kim, S. Y.; Suh, D.-Y., Divergent evolution of the thiolase superfamily and chalcone synthase family. *Molecular Phylogenetics and Evolution* **2008**, *49* (3), 691-701.
164. Mohd Azhar, S. H.; Abdulla, R.; Jambo, S. A.; Marbawi, H.; Gansau, J. A.; Mohd Faik, A. A.; Rodrigues, K. F., Yeasts in sustainable bioethanol production: A review. *Biochemistry and Biophysics Reports* **2017**, *10*, 52-61.
165. Bekatorou, A.; Psarianos, C.; Koutinas, A. A., Production of food grade yeasts. *Food Technology & Biotechnology* **2006**, *44* (3), 407-415.
166. Ferreira, I. M. P. L. V. O.; Pinho, O.; Vieira, E.; Tavela, J. G., Brewer's *Saccharomyces* yeast biomass: characteristics and potential applications. *Trends in Food Science & Technology* **2010**, *21* (2), 77-84.
167. Liu, L.; Guan, N.; Li, J.; Shin, H.-d.; Du, G.; Chen, J., Development of GRAS strains for nutraceutical production using systems and synthetic biology approaches: advances and prospects. *Critical Reviews in Biotechnology* **2017**, *37* (2), 139-150.
168. Zhang, W.; Du, G.; Zhou, J.; Chen, J., Regulation of Sensing, Transportation, and Catabolism of Nitrogen Sources in *Saccharomyces cerevisiae*. *Microbiology and Molecular Biology Reviews* **2018**, *82* (1), 00040-17.
169. Lee, J. J.; Cooray, S. T.; Mark, R.; Chen, W. N., Effect of sequential twin screw extrusion and fungal pretreatment to release soluble nutrients from soybean residue for carotenoid production. *Journal of the Science of Food and Agriculture* **2019**, *99* (5), 2646-2650.
170. Holden, H. M.; Rayment, I.; Thoden, J. B., Structure and Function of Enzymes of the Leloir Pathway for Galactose Metabolism. *Journal of Biological Chemistry* **2003**, *278* (45), 43885-43888.
171. Koschwanez, J. H.; Foster, K. R.; Murray, A. W., Sucrose Utilization in Budding Yeast as a Model for the Origin of Undifferentiated Multicellularity. *PLOS Biology* **2011**, *9* (8), 1001122-1001132.

172. Jung, J.-Y.; Yun, H. S.; Lee, J.; Oh, M.-K., Production of 1,2-propanediol from glycerol in *Saccharomyces cerevisiae*. *Journal of Microbiology and Biotechnology* **2011**, *21* (8), 846-853.
173. Gupta, S.; Lee, J. J. L.; Chen, W. N., Analysis of Improved Nutritional Composition of Potential Functional Food (Okara) after Probiotic Solid-State Fermentation. *Journal of Agricultural and Food Chemistry* **2018**, *66* (21), 5373-5381.
174. David, F.; Siewers, V., Advances in yeast genome engineering. *Federation of European Microbiological Societies Yeast Research* **2015**, *15* (1), 1-14.
175. Winkel, B. S. J., Metabolic Channelling in Plants. *Annual Review of Plant Biology* **2004**, *55* (1), 85-107.
176. Zhang, Y., et al., Using unnatural protein fusions to engineer resveratrol biosynthesis in yeast and Mammalian cells. *Journal of the American Chemical Society* **2006**, *128* (40), 13030-13031.
177. Nixon, A. E.; Ostermeier, M.; Benkovic, S. J., Hybrid enzymes: manipulating enzyme design. *Trends in Biotechnology* **1998**, *16* (6), 258-264.
178. Chiang, C.-J.; Lin, L.-J.; Wang, Z. W.; Lee, T.-T.; Chao, Y.-P., Design of a noncovalently linked bifunctional enzyme for whole-cell biotransformation. *Process Biochemistry* **2014**, *49* (7), 1122-1128.
179. Tian, L.; Dixon, R. A., Engineering isoflavone metabolism with an artificial bifunctional enzyme. *Planta* **2006**, *224* (3), 496-507.
180. Xu, J.; Du, W.; Zhao, X.; Zhang, G.; Liu, D., Microbial oil production from various carbon sources and its use for biodiesel preparation. *Biofuels, Bioproducts and Biorefining* **2013**, *7* (1), 65-77.
181. Chi, Z.; Zheng, Y.; Jiang, A.; Chen, S., Lipid Production by Culturing Oleaginous Yeast and Algae with Food Waste and Municipal Wastewater in an Integrated Process. *Applied Biochemistry and Biotechnology* **2011**, *165* (2), 442-453.
182. Mirabella, N.; Castellani, V.; Sala, S., Current options for the valorization of food manufacturing waste: a review. *Journal of Cleaner Production* **2014**, *65*, 28-41.
183. Cheirsilp, B.; Radchabut, S., Use of whey lactose from dairy industry for economical kefir production by *Lactobacillus kefirifaciens* in mixed cultures with yeasts. *New Biotechnology* **2011**, *28* (6), 574-580.
184. Hang, Y. D.; Splittstoesser, D. F.; Landschoot, R. L., Sauerkraut waste: a favorable medium for cultivating yeasts. *Applied microbiology* **1972**, *24* (6), 1007-1008.
185. Ubalua, A., Cassava wastes: Treatment options and value addition alternatives. *African journal of biotechnology* **2007**, *6* (18), 2065-2073.
186. Dharumadurai, D.; Subramaniyan, L.; Subhasish, S.; Nooruddin, T.; Annamalai, P., Production of single cell protein from pineapple waste using yeast. *Innovative Romanian Food Biotechnology* **2011**, *8*, 26-32.
187. Nikhil, V., *Process standardization for the production of Baker's yeast from mango sones*. College of Horticulture, MSc thesis: Vellanikkara, 2013.
188. Zhou, Y.-M.; Chen, Y.-P.; Guo, J.-S.; Shen, Y.; Yan, P.; Yang, J.-X., Recycling of orange waste for single cell protein production and the synergistic and antagonistic effects on production quality. *Journal of Cleaner Production* **2019**, *213*, 384-392.
189. Win, S. S.; Impoolsup, A.; Noomhorm, A., Growth kinetics of *Saccharomyces cerevisiae* in batch and fedbatch cultivation using sugarcane molasses and glucose syrup from cassava starch. *Journal of Industrial Microbiology* **1996**, *16* (2), 117-123.

190. Jin, B.; Yin, P.; Ma, Y.; Zhao, L., Production of lactic acid and fungal biomass by *Rhizopus* fungi from food processing waste streams. *Journal of Industrial Microbiology and Biotechnology* **2005**, *32* (11), 678-686.
191. Soccol, C. R.; Dalmas Neto, C. J.; Soccol, V. T.; Sydney, E. B.; da Costa, E. S. F.; Medeiros, A. B. P.; Vandenberghe, L. P. d. S., Pilot scale biodiesel production from microbial oil of *Rhodospiridium toruloides* DEBB 5533 using sugarcane juice: Performance in diesel engine and preliminary economic study. *Bioresource Technology* **2017**, *223*, 259-268.
192. Nofemele, Z.; Shukla, P.; Trussler, A.; Permaul, K.; Singh, S., Improvement of ethanol production from sugarcane molasses through enhanced nutrient supplementation using *Saccharomyces cerevisiae*. *Journal of Brewing and Distilling* **2012**, *3* (2), 29-35.
193. Chandel, A. K.; Lakshmi Narasu, M.; Chandrasekhar, G.; Manikyam, A.; Venkateswar Rao, L., Use of *Saccharum spontaneum* (wild sugarcane) as biomaterial for cell immobilization and modulated ethanol production by thermotolerant *Saccharomyces cerevisiae* VS3. *Bioresource Technology* **2009**, *100* (8), 2404-2410.